Universidade Federal do Piauí
Centro de Ciências da Natureza
Pós-Graduação em Matemática
Mestrado em Matemática

Proximal point method for solving quasi-equilibrium problems with application to economics

Francisco Deyvid de Sousa Andrade

Teresina - 2025

Francisco Deyvid de Sousa Andrade

Dissertação de Mestrado:

Proximal point method for solving quasi-equilibrium problems with application to economics

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:
Prof. Dr. João Carlos de Oliveira Souza

Teresina - 2025

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Proximal point method for solving quasi-equilibrium problems with application to economics

Francisco Deyvid de Sousa Andrade

Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Dissertação aprovada em 31 de julho de 2025.

Banca Examinadora:

Documento assinado digitalmente
JOAO CARLOS DE OLIVEIRA SOUZA
Data: 15/08/2025 12:48:28-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. João Carlos de Oliveira Souza – Orientador

Documento assinado digitalmente
JURANDIR DE OLIVEIRA LOPES
Data: 14/08/2025 16:15:57-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Jurandir de Oliveira Lopes - UFPI

Documento assinado digitalmente
PEDRO JORGE SOUSA DOS SANTOS
Data: 15/08/2025 08:16:05-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Pedro Jorge Sousa dos Santos - UFDPar

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Sistema de Bibliotecas UFPI - SIBi/UFPI
Biblioteca Setorial do CCN

A553p Andrade, Francisco Deyvid de Sousa.
Proximal point method for solving quasi-equillibrium problems with application to economics / Francisco Deyvid de Sousa Andrade. -- 2025.
93 f.

Dissertação (Mestrado) - Universidade Federal do Piauí, Centro de Ciências da Natureza, Programa de Pós-Graduação em Matemática, Teresina, 2025.
"Orientador: Prof. Dr. João Carlos de Oliveira Sousa".

1. Método do Ponto Proximal. 2. Problemas de quase-equilíbrio. 3. Distância de Bregman. I. Sousa, João Carlos de Oliveira. II. Titulo.

CDD 516.36

Bibliotecária: Caryne Maria da Silva Gomes - CRB3/1461

Dedico este trablaho a todos que eu amo.

Agradecimentos

Agradeço aos meus pais Domingos de Sousa Andrade e Rosilene de Sousa Barros e a minha irmã Maria dos Remédios de Sousa Barros por sempre me apoiarem e me incentivarem, além de todo o carinho do mundo.

Agradeço a todos os meus amigos, em especial meus amigos do 404: Abel, Cheila, Miranda e Marcelo por sempre estarem comigo e pelos bons momentos juntos. Agradeço também em especial aos meus amigos Adriano e Natali pelo companherismo, pela grande amizade e pelos conselhos duvidosos. Juntamente agradeço aos meus amigos Vinição, Pablo, Euclides Jr. e aos dois Carlos Eduardos.

Agradeço em especial a minha namorada Luana Sousa pelo amor, o carinho e todo o apoio que tem me dado todo esse tempo.

Agradeço ao meu orientador professor Dr. João Carlos de Oliveira Sousa por ter aceitado me orientar e por tudo que fez por mim. Agradeço aos professores Dr. Jurandir de Oliveira Lopes e Dr. Pedro Jorge Sousa Santos por terem aceitado participar da banca.

Agradeço a todos os professores do departamento de matématica da UFPI em geral, em especial aqueles que eu gosto.

E agradeço a FAPEPI pelo apoio financeiro.

"*Enquanto esperamos pela vida, a vida passa*".

Sêneca.

Resumo

Neste trabalho propomos um estudo sobre o método do ponto proximal para resolver problemas de quasi-equilíbrio em duas versões: exata e inexata. A versão exata, a qual é uma generalização para o contexto dos problemas de quasi-equilíbrio do método do ponto proximal para problemas de equilibrio; e a versão inexata, é uma extensão da versão exata, onde é considerada uma perturbação controlada da função parametrizada. Para isso iniciaremos analizando uma versão método do ponto proximal para problemas de equilíbrio utilizando distancias de Bregman e uma versão alternativa com uma regulatização do tipo Bregman. Além disso, evidenciamos a eficiência do algoritimo realizando experimentos numéricos. Por fim, mostraremos uma aplicação dos conceitos e resultados de quasi-equilíbrio aplicados no modelo de duopólio de Cournot.

Palavras chave: Método do Ponto proximal; Problemas de equilibrio; Problemas de quase-equilibrio; Distâncias de Bregman, Modelo do duopólio de Cournot.

Abstract

In this work, we propose a study on the proximal point method to solve quasi-equilibrium problems in two versions: exact and inexact. The exact version, which is a generalization to the context of quasi-equilibrium problems of the proximal point method for equilibrium problems; and the inexact version, which is an extension of the exact version, where a controlled perturbation of the parameterized function is considered. For this, we will start by analyzing a version of the proximal point method for equilibrium problems using Bregman distances and an alternative version with Bregman-type regularization. Furthermore, we demonstrate the efficiency of the algorithm by conducting numerical experiments. Finally, we will show an application of the concepts and results of quasi-equilibrium applied in the Cournot duopoly model.

Key words: Proximal point method; Equilibrium problems; Quasi-equilibrium problems; Bregman distance, Cournot duopoly model.

Contents

Introduction

In this work, we will study a quasi-equilibrium problem (we will denote it by QEP) that consist in find $x^* \in C(x^*)$ in such a way that given a closed, convex, and non-empty subset $X \subset \mathbb{R}^n$, and given a function $f : X \times X \to \mathbb{R}$, one has

$$f(x^*, y) \geqslant 0, \qquad \forall \;\; y \in C(x^*);$$

where $C : X \rightrightarrows X$ is a multi-valued map that associates to each $x \in X$ a closed, convex, and nonempty subset $C(x) \subset X$. The function f satisfies some properties:

$(P1):\;\; f(x, x) = 0, \;\; \forall\; x \in X;$

$(P2):\;\; f(\cdot, y)$ is upper semicontinuous for all $y \in X$;

$(P3):\;\; f(x, \cdot)$ is lower semicontinuous and convex for all $x \in X$.

If f satisfies the condition $(P1)$, it is called bifunction. Also, we will consider some other properties that can be seen in Assumptions 1 and Assumptions 2 in Chapter 3 and Chapter 4, respectively.

To this end, we start to seen the definition of the equilibrium problem (denoted by EP), which consists of the following: given a subset $X \subset \mathbb{R}^n$ and a function $f : X \times X \to \mathbb{R}$, the aim is to find $x^* \in X$, such that

$$f(x^*, y) \geqslant 0, \quad \forall\; y \in X.$$

Related to this, there exists a so-called equilibrium problem dual, where consists in finding a $y^* \in X$, such that

$$f(x, y^*) \leqslant 0, \quad \forall\; x \in X.$$

Hence, note that when $C(x) \equiv X$ for all $x \in X$, we have that a quasi-equilibrium problem becomes an equilibrium problems, i.e., QEP is a generalization to the equilibrium problems, since there are some problems that can be modeled as a quasi-equilibrium problem, but cannot be modeled as an equilibrium problem, see for instance, Facchinei et al. [11], You and Li [42] and Han et al. [13].

There exist several methods to solve a QEP(or, in particular,EP), and in this work we study the proximal point method (denoted by PPM), which was first popularized, in its classical version, by Rockafellar [33] for the context of monotone operators. Iusem and Sosa [20] shown a version for the context of equilibrium problems, where they solve an equilibrium problem for a regularized function $\mathsf{EP}(f_k, X)$ given by

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

where $\{\gamma_k\}$ is a bounded positive auxiliary sequence of parameters. Hence, in each iteration, it is solved a regularized equilibrium problem (called subproblem) and it is proved that it approximates the solution of the original problem.

In this work, we will see the generalization of this method that was proposed by Burachik and Kassay in [5], where it uses Bregman distances instead of Euclidean distances in the regularization of the function, that is given by

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

where $\{\gamma_k\}$ is a bounded and positive sequence of parameters, and φ is a Bregman function.

Moreover, still in the context of equilibrium problems, we will see a variation of the PPM that is a separable Bregman regularization and has been proposed by Cruz Neto et al. [2], consist in regularizing the function in the following way:

$$f_k(x, y) = f(x, y) + \gamma_k (D_\varphi(y, x^k) - D_\varphi(x, x^k)).$$

One of the first extension of the PPM for quasi-equilibrium was proposed by Santos and Souza in [34], which is the main result analyzed in our work adding to its inexact version that was proposed by E. Junior et al. [18]. In the inexact version we study the convergence of method using the following regularized function:

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

For this regularization, we will consider two estimates for e^k:

$$(\mathsf{E1}) : \|e^k\| \leqslant \|x^{k+1} - x^k\|;$$

$$(\mathsf{E2}) : \|x^{k+1} - x^k - e^k\| \leqslant \max\{\|e^k\|, \|x^{k+1} - x^k\|\}.$$

Also, we will present an alternative method that is an extension of the method proposed by Konnov [23] for the context of quasi-equilibrium problems, which considers the point $x^{k+1} \in X$ a point in the neighborhood of the solution x^*, not necessarily on $C(x^k)$, and that is controlled by a summable sequence of error $\{\varepsilon_k\}$.

Currently, there are several works in the area of equilibrium and quasi-equilibrium problems. In turn, it has attracted the attention of many researchers due to its wide applicability in problems in diverse areas, such as economics, computing and mathematics. In mathematics, it can highlight applicability in vector and scalar optimization problems, variational inequality and Nash equilibrium problems; see for instance, Muu and Quoc [26] and Facchinei et al. [11]. As cited, QEP has many application in economics; in this work, we will apply the quasi-equilibrium problem to the Cournot duopoly model as proposed [18].

This work was separated into six chapters; in the first one, we will show basic concepts that will be necessary along of the text. In the second one we will show the proximal point method in its classical version and the Bregman version. In the Chapter 3, it is presented concepts and results about equilibrium problems. In Chapter 4, we will introduce quasi-equilibrium problems and the proximal point method in exact and inexact versions and we will show numerical experiments. Finally, in the last chapter, we will be shown an application of the quasi-equilibrium problems to the Cournot Model.

Chapter 1

Basic Concepts

In this chapter, we will see some basic results that will be necessary along this work. We will introduce some basic concepts about convex analysis that can be seen in Izmailov and Solodov [21], and next some results about Bregman distances that have as their base the works Chen and Teboulle [4] and Solodov and Svaiter [36], and finally a few concepts about Fejér convergence. All concepts here are wildly used in literature, and therefore we do not present their proof.

1.1 Convex Analysis

.

In this section, it is presented concepts about convex analysis. See [21] for more details.

First of all, consider the problem:

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

where $f : \mathbb{R}^n \to \mathbb{R}$. Next we will study the conditions that must be satisfied when a given $\bar{x} \in \mathbb{R}^n$ is a local minimizer of the problem above. This type of condition is called an optimality necessary condition.

Theorem 1 (Necessary condition of first order)**.** Suppose that the function $f : \mathbb{R}^n \to \mathbb{R}$ is smooth at the point $\bar{x} \in \mathbb{R}^n$. If $\bar{x}$ is a local minimizer unrestricted of f, then $\nabla f(\bar{x}) = 0$.

Proof. See [21, Theorem 1.3.1.]. $\square$

Theorem 2 (Weierstrass Theorem)**.** Let $D \subset \mathbb{R}^n$ a compact and nonempty set, and let $f : D \to \mathbb{R}$ be a continuous function. Then, the issue of minimize and maximize f on D has global solutions.

Proof. See [21, Theorem 1.2.1.]. □

Following we will present some concepts about convex sets and functions.

1.1.1 Convex sets and functions

Definition 1. A set $D \subseteq \mathbb{R}^n$ is said to be convex if for all elements $x, y \in D$ and $\alpha \in [0, 1]$, is valid that $\alpha x + (1 - \alpha)y \in D$.

The point $\alpha x + (1 - \alpha)y \in D$, where $\alpha \in [0, 1]$, is called a *convex combination* of x and y (with parameter α).

The empty set, all space $\mathbb{R}^n$ and all set that has only one element are trivially convex.

Definition 2. Let $D \subseteq \mathbb{R}^n$ be a convex set. A function $f : D \to \mathbb{R}$ is said to be *convex* if for any $x, y \in D$ and $\alpha \in [0, 1]$, it holds

$$f(\alpha x + (1 - \alpha)y) \leqslant \alpha f(x) + (1 - \alpha)f(y). \tag{1.1}$$

When inequality is strict in (1.1) for any $x, y \in D$ and $\alpha \in (0, 1)$, then f is called *strictly convex*.

A function is said to be strongly convex with module $\gamma > 0$ when for any $x \in D, y \in D$ and $\alpha \in (0, 1)$, one has

$$f(\alpha x + (1 - \alpha)y \leqslant \alpha f(x) + (1 - \alpha)f(y) - \alpha\gamma(1 - \alpha)\|x - y\|^2.$$

Note that when the function is strongly convex with module $\gamma > 0$, then it is a strictly convex function and so is convex. However, there exist functions that are convex but not strongly convex; for instance, $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = e^x$ is strictly convex and not strongly convex.

Definition 3. (Epigraph) The epigraph of a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is a set

$$E_f := \{(x, c) \in D \times \mathbb{R} | f(x) \leqslant c\}.$$

The relation between convexity of the epigraph and convexity of the functions is given by following theorem.

Theorem 3. Let $D \subset \mathbb{R}^n$ a convex set. A function $f : D \to \mathbb{R}$ is convex on D if and only if, the epigraph of f is convex set on $\mathbb{R}^n \times \mathbb{R}$.

Proof. See [21, Theorem 3.1.15]. □

Theorem 4. A set $D \subseteq \mathbb{R}^n$ is convex if, and only if, for any $p \in \mathbb{N}, x^i \in D$ and $\alpha_i \in [0,1], i = 1, 2, ..., p$ such that $\sum_{i=1}^{p} \alpha_i = 1$, the convex combination $\sum_{i=1}^{p} \alpha_i x^i$ belongs to D.

Proof. See [21, Theorem 3.2.8]. □

The importance of the convexity is evidenced in the following result.

Theorem 5 (Minimization convex theorem)**.** Let $D \subseteq \mathbb{R}^n$ be a convex set and $f : D \to \mathbb{R}$ a convex function in D. Then every local minimizer in the problem: $\min f$ subject to $x \in D$, is a global minimizer. Also, a set of minimizers is convex. And, if f is strictly convex, it doesn't have more than one minimizer.

Proof. See [21, Theorem 3.1.6]. □

Definition 4. (Level set): The *level set* of a function $f : D \to \mathbb{R}$ associated to $c \in \mathbb{R}$ is a set given by

$$L_{f,D}(c) = \{x \in D |\, f(x) \leqslant c\}.$$

Definition 5. Let $D \in \mathbb{R}^n$ be a convex set. We say that $f : D \to \mathbb{R}$ is *quasi-convex* in D when the level sets $L_{f,D}(c)$ are convex for all $c \in \mathbb{R}$.

All convex functions are quasi-convex functions, and it may be seen by the following result. On the other hand, the function $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$ is quasi-convex because all level sets are convex, but it is not convex.

Theorem 6. Suppose that a set $D \subset \mathbb{R}^n$ is convex and a function $f : D \to \mathbb{R}$ is convex on D. Then, the level set $L_{f,D}(c) = \{x \in D;\ f(x) \leqslant c\}$ is convex for every $c \in \mathbb{R}$.

Proof. See [21, Theorem 3.4.10]. □

Definition 6. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. f is said to be locally Lipschitz continuous if there exists $\bar{x} \in \mathbb{R}^n$ and f is Lipschitz continuous in a neighborhood V of $\bar{x}$, i.e., there exists $\gamma \in \mathbb{R}$ such that,

$$|f(x) - f(y)| \leqslant \gamma \|x - y\|, \quad \forall\ x, y \in V$$

γ is called a Lipschitz constant.

Theorem 7. Let $D \subseteq \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$ a convex function in D. Then f is locally Lipschitz-continuous in D. In particular, f is continuous in D.

Proof. See [21, Theorem 3.4.17]. □

Definition 7. Let $D \subseteq \mathbb{R}^n$ be a convex set and $\bar{x} \in D$. The *normal cone* (cone of normal directions) at a point $\bar{x}$ in relation to D is

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

The (orthogonal) projection of a point $y \in \mathbb{R}^n$ onto a set $D \subset \mathbb{R}^n$ is the global solution of the problem

$$\min_{x \in D} \|x - y\|.$$

In other words, the projection of y onto D is a point in D that is closer to y.

Next we will show a famous projection theorem.

Theorem 8. Let $D \subset \mathbb{R}^n$ be a convex and closed set. Then for all $x \in \mathbb{R}^n$, the projection of x onto X, denoted by $P_D(x)$, exists and it is unique. Moreover, $\bar{x} = P_D(x)$ if and only if

$$x \in D, \quad \langle x - \bar{x}, y - \bar{x} \rangle \leqslant 0, \quad \forall\, y \in D,$$

or equivalently, $x - \bar{x} \in \mathcal{N}_D(\bar{x})$.

Proof. See [21, Theorem 3.2.32]. □

The following theorem provides a characterization of convex differentiable functions.

Theorem 9. Let $\Omega \subseteq \mathbb{R}^n$ be an open and convex set and $f : \Omega \to \mathbb{R}$ a differentiable function in Ω. Then the following properties are equivalents:

(a) f is convex in Ω.

(b) For every $x \in \Omega$, and every $y \in \Omega$

$$f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle.$$

(c) For every $x \subset \Omega$, and every $x \in \Omega$

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geqslant 0.$$

When f is twice differentiable in Ω, the above properties are also equivalent to

(d) The Hessian matrix of f is positive semi-defined at a every point of Ω :

$$\langle \nabla^2 f(x) d, d \rangle \geqslant 0, \quad \forall\, x \in \Omega, \forall\, d \in \mathbb{R}^n.$$

Proof. See [21, Theorem 3.4.30]. □

Theorem 10. (Sufficient condition of first order) Let $D \subseteq \mathbb{R}^n$ be a convex set, and $f : D \to \mathbb{R}$ is a differentiable function at a point $\bar{x} \in D$. If $\bar{x}$ is a local minimizer of f on D, then

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geqslant 0, \quad \forall x \in D,$$

or equivalently,

$$-\nabla f(\bar{x}) \in \mathcal{N}_D(\bar{x}).$$

Proof. See [21, Theorem 3.1.12]. □

Theorem 11. (Necessary and sufficient conditions for convex minimization) Let $D \subset \mathbb{R}^n$ be a convex set and $f : \Omega \to \mathbb{R}$ be a convex and differentiable function on the open set Ω containing D. Then, $\bar{x}$ is a minimizer of f in D if and only if

$$\bar{x} \in D, \;\; \langle \nabla f(\bar{x}), x - \bar{x} \rangle \geqslant 0, \;\; \forall\, x \in D, \tag{1.2}$$

or equivalently,

$$-\nabla f(\bar{x}) \in \mathcal{N}_D(\bar{x}).$$

Moreover, the condition (1.2) is equivalent to

$$\bar{x} \in D, \quad \langle \nabla f(x), x - \bar{x} \rangle \geqslant 0, \quad \forall\, x \in D.$$

If D is closed, (1.2) and (1.3) are also equivalent to the following condition:

$$\bar{x} = P_D(\bar{x}) - \alpha \nabla f(\bar{x})$$

for some $\alpha > 0$. Where $P_D(\bar{x})$ is the orthogonal projection of $\bar{x}$ onto D.

Proof. See [21, Theorem 3.4.36]. □

1.1.2 Subdifferential

Definition 8. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. We say that $y \in \mathbb{R}^n$ is a *subgradient* of f at a point $x \in \mathbb{R}^n$ if

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

A set of all the subgradients of f at a point x is called the *subdifferential* of f, denoted to $\partial f(x)$.

In other words, the subgradient define a linear approximation of f whose the graphic stay below of the graphic of f and the value coincide at the point x.

Remark 1. We have the following definitions equivalent to the subdifferential:

$$\begin{aligned}\partial f(x) &= \{y \in \mathbb{R}^n | \ \ f(z) \geqslant f(x) + \langle y, z - x\rangle, \ \ \forall\, z \in \mathbb{R}^n\} \\ &= \{y \in \mathbb{R}^n | \ \ f'(x; d) \geqslant d, \ \ \forall\ d \in \mathbb{R}^n\}.\end{aligned}$$

Where, $f'(x : d)$ is the directional derivative, i.e., $f'(x : d) = \lim\limits_{\alpha \to 0+} \dfrac{f(x + \alpha d) - f(x)}{\alpha}$.

Theorem 12. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then for all $x \in \mathbb{R}^n$, the set $\partial f(x)$ is convex, compact and nonempty. Moreover, for all $d \in \mathbb{R}^n$, one has

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

Proof. See [21, Theorem 3.4.52]. □

Next, we will see few important rules for the subdifferential calculus.

Proposition 1. Let $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, ..., p$, be convex functions. Then

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

Proof. See [21, Theorem 3.4.59]. □

Proposition 2. A convex function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point $x \in \mathbb{R}^n$ if, and only if, the set $\partial f(x)$ contains a unique element, in this case $\partial f(x) = \{\nabla f(x)\}$.

Proof. See [21, Theorem 3.4.53]. □

The importance of the subdifferential in the theory of optimization could be seen in the following result.

Theorem 13. (Optimally condition for minimizer a convex function on a convex set)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $D \subset \mathbb{R}^n$ a convex set. Then $\bar{x} \in \mathbb{R}^n$ is a minimizer of f in D if, and only if,

$$\exists \ \ y \in \partial f(\bar{x}) \text{ such that } \langle y, x - \bar{x}\rangle \geqslant 0, \ \ \forall \ \ x \in D,$$

or, equivalently,

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

In particular, $\bar{x}$ is a minimizer of f in $\mathbb{R}^n$ if, and only if, $0 \in \partial f(\bar{x})$.

Proof. See [21, Theorem 3.4.54]. $\square$

Here, we will present a simple example of the subdifferential calculus.

Example 1. Let $f : \mathbb{R} \to \mathbb{R}, f(x) = |x|$. At point $\bar{x} = 0$ is the single minimizer unrestricted of f. It is easy to see that

$$\partial f(x) = \begin{cases} -1, & \text{if} \quad x < 0, \\ [-1, 1], & \text{if} \quad x = 0, \\ 1, & \text{if} \quad x > 0. \end{cases}$$

In particular, $0 \in \partial f(\bar{x})$ and $0 \notin \partial f(x)$ for all $x \neq \bar{x}$.

Next, we will cite a important proposition that will be used in your work.

Proposition 3. Let $x, y, z \in \mathbb{R}^n$ we have that

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

Proof. Indeed,

$$\|y - z\|^2 = \langle y - z, y - z \rangle, \|x - z\|^2 = \langle x - z, x - z \rangle \quad \text{and} \quad \|x - y\|^2 = \langle x - y, x - y \rangle,$$

hence,

$$\begin{aligned} \|y - z\|^2 - \|x - z\|^2 - \|x - y\|^2 &= \langle y - z, y - z \rangle - \langle x - z, x - z \rangle - \langle x - y, x - y \rangle \\ &= \langle y - z, y \rangle - \langle y - z, z \rangle - \langle x - z, x \rangle + \langle x - z, z \rangle - \langle x - y, x \rangle + \langle x - y, y \rangle \\ &= \langle x - z, y - x \rangle + \langle x - z, y - x \rangle \\ &= 2\langle x - z, y - x \rangle. \end{aligned}$$

$\square$

1.2 Bregman Distances

In this section we will present some results about Bregman distances, that will be used in the next chapters. The results here cited can be seen in [4] and [36] for more details.

Definition 9. Let $S \subset \mathbb{R}^n$ be an open and convex set, and $\bar{S}$ its closure. Consider $\varphi : \bar{S} \to \mathbb{R}$ and $D_\varphi : \bar{S} \times S \to \mathbb{R}$ given by

$$D_\varphi(x, y) = \varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y \rangle.$$

φ is called Bregman function, and D_φ is the Bregman distance induced by φ if it satisfies the following conditions:

(B1) φ is continuously differentiable on S.

(B2) φ is strictly convex and continuous on $\bar{S}$.

(B3) For all $\delta \in \mathbb{R}$, the partial level sets given by

$$\Gamma_1(y, \delta) = \{x \in \bar{S} : D_\varphi(x, y) \leqslant \delta\},$$

$$\Gamma_2(x, \delta) = \{y \in \bar{S} : D_\varphi(x, y) \leqslant \delta\}$$

are bounded for all $y \in S$ and all $x \in \bar{S}$, respectively.

(B4) If $\{y^k\} \subset S$ converges to y^*, then $D_\varphi(y^*, y^k)$ converges to 0.

(B5) If $\{x^k\} \subset \bar{S}$ and $\{y^k\} \subset S$ are sequences such that $\{x^k\}$ is bounded, $\lim_{k\to\infty} y^k = y^*$ and $\lim D_\varphi(x^k, y^k) = 0$, then $\lim_{k\to\infty} x^k = y^*$.

Now we introduce two subclasses of Bregman functions:

A Bregman function φ is said to be *boundary coercive* if:

(B6) If $\{y^k\} \subset S$ is such that $\lim_{k\to\infty} y^k = y \in \partial S$, then $\lim_{k\to\infty}\langle \nabla\varphi(y^k), x - y^k\rangle = -\infty$, for all $x \in S$.

A Bregman function φ is said to be *zone coercive* if:

(B7) For every $y \in \mathbb{R}^n$, there exists $x \in S$ such that $\nabla\varphi(x) = y$.

Moreover, we will consider the following condition,

(B8) If $\{x^k\}, \{y^k\} \in S$ are sequences bounded, such that $\lim_{k\to\infty} \|x^k - y^k\| = 0$, then $\lim_{k\to\infty} \|\nabla\varphi(x^k) - \nabla\varphi(y^k)\| = 0$.

Remark 2. $D_\varphi(x, y) \geqslant 0$ for all $x \in \bar{S}$ and $y \in S$. And $D_\varphi(x, y) = 0$ if and only if $x = y$. This result comes from definition and of the fact φ is strictly convex.

S is called the *zone* of φ. And it has been proved in [15] that when $S = \mathbb{R}^n$ a sufficient condition for a convex and differentiable function φ to be a Bregman function is $\lim_{\|x\|\to\infty} \dfrac{\varphi(x)}{\|x\|} = \infty$.

Definition 10. A Bregman function φ is said to be separable when $\varphi(x) = \sum_{i=1}^{n} \varphi_i(x)$ with φ_i scalar functions. In this case, D_φ is said to be separable.

The Bregman distance can be seen as a generalization of the Euclidian distance, and this is evidenced in the following example.

Example 2. If $S = \mathbb{R}^n$ and $\varphi(x) = \langle Ax, x\rangle$ with $A \in \mathbb{R}^{n\times n}$ symmetric and positive definite, and its respective Bregman distance is $D_\varphi(x, y) = \langle A(x - y), x - y\rangle$. Indeed

$$\begin{aligned}
D_\varphi(x, y) &= \varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y\rangle \\
&= \langle Ax, x\rangle - \langle Ay, y\rangle - \langle 2Ay, x - y\rangle \\
&= \langle Ax, x\rangle - \langle Ay, y\rangle - 2(\langle Ay, x\rangle - \langle Ay, y\rangle) \\
&= \langle Ax, x\rangle + \langle Ay, y\rangle - 2\langle Ay, x\rangle \\
&= \langle A(x - y), x - y\rangle.
\end{aligned}$$

Remark 3. Note that when $A = I$, we have that $\varphi(x) = \|x\|^2$ and $D_\varphi(x, y) = \|x - y\|^2$.

Similarly as there exists orthogonal projection for Euclidean distances, there exists a called Bregman projection, that one can see in the following proposition.

Proposition 4. If φ satisfies $(B1) - (B3)$ and $C \subset \mathbb{R}^n$ is closed and convex, then for all $\bar{x} \in \mathbb{R}^n$ there exists a unique solution $\hat{x}$ of the problem $\min_{x\in C} D_\varphi(x, \bar{x})$ which satisfies $\langle \varphi(\bar{x}) - \varphi(\hat{x}), y - \hat{x}\rangle \leqslant 0$ for all $y \in C$.

The point $\hat{x}$ is said to be the *Bregman projection* of $\bar{x}$ onto C.

Proof. See [29, p. 70]. □

The proposition below present important results that will be used in our work, in the next chapter.

Proposition 5. If $\varphi : S \to \mathbb{R}$ is a Bregman function with zone S, then

i) (Three points equality) $D_\varphi(x, y) - D_\varphi(x, z) - D_\varphi(z, y) = \langle \nabla\varphi(y) - \nabla\varphi(z), z - x\rangle$ for all $x \in \bar{S}$, for all $z \in S$.

ii) $\nabla_1 D_\varphi(x,y) = \nabla\varphi(x) - \nabla\varphi(y)$ for all $x, y \in S$.

iii) $D_\varphi(\cdot, y)$ is strictly convex for all $y \in S$.

iv)(Four points equality) For any $x, z \in S$ and $y, s \in \bar{S}$ it holds that

$$D_\varphi(s,z) - D_\varphi(s,x) - D_\varphi(y,z) + D_\varphi(y,x) = \langle \nabla\varphi(x) - \nabla\varphi(z), s - y\rangle.$$

Proof. It follows directly by Definition 9 and properties of the Bregman function. □

1.3 Fejér and quasi-Fejér Convergence

Now, we recall some important results concerning sequences in equilibrium and quasi-equilibrium problems.

Definition 11. A sequence $\{z^k\}$ is Fejér convergent to a nonempty set U if, for all $k \in \mathbb{N}$,

$$\|z^{k+1} - x\| \leqslant \|z^k - x\|, \quad \forall x \in U.$$

Definition 12. A sequence $\{z^k\}$ is said to be quasi-Fejér convergent to a set U if, for each $u \in U$, there exists a non-negative sequence $\{\varepsilon_k\}$ with $\sum_{k=0}^{\infty} \varepsilon_k < \infty$ such that

$$\|z^{k+1} - u\|^2 \leqslant \|z^k - u\|^2 + \varepsilon_k, \quad \forall k \in \mathbb{N}, \ \ \forall u \in U.$$

An important property of sequences that are Fejér and quasi-Fejér convergent is that they are bounded. Furthermore, in some cases, it is possible to show that all sequences converges. This fact is evidenced in the next results.

Lemma 1. Let U be a nonempty set, and assume that $\{z^k\}$ is Fejér convergent to U. Then $\{z^k\}$ is bounded. Moreover, if a cluster point z of $\{z^k\}$ belongs to U, then $\lim_{k\to\infty} z^k = z$.

Proof. See [38, Proposition 2.1]. □

Lemma 2. Let U be a nonempty set, and assume that $\{z^k\}$ is quasi-Fejér convergent to U. Then $\{z^k\}$ is bounded. Moreover, if a cluster point z of $\{z^k\}$ belongs to U, then $\lim_{k\to\infty} z^k = z$.

Proof. See [32, Proposition 1]. □

Lemma 3. Let $\{\nu_k\}, \{\gamma_k\}$ and $\{\beta_k\}$ be non-negative sequences for real numbers satisfying $\nu_{k+1} \leqslant (1+\gamma_k)\nu_k + \beta_k$, such that $\sum_{k=0}^{+\infty} \gamma_k < +\infty, \sum_{k=0}^{+\infty} \beta_k < +\infty$. Then, the sequence $\{\nu_k\}$ converges.

Proof. See [18, Lemma 2]. □

Chapter 2

Proximal Point Method

Throughout this chapter let us see the proximal point algorithm, that was popularized by Rockafellar [33]. We will use this concept to solve optimization problems of the type

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

The method, presented in Algorithm 1, create a sequence $\{x^k\}$ that under reasonable hypotheses converges to a minimizer of the function f. Also, we will see an extension of this method using Bregman distance, based on Chen [4].

2.1 The proximal point for optimization in $\mathbb{R}^n$

Algorithm 1 Proximal Point Method

1. Choose an initial $x^0 \in \mathbb{R}^n$ and a sequence of bounded and positive parameters $\{\gamma_k\}$;
2. The sequence $\{x^k\}$ is generated in the following way: given x^k, compute

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

3. If $x^k = x^{k+1}$, stop. Otherwise, take $k = k + 1$ and return to the previous step.

Theorem 14. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and continuously differentiable. Assume that the U set of minimizers of f on $\mathbb{R}^n$ is nonempty. Then the sequence $\{x^k\}$ generates for Algorithm 1 converges to a point $x^* \in U$.

Proof. First of all, we will show that the algorithm is well-defined.

Let $f_k(x) = f(x) + \gamma_k\|x - x^k\|^2$ be. Since f attains a minimum, then f is bounded below, and it follows immediately that $\lim_{\|x\|\to+\infty} f_k(x) = +\infty$. From continuity of f, we have that f_k is continuous, then minimize $f_k(x)$ reduces to a compact set; thus, the minimum is attained. Since f is convex and $\gamma_k\|x-x^k\|^2$ is strictly convex, we have that f_k is strictly convex, and so this minimum is unique. Therefore, x^{k+1} is uniquely determined. Moreover, if we have that $x^k = x^{k+1}$, then stop because we are in the solution.

Now, we will see that $\{x^k\}$ is Fejér convergent to the set of minimizers U. To this end, we will see that $\|x^{k+1} - \bar{x}\|^2 \leqslant \|x^k - \bar{x}\|^2 + \|x^{k+1} - x^k\|^2, \quad \forall\ k \in \mathbb{N}, \quad \forall\ \ \bar{x} \in U$.

Hence,

$$\begin{aligned}\|x^k - \bar{x}\|^2 &= \|x^k - x^{k+1} + x^{k+1} - \bar{x}\|^2 \\ &= \|x^k - x^{k+1}\|^2 + \|x^{k+1} - \bar{x}\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - \bar{x}\rangle.\end{aligned} \tag{2.1}$$

Since x^{k+1} is a minimizer of f_k, then

$$\nabla f_k(x^{k+1}) = 0. \tag{2.2}$$

By definition of f_k,

$$\nabla f_k(x^{k+1}) = \nabla f(x^{k+1}) + 2\gamma_k(x^{k+1} - x^k).$$

Thus, from (2.2), we obtain

$$\nabla f(x^{k+1}) + 2\gamma_k(x^{k+1} - x^k) = \nabla f_k(x^{k+1}) = 0.$$

That implies

$$2(x^k - x^{k+1}) = \frac{1}{\gamma_k}\nabla f(x^{k+1}). \tag{2.3}$$

From (2.1) and (2.3), we get

$$\|x^k-\bar{x}\|^2-\|x^{k+1}-x^k\|^2-\|x^{k+1}-\bar{x}\|^2 = 2\langle x^k-x^{k+1}, x^{k+1}-\bar{x}\rangle = \frac{1}{\gamma_k}\langle\nabla f(x^{k+1}), x^{k+1}-\bar{x}\rangle.$$

Now, using the fact that f is convex and differentiable, then by Theorem 9(b), we have that

$$\frac{1}{\gamma_k}\langle\nabla f(x^{k+1}), x^{k+1} - \bar{x}\rangle \geqslant \frac{1}{\gamma_k}[f(x^{k+1}) - f(\bar{x})].$$

And as $\bar{x} \in U$ is a minimum of f, then $f(x^{k+1}) - f(\bar{x}) \geqslant 0$. Therefore,

$$\|x^k-\bar{x}\|^2-\|x^{k+1}-x^k\|^2-\|x^{k+1}-\bar{x}\|^2 = \frac{1}{\gamma_k}\langle\nabla f(x^{k+1}), x^{k+1}-\bar{x}\rangle \geqslant \frac{1}{\gamma_k}[f(x^{k+1})-f(\bar{x})] \geqslant 0.$$

i.e.,

$$\|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2 - \|x^{k+1} - \bar{x}\| \geqslant 0.$$

Now, we establish the useful fact that is $\lim_{k\to\infty}(x^{k+1} - x^k) = 0$ to be used in the next step, which we will prove that any cluster point of $\{x^k\}$ belongs to U.

From the previous step

$$0 \leqslant \|x^{k+1} - x^k\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|. \tag{2.4}$$

Using the fact that $\{x^k\}$ is Fejér convergent, then $\{\|x^k - \bar{x}\|\}$ is decreasing and non-negative, and so, it is convergent, i.e., taking the limit in (2.4) as $k \to \infty$; one has, $\|x^k - \bar{x}\| - \|x^{k+1} - \bar{x}\| \to 0$, thus

$$\lim_{k\to\infty}(x^{k+1} - x^k) = 0. \tag{2.5}$$

Finally, we will prove that $\{x^k\}$ has clusters point and all of them belong to U.

We seen that $\{x^k\}$ is Fejér convergent, and by Lemma 1, it ensures that a cluster point exists. Let $\bar{x}$ be a cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ a subsequence of $\{x^k\}$, such that $\lim\limits_{j\to\infty} x^{k_j} = \bar{x}$. By (2.3)

$$\nabla f(x^{k_j+1}) = 2\gamma_{k_j}(x^{k_j} - x^{k_j+1}). \tag{2.6}$$

Thus, from (2.5), we obtain $\lim\limits_{j\to\infty} x^{k_j+1} = \lim\limits_{j\to\infty} x^{k_j} = \bar{x}$. Now, taking the limit in (2.6) as $j \to \infty$ and using that $\gamma_k \leqslant \tilde{\gamma}$, we have that

$$\lim_{j\to\infty} \nabla f(x^{k_j+1}) = \lim_{j\to\infty} 2\gamma_{k_j}(x^{k_j} - x^{k_j+1}) = 0.$$

Thus, $\nabla f(\bar{x}) = 0$. That means $\bar{x} \in U$. Therefore, there exists $x^* \in U$ such that $x^* = \lim\limits_{k\to\infty} x^k$. □

2.2 The proximal point method with Bregman distances

Now, we will present the proximal point method using Bregman distance that will be used to solve problems as follows:

$$\min_{x\in\bar{S}} f(x), \tag{2.7}$$

Algorithm 2 Proximal point with Bregman distance

1. Choose x^0 on S, and a sequence of parameters of $\{\gamma_k\}$, where $0 < \gamma_k < \tilde{\gamma}$.

2. Compute $x^{k+1} = \text{argmin}_{x \in \bar{S}}\{f(x) + \gamma_k D_\varphi(x, x^k)\}$.

3. If $x^k = x^{k+1}$, stop. Otherwise, take $k = k+1$ and return to the previous step.

Where φ is a Bregman function with zone S, and $\tilde{\gamma} > 0$ is any.

which $S \subset \mathbb{R}^n$ is an open and convex set, $\bar{S}$ its closure, and f is a continuous and convex function on $\bar{S}$.

We assume that problem (2.7) has solutions, so that f is bounded below on $\bar{S}$.

Theorem 15. If problem (2.7) has solutions and φ is a Bregman function boundary coercive with respect to S. Then, the sequence $\{x^k\}$ generated by Algorithm 2 converge to x^* a solution of problem (2.7).

Proof. We go to separate the proof in four steps; in the first one, we go to show that the sequence is well-defined, and it is contained on S. In the second one, we will prove that the sequence is decreasing; in the next, we will see that $\lim(x^{k+1} - x^k) = 0$. This is a useful fact that will be used in the last one, in which we will prove that the sequence has cluster points and all belong to the set of solutions of the problem.

Step 1. The sequence is well-defined and contained in S.

Let β be a lower bound for f on S, and $f_k(x) = f(x) + \gamma_k D_\varphi(x, x^k)$, then

$$f_k(x) \geqslant \beta + \gamma_k D_\varphi(x, x^k).$$

Hence, f_k is lower bound, and by (B3) the partial level sets of D_φ are bounded, as f_k is continuous (by continuously of f and D_φ) the level sets are closed. Then, minimize f_k reduces to a compact set, and therefore the minimum is attained. Moreover, f_k is strictly convex, because f is convex and φ is strictly convex, then the minimum is uniquely determined.

Now, we will prove that $x^{k+1} \in S$. Since x^{k+1} is a minimizer of f_k; then, from the Theorem 13, we have that

$$0 \in \partial f_k(x^{k+1}). \tag{2.8}$$

Thus, from (2.8) and using the definition of f_k and subdifferential calculus rules,

$$0 \in \partial f(x) + \gamma_k \partial(D_\varphi(x, x^k))(x^{k+1}),$$

i.e., since D_φ is differentiable, then by Proposition 2,

$$\partial(D_\varphi(x, x^k))(x^{k+1}) = \nabla_x(D_\varphi(x, x^k))(x^{k+1}),$$

thus, from the Proposition 5, we have that

$$0 \in \partial f(x) + \gamma_k(\nabla\varphi(x) - \nabla\varphi(x^k)), \tag{2.9}$$

then, taking into account that $\nabla\varphi(x) = \partial\varphi(x)$ we can rewrite (2.9) as

$$0 \in (\partial f + \gamma_k\varphi)(x) - \gamma_k\nabla\varphi(x^k),$$

that means

$$\gamma_k\nabla\varphi(x^k) \in (\partial f + \gamma_k\varphi)(x). \tag{2.10}$$

We want to show that if is valid (B6), then $\partial(f+\gamma_k\varphi)(x) = \emptyset$ for all $x \in \partial S$ (boundary of S), which implies, taking into account (2.10), that $x^{k+1} \in S$.

Hence, recall (B6): let $\{y^k\} \subset S$ and $\lim y^k = y \in \partial S$. Then $\lim_{k\to\infty}\langle\nabla\varphi(y^k), x - y^k\rangle = -\infty, \quad \forall x \in S$. Thus, suppose that there exists $\xi \in \partial(f+\gamma_k\varphi)(x)$. Take $z \in S$ and $x \in \partial S$ (boundary of S). From convexity of S there exists $y^l \in S, \varepsilon_l \in (0,1)$, such that

$$y^l = (1-\varepsilon_l)x + \varepsilon_l z. \tag{2.11}$$

Hence, if $\lim\limits_{l\to\infty} \varepsilon_l = 0$, then $\lim\limits_{l\to\infty} y^l = x$. Thus,

$$y^l - x = \varepsilon_l(z - x) \tag{2.12}$$

and, from (2.12) and (2.11)

$$\varepsilon_l\langle\xi, z - x\rangle = \langle\xi, y^l - x\rangle.$$

Since $\xi \in \partial(f + \gamma_k\varphi)(x)$, by definition of subdifferential and definition of $f + \gamma_k\varphi$, we have that

$$\begin{aligned}\langle\xi, y^l - x\rangle &\leqslant f(x + \gamma_k\varphi)(y^l) - f(x + \gamma_k\varphi)(x)\\ &= f(y^l) - f(x) + \gamma_k(\varphi(y^l) - \varphi(x)).\end{aligned} \tag{2.13}$$

Now, using the fact that φ is convex and differentiable on S, by Theorem 9 and (2.13), we obtain that

$$f(y^l) - f(x) + \gamma_k(\varphi(y^l) - \varphi(x)) \leqslant f(y^l) - f(x) + \gamma_k\langle\nabla\varphi(y^l), y^l - x\rangle. \tag{2.14}$$

Thus, from (2.12),(2.13) and (2.14), we have

$$\langle \xi, y^l - x\rangle \leqslant f(y^l) - f(x) + \gamma_k \frac{\varepsilon_l}{1-\varepsilon_l}\langle \nabla\varphi(y^l), z - y^l\rangle. \tag{2.15}$$

Then, by convexity of f and (2.11),

$$f(y^l) = f((1-\varepsilon_l)x + \varepsilon_l z) \leqslant (1-\varepsilon_l)f(x) + \varepsilon_l f(z) = f(x) + \varepsilon_l(f(z) - f(x)). \tag{2.16}$$

Hence, combining (2.15) and (2.16)

$$\varepsilon_l\langle \xi, z - x\rangle \leqslant \varepsilon_l(f(z) - f(x)) + \gamma_k \frac{\varepsilon_l}{1-\varepsilon_l}\langle \nabla\varphi(y^l), z - y^l\rangle,$$

so that,

$$\frac{1-\varepsilon_l}{\gamma_k}(f(x) - f(z)) + \langle \xi, z - x\rangle) \leqslant \langle \nabla\varphi(y^l), z - y^l\rangle.$$

Thus, taking the limit as $l \to \infty$ in both sides, we have, from (B6) $\lim\limits_{l\to\infty} \langle \nabla\varphi(y^l), z - y^l\rangle = -\infty$, while in the left hand side $\lim\limits_{l\to\infty} \varepsilon_l = 0$, then we have a finite limit. This contradiction implies that $\partial(f + \gamma_k\varphi) = \emptyset$ for all $x \in \partial S$. Therefore $x^{k+1} \in S$.

Step 2. We will check that $D_\varphi(\bar{x}, x^{k+1}) \leqslant D_\varphi(\bar{x}, x^k) - D_\varphi(x^{k+1}, x^k)$ for all $\bar{x}$ solution of problem (2.7).

From the Algorithm (2), and from the Theorem 13, $0 \in \partial[f + \gamma_k D_\varphi(\cdot, x^k)](x^{k+1})$. By proposition 5(ii)

$$0 \in \partial f(x^{k+1}) + \gamma_k(\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)).$$

This implies,

$$\gamma_k(\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)) \in \partial f(x^{k+1}). \tag{2.17}$$

Define $y^k = \gamma_k(\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k))$. From Proposition 5(i) with $x = \bar{x}, y = x^k$ and $z = x^{k+1}$ and by definition of y^k we have that

$$D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}) - D_\varphi(x^{k+1}, x^k) = \frac{1}{\gamma_k}\langle y^k, x^{k+1} - \bar{x}\rangle. \tag{2.18}$$

From (2.17) $y^k \in \partial f(x^{k+1})$, then

$$\langle y^k, \bar{x} - x^{k+1}\rangle \leqslant f(\bar{x}) - f(x^{k+1}), \tag{2.19}$$

and so

$$\langle y^k, x^{k+1} - \bar{x}\rangle \geqslant f(x^{k+1}) - f(\bar{x}).$$

Therefore, combining (2.19) with (2.18), and by the fact that $\bar{x}$ is a minimizer, we have

$$D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}) - D_\varphi(x^{k+1}, x^k) \geqslant \frac{1}{\gamma_k}(f(x^{k+1}) - f(\bar{x})) \geqslant 0. \tag{2.20}$$

Since $D_\varphi(x^{k+1}, x^k) \geqslant 0$, from (2.20), we obtain

$$D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}) \geqslant 0.$$

Step 3. We will prove that $\{x^k\}$ is bounded and $\lim_{j\to\infty} x^{k_j} = \hat{x}$ implies $\lim_{j\to\infty} x^{k_j+1} = \hat{x}$.

Hence, from the Step 2, $\{D_\varphi(\bar{x}, x^k)\}$ is sequence decreasing and nonnegative, i.e., it is convergent. Moreover, from (2.20)

$$0 \leqslant D_\varphi(x^{k+1}, x^k) \leqslant D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}), \tag{2.21}$$

taking the limit in (2.21) as $k \to \infty$,

$$\lim_{k\to\infty} D_\varphi(x^{k+1}, x^k) = 0. \tag{2.22}$$

Since $D_\varphi(\bar{x}, x^k)$ is decreasing, we have that $D_\varphi(\bar{x}, x^k) \leqslant D_\varphi(\bar{x}, x^0)$. By (B3) the set $\{x^k \in S; D_\varphi(\bar{x}, x^k) \leqslant \delta\}$ is bounded for all $\bar{x} \in S$, i.e., the sequence $\{x^k\}$ is bounded. Then, if $\{x^{k_j}\}$ a subsequence of $\{x^k\}$ converges to $\hat{x}$, then by (2.22), $\lim_{j\to\infty} D_\varphi(x^{k_j+1}, x^{k_j}) = 0$. Therefore, from (B5) we have $\lim x^{k_j+1} = \hat{x}$.

Step 4. Now, to prove that $\{x^k\}$ has a cluster points and all of them are solutions for Algorithm 2.

Take $\bar{x}$ a solution of a problem (2.7). Let $\hat{x}$ be a cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ a subsequence such that $\lim x^{k_j} = \hat{x}$. The existence of $\hat{x}$ follows from the Step 3, which ensures that $\lim x^{k_j+1} = \hat{x}$. Furthermore, from (2.20) in Step 2, we have

$$0 \leqslant \frac{1}{\tilde{\gamma}}(f(x^{k_j+1}) - f(\bar{x})) \leqslant \frac{1}{\gamma_k}(f(x^{k_j+1}) - f(\bar{x})) \leqslant D_\varphi(\bar{x}, x^{k_j}) - D_\varphi(\bar{x}, x^{k_j+1}) - D_\varphi(x^{k_j+1}, x^{k_j}). \tag{2.23}$$

Letting $j \to \infty$, the rightmost term in (2.23) goes to 0, and then we obtain

$$0 \leqslant \frac{1}{\tilde{\gamma}}(\lim f(x^{k_j+1}) - f(\bar{x})) \leqslant 0,$$

i.e., since $\tilde{\gamma} > 0$, then $f(\hat{x}) = f(\bar{x})$. Hence, since $\bar{S}$ is closed and $\{x^k\} \subset S$ for all $k \in \mathbb{N}$, we have that $\hat{x} \in S$ and solves the problem (2.7).

To finish the proof we need a Fejér convergence to Bregman distance. This fact holds, but we can proceed directly: Let $\hat{x}$ a cluster point of $\{x^k\}$ such that $\lim x^{k_j} = \hat{x}$. From (B4), $\lim\limits_{k\to\infty} D_\varphi(\hat{x}, x^{k_j}) = 0$. From Step 4, $\hat{x}$ solves the problem (2.7) and from Step 2 $\{D_\varphi(\hat{x}, x^k)\}$ is a sequence decreasing and nonnegative which has a subsequence goes to 0, then all sequence goes to 0 when $k \to \infty$, then, by (B4) again we have, $\lim\limits_{k\to\infty} x^k = \hat{x}$.

□

Chapter 3

Equilibrium Problems

The study proposed in this chapter has as its main works those that were presented by Burachik and Kassay [5] and an alternative version shown by Cruz Neto et al. [2].

Hence, let $X \subset \mathbb{R}^n$ be a convex, closed and nonempty set and $f : X \times X \to \mathbb{R}$. An equilibrium problem, denoted by $\mathsf{EP}(f, X)$, consists in finding $x^* \in X$ such that

$$f(x^*, y) \geqslant 0 \quad \text{for all} \quad y \in X.$$

A problem related is finding $y^* \in X$, such that $f(x, y^*) \leqslant 0$ for all $x \in X$, this problem is called dual problem. The set of solutions of $\mathsf{EP}(f, X)$ will be defined by $S_{EP}(f, X)$ and of dual problem by $S^d_{EP}(f, X)$.

We will consider some conditions with respect to the function f that will be shown following:

Assumptions 1.

(P1) $f(x, x) = 0$ for all $x \in X$.

(P2) $f(\cdot, y) : X \to \mathbb{R}$ is upper semi-continuous for all $y \in X$.

(P3) $f(x, \cdot) : X \to \mathbb{R}$ is lower semi-continuous and convex for all $x \in X$.

(P4) f is monotone, i.e., $f(x, y) + f(y, x) \leqslant 0$ for all $x, y \in X$.

(P5) For any sequence $\{z^n\} \subset X$ with $\lim\limits_{n\to\infty} \|z^n\| = \infty$ there exists $u \in X$ and $n_0 \in \mathbb{N}$ such that $f(z^n, u) \leqslant 0$ for all $n \geqslant n_0$.

Moreover, we will need of others conditions of monotonicity of f. Then, we are going to consider:

$(\mathsf{P4}^{\bullet})$ (Sub-monotonicity) There exists $\theta \geqslant 0$, such that

$$f(x,y) + f(y,x) \leqslant \theta\langle \nabla\varphi(x) - \nabla\varphi(y), x - y\rangle \quad \text{for all} \quad x, y \in X.$$

Proposition 6. Under $(\mathsf{P1}) - (\mathsf{P4})$ holds for f, then $S_{EP}(f,X) = S^d_{EP}(f,X)$.

Proof. Take $x^* \in S_{EP}(f,X)$, then $f(x^*,y) \geqslant 0$, for all $y \in X$. From $(\mathsf{P4})$, we have $f(y,x^*) \leqslant -f(x^*,y) \leqslant 0$. Then $x^* \in S^d_{EP}(f,X)$, and $S_{EP}(f,X) \subset S^d_{EP}(f,X)$.

Now, take $x^* \in S^d_{EP}(f,X)$ and $w \in X$. Define, for $t \in (0,1)$,

$$w_t = (1-t)x^* + tw.$$

Since X is convex, then $w_t \in X$. By the fact $x^* \in S^d_{EP}(f,X)$, we get, $f(w_t,x^*) \leqslant 0$. Hence,

$$f(w_t,x^*) + tf(x^*,w) - tf(x^*,w) \leqslant 0. \tag{3.1}$$

Recall that $f(x,x) = 0, \quad \forall x \in X$, thus

$$f(w_t,x^*) + tf(x^*,w) - \frac{1}{1-t}f(w_t,w_t) \leqslant tf(x^*,w).$$

We have that $f(w_t,\cdot)$ is convex, then substituting $w_t = (1-t)x^* + tw$, we have

$$\begin{aligned} f(w_t,w_t) &= f(w_t,(1-t)x^* + tw) \\ &\leqslant (1-t)f(w_t,x^*) + tf(w_t,w). \end{aligned} \tag{3.2}$$

Substituting (3.2) into (3.1),

$$f(w_t,x^*) + tf(x^*,w) - \frac{1}{1-t}[(1-t)f(w_t,x^*) + tf(w_t,w)] \leqslant tf(x^*,w).$$

Then, one has

$$tf(x^*,w) - \frac{t}{1-t}f(w_t,w) \leqslant tf(x^*,w),$$

dividing all equation above by $t > 0$,

$$f(x^*,w) - \frac{1}{1-t}f(w_t,w) \leqslant f(x^*,w).$$

Thus, taking the limit as $t \to 0^+$ in the inequality above, we have $w_t \to x^*$, and so

$$f(x^*,w) - f(x^*,w) \leqslant f(x^*,w).$$

That means $f(x^*,w) \geqslant 0$. Hence, since w is taking any in X, we get $x^* \in S_{EP}(f,X)$ and $S^d_{EP}(f,X) \subset S_{EP}(f,X)$. Therefore $S_{EP}(f,X) = S^d_{EP}(f,X)$.

□

Theorem 16. Assume that f satisfies $(\mathsf{P1})-(\mathsf{P4})$. Then $\mathsf{EP}(\mathsf{f}, \mathsf{X})$ has solution if and only f satisfies (P5).

Proof. See [17, Theorem 4.3]. □

3.1 Proximal point method with Bregman distance for equilibrium problems

Throughout the section, we are going to study the work proposed by Burachik and Kassay [5]. We will consider $X \subset \mathbb{R}^n$ a closed, convex and nonempty set and $f : X \times X \to \mathbb{R}$ be a bifunction such that satisfies the conditions $(P1), (P2)$ and $(P3)$ in Assumptions 1. Let $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a convex proper and lower semicontinuous function. Denote its domain by $\mathcal{D} := \mathrm{dom}(\varphi)$ and its interior by $\mathrm{int}\mathcal{D}$. Assume from now on that $\mathrm{int}\mathcal{D} \neq \emptyset$ and φ is differentiable on $\mathrm{int}\mathcal{D}$. The Bregman distance associate to φ is the function $D_\varphi : \mathcal{D} \times \mathrm{int}\mathcal{D} \to \mathbb{R}$ defined as in the Section 1.2, and suppose $X \subset \mathrm{int}\mathcal{D}$.

Fix $\gamma > 0$ and $\bar{x} \in X \subset \mathrm{int}\mathcal{D}$. The regularization of f will be denoted by $\tilde{f} : X \times X \to \mathbb{R}$ and defined by

$$\tilde{f}(x, y) = f(x, y) + \gamma\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle. \tag{3.3}$$

Proposition 7. Take $\bar{x} \in X \subset \mathtt{int}\mathcal{D}$. Suppose that f satisfies assumptions $(P1) - (P3)$ and $(P4^\bullet)$ with $\theta < \gamma$. Then $\tilde{f}$ satisfies assumptions $(P1) - (P4)$. Furthermore, if for any sequence $\{z^n\} \subset X$ such that $\lim_{n\to\infty} \|z^n\| = +\infty$, we have

$$(P6): \quad \liminf_{n\to\infty}[f(\bar{x}, z^n) + (\gamma - \theta)\langle \nabla\varphi(\bar{x}) - \nabla\varphi(z^n), \bar{x} - z^n\rangle] > 0,$$

then $\tilde{f}$ satisfies assumption (P5).

Proof. $\tilde{f}$ satisfies $(P1)$, indeed,

$$\tilde{f}(x, x) = f(x, x) + \gamma\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), x - x\rangle = f(x, x) = 0.$$

To show that condition $(P2)$ holds for $\tilde{f}$, we will show that the map $x \mapsto \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle$, is continuous at every $x \in X$. To this end, let $\{x^n\} \subset X$ be a sequence converging to $x \in X$. Then, $\langle \nabla\varphi(\bar{x}), x^n\rangle \to \langle \nabla\varphi(\bar{x}), x\rangle$, while $\langle \nabla\varphi(x^n), y - x^n\rangle \to \langle \nabla\varphi(x), y - x\rangle$ as the gradient of φ is continuous (by differentiability of φ). Thus, we obtain

$$\langle \nabla\varphi(x^n) - \nabla\varphi(\bar{x}), y - x^n\rangle \to \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle.$$

Hence, if $(\mathsf{P2})$ holds for f, then also holds for $\tilde{f}$.

To prove $(\mathsf{P3})$, we know that the map $y \mapsto s(y) = \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle$ is continuous and convex. Since $f(x,\cdot)$ is convex, we get that $\tilde{f}(x,\cdot)$ also is convex.

Now, we claim that $\tilde{f}$ verifies $(\mathsf{P4})$. Indeed, by condition $(\mathsf{P4}^\bullet)$, $f(x,y) + f(y,x) \leqslant \theta\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle$. Thus,

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

that implies

$$\begin{aligned} f(x,y) + f(y,x) - \gamma\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle &\leqslant \\ \theta\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle - \gamma\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle &= \\ (\theta - \gamma)\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle &\leqslant 0. \end{aligned}$$

Because, by hypothesis $\theta - \gamma < 0$.

Finally, we will prove that $\tilde{f}$ satisfies $(\mathsf{P5})$. From (3.3), we have that

$$\begin{aligned} \tilde{f}(z^n,\bar{x}) &= f(z^n,\bar{x}) + \gamma\langle \nabla\varphi(z^n) - \nabla\varphi(\bar{x}), \bar{x} - z^n\rangle \\ &= f(z^n,\bar{x}) - \gamma\langle \nabla\varphi(\bar{x}) - \nabla\varphi(z^n), \bar{x} - z^n\rangle. \end{aligned} \tag{3.4}$$

From $(\mathsf{P4}^\bullet)$, we have $f(z^n,\bar{x}) + f(\bar{x},z^n) \leqslant \theta\langle \nabla\varphi(\bar{x}) - \nabla\varphi(z^n), \bar{x} - z^n\rangle$. Then, from (3.4)

$$\begin{aligned} \tilde{f}(z^n,\bar{x}) &\leqslant -f(\bar{x},z^n) - (\gamma - \theta)\langle \nabla\varphi(\bar{x}) - \nabla\varphi(z^n), \bar{x} - z^n\rangle \\ &= -(f(\bar{x},z^n) + (\gamma - \theta)\langle \nabla\varphi(\bar{x}) - \nabla\varphi(z^n), \bar{x} - z^n\rangle). \end{aligned}$$

From assumption (P6) there exists $n_0 \in \mathbb{N}$ such that the expression in parentheses in last equality is nonnegative for all $n \geqslant n_0$. Thus, following immediately that condition $(\mathsf{P5})$ holds for $\tilde{f}$. □

Corollary 1. Consider all assumptions in Proposition 7 be valid. Then the following assertions hold:

(i) If condition $\mathsf{P6}$ holds, then $\mathsf{EP}(\tilde{f}, X)$ admits as least one solution.

(ii) If φ is strictly convex, then $\mathsf{EP}(\tilde{f}, X)$ admits at most one solution.

Proof. The proof the item (i) following from the Proposition 7, because we have that $\tilde{f}$ satisfies condition $(\mathsf{P5})$, and by the Theorem 16 we obtain that $\mathsf{EP}(\tilde{f}, X)$ has a solution.

To prove item (ii), assume that both $\hat{x}$ and $\tilde{x}$ solve $\mathsf{EP}(\tilde{f}, X)$. Thus

$$\begin{aligned} 0 \leqslant \tilde{f}(\hat{x},\tilde{x}) = f(\hat{x},\tilde{x}) + \gamma\langle \nabla\varphi(\hat{x}) - \nabla\varphi(\bar{x}), \tilde{x} - \hat{x}\rangle; \\ 0 \leqslant \tilde{f}(\tilde{x},\hat{x}) = f(\tilde{x},\hat{x}) + \gamma\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\bar{x}), \hat{x} - \tilde{x}\rangle. \end{aligned}$$

Adding both equations above, we obtain

$$0 \leqslant f(\hat{x}, \tilde{x}) + f(\tilde{x}, \hat{x}) - \gamma\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x}), \tilde{x} - \hat{x}\rangle. \tag{3.5}$$

Since (P4$^\bullet$) holds for f, we have $f(\hat{x}, \tilde{x}) + f(\tilde{x}, \hat{x}) \leqslant \theta\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x})\rangle$, substituting this into (3.5), we get

$$0 \leqslant \theta\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x}), \tilde{x} - \hat{x}\rangle - \gamma\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x}), \tilde{x} - \hat{x}\rangle = (\theta - \gamma)\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x}), \tilde{x} - \hat{x}\rangle,$$

thus, since $\theta - \gamma < 0$,

$$0 \leqslant (\theta - \gamma)\langle \nabla\varphi(\tilde{x}) - \nabla\varphi(\hat{x}), \tilde{x} - \hat{x}\rangle \leqslant 0. \tag{3.6}$$

Hence, by the fact that φ is strictly convex, then $\nabla\varphi$ is strictly monotone, thus, we obtain from (3.6) that $\tilde{x} = \hat{x}$. Therefore, $\mathsf{EP}(\tilde{f}, X)$ admits at most one solution. □

3.2 Bregman proximal method for solving EP

Algorithm 3 Proximal point method + Bregman distance

Step 1: Take a sequence of regularization parameters $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$, for some $\bar{\gamma} > 0$ and θ is a constant of sub-monotonicity of f.

Step 2: Choose $x_0 \in X$ and construct the sequence $\{x^k\}$ as follows.

Step 3: Given $x^k \in X$, x^{k+1} is the unique solution of the problem $\mathsf{EP}(f_k, X)$, where $f_k : X \times X \to \mathbb{R}$ is given by

$$f_k(x, y) = f(x, y) + \gamma_k\langle \nabla\varphi(x) - \nabla\varphi(x^k), y - x\rangle. \tag{3.7}$$

Remark 4. Assume that φ is strictly convex, and that is valid all assumptions in Proposition 7.

Definition 13. We said that a sequence $\{z^n\} \subset X$ is asymptotically solving for the problem $\mathsf{EP}(f, X)$ if $\liminf_{n\to\infty} f(z^n, y) \geqslant 0$, for all $y \in X$.

Theorem 17. Consider the problem $\mathsf{EP}(f, X)$, the following assertions holds for all $x^0 \in X$:

(i) The sequence $\{x^k\}$ generated by Algorithm 3 is well-defined.

(ii) If, $S^d(f, X) \neq \emptyset$, then the sequence $\{x^k\}$ is bounded and $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$. Moreover, $\{x^k\}$ is an asymptotically solving sequence of $\mathsf{EP}(f, X)$.

(iii) If additionally $f(\cdot, y)$ is semicontinuous for all $y \in X$, then the cluster points of $\{x^k\}$ solve $\mathsf{EP}(f, X)$.

Proof. Item (i) follows directly from Corollary 1. Indeed, for each k, $f_k(x, y) = f(x, y) + \gamma_k\langle\nabla\varphi(x) - \nabla\varphi(x^k), y - x\rangle$ satisfies (P6) then EP(f_k, X) has as least one solution, and as φ is strictly convex, then EP(f_k, X) has at most one solution. Thus, for each k, $x^{k+1} \in S(f_k, X)$ is unique determined.

To prove item (ii), since $S^d(f, X) \neq \emptyset$, take $x^* \in S^d(f, X)$ any, then $f(x, x^*) \leqslant 0$ for all $x \in X$. Now, by definition of algorithm $x^{k+1} \in S(f_k, X)$, thus we have $f_k(x^{k+1}, y) \geqslant 0$ for all $y \in X$. Then taking $y = x^*$, we have

$$0 \leqslant f_k(x^{k+1}, x^*) = f(x^{k+1}, x^*) + \gamma_k\langle\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k), x^* - x^{k+1}\rangle,$$

hence, using that $f(x, x^*) \leqslant 0$, in particular, for $x = x^{k+1}$, i.e., $f(x^{k+1}, x^*) \leqslant 0$, and using that $\gamma_k > 0$, we obtain

$$0 \leqslant \langle\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k), x^* - x^{k+1}\rangle.$$

Now, from the property of three points for Bregman distances, we have

$$0 \leqslant \langle\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k), x^* - x^{k+1}\rangle = D_\varphi(x^*, x^k) - D_\varphi(x^{k+1}, x^k) - D_\varphi(x^*, x^{k+1}).$$

That means,

$$D_\varphi(x^{k+1}, x^k) + D_\varphi(x^*, x^{k+1}) \leqslant D_\varphi(x^*, x^k). \tag{3.8}$$

By the fact that $D_\varphi(x^{k+1}, x^k)$ is non-negative, it follows that the sequence $\{D_\varphi(x^*, x^k)\}$ is decreasing and non-negative, it is convergent, in particular, it is bounded. Hence, by Condition (B3) in Section 1.2, we have that the sequence $\{x^k\}$ is bounded. To prove the second assertion of item (ii), we have from (3.8) that

$$0 \leqslant D_\varphi(x^{k+1}, x^k) \leqslant D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1}).$$

Taking the limit as $k \to \infty$ in the inequality above we obtain that the right-hand side converges to 0, then it follows that

$$\lim_{k\to\infty} D_\varphi(x^{k+1}, x^k) = 0. \tag{3.9}$$

From (3.9) and Condition (B5), $\lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} x^k$, and so

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{3.10}$$

Now, by (3.10) and using Condition (B8), we have

$$\lim_{k\to\infty} \|\nabla\varphi(x^k) - \nabla\varphi(x^{k+1})\| = 0. \tag{3.11}$$

Fix $y \in X$ any. Since $x^{k+1} \in S_{EP}(f_k, X)$ that implies $f_k(x^{k+1}, y) \geqslant 0$; thus, using definition of f_k and Cauchy-Schwartz inequality, one has

$$\begin{aligned} 0 &\leqslant f(x^{k+1}, y) + \gamma_k\langle\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k), y - x^{k+1}\rangle \\ &\leqslant f(x^{k+1}, y) + \gamma_k\|\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)\|\|y - x^{k+1}\|. \end{aligned} \tag{3.12}$$

Letting $k \to \infty$ in inequality (3.12), since $\{\gamma_k\}$ is bounded for $\bar{\gamma}$ and $\{x^k\}$ is bounded by (3.9); from (3.11) we get $\lim_{k\to\infty} \gamma_k\|\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)\|\|y - x^{k+1}\| = 0$, thus we have

$$0 \leqslant \liminf_{k\to\infty} f(x^k, y). \tag{3.13}$$

Therefore, since $y \in X$ is taking any, we have that $\{x^k\}$ is asymptotically solving for $EP(f, X)$.

To prove item (iii), in view of (ii), we have that $\{x^k\}$ is bounded and it has cluster points, all of which belong to X, which is closed and convex. Let $\hat{x}$ be one of them. Let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ convergent to $\hat{x}$. Under semi-continuity of $f(\cdot, y)$, we have, in view of (3.13),

$$f(\hat{x}, y) \geqslant \limsup_{j\to\infty} f(x^{k_j}, y) \geqslant \liminf_{j\to\infty} f(x^{k_j}, y) \geqslant 0, \quad \forall\, y \in X;$$

and hence, $\hat{x} \in S(f, X)$. □

Proposition 8. Let $\{x^k\}$ be the sequence generated by the Algorithm 3, if $S_{EP}(f, X) = S^d_{EP}(f, X)$, then the sequence converges to solution of the problem $EP(f, X)$.

Proof. Since f satisfies conditions $(P1) - (P4)$, then $S_{EP}(f, X) = S^d_{EP}(f, X)$. Hence, in view of the Theorem 17 it remains to check that there exists only one cluster point of $\{x^k\}$. Let x' and x'' be two cluster points of $\{x^k\}$ and consider the subsequences x^{k_l} and x^{k_j} converging to x' and x'' respectively. Then, it follows that both x' and x'' belongs to $S(f, X) = S^d(f, X)$. Thus, from (3.8) both $D_\varphi(x', x^k)$ and $D_\varphi(x'', x^k)$ are convergent, say, to $\sigma \geqslant 0$ and $\delta \geqslant 0$. Thus, by the Proposition 5

$$\langle\nabla\varphi(x^{k_l}) - \nabla\varphi(x^{k_j}), x' - x''\rangle = [D_\varphi(x', x^{k_j}) - D_\varphi(x', x^{k_l})] - [D_\varphi(x'', x^{k_j}) - D_\varphi(x'', x^{k_l})].$$

Taking a limit as $k \to \infty$ in both sides in expression above, one has

$$\langle \nabla\varphi(x') - \nabla\varphi(x''), x' - x'' \rangle = (\sigma - \sigma) - (\delta - \delta) = 0.$$

Therefore, by strictly monotonicity of $\nabla\varphi$, $x' = x''$. $\square$

Remark 5. It is worth mentioning that, when we have $\varphi(x) = \frac{1}{2}\|x\|^2$, the regularization of f_k in (3.7) becomes

$$f_k(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle. \tag{3.14}$$

Hence, the Algorithm 3 becomes classical proximal point for equilibrium problems proposed by Iusem and Sosa in [20], which it is presented following.

Algorithm 4 Classical Proximal Point for EP

Step 1. Take a positive and bounded sequence of regularization parameters $\{\gamma_k\}$.

Step 2. Choose a initial $x^0 \in X$ and construct the sequence $\{x^k\}$ as follow:

Step 3. Given x^k. compute x^{k+1} as the unique solution of $\mathsf{EP}(f_k, X)$, where f_k is given by (3.14).

3.3 Separable Bregman regularization

In this section we will introduce a new Bregman regularization that was presented by Cruz Neto et al. [2]. Where, they consider a regularization, which the regularization process of being separable with respect to the variables x and y. Hence, we will shown existence and uniqueness results of solutions for the regularized problem as well as its convergence to a solution of the original equilibrium problem.

Let $X \subset \mathbb{R}^n$ be a closed and convex set and let $f : X \times X \to \mathbb{R}$ be a function such that satisfies Conditions $(P1) - (P5)$ in Assumptions 1 and let D_φ be the Bregman distance associate to function φ satisfying $X \subset \text{int}\mathcal{D}$, as $\mathcal{D} = \mathsf{dom}\ \varphi$.

Remark 6. The assumption $S(f, X) \neq \emptyset$ was taken as a standard assumption. Furthermore, we will consider the following condition:

(B5$^\bullet$): If $\{y^k\} \subset \mathcal{D}$ and $\lim_{k\to+\infty} \|y^k\| = +\infty$, then $\lim_{k\to\infty} D_\varphi(y^k, u) = +\infty$ for any $u \in \text{int}\mathcal{D}$.

Which is a form of coercivity weaker than the form used by Iusem and Narsi [19].

$$\lim_{\|x\|\to+\infty}[\varphi(x)-\rho\|x-z\|]=+\infty,\quad \forall\, z\in X \text{ and } \rho\geqslant 0. \tag{3.15}$$

Indeed, condition above implies $(B5^{\bullet})$ and one can verify that $\varphi(t)=-\log t$ verifies $(B5^{\bullet})$ and does not satisfy condition (3.15).

Fix $\gamma>0$ and $\bar{x}\in X\subset \operatorname{int}\mathcal{D}$. The new regularization of f, in this section, will be denoted by

$$\tilde{f}(x,y)=f(x,y)+\gamma(D_{\varphi}(y,\bar{x})-D_{\varphi}(x,\bar{x})).$$

Proposition 9. Take $\bar{x}\in X$ and assume that f satisfies $(P1)-(P4)$. Then $\tilde{f}$ satisfies $(P1)-(P4)$. Moreover, if it holds the condition $(B5^{\bullet})$, then $\tilde{f}$ satisfies condition $(P5)$.

Proof. It is easy to see that $\tilde{f}$ satisfies $(P1)$. Indeed,

$$\tilde{f}(x,x)=f(x,x)+\gamma(D_{\varphi}(x,\bar{x})-D_{\varphi}(x,\bar{x}))=0.$$

Now, $\tilde{f}$ satisfies $(P2)$, because since $X\subset \operatorname{int}\mathcal{D}$ we get that φ is continuous, and so since f is upper semicontinuous we have that $\tilde{f}$ will also be.

The same argument is valid for the case where f is lower semicontinuous. Moreover, by the fact that $D_{\varphi}(\cdot,y)$ is convex and f satisfies $(P3)$, we obtain that $\tilde{f}(x,\cdot)$ is convex and lower semicontinuous, and so satisfies $(P3)$.

Now, note that $\tilde{f}(x,y)+\tilde{f}(y,x)=f(x,y)+f(y,x)$. Since, $(P4)$ holds for f, then also holds for $\tilde{f}$.

It remains to prove that $(P5)$ holds for $\tilde{f}$. Since $S_{EP}(f,X)\neq\emptyset$, take $u\in S_{EP}(f,X)$ and $\{z^n\}\subset X$, such that $\lim_{k\to\infty}\|z^n\|=+\infty$. Then,

$$\tilde{f}(z^n,u)=f(z^n,u)+\gamma(D_{\varphi}(u,\bar{x})-D_{\varphi}(z^n,\bar{x})).$$

Note that, $f(u,z^n)\geqslant 0$ (because, $u\in S_{EP}(f,X)$), from (P4), $f(z^n,u)\leqslant 0$. Hence,

$$\tilde{f}(z^n,u)=f(z^n,u)+\gamma(D_{\varphi}(u,\bar{x})-D_{\varphi}(z^n,\bar{x}))\leqslant\gamma(D_{\varphi}(u,\bar{x})-D_{\varphi}(z^n,\bar{x})). \tag{3.16}$$

Using $(B5^{\bullet})$, we have $D_{\varphi}(z^n,\bar{x})\to+\infty$ when $n\to\infty$. Then, taking the limit as $n\to\infty$ the most right in expression (3.16) goes to $-\infty$, hence, $f(z^n,u)\leqslant 0$ for enough large $n\in\mathbb{N}$.

□

Proposition 10. Let all assumptions of Proposition 9 be valid, then $\mathsf{EP}(\tilde{f}, X)$ admits solution. Moreover, if φ be strictly convex, $\mathsf{EP}(\tilde{f}, X)$ admits at most one solution.

Proof. First assertion comes from the Proposition 9 adding Theorem 16.

To prove second assertion, take $x^* \in S_{EP}(\tilde{f}, X)$ and $\bar{x} \in X$, then

$$\tilde{f}(x^*, y) = f(x^*, y) + \gamma(D_\varphi(y, \bar{x}) - D_\varphi(x^*, \bar{x})) \geqslant 0.$$

Define $g_{x^*} : X \to \mathbb{R}$, given by $g_{x^*}(y) = \tilde{f}(x^*, y)$, and note that $g_{x^*}(y) \geqslant 0$ for all $y \in X$, and $g_{x^*}(x^*) = 0$. It shows that x^* is a minimizer of g_{x^*}. From the strict convexity of φ and $D_\varphi(\cdot, \bar{x})$ and by the convexity of $\tilde{f}(x, \cdot)$, we have $g_{x^*}(\cdot)$ is strictly convex, then has a unique minimizer. Now, take $\hat{x} \in S_{EP}(\tilde{f}, X)$, such that $\hat{x} \neq x^*$, this implies that $g_{x^*}(x^*) < g_{x^*}(\hat{x})$, i.e., $\tilde{f}(x^*, x^*) < \tilde{f}(x^*, \hat{x})$, and hence

$$0 < \tilde{f}(x^*, \hat{x}). \tag{3.17}$$

Since, $\hat{x} \in S_{EP}(\tilde{f}, X)$, we have

$$\tilde{f}(\hat{x}, x^*) \geqslant 0. \tag{3.18}$$

Then, adding (3.17) with (3.18)

$$0 < \tilde{f}(\hat{x}, x^*) + \tilde{f}(x^*, \hat{x}),$$

which is a contradiction due (P4): $f(\hat{x}, x^*) + f(x^*, \hat{x}) \leqslant 0$. Therefore there exists a unique $x^* \in S_{EP}(\tilde{f}, X)$. □

Next, we present the proximal point method with a separable Bregman regularization for solving an $\mathsf{EP}(f, X)$.

Algorithm 5 New Bregman algorithm for solving EP

Step 1. Take a bounded and positive sequence of parameters $\{\gamma_k\}$ and choose $x^0 \in X$.

Step 2. Given x^k, compute

$$x^{k+1} \in S_{EP}(f_k, X); \tag{3.19}$$

where, $f_k(x, y) = f(x, y) + \gamma_k(D_\varphi(y, x^k) - D_\varphi(x, x^k))$.

Remark 7. Since φ be strictly convex and $\tilde{f}$ satisfies conditions $(\mathsf{P1}) - (\mathsf{P5})$ (Proposition 9). Then, by Corollary 1 and Proposition 10 we have that each subproblem has a unique solution. Hence, the Algorithm 5 is well-defined.

Proposition 11. Let $\{x^k\}$ be a sequence generated by algorithm, then for all $k \in \mathbb{N}$ the following assertions holds:

$$f(x^{k+1}, y) + \gamma_k(D_\varphi(y, x^k) - D_\varphi(y, x^{k+1})) \geqslant 0, \quad \forall y \in X, \tag{3.20}$$

and

$$D_\varphi(\bar{x}, x^{k+1}) \leqslant D_\varphi(\bar{x}, x^k), \tag{3.21}$$

for all $k \in \mathbb{N}$ and for every $\bar{x} \in S_{EP}(f, X)$. Moreover, $\sum_{k=0}^{\infty} D_\varphi(x^{k+1}, x^k) < \infty$ and, in particular, $\lim_{k\to\infty} D_\varphi(x^{k+1}, x^k) = 0$.

Proof. By definition of Algorithm 5, we have that $x^{k+1} \in S_{EP}(f_k, X)$, and $f_k(x^{k+1}, y) \geqslant 0$, for all $y \in X$, i.e.,

$$f_k(x^{k+1}, y) = f(x^{k+1}, y) + \gamma_k(D_\varphi(y, x^k) - D_\varphi(x^{k+1}, x^k)) \geqslant 0, \quad \forall y \in X.$$

Note that $f_k(x^{k+1}, x^{k+1}) = 0$ and $f_k(x^{k+1}, y) \geqslant 0$, for all $y \in X$, hence

$$x^{k+1} \in \operatorname{argmin}_{y \in K} f_k(x^{k+1}, y).$$

By Proposition 9, $f_k(x, \cdot)$ is convex. Thus, from definition of subdifferential (Definition 8) we have that there exists $w^{k+1} \in \partial f_k(x^{k+1}, x^{k+1})$, such that

$$\langle w^{k+1}, y - x^{k+1} \rangle \geqslant f_k(x^{k+1}, y) - f_k(x^{k+1}, x^{k+1}) \geqslant f_k(x^{k+1}, y).$$

Now, from (3.19), $f_k(x^{k+1}, y) \geqslant 0$, thus

$$\langle w^{k+1}, y - x^{k+1} \rangle \geqslant 0. \tag{3.22}$$

Using that φ is (strictly) convex and the definition of f_k, we obtain

$$w^{k+1} = v^{k+1} + \gamma_k \nabla_1 D_\varphi(x^{k+1}, x^k), \tag{3.23}$$

where, $v^{k+1} \in \partial f(x^{k+1}, x^{k+1})$ and ∇_1 denotes the gradient with respect to the first argument of the D_φ at the point $y = x^{k+1}$. Then, combining (3.23) with (3.22), one has

$$\langle v^{k+1} + \gamma_k \nabla_1(D_\varphi(x^{k+1}, x^k), y - x^{k+1} \rangle \geqslant 0. \tag{3.24}$$

By the Proposition 5 (ii), we have $\nabla_1 D_\varphi(x^{k+1}, x^k) = \nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)$. Thus, from (3.24)

$$\langle v^{k+1} + \gamma_k(\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)), y - x^{k+1} \rangle \geqslant 0.$$

That implies in

$$\langle \nu^{k+1}, y - x^{k+1} \rangle + \gamma_k \langle \nabla \varphi(x^{k+1}) - \nabla \varphi(x^k), y - x^{k+1} \rangle \geqslant 0. \tag{3.25}$$

Now, note that $f(x^{k+1}, \cdot)$ is convex and $\nu^{k+1} \in \partial f(x^{k+1}, x^{k+1})$, then

$$\langle \nu^{k+1}, y - x^{k+1} \rangle \leqslant f(x^{k+1}, y) - f(x^{k+1}, x^{k+1}) = f(x^{k+1}, y). \tag{3.26}$$

Substituting (3.26) into (3.25), one has

$$f(x^{k+1}, y) + \gamma_k \langle \nabla \varphi(x^{k+1}) - \nabla \varphi(x^k), y - x^{k+1} \rangle \geqslant 0. \tag{3.27}$$

From three points equality (Proposition 5(i)), we obtain

$$f(x^{k+1}, y) + \gamma_k (D_\varphi(y, x^k) - D_\varphi(y, x^{k+1}) - D_\varphi(x^{k+1}, x^k)) \geqslant 0,$$

since $D_\varphi(x^{k+1}, x^k) \geqslant 0$ and $\gamma_k \geqslant 0$, one follows that is valid (3.20).

It remains to prove (3.21), i.e., $D_\varphi(\bar{x}, x^{k+1}) \leqslant D_\varphi(\bar{x}, x^k)$, for all $\bar{x} \in S(f, X)$. Then note that taking $y = \bar{x} \in S(f, X)$ in (3.27), one has

$$f(x^{k+1}, \bar{x}) + \gamma_k \langle \nabla \varphi(x^{k+1}) - \nabla \varphi(x^k), \bar{x} - x^{k+1} \rangle \geqslant 0.$$

Using again property of three points, we obtain

$$f(x^{k+1}, \bar{x}) + \gamma_k (D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}) - D_\varphi(x^{k+1}, x^k)) \geqslant 0.$$

Now, note that $\bar{x} \in S(f, X)$ then $f(\bar{x}, \tilde{y}) \geqslant 0$ for all $\tilde{y} \in X$, by monotonicity of f, $f(\tilde{y}, \bar{x}) \leqslant 0$, in particular, for $\tilde{y} = x^{k+1}$, i.e., $f(x^{k+1}, \bar{x}) \leqslant 0$. Thus, as $\gamma_k > 0$ we have

$$D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}) - D_\varphi(x^{k+1}, x^k) \geqslant 0.$$

Since $D_\varphi(x^{k+1}, x^k) \geqslant 0$, then

$$D_\varphi(\bar{x}, x^{k+1}) \leqslant D_\varphi(\bar{x}, x^k), \quad \forall \bar{x} \in S, \quad \forall k \in \mathbb{N}.$$

To prove last assertion use the fact that

$$D_\varphi(x^{k+1}, x^k) \leqslant D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}).$$

Hence, summing from $k = 0$ to $k = m$,

$$\sum_{k=0}^{m} D_\varphi(x^{k+1}, x^k) \leqslant \sum_{k=0}^{m} D_\varphi(\bar{x}, x^k) - D_\varphi(\bar{x}, x^{k+1}),$$

this implies in

$$\sum_{k=0}^{m} D_{\varphi}(x^{k+1}, x^k) \leqslant D_{\varphi}(\bar{x}, x^0) - D_{\varphi}(\bar{x}, x^{m+1}) \leqslant D_{\varphi}(\bar{x}, x^0).$$

Finally, letting m goes to $+\infty$, we have

$$\sum_{k=0}^{+\infty} D_{\varphi}(x^{k+1}, x^k) \leqslant D_{\varphi}(\bar{x}, x^0) < +\infty.$$

In particular, $\lim_{k\to+\infty} D_{\varphi}(x^{k+1}, x^k) = 0$. □

Corollary 2. The sequence $\{x^k\}$ generated by algorithm is bounded.

Proof. From the last proposition $D_{\varphi}(\bar{x}, x^{k+1}) \leqslant D_{\varphi}(\bar{x}, x^k)$, i.e.,$\{D_{\varphi}(\bar{x}, x^k)\}$ is non increasing and nonnegative, in particular, it is convergent, and in particular it is bounded, combining this fact with condition $(B3)$ we obtain that $\{x^k\}$ is bounded. □

Theorem 18. Every cluster point of the sequence $\{x^k\}$ generated by Algorithm 5 is a solution of $EP(f, X)$.

Proof. In the last corollary we get that $\{x^k\}$ is bounded, then, there exists cluster points of $\{x^k\}$. Let x^* be a cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$, such that $\lim_{j\to\infty} x^{k_j} = x^*$. By the Proposition 11, $\lim_{k\to\infty} D_{\varphi}(x^{k+1}, x^k) = 0$, then

$$\lim_{j\to\infty} D_{\varphi}(x^{k_j+1}, x^{k_j}) = 0,$$

due $(B5)$, we have

$$\lim_{j\to\infty} x^{k_j+1} = \lim_{j\to\infty} x^{k_j} = x^*.$$

On other hand, from the last Proposition in (3.20),

$$f(x^{k+1}, y) + \gamma_k(D_{\varphi}(y, x^k) - D_{\varphi}(y, x^{k+1})) \geqslant 0, \quad \forall\, y \in X.$$

Hence,

$$\gamma_{j_k}(D_{\varphi}(y, x^{k_j+1}) - D_{\varphi}(y, x^{k_j})) \leqslant f(x^{k_j+1}, y). \tag{3.28}$$

Furthermore,

$$D_{\varphi}(y, x^{k_j}) - D_{\varphi}(y, x^{k_j+1}) = \langle \nabla\varphi(x^{k_j}) - \nabla\varphi(x^{k_j+1}), x^{k_j+1} - y\rangle + D_{\varphi}(x^{k_j+1}, x^{k_j}). \tag{3.29}$$

From (3.29) and (3.28), we have

$$\langle \nabla\varphi(x^{k_j}) - \nabla\varphi(x^{k_j+1}), x^{k_j+1} - y\rangle + D_{\varphi}(x^{k_j+1}, x^{k_j}) \leqslant f(x^{k_j+1}, y).$$

Letting $j \to +\infty$ we have that the most left side of the above inequality goes to 0, because $x^{k_j} \to x^*$ and $x^{k_j+1} \to x^*$. Then

$$\begin{aligned} 0 &\leqslant \lim f(x^{k_j+1}, y) \\ &\leqslant \limsup f(x^{k_j+1}, y) \\ &\leqslant f(x^*, y), \qquad \forall\, y \in X, \end{aligned}$$

where the last inequality follows from assumption (P2). Therefore, we have $x^* \in S(f, X)$.

□

Theorem 19. The sequence $\{x^k\}$ generated by Algorithm 5 converges to a solution of $EP(f, X)$.

Proof. By Corollary 2 we have the sequence $\{x^k\}$ is bounded and convergent, and by Theorem 3.27, we have that every cluster point is solution of $EP(f, X)$, then it is sufficient to proof that there exists only one cluster point of $\{x^k\}$. Thus, let $\bar{x}$ and $\hat{x}$ be two cluster points of $\{x^k\}$, and $\{x^{k_j}\}$ and $\{x^{k_l}\}$ be subsequences that converges to $\bar{x}$ and $\hat{x}$ respectively. From Theorem 18 follows that $\hat{x}, \bar{x} \in S(f, X)$. Since f satisfies (P4), then $S_{EP}(f, X) = S^d_{EP}(f, X)$. Hence, by Proposition 11 $\{D_\varphi(\hat{x}, x^k)\}$ and $\{D_\varphi(\bar{x}, x^k)\}$ are convergent, say, to $\gamma \geqslant 0$ and $\mu \geqslant 0$, respectively. Then, by the Proposition 5, we obtain

$$\langle \nabla\varphi(x^{k_j}) - \nabla\varphi(x^{k_l}), \hat{x} - \bar{x} \rangle = [D_\varphi(\hat{x}, x^{k_j}) - D_\varphi(\hat{x}, x^{k_l})] - [D_\varphi(\bar{x}, x^{k_j}) - D_\varphi(\bar{x}, x^{k_j})].$$

Taking the limit on above equality, we have

$$\langle \nabla\varphi(\bar{x}) - \nabla\varphi(\hat{x}), \hat{x} - \bar{x} \rangle = (\gamma - \gamma) - (\mu - \mu) = 0.$$

Therefore, due the strict monotonicity of $\nabla\varphi$ we have that $\hat{x} = \bar{x}$.

□

Chapter 4

Quasi-equilibrium Problems

In the last chapter, we saw that a Equilibrium problem consist in finding $x^* \in X \subset \mathbb{R}^n$ such that

$$f(x^*, y) \geqslant 0, \quad \forall\ y \in X.$$

The quasi-equilibrium problem (shortly QEP), consists in finding $x^* \in C(x^*)$ such that

$$f(x^*, y) \geqslant 0, \quad \forall\ y \in C(x^*);$$

where, C is a multi-valued map from X into itself. $C : X \rightrightarrows X$ associates to each $x \in X$ a closed, convex and nonempty subset $C(x)$ on X. Furthermore, we will consider the continuity of mapping C, in the sense of Mosco (M-continuous), that will be shown below. This assumptions were consider in others algorithms for QEP, see for instance Bigi and Passacantando [3], Strodiot et al. [37] and Van et al. [40] .

Definition 14. C is said M-continuous if
(i) For any $\{x^k\}, \{y^k\} \subset X$ with $y^k \in C(x^k)$, $x^k \to x$ and $y^k \to y$ implies that $y \in C(x)$.
(ii) For any sequence $\{x^k\} \subset X$ with $x^k \to x$ and for each $y \in C(x)$ there exists a sequence $\{y^k\} \subset X$ with $y^k \in C(x^k)$ such that $y^k \to y$.

The solution set of QEP will be denoted by $S_{QEP}(f, C)$. Moreover, we will consider a set $S^* \subset S_{QEP}$, where

$$S^* = \{x \in \bigcap_{z \in X} C(z); f(x, y) \geqslant 0, \quad \forall\ y \in \bigcup_{z \in X} C(z)\} \tag{4.1}$$

and we will assume that $S^* \neq \emptyset$. This assumptions was considered for convergence of extragradient algorithms, see Strodiot et al. [37], and generalized Nash equilibrium , see Han et al. [13].

Remark 8. "The solution set of the equilibrium problem is nonempty", has been assumed as a mild assumption. In QEP the assumption $S^* \neq \emptyset$ may be viewed as a natural extension to QEP of the assumption $S_{QEP} \neq \emptyset$, because if $C(x) = X$, for all $x \in X$, then $\bigcap_{z\in X} C(z) = \bigcup_{z\in X} C(z) = X$ and $S^* = S_{EP}(f, X)$. As $S^* \subset S_{QEP}$, then when $S^* \neq \emptyset$ guarantees that $S_{QEP} \neq \emptyset$.

From now on, we will consider $f : X \times X \to \mathbb{R}$ a bifunction that satisfies the following properties.

Assumptions 2.

(Q1) $f(x, x) = 0$, for all $x \in X$;

(Q2) $f(\cdot, \cdot)$ is jointly continuous on $X \times X$ (or the graph of f is sequentially closed), in the sense that, if $x, y \in X, \{x^k\}$ and $\{y^k\}$ are sequences in X converging to x and y respectively, then $f(x^k, y^k)$ converges to $f(x, y)$;

(Q3) $f(x, \cdot) : X \to \mathbb{R}$ is lower semicontinuous and convex for all $x \in X$;

(Q4) f is monotone, i.e., for each pair of points $x, y \in X$, we have $f(x, y) + f(y, x) \leqslant 0$;

(Q5) For any sequence $\{z^n\} \subset X$ such that $\lim_{n\to\infty} \|z^n\| = +\infty$ there exists $u \in X$ and $n_0 \in \mathbb{N}$ such that $f(z^n, u) \leqslant 0$, for $n \geqslant n_0$.

Note that quasi-equilibrium problems are a generalization of equilibrium problems. That in turn, they are a generalization of other problems, for instance, the minimization problem, the variational inequalities and Nash equilibrium and etc., see for more details, Nikaido and Isoda [27], Konnov [24], Facchinei et al. [11], You and Li [42] and Han et al. [13].

Next, it is presented a QEP version of a result quite useful in equilibrium problems.

Proposition 12. Let $\bar{x} \in X$ be a arbitrary point and $\hat{x}, x^* \in X$, such that $\hat{x} \in S_{EP}(\hat{f}, C(\bar{x}))$ and $x^* \in S^d_{EP}(f, C(\bar{x}))$, where $\hat{f}$ is denoted by $\hat{f}(x, y) = f(x, y) + \gamma\langle x - \bar{x}, y - x\rangle$, for some $\gamma > 0$. If f satisfies $(Q1) - (Q3)$ in assumptions 2, then $\|\hat{x} - x^*\|^2 + \|\bar{x} - \hat{x}\|^2 \leqslant \|\bar{x} - x^*\|^2$.

Proof. Take $\hat{x} \in S_{EP}(\hat{f}, C(\bar{x}))$, then $\hat{x} \in C(\bar{x})$, and

$$0 \leqslant \hat{f}(\hat{x}, y) = f(\hat{x}, y) + \gamma\langle \hat{x} - \bar{x}, y - \hat{x}\rangle, \quad \forall\, y \in C(\bar{x}),$$

i.e.,

$$-f(\hat{x}, y) \leqslant \gamma\langle \hat{x} - \bar{x}, y - \hat{x}\rangle, \quad \forall\, y \in C(\bar{x}). \tag{4.2}$$

Now, since $x^* \in S^d_{EP}(f, C(\bar{x}))$, then $x^* \in C(\bar{x})$ and $f(y, x^*) \leqslant 0$, for all $y \in C(\bar{x})$. In particular, for $y = \hat{x}$, then

$$f(\hat{x}, x^*) \leqslant 0. \tag{4.3}$$

Making $y = x^*$ in (4.2), we obtain

$$-f(\hat{x}, x^*) \leqslant \langle \hat{x} - \bar{x}, x^* - \hat{x}\rangle. \tag{4.4}$$

Combining (4.4) with (4.3), we have

$$0 \leqslant \gamma\langle \hat{x} - \bar{x}, x^* - \hat{x}\rangle.$$

Recall the Proposition 3, thus

$$0 \leqslant \gamma\langle \hat{x} - \bar{x}, x^* - \hat{x}\rangle = \frac{\gamma}{2}(\|\hat{x} - x^*\|^2 - \|\bar{x} - \hat{x}\|^2 - \|\bar{x} - x^*\|^2).$$

Furthermore, as $\gamma > 0$, then

$$\|\hat{x} - x^*\|^2 - \|\bar{x} - \hat{x}\|^2 \leqslant \|\bar{x} - x^*\|^2.$$

□

Let $f : X \times X \to \mathbb{R}$ be a bifunction. We denote by $\partial_2 f$ the subdifferential of f (in the sense of convex analysis) with respect to its second argument computed at a point $(x, x) \in X \times X$. In view (Q1) and (Q3), one has

$$\begin{aligned} \partial_2 f(x, x) &= \{v \in \mathbb{R}^n : f(x, y) \geqslant f(x, x) + \langle v, y - x\rangle, \quad \forall y \in X\} \\ &= \{v \in \mathbb{R}^n : f(x, y) \geqslant \langle v, y - x\rangle, \quad \forall y \in X\}. \end{aligned}$$

It is called subdifferential diagonal; see Iusem in [16] for more details. Given $x \in X$ we denote by

$$\Upsilon(x) = \|x - \underset{y \in C(x)}{\operatorname{argmin}}\{f(x, y) + \frac{1}{2}\|y - x\|^2\}\|.$$

The next proposition will be used in numerical experiments. It measure the quality of a candidate to solution of QEP.

Proposition 13. A point $x^* \in S_{QEP}(f, C)$ if and only if $\Upsilon(x^*) = 0$.

Proof. Take $x^* \in S_{QEP}(f, C)$, then $x^* \in C(x^*)$ and $f(x^*, y) \geqslant 0$, for all $y \in C(x^*)$. Thus, as $\frac{1}{2}\|y - x^*\|^2$ is positive, we have that

$$f(x^*, y) + \frac{1}{2}\|y - x^*\|^2 \geqslant f(x^*, y) \geqslant 0. \tag{4.5}$$

Now, since $f(x, x) = 0$, for all $x \in X$, then from (4.5), we have

$$f(x^*, y) + \frac{1}{2}\|y - x^*\|^2 \geqslant f(x^*, x^*) + \frac{1}{2}\|x^* - x^*\|^2, \quad \forall y \in C(x^*).$$

Hence, $x^* = \operatorname{argmin}_{y \in C(x^*)}\{f(x^*, y) + \frac{1}{2}\|y - x^*\|^2\}$ and $\Upsilon(x^*) = 0$.

Now, suppose that $\Upsilon(x^*) = 0$, i.e., $x^* = \operatorname{argmin}_{y \in C(x^*)}\{f(x^*, y) + \frac{1}{2}\|y - x^*\|^2\}$, then $x^* \in C(x^*)$, and from the first order optimality condition (Theorem 13),

$$0 \in \partial_2 f(x^*, x^*) + \mathcal{N}_{C(x^*)}(x^*),$$

where, $\mathcal{N}_{C(x^*)}$ is the normal cone of the set $C(x^*)$ at x^*. Then, there exists $s \in \partial_2 f(x^*, x^*)$, such that $0 \in s + \mathcal{N}_{C(x^*)}(x^*)$, i.e., $-s \in \mathcal{N}_{C(x^*)}(x^*)$. By definition of normal cone, we have

$$\langle y - x^*, -s \rangle \leqslant 0, \quad \forall y \in C(x^*),$$

thus,

$$\langle y - x^*, s \rangle \geqslant 0, \quad \forall y \in C(x^*). \tag{4.6}$$

On the other hand, $s \in \partial_2 f(x^*, x^*)$, and from (Q3) $f(x^*, \cdot)$ is convex, and from the definition of subdifferential diagonal, we obtain

$$f(x^*, y) \geqslant f(x^*, x^*) + \langle s, y - x^* \rangle = \langle s, y - x^* \rangle \geqslant 0, \quad \forall\, y \in C(x^*).$$

The last inequality comes from (4.6), and therefore, $x^* \in S_{QEP}(f, C)$. $\square$

4.1 Proximal Point Method for solving QEP

Next, we are going to present the proximal point method for solving a quasi-equilibrium problem, that was proposed by Santos and Souza [34], which was the first work that extended the proximal point for context the of quasi-equilibrium. Furthermore, we will present its well-definition and convergence analysis for the solution of QEP. Assuming from now on that f satisfies $(Q1) - (Q5)$ and the set S^* given by (4.1) is nonempty. The algorithm will be defined below.

Algorithm 6 Proximal Point Algorithm for QEP

Step 1. Take a bounded auxiliary sequence of positives parameters $\{\gamma_k\}$ and choose $x^0 \in X$.

Step 2. Given x^k, compute

$$x^{k+1} \in S_{EP}(f_k, C_k), \tag{4.7}$$

where, $f_k(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle$ and $C_k = C(x^k)$.

Step 3. If $x^{k+1} = x^k$, stop and return x^k. Otherwise set $k = k + 1$ and return to previous step.

Remark 9. It is worth mentioning that the method does not need to start at a fixed point of C, i.e., $x^0 \in C(x^0)$. Moreover, we do not assume that $x \in C(x)$ for all $x \in X$, as done in some existing works on extragradient algorithms for QEP, see for instance [37] and gap functions; see [3].

Proposition 14. If f satisfies $(Q1) - (Q5)$, then f_k satisfies $(Q1) - (Q5)$.

Proof. Suppose that f satisfies $(Q1)$ then, $f_k(x, x) = f(x, x) + \gamma \langle x - x^k, x - x \rangle = 0$.

To prove that f_k satisfies $(Q2)$, we have since f satisfies $(Q2)$ and $\{x^n\}, \{y^n\} \subset X$, such that $x^n \to x$ and $y^n \to y$, then

$$\begin{aligned}
\lim_{n\to\infty} f_k(x^n, y^n) &= \lim_{n\to\infty} (f(x^n, y^n) + \gamma_k \langle x^n - x^k, y^n - x^n \rangle) \\
&= f(x, y) + \gamma_k \langle x - x^k, y - x \rangle = f_k(x, y).
\end{aligned}$$

Since $(Q3)$ holds for f, i.e., $f(x, \cdot)$ is convex, and having in mind that inner product is convex; then, as by definition $f_k(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle$ we get that f_k satisfies assertion $(Q3)$.

Now, $(Q4)$ holds for f_k, indeed

$$\begin{aligned}
f_k(x, y) + f_k(y, x) &= f(x, y) + \gamma \langle x - x^k, y - x \rangle + f(y, x) + \langle y - x^k, x - y \rangle \\
&= f(x, y) + f(y, x) + \gamma_k \langle x - y, y - x \rangle \\
&= f(x, y) + f(y, x) - \gamma_k \|x - y\|^2 \leqslant 0,
\end{aligned}$$

because, $f(x, y) + f(y, x) \leqslant 0$, $\|x - y\|^2 \geqslant 0$ and γ_k is bounded and positive.

To prove that $(Q5)$ holds for f_k, take $u = P(x^k)$, where P is orthogonal projection from x^k upon X. Thus,

$$\begin{aligned} f_k(z^n, u) &= f(z^n, u) + \gamma_k\langle z^n - x^k, u - z^n\rangle \\ &= f(z^n, u) + \gamma_k\langle z^n - x^k + u - u, u - z^n\rangle \\ &= f(z^n, u) + \gamma_k\langle z^n - u, u - z^n\rangle + \gamma_k\langle u - x^k, u - z^n\rangle \\ &= f(z^n, u) - \gamma_k\|z^n - u\|^2 + \gamma_k\langle u - x^k, u - z^n\rangle \end{aligned} \tag{4.8}$$

from the obtuse angle property for orthogonal projection (Theorem 8), we have $\langle u - x^k, u - z^n\rangle \leqslant 0$, substituting this into (4.8) and taking into account that f is monotone, i.e., $f(z^n, u) \leqslant -f(u, z^n)$, we obtain

$$f_k(z^n, u) \leqslant -f(u, z^n) - \gamma_k\|z^n - u\|^2. \tag{4.9}$$

Now, define $g_x : X \to \mathbb{R}$ such that, $g_x(y) = f(x, y)$. Take $\hat{x} \in ri(X)$, so that $\hat{x}$ belong to the interior of effective domain of g_x. Since, $f(x, \cdot)$ is convex, we have g_x is convex. Then, ∂g_x is convex, closed and nonempty. Thus, take $\nu \in \partial g_u(\hat{x})$. By the definition of subdifferential

$$\langle \nu, z^n - \hat{x}\rangle \leqslant g_u(z^n) - g_u(\hat{x}) = f(u, z^n) - f(u, \hat{x}).$$

Then, using Cauchy-Schwartz and triangular inequality, we have

$$\begin{aligned} -f(u, z^n) &\leqslant \langle \nu, \hat{x} - z^n\rangle - f(u, \hat{x}) \\ &\leqslant \|\nu\|\|\hat{x} - z^n\| - f(u, \hat{x}) \\ &\leqslant \|\nu\|\|z^n - u\| + \|\nu\|\|u - \hat{x}\| - f(u, \hat{x}). \end{aligned} \tag{4.10}$$

Substituting (4.10) into (4.9), we get

$$\begin{aligned} f_k(z^n, u) &\leqslant \|\nu\|\|z^n - u\| + \|\nu\|\|u - \hat{x}\| - f(u, \hat{x}) - \gamma_k\|z^n - u\|^2 \\ &= \|z^n - u\|(\|\nu\| - \gamma_k\|z^n - u\|) + \|\nu\|\|u - \hat{x}\| - f(u, \hat{x}). \end{aligned}$$

Since $\|z^n\| \to \infty$, then $\|z^n - u\| \to \infty$. Hence, as $\gamma_k > 0$, one has $\lim\limits_{n\to\infty} f_k(z^n, u) = -\infty$. Therefore, $f_k(z^n, u) \leqslant 0$ for large enough n.

□

Remark 10. If f satisfies (Q1)−(Q5) then f_k satisfies (Q1)−(Q5), then by the Theorem 16 $S_{EP}, (f_k, C_K)$ is nonempty for each $k \in \mathbb{N}$. Therefore, the Proposition 16 guarantees well-definition of the sequence of algorithm, taking into account f_k satisfies (Q1) − (Q5) and $C_k = C(x^k)$ is a nonempty, closed and convex set (fixed at each step).

4.1.1 Convergence Analysis

Proposition 15. Let $\{x^k\}$ be the sequence generated by algorithm PPM. If $x^{k+1} = x^k$ for some $k \in \mathbb{N}$, then x^k is a solution of QEP.

Proof. From definition of the algorithm, we have that $x^{k+1} \in S_{EP}(f_k, C_k)$; that implies $x^{k+1} \in C(x^k)$ and $f_k(x^{k+1}, y) = f(x^{k+1}, y) + \gamma_k\langle x^{k+1} - x^k, y - x^{k+1}\rangle \geqslant 0, \ \ \forall\, y \in C(x^k)$. Since $x^{k+1} = x^k$, we obtain that $x^k \in C(x^k)$ and $\langle x^{k+1} - x^k, y - x^{k+1}\rangle = 0$, thus $f(x^k, y) \geqslant 0$, for all $y \in C(x^k)$. Therefore, $x^k \in S_{QEP}(f, C)$.

□

In order to demonstrate that $\{x^k\}$ converges to the set of solutions of QEP, we will apply the result presented in Lemma 1, i.e., since $\{x^k\}$ is Fejér convergent and has a cluster point that belongs to the set, we obtain that $\{x^k\}$ converges to the set. Hence, we will demonstrate this in the results that follows.

Proposition 16. Let $\{x^k\}$ be the sequence generated by the Algorithm 6, then the following assertions hold:
(i) $\{x^k\}$ is Fejér convergent to S^*;
(ii) $\{x^k\}$ is bounded;
(iii) $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.

Proof. To prove item (i), take any $x^* \in S^* \subset S_{EP}(f, C(z))$, for all $z \in X$. Since, f is monotone $S_{EP}(f, C(z)) = S^d_{EP}(f, C(z))$, in particular, for $z = x^k$, then $x^* \in S_{EP}(f, C(x^k))$, for all $k \in \mathbb{N}$. From definition of the algorithm, we have that $x^{k+1} \in S_{EP}(f_k, C(x^k))$. Thus, applying Proposition 12 with $\hat{f} = f_k,\ \bar{x} = x_k$ and $\hat{x} = x^{k+1}$ we obtain

$$\|x^{k+1} - x^*\|^2 + \|x^{k+1} - x^k\|^2 \leqslant \|x^k - x^*\|^2, \ \ \forall\, k \in \mathbb{N}, \tag{4.11}$$

consequently, $\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2$, for all $k \in \mathbb{N}$. Since x^* was take arbitrarily in S^*, we have that $\{x^k\}$ is Fejér convergent to S^*.

For item (ii), note that $\{\|x^k - x^*\|\}$ is non-increasing and convergent, hence, $\|x^k - x^*\| \leqslant \|x^0 - x^*\|$. Thus, $\{x^k\}$ is contained in a ball of center x^* and rate $\|x^0 - x^*\|$. So $\{x^k\}$ is bounded.

To prove item (iii) note that from item (ii), $\{\|x^k - x^*\|\}$ is decreasing and convergent; thus, by inequality (4.11), we have

$$0 \leqslant \|x^{k+1} - x^k\|^2 \leqslant \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2, \ \ \forall\, k \in \mathbb{N}. \tag{4.12}$$

Taking the limit as $k \to \infty$ in both sides in (4.12), we obtain

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0.$$

□

Theorem 20. Every cluster point of $\{x^k\}$ belong to $S_{QEP}(f, C)$.

Proof. Let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ that converges to $\hat{x}$. From the definition of algorithm, we have that $x^{k_j+1} \in C(x^{k_j})$. By Proposition 16(iii),

$$\lim_{j\to\infty} \|x^{k_j+1} - x^{k_j}\| = 0,$$

hence, $\lim_{j\to\infty} x^{k_j+1} = \hat{x}$. Thus, from M-continuity of C, we have that $\hat{x} \in C(\hat{x})$ and given $y \in C(\hat{x})$, there exists a sequence $\{y^{k_j}\}$ that converges to y and $y^{k_j} \in C(x^{k_j})$. Now, from (4.7), $x^{k_j+1} \in S_{EP}(f_{k_j}, C_{k_j})$ and we obtain

$$0 \leqslant f_{k_j}(x^{k_j+1}, z), \quad \forall \;\; z \in C(x^{k_j}),$$

which means that, in particular, for $z = y^{k_j} \in C(x^{k_j})$,

$$0 \leqslant f_{k_j}(x^{k_j+1}, y^{k_j}).$$

Then, by the definition of f_k and by Cauchy-Schwartz inequality,

$$\begin{aligned} 0 \leqslant f_{k_j}(x^{k_j+1}, y^{k_j}) &= f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\langle x^{k_j+1} - x^{k_j}, y^{k_j} - x^{k_j+1}\rangle \\ &\leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|x^{k_j+1} - x^{k_j}\|\|y^{k_j} - x^{k_j+1}\|, \quad \forall j \in \mathbb{N}. \end{aligned} \tag{4.13}$$

Thus, having in mind that $\{\gamma_{k_j}\}$, $\{x^{k_j}\}$ and $\{y^{k_j}\}$ are bounded sequences, $x^{k_j+1} \to \hat{x}$ and $y^{k_j} \to y$. Since f satisfies (Q2) and letting $j \to \infty$ in (4.13), we have

$$0 \leqslant f(\hat{x}, y).$$

As $y \in C(\hat{x})$ was taking arbitrarily, this means that $f(\hat{x}, y) \geqslant 0$, for all $y \in C(\hat{x})$. Therefore, $\hat{x} \in S_{QEP}(f, C)$.

□

Remark 11. Note that $S^* \subset S_{QEP}$ and from Proposition 16 we get $\{x^k\}$ is Fejér convergent to S^*. But, the Theorem 20 does not guarantee that cluster points of $\{x^k\}$ belong to S^*, so we cannot apply the Lemma 1 in order to obtain convergence of the whole sequence to a point of $S_{QEP}(f, C)$. But the next result gives a sufficient condition to overcome this drawback.

Corollary 3. If f is strictly monotone, then $\{x^k\}$ converges to a solution of QEP.

Proof. Statement 1: If f is strictly monotone, then $S_{QEP}(f, C) = S^* = \{x^*\}$, where x^* is a cluster point of the sequence $\{x^k\}$. The result follows directly from:
(i) By Proposition 16, we have that $\{x^k\}$ is Fejér convergent, bounded and $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.
(ii) Every cluster point of $\{x^k\} \in S_{QEP}(f, C)$. Since $\{x^k\}$ is bounded, it admits a convergent subsequence. As it has a single point of accumulation, the whole sequence is convergent to this point.
Statement 2: $S^* \subset S_{QEP}(f, C)$.

Let $x^* \in S^*$, that means, $x^* \in \bigcap_{z\in X} C(z)$, for all $z \in X$, we have, in particular, that $x^* \in C(x^*)$. Moreover, since $f(x^*, y) \geqslant 0, \quad \forall y \in \bigcup_{z\in X} C(z)$ we have $f(x^*, y) \geqslant 0, \quad \forall \; y \in C(x^*)$. Thus, $S^* \subset S_{QEP}(f, C)$.

Finally, take arbitrary points $\hat{x} \in S_{QEP}(f, C)$ and $x^* \in S^*$. Then:

$$\hat{x} \in C(\hat{x}) \text{ and } f(\hat{x}, y) \geqslant 0, \quad \forall y \in C(\hat{x}); \tag{4.14}$$

$$x^* \in \bigcap_{z\in X} C(z) \quad \text{and} \quad f(x^*, y) \geqslant 0, \quad \forall y \in \bigcup_{z\in X} C(z). \tag{4.15}$$

Choosing $y = x^*$ in (4.14) (because, as $x^* \in \bigcap\limits_{z\in X} C(z)$, in particular, $x^* \in C(\hat{x})$), we obtain

$$f(\hat{x}, x^*) \geqslant 0. \tag{4.16}$$

On other hand, letting $y = \hat{x}$ in (4.15), we have $f(x^*, \hat{x}) \geqslant 0$ and taking account that f is monotone, we get

$$f(\hat{x}, x^*) \leqslant 0. \tag{4.17}$$

Thus, we conclude from (4.16) and (4.17) that $f(\hat{x}, x^*) = 0$. Whether holds that $\hat{x} \neq x^*$, as $x^* \in S^* \subset S_{QEP}$, we have $f(x^*, \hat{x}) \geqslant 0$, and taking account $f(\hat{x}, x^*) = 0$ and having in mind that f is strictly monotone, this implies in

$$0 \leqslant f(x^*, \hat{x}) = f(x^*, \hat{x}) + f(\hat{x}, x^*) < 0.$$

Which is a contradiction. Therefore, $x^* = \hat{x}$, and $S_{QEP} = S^* = \{x^*\}$. $\square$

4.2 Inexact proximal point for QEP

Now we will be show two inexact versions of the proximal point for QEPs that have been presented by E. Junior et al. [18]. In the first one is an inexact version of the algorithm that was proposed by Santos and Souza in [34], and in the other one, it is a version for QEPs of the algorithm proposed by Konnov [23].

First, we will shown a variation of the equilibrium problem, proposed by Hung and Muu [43], where they consider an inexact version of $\mathrm{EP}(f, X)$ as follows: given $\delta \geqslant 0, x^* \in X$ is called a δ-solution of the EP if

$$f(x^*, y) \geqslant -\delta, \quad \forall\, y \in X.$$

Before we introduce the inexact proximal point, we will show an important result that will be used in convergence analysis.

Let $\bar{x} \in X$, $e \in \mathbb{R}^n$ and $\gamma > 0$ be arbitrary. Given an equilibrium bifunction f, consider the following inexactly regularized equilibrium bifunction:

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

Proposition 17. Given $\delta > 0$ and $\bar{x} \in X$, let $\hat{x}, x^* \in C(\bar{x})$ be such that $\hat{x} \in S_{\mathrm{EP}_\delta}(\tilde{f}, C(\bar{x}))$ and $x^* \in S^d_{\mathrm{EP}}(f, C(\bar{x}))$, where

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

then,

$$\|\bar{x} - \hat{x}\|^2 + \|\hat{x} - x^*\|^2 \leqslant \|\bar{x} - x^*\|^2 + \frac{2\delta}{\gamma} - \frac{2}{\gamma}\langle e, x^* - \hat{x}\rangle.$$

Moreover, if $\|x^* - \hat{x} - e\| \leqslant \max\{\|e\|, \|x^* - \hat{x}\|\}$, then

$$\|\bar{x} - \hat{x}\|^2 + \|\hat{x} - x^*\|^2 \leqslant \|\bar{x} - x^*\|^2 + \frac{2\delta}{\gamma}.$$

Proof. Since $\hat{x} \in S_{\mathrm{EP}_\delta}(\tilde{f}, C(\bar{x}))$, we have that $\hat{x} \in C(\bar{x})$ and $\tilde{f}(\hat{x}, y) \geqslant -\delta$, for all $y \in C(\bar{x})$. That means,

$$-\delta \leqslant \tilde{f}(\hat{x}, y) = f(\hat{x}, y) + \gamma\langle \hat{x} - \bar{x}, y - \hat{x}\rangle - \langle e, y - \hat{x}\rangle, \quad \forall y \in C(\bar{x}). \tag{4.18}$$

Since $x^* \in S^d_{\mathrm{EP}}(f, C(\bar{x}))$, we have that $x^* \in C(\bar{x})$ and $f(x, x^*) \leqslant 0$, for all $x \in C(\bar{x})$. Taking $x = \hat{x}$, we obtain $f(\hat{x}, x^*) \leqslant 0$; and taking $y = x^*$ in (4.18), one has

$$-\delta \leqslant f(\hat{x}, x^*) + \gamma\langle \hat{x} - \bar{x}, x^* - \hat{x}\rangle - \langle e, x^* - \hat{x}\rangle,$$

using the fact that $f(\hat{x}, x^*) \leqslant 0$, we have

$$-\delta \leqslant \gamma\langle \hat{x} - \bar{x}, x^* - \hat{x}\rangle - \langle e, x^* - \hat{x}\rangle. \tag{4.19}$$

Taking account the Proposition 3, from (4.19), we obtain

$$\frac{\gamma}{2}(\|\bar{x} - \hat{x}\|^2 - \|\hat{x} - x^*\|^2 - \|\bar{x} - x^*\|^2) - \langle e, x^* - \hat{x}\rangle \geqslant -\delta.$$

Then, since $\gamma > 0$, we prove the first assertion

$$\|\bar{x} - \hat{x}\|^2 + \|\hat{x} - x^*\|^2 \leqslant \|\bar{x} - x^*\|^2 + \frac{2\delta}{\gamma} - \frac{2}{\gamma}\langle e, x^* - \hat{x}\rangle. \tag{4.20}$$

For second assertion, note that

$$\langle e, x^* - \hat{x}\rangle = \frac{1}{2}(\|x^* - \hat{x}\|^2 + \|e\|^2 - \|x^* - \hat{x} - e\|^2). \tag{4.21}$$

Since, $\|x^* - \hat{x} - e\| \leqslant \max\{\|e\|, \|x^* - \hat{x}\|\}$, from (4.21), it follows that $\langle e, x^* - \hat{x}\rangle \geqslant 0$. Hence, using this in (4.20) and by the fact that $\gamma > 0$, we prove the second assertion. $\square$

Definition 15. Let $S_\# \subset S_{QEP}(f, C)$ be a subset given by

$$S_\# = \{x \in C(x);\ f(x, y) \geqslant 0, \quad \forall\, y \in X\};$$

and

$$S^* = \{x \in \bigcap_{z \in X} C(z);\ f(x, y) \geqslant 0, \quad \forall\, y \in \bigcup_{z \in X} C(z)\}.$$

Remark 12. This sets will play an important role in our convergence analysis, and it has been considered in the convergence analysis of different algorithms in quasi-equilibrium problems, for instance, in [8], [13], [25], [37], [35], [40] and [41].

Next, we are going to present some examples where the sets S^* and $S_\#$ are nonempty.

Example 3. Consider $x^* \in S_{EP}(f, X)$ and let C be a point-set mapping given by

$$C(x) = \{z \in X; \|z - x^*\| \leqslant \|x\|\}$$

i.e., for each $x \in X, C(x)$ is a closed ball with center at the point x^* and ratio $\|x\|$. It is easy to see that $x^* \in \bigcap_{x \in X} C(x)$ and, in particular, $x^* \in C(x^*)$. By the fact that $x^* \in S_{EP}(f, X)$ we have that $x^* \in S^*$ and $x^* \in S_\#$. Therefore, S^* and $S_\#$ is nonempty, because it follows from their existence, which results in an equilibrium problem.

Example 4. (Minimization problem). Let $\varphi : X \subset \mathbb{R}^n \to \mathbb{R}$ be a function. The minimization problem

$$\min_{x \in X} \varphi(x),$$

can be modeled as a QEP, because taking

$$C(x) = \{z \in X : \varphi(z) \leqslant \varphi(x)\} \text{ and } f(x, y) = \varphi(y) - \varphi(x).$$

In this case, clearly $x \in C(x)$, for every $x \in X$ and $\bar{x} = \text{argmim}_{x \in X} \varphi(x) = S^* = S_{\#}$. Indeed, let $\bar{x} = \text{argmin}_{x \in X} \varphi(x)$ be then $\bar{x} \in C(x)$, for all $x \in X$, because

$$\varphi(\bar{x}) \leqslant \varphi(x), \quad \forall\, x \in X.$$

Moreover,

$$f(\bar{x}, y) = \varphi(y) - \varphi(\bar{x}) \geqslant 0, \quad \forall\, y \in X.$$

Therefore, $\bar{x} \in S^*$ and $\text{argmin}_{x \in X} \varphi(x) \subset S^*$. On other hand, let $x^* \in S^*$ be, then

$$x^* \in C(x), \quad \forall\, x \in X.$$

This means $\varphi(x^*) \leqslant \varphi(x)$, for all $x \in X$. Then, $x^* \in \text{argmin}_{x \in X} \varphi(x)$ and $S^* \subset \text{argmin}_{x \in X} \varphi(x)$. Hence, $S^* = \text{argmin}_{x \in X} \varphi(x)$. Likewise, $S_{\#} = \text{argmin}_{x \in X} \varphi(x)$. Because, let $\bar{x} \in \text{argmin}_{x \in X} \varphi(x)$ be, then

$$\bar{x} \in C(\bar{x}) \text{ and } f(\bar{x}, y) = \varphi(y) - \varphi(\bar{x}) \geqslant 0, \quad \forall\, y \in X.$$

Therefore, $\bar{x} \in S_{\#}$ and $\text{argmin}_{x \in X} \varphi(x) \subset S_{\#}$. Now, let $\hat{x} \in S_{\#}$, then

$$\hat{x} \in C(\hat{x}) \text{ and } f(\hat{x}, y) = \varphi(y) - \varphi(\hat{x}) \geqslant 0.$$

And so, we have $\varphi(y) \geqslant \varphi(\hat{x})$, for all $y \in X$. This implies that $\hat{x} \in \text{argmin}_{x \in X} \varphi(x)$ and $S_{\#} \subset \text{argmin}_{x \in X} \varphi(x)$. Therefore, so that S^* and $S_{\#}$ be nonempty depends on $\text{argmin}_{x \in X} \varphi(x)$ be nonempty.

Now, will be proposed two inexact version of the proximal algorithm which we will consider two kinds of inexactness. In the first one, our method will solve the perturbed regularized equilibrium problem. In this case, we will suppose that, at each iteration, the error e^k committed is controlled, i.e., it satisfies one the following conditions, for all $k \in \mathbb{N}$:

(E1) $\|e^k\| \leqslant \|x^{k+1} - x^k\|$;

(E2) $\|x^{k+1} - x^k - e^k\| \leqslant \max\{\|x^{k+1} - x^k\|, \|e^k\|\}$.

Algorithm 7 Inexact PPM for quasi-equilibrium problems

Step 1: Take a bounded auxiliary sequence of positive parameters $\{\gamma_k\}$ such that $0 < a \leqslant \gamma_k \leqslant b$, for all $k \in \mathbb{N}$, and choose $x^0 \in X$.

Step 2: Given $x^k \in X$, compute

$$x^{k+1} \in S_{EP_{\delta_k}}(f_k^e, C_k), \quad \text{with} \quad \sum_{k=1}^{\infty} \delta_k < \infty,$$

where $f_k^e(x,y) = f(x,y) + \gamma_k \langle x - x^k, y - x\rangle - \langle e^k, y - x\rangle, C_k = C(x^k)$ and the error $\{e^k\}$ satisfies $(E1)$ or $(E2)$.

Remark 13. Note that, if $e^k = 0$ and $\delta_k = 0$, for all $k \in \mathbb{N}$, the algorithm becomes the exact version proposed by Souza and Santos in [34]. Since x^{k+1} is a solution to an equilibrium problem and from [19, Proposition 3.1] the set $S_{EP}(f_k^e, C_k)$ is non-empty, then the Algorithm 7 is well-defined. Moreover, the assumptions $(E1)$ and $(E2)$ for the error $\{e^k\}$ are well know in the literature of inexact proximal point methods, and has been used in other works, for instance [6],[14], [31],[22] and [35].

In the second algorithm, we will consider as the next iteration $x^{k+1} \in X$ a point close enough to the exact solution controlled by summable sequence of errors $\{\varepsilon_k\}$. This inexact version takes as the next iteration any point in an $\varepsilon-$neighborhood (not necessarily in $C(x^k)$) of the exact solution of the subproblem. This method will be shown in the following algorithm.

Algorithm 8 New inexact proximal point method

Step 1: Take a bounded auxiliary sequence of positive parameters $\{\gamma_k\}$ such that $0 < a \leqslant \gamma_k \leqslant b$, for all $k \in \mathbb{N}$, and choose $x^0 \in X$.

Step 2: Given $\varepsilon_k \geqslant 0, x^k \in X$ and $v^k \in S_{EP}(f_k, C_k)$, compute $x^{k+1} \in X$ such that

$$\|x^{k+1} - v^k\| \leqslant \varepsilon_k \quad \text{with} \quad \sum_{k=0}^{\infty} \varepsilon_k < \infty,$$

where $f_k(x,y) = f(x,y) + \gamma_k \langle x - x^k, y - x\rangle$ and $C_k = C(x^k)$.

Remark 14. If the sequence of error $\{\varepsilon_k\}$ is taken equal to zero for all $k \in \mathbb{N}$, in the algorithm above, then $x^{k+1} = v^k \in S_{EP}(f_k, C_k)$ is the exact solution of the subproblems; thus, this algorithm becomes the proximal point method of Souza and Santos [34] that was also shown in the section 4.1. Hence, the well-definition of the Algorithm 8 comes from the well-definition of the exact method. On the other hand, if $C_k = X$ for all $k \in \mathbb{N}$, then Algorithm 8 becomes the inexact proximal method proposed by Konnov [23] for solving equilibrium problems.

Remark 15. Note that, from exact method the algorithm stops when $v^k = v^{k-1}$, then since $\|x^{k+1} - v^k\| < \varepsilon_k$ and $\|x^{k+1} - v^{k-1}\| < \varepsilon_{k-1}$. We can verified heuristically that $x^{k+1} = x^k$ is a good stopping criterion for the algorithm.

4.2.1 Convergence analysis

We present the convergence results. In the case of Algorithm 7 we will consider the error satisfying each one of the assumptions $(E1)$ and $(E2)$ separately. We will start with assumption (E1).

Proposition 18. Let $\{x^k\}$ be the sequence generated by Algorithm 7. We suppose that following assumptions hold:

$$\sum_{k=0}^{\infty} \|e^k\| < \infty; \tag{4.22}$$

$$\sum_{k=0}^{\infty} |\langle e^k, x^{k+1}\rangle| < \infty. \tag{4.23}$$

One has,

(i) If $S^* \neq \emptyset$, then $\{x^k\}$ is quasi-Fejér convergent to S^*.

(ii) If $S_\# \neq \emptyset$ and $S_\# \subset C_k$, for all $k \in \mathbb{N}$, then $\{x^k\}$ is quasi-Fejér convergent to $S_\#$.

Therefore, if item (i) or (ii) holds, then $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.

Proof. To prove item (i) take $x^* \in S^*$. Since $S^* \subset S_{EP}(f, C(z))$ for all $z \in X$, this means that $x^* \in S_{EP}(f, C(z))$ and $f(x^*, y) \geqslant 0$, for all $y \in C(z)$. Since, $(Q4)$ holds we have that $f(y, x^*) \leqslant 0$, for all $y \in C(z)$. This implies that $x^* \in S^d_{EP}(f, C(z))$, for all $z \in X$, in particular, for $z = x^k$. Thus, $x^* \in S^d_{EP}(f, C_k)$. Now, from definition of Algorithm 7 we have that $x^{k+1} \in S_{EP_\delta}(f^e_k, C_k)$. Hence, we can use the Proposition 17 with $\tilde{f} = f^e_k, \hat{x} = x^{k+1}, \bar{x} = x^k$, we have

$$\|x^{k+1} - x^*\|^2 + \|x^{k+1} - x^k\|^2 \leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k} - \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle, \quad \forall k \in \mathbb{N}. \tag{4.24}$$

Since $\|x^{k+1} - x^k\|^2 \geqslant 0$, this means that

$$\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k} - \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle, \quad \forall k \in \mathbb{N}, \tag{4.25}$$

using Cauchy-Schwartz inequality and the fact that $\gamma_k \geqslant a$ in (4.25), we obtain

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k} + \frac{2}{\gamma_k}(|\langle e^k, x^{k+1}\rangle| + \|e^k\|\|x^*\|) \\ &\leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{a} + \frac{2}{a}(|\langle e^k, x^{k+1}\rangle| + \|e^k\|\|x^*\|). \end{aligned} \tag{4.26}$$

Finally, taking $\varepsilon_k = \frac{2\delta_k}{a} + \frac{2}{a}(|\langle e^k, x^{k+1}\rangle| + \|e^k\|\|x^*\|)$, since (4.22) and (4.23) holds, together with the fact $\sum_{k=0}^{\infty} \delta_k < \infty$, we have that $\sum_{k=0}^{\infty} \varepsilon_k < \infty$. Therefore, by the Definition 12 and (4.26) $\{x^k\}$ is quasi-Fejér convergent to S^*.

To prove item (ii), note that if $x^* \in S_{\#}$, then $x^* \in C(x^*) \subset X$ and $f(x^*, y) \geqslant 0$, for all $y \in X$. From the monotonicity of f, we have that $f(y, x^*) \leqslant 0$ hence $x^* \in S^d_{EP}(f, X)$. Combining this fact with $S_{\#} \subset C_k$ for all $k \in \mathbb{N}$, we have that $f(x^*, y) \geqslant 0$ for all $y \in C_k$; thus, we obtain that $x^* \in S^d_{EP}(f, C_k)$ and using the same reasoning of the item (i) we prove the second assertion.

To prove the last assertion, note that $\{\gamma_k\}$ is bounded, the sequence $\{\delta_k\}$ is summable and assertions (4.23) and (4.22) hold; then,

$$\sum_{k=0}^{\infty} \frac{2\delta_k}{\gamma_k} + \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle < \infty,$$

that implies,

$$\lim_{k\to\infty} \frac{2\delta_k}{\gamma_k} + \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle = 0. \tag{4.27}$$

From (4.26), we obtain

$$\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k} - \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle. \tag{4.28}$$

Then, by Lemma 3 and (4.28), it follows that the sequence $\{\|x^k - x^*\|\}$ converges, for all $x^* \in S^*$ or $x^* \in S_{\#}$. Now, from (4.24) one has

$$0 \leqslant \|x^{k+1} - x^k\|^2 \leqslant \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 + \frac{2\delta_k}{\gamma_k} - \frac{2}{\gamma_k}\langle e^k, x^* - x^{k+1}\rangle, \ \ \forall k \in \mathbb{N}. \tag{4.29}$$

Taking the limit as $k \to \infty$ in (4.29), taking into account (4.27) and using the fact that $\{\|x^k - x^*\|\}$ converges, we obtain $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.

□

Remark 16. Assumptions (4.22) and (4.23) are well-known for inexact proximal algorithms in different contexts, for instance in [6], [10], [22] and [30] . In the case of (4.23) which involves not only the error sequence but also the iteration x^{k+1} which is determined after using the algorithm. However, as remarked by Eckstein [10], the condition (4.23) should not be difficult to enforce in practice. For instance, one could stop the process of computing x^{k+1} as soon as $\|e^k\| \leqslant \gamma\beta^k$ and $|\langle e^k, x^{k+1}\rangle| \leqslant \tilde{\gamma}\beta^k$, where $\gamma, \tilde{\gamma} \geqslant 0$ and $\beta \in [0,1)$ hence both are summable and since $\{x^k\}$ is bounded, then (4.23) is a simple consequence of (4.22).

Remark 17. Note that if $\{x^k\}$ is generated by Algorithm 7 with the error given by $(\mathsf{E2})$, then we do not need of the assumptions (4.22) and (4.23). Because, from the second part of Proposition 17 with $\tilde{f} = f^e_k, \hat{x} = x^{k+1}$ and $\bar{x} = x^k$, we have that

$$\|x^{k+1} - x^*\|^2 + \|x^{k+1} - x^k\|^2 \leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k}, \quad \forall k \in \mathbb{N}, \tag{4.30}$$

and this implies that $\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 + \frac{2\delta_k}{\gamma_k}$, since $\sum_{k=0}^{\infty} \delta_k < \infty$ and γ_k is bounded we can take $\varepsilon_k = \frac{2\delta_k}{\gamma_k}$, and therefore, $\{x^k\}$ is quasi-Fejér convergent to S^*. Moreover, $\{\|x^k - x^*\|\}$ is non-increasing and bounded from below, then it is converges for all $x^* \in S^*$. Thus, taking a limit in (4.30) as $k \to +\infty$, we obtain that $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$. Furthermore, if we replace S^* by $S_\#$ this remark also is true.

Theorem 21. Let $\{x^k\}$ be the sequence generated by Algorithm 7 and suppose that $S^* \neq \emptyset$ or $S_\# \neq \emptyset$ with $S_\# \subset C_k$, for all $k \in \mathbb{N}$. In addition, if criterion $(\mathsf{E1})$ is used, we suppose the assumptions (4.22) and (4.23) hold. Otherwise, if $(\mathsf{E2})$ is used, we assume that $\lim_{k\to\infty} \|e^k\| = 0$. Then, every cluster point of $\{x^k\}$ is a solution to the $\mathsf{QEP}(f, C)$.

Proof. If S^* is non-empty (respectively, $S_\#$ is non-empty with $S_\# \subset C_K$), from Proposition 18, we have that $\{x^k\}$ is quasi-Fejér convergent to S^* (respectively, $S_\#$). Thus, from Lemma 1 $\{x^k\}$ is bounded. Let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$, by the Algorithm 7, one has, $x^{k_j+1} \in C(x^{k_j})$, and by M-continuity of C, we obtain that $x^* \in C(x^*)$. From Proposition 18, we have $\lim_{j\to\infty} \|x^{k_j+1} - x^{k_j}\| = 0$, i.e., $\lim_{j\to\infty} x^{k_j+1} = \lim_{j\to\infty} x^{k_j} = x^*$.

Now, for any $y \in C(x^*)$, again from the M-continuity, there exists a sequence $\{y^{k_j}\}$ such that $y^{k_j} \in C(x^{k_j})$ and $y^{k_j} \to y$. Since, $x^{k_j+1} \in S_{\mathsf{EP}_{\delta_{k_j}}}(f^e_{k_j}, C_{k_j})$ we have, $f^e_{k_j}(x^{k_j+1}, z) \geqslant -\delta_{k_j}$, for all $z \in C(x^{k_j})$, that implies in,

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, z) + \gamma_{k_j}\langle x^{k_j+1} - x^{k_j}, z - x^{k_j+1}\rangle - \langle e^{k_j}, z - x^{k_j+1}\rangle, \quad \forall z \in C(x^{k_j}). \tag{4.31}$$

First, we will prove that results holds using that $\{e^k\}$ satisfies condition $(\mathsf{E1})$. Since (4.31) holds for all $z \in C(x^{k_j})$, taking $z = y^{k_j}$, one has

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\langle x^{k_j+1} - x^{k_j}, y^{k_j} - x^{k_j+1}\rangle - \langle e^{k_j}, y^{k_j} - x^{k_j+1}\rangle,$$

from Cauchy-Schwartz inequality,

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|x^{k_j+1} - x^{k_j}\|\|y^{k_j} - x^{k_j+1}\| + \|e^{k_j}\|\|x^{k_j+1} - y^{k_j}\|.$$

Now, due assertion $(E1)$, we get

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|x^{k_j+1} - x^{k_j}\|\|y^{k_j} - x^{k_j+1}\| + \|x^{k_j+1} - x^{k_j}\|x^{k_j+1} - y^{k_j}\|.$$

Hence, taking the limit as $j \to \infty$ in the last inequality above, since $\{\gamma_k\}$, $\{x_k\}$ and $\{y_k\}$ are bounded, we have that

$$f(x^*, y) \geqslant 0, \quad \forall y \in C(x^*),$$

have in mind that C is M-continuous. This means that $x^* \in S_{QEP}(f, C)$.

Now, we will prove that if the error $\{e^k\}$ satisfies condition $(E2)$, the results holds. From $f^e_{k_j}(x^{k_j+1}, z) \geqslant -\delta_{k_j}$, with $z = y^{k_j}$, and using Cauchy-Schwartz inequality, we get

$$\begin{aligned} -\delta_{k_j} &\leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\langle x^{k_j+1} - x^{k_j} - e^{k_j}, y^{k_j} - x^{k_j+1}\rangle \\ &\leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|x^{k_j+1} - x^{k_j} - e^{k_j}\|\|y^{k_j} - x^{k_j+1}\|. \end{aligned} \tag{4.32}$$

By condition $(E2)$, if $\|x^{k_j+1} - x^{k_j} - e^{k_j}\| \leqslant \|x^{k_j+1} - x^{k_j}\|$, then from (4.32)

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|x^{k_j+1} - x^{k_j}\|\|y^{k_j} - x^{k_j+1}\|,$$

otherwise, if $\|x^{k_j+1} - x^{k_j} - e^{k_j}\| \leqslant \|e^{k_j}\|$, then from (4.32)

$$-\delta_{k_j} \leqslant f(x^{k_j+1}, y^{k_j}) + \gamma_{k_j}\|e^{k_j}\|\|y^{k_j} - x^{k_j+1}\|.$$

Taking the limit as $j \to \infty$, and taking into account that the sequences $\{\gamma_k\}$,$\{x_k\}$ and $\{y_k\}$ are bounded, and the fact that the sequence $\{\delta_k\}$ is summable, i.e., $\lim_{k\to\infty} \delta_k = 0$; also from Proposition 18, $\lim_{k\to\infty} \|x^{k_j+1} - x^{k_j}\| = 0$ and $\lim_{k\to\infty} \|e^k\| = 0$, we obtain that

$$f(x^*, y) \geqslant 0, \quad \forall y \in C(x^*).$$

This, means that $x^* \in S_{QEP}(f, C)$. □

Now, we will show the convergence results for Algorithm 8. Here, we will consider only the case where S^* is non-empty.

Proposition 19. Let $\{x^k\}$ be the sequence generated by Algorithm 8. Then, the following assertions holds:

(i) $\{x^k\}$ is quasi-Fejér convergent to S^*;

(ii) $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$ and $\lim_{k\to\infty} \|v^k - x^k\| = 0$.

Proof. Item (i), take $x^* \in S^*$ any. Thus, $f(x^*, \bar{y}) \geqslant 0$, for all $y \in \bigcup_{z \in X} C(z)$; due Q4, $f(\bar{y}, x^*) \leqslant 0$, for all $y \in C(z), z \in X$. In particular, for $\bar{y} = v^k$, i.e., $f(v^k, x^*) \leqslant 0$, for all $k \in \mathbb{N}$. Now, by definition of Algorithm 8, $v^k \in S_{EP}(f_k, C_k)$, then

$$f_k(v^k, y) = f(v^k, y) + \gamma_k \langle v^k - x^k, y - v^k \rangle \geqslant 0, \quad \forall \ \ y \in C_k.$$

Setting, $y = x^*$,

$$f(v^k, x^*) + \gamma_k \langle v^k - x^k, x^* - v^k \rangle \geqslant 0,$$

this implies,

$$\langle v^k - x^k, x^* - v^k \rangle \geqslant -\frac{f(v^k, x^*)}{\gamma_k} \geqslant 0. \tag{4.33}$$

Because, $\gamma_k > 0$ and $f(v^k, x^*) \leqslant 0$.

Thus,

$$\begin{aligned} \|x^k - x^*\|^2 &= \|(x^k - v^k) + (v^k - x^*)\|^2 \\ &= \|x^k - v^k\|^2 + \|v^k - x^*\|^2 + 2\langle x^k - v^k, v^k - x^* \rangle \\ &\geqslant \|x^k - v^k\|^2 + \|v^k - x^*\|^2 \\ &\geqslant \|v^k - x^*\|^2. \end{aligned} \tag{4.34}$$

Where, in the first inequality comes from (4.33) and last inequality comes from of the fact that $\|v^k - x^*\|^2 \geqslant 0$. Then,

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|x^{k+1} - v^k + v^k - x^*\| \\ &\leqslant \|x^{k+1} - v^k\| + \|v^k - x^*\| \\ &\leqslant \varepsilon_k + \|x^k - x^*\|. \end{aligned} \tag{4.35}$$

We used triangular inequality in the first inequality, in the second inequality we used (4.34) and condition $\|x^{k+1} - v^k\| \leqslant \varepsilon_k$. Therefore, from (4.35) we obtain that $\{x^k\}$ is quasi-Fejér convergent to S^* and we prove the first assertion.

To prove item (ii), from Lemma 1 we have that the sequence $\{\|x^k - x^*\|\}$ converges, and since x^* is any, $\{\|x^k - x^*\|\}$ converges for all $x^* \in S^*$. Let $\mu > 0$ be the limit of the $\{\|x^k - x^*\|\}$. Note that, from (4.35) and (4.34)

$$\|x^{k+1} - x^*\| \leqslant \|v^k - x^*\| + \varepsilon_k \leqslant \|x^k - x^*\| + \varepsilon_k.$$

Then taking the limit as $k \to \infty$, and taking into account that $\lim_{k\to\infty} \varepsilon_k = 0$, we obtain that $\lim_{k\to\infty} \|v^k - x^*\| = \mu$.

Now, note that from (4.34) we have

$$\|x^k - x^*\|^2 \geqslant \|x^k - v^k\|^2 + \|v^k - x^*\|^2 \geqslant \|v^k - x^*\|^2.$$

Letting k goes to $+\infty$, and taking into account that $\lim_{k\to\infty} \|v^k - x^*\| = \lim_{k\to\infty} \|x^k - x^*\| = \mu$, we obtain that $\lim_{k\to\infty} \|x^k - v^k\| = 0$.

On other hand, using triangular inequality and the fact that $\|x^{k+1} - v^k\| \leqslant \varepsilon_k$, we have

$$0 \leqslant \|x^{k+1} - x^k\| \leqslant \|x^{k+1} - v^k\| + \|v^k - x^k\| \leqslant \varepsilon_k + \|v^k - x^k\|.$$

Taking the limit in inequality above as $k \to \infty$, taking into account that $\lim_{k\to\infty} \|x^k - v^k\| = 0$ and $\lim_{k\to\infty} \varepsilon_k = 0$, we obtain that

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0,$$

and we finish the proof. □

Theorem 22. Let $\{x^k\}$ be the sequence generated by Algorithm 8. Then every cluster point of $\{x^k\}$ is a solution to the $\mathsf{QEP}(f, C)$.

Proof. From Proposition 19, we have that $\{x^k\}$ is quasi-Fejér convergent to S^*. Thus from Lemma 1 $\{x^k\}$ is bounded. Let $\{x^{k_j}\}$ be subsequence of $\{x^k\}$ that converges to x^*, since Proposition 16 (ii) holds, we have $\lim_{j\to\infty} \|x^{k_j+1} - x^{k_j}\| = 0$, and $\lim_{j\to\infty} \|v^{k_j} - x^{k_j}\| = 0$ we obtain that $x^{k_j+1} \to x^*$, and $v^{k_j} \to x^*$. Now, by definition of the algorithm $v^k \in S_{\mathsf{EP}}(f_k, C_k)$, for all $k \in \mathbb{N}$, and in particular, $v^{k_j} \in C(x^{k_j})$. Therefore, since C is M-continuous, we obtain that $x^* \in C(x^*)$. On other hand,

$$\begin{aligned} 0 &\leqslant f(v^{k_j}, y) + \gamma_{k_j} \langle v^{k_j} - x^{k_j}, y - v^{k_j} \rangle \\ &\leqslant f(v^{k_j}, y) + \gamma_{k_j} \|v^{k_j} - x^{k_j}\| \|y - v^{k_j}\|, \quad \forall y \in C(x_{k_j}). \end{aligned} \tag{4.36}$$

Since the sequences $\{\gamma_k\}$,$\{x_k\}$ and $\{v_k\}$ are bounded, and $\lim_{j\to\infty} \|x^{k_j} - v^{k_j}\| = 0$. Taking the limit as $k \to \infty$ in (4.36) we have,

$$f(x^*, y) \geqslant 0, \quad \forall y \in C(x^*),$$

having in mind that C is M-continuous. This means that $x^* \in S_{\mathsf{QEP}}(f, C)$. □

4.3 Numerical experiments

In this section, we illustrate the performance of the proposed method on one test problem from Santos and Souza [34]. The first example is a QEP version which is originally an equilibrium problem and we replace the constraint set by a point-to-set mapping C. The second example is a quasi-variational inequality formulation of a generalized Nash game, both given by Strodiot et al. [37]. The algorithm is coded in JULIA v.1.11 on a 8 GB RAM Intel Core i5-8265U, 1.60GHz and GPU NVIDIA 2GB to obtain the numerical results. The stopping rule is $\|x^{k+1} - x^k\| \leqslant \mathtt{tol}$. We take $\gamma_k = \gamma$, for all $k \in \mathbb{N}$. We solve the subproblem using the regularized method in Muu and Quoc [26], they consider the following iterative method for solving equilibrium problems: for any point $x^0 \in X$ and $\gamma > 0$, given x^k define $x^{k+1} \in X$ such that

$$x^{k+1} = \mathrm{argmin}_{y \in X}\{\gamma f(x^k, y) + \frac{1}{2}\|y - x^k\|^2\}. \tag{4.37}$$

The solutions of the subproblem in (4.37) are computed by a built-in JULIA using package Convex.jl and solver: "ECOS.Optimizer". It is worth mentioning that in [34] the numerical experiments were made using the programming language MATLAB. Here we will show that if we use other programming language, the results still holds. Next we present examples and results.

Example 5. [34, Example 1] Consider the 2-dimensional non-smooth quasi-equilibrium problem defined by the bifunction

$$f(x, y) = |y_1| - |x_1| + y_2^2 - x_2^2$$

and the multi-valued mapping C given by

$$C(x) = \{y \in \mathbb{R}^2_+; y_1 + y_2 = 1 + \frac{|x_1|}{1 + |x_1|}\}.$$

One can check that f is monotone and the solution set is the single point $x^* = (1, \frac{1}{2})$.

In Table 4.1 we present the result of algorithm with $x_0 = (0, 0)$ and the values for $\gamma = 1.0$, $\gamma = 3.5$ and $\gamma = 5.5$. The columns present the value of γ, the CPU time in seconds and the number of iterations until the stopping rules are satisfied and the accuracy $\Upsilon^* = \Upsilon(x^k)$(Proposition 13). In table 4.2 we report the results of 100 running of the method with random initial points in the box $[-5, 5]^2$. The columns present the

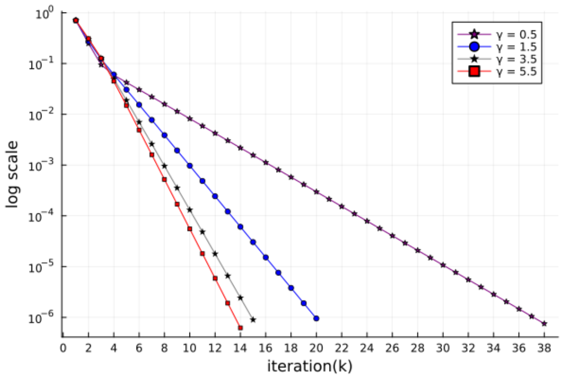


Figure 4.1: Accuracy $\Upsilon(x^k)$ in each iteration (using log scale) for Ex. 5.

minimum, maximum and mean of iterations and CPU time until the stopping rule are satisfied as well as the average of the accuracy Υ^*, in the last iterate x^k.

Table 4.1: Results for Example 5 with fix $x_0 = (0,0)$, stop rules $\mathtt{tol} = 10^{-6}$, and values of γ given by $\gamma = 0.5,\ \gamma = 1.5,\ \gamma = 3.5$ and $\gamma = 5.5$.

γ	CPU(s)	iter.(k)	Υ^*
0.5	0.2360	38	9.7908111439011e-7
1.5	0.0659	20	9.3823047066381e-7
3.5	0.0600	15	8.7296893006286e-7
5.5	0.0589	14	6.1999109132859e-7

Figure 4.1 and Table 4.1 demonstrate that a larger value of γ results in fewer iterations and higher accuracy.

Table 4.2: Results for Example 5 with 100 random points in $[-5,5]^2, \gamma = 1$ and $\mathtt{tol}_{qep} = 10^{-6}$.

min iter.(k)	max iter(k)	aver. iter(k)	min CPU(s)	aver. CPU(s)	aver. Υ^*
17	29	25.76	0.017	0.237	7.4934904e-7

Example 6. [34, Example 3] Consider the 2-dimensional quasi-variational inequality problem given by $F(x) = (2x_1 + \frac{8}{3}x_2 - 34,\ 2x_2 + \frac{5}{4}x_1 - 24.25)$ and the valued mapping defined by $C(x) = C_1(x_2) \times C_2(x_1)$, where $C_1(x_2) = \{y_1 \in \mathbb{R}; 0 \leqslant y_1 \leqslant 10, y_1 \leqslant 15 - x_2\}$ and $C_2(x_1) = \{y_2 \in \mathbb{R}; 0 \leqslant y_2 \leqslant 10, y_2 \leqslant 15 - x_1\}$. Its solution set is the point $(5,9)$

and the line segment [(9,6), (10,5)]. One can check the that the equilibrium bifunction is strictly monotone .

The numerical results that are presented in the Tables 4.3 and 4.4 we report the performance of our method for the same initial points [37], in the columns we can see the solution found, the number of iterations and CPU time until stopping rules is satisfied and the accuracy. In Table 4.4 we present the performance for 100 running using random initial points.

The Figures 4.2a, 4.2b and 4.2c present the accuracy (in log scale) at each iteration using the the initial points $x_0 = (0,0)$, $x_0 = (5,5)$ and $x_0 = (-2.264513, 6.911972)$ as well as the values of γ given by $\gamma = 0.15, \gamma = 0.25, \gamma = 0.45, \gamma = 0.5$.

Then, it is possible to conclude that the algorithm's performance is strongly dependent on the value of γ rather than the value of x_0.

Table 4.3: Example 6 with the same initial point as in [34], $\mathtt{tol} = 10^{-4}$ and $\gamma = 0.45$.

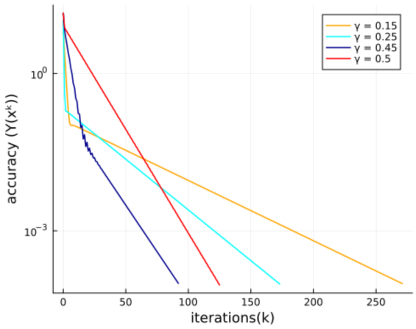
starting point	solution found	iterations(k)	CPU time	Accuracy Υ^*
(0,0)	(5,9)	93	0.34100	9.963879476514962e-5
(10,0)	(10,5)	2	0.01200	1.475667130978465e-7
(10,10)	(5,9)	92	0.44800	9.963878781336826e-5
(0,10)	(5,9)	92	0.32600	9.382028252290837e-5
(5,5)	(5,9)	99	0.35199	9.285729671393152e-5

Table 4.4: Results for Example 6 with 100 random points in the box $[-10, 10]^2$, $\mathtt{tol} = 10^{-4}$ and $\gamma = 0.45$.

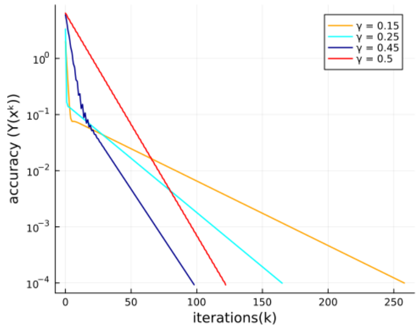
min iter(k)	max iter(k)	aver. iter(k)	min CPU(s)	aver. CPU(s)	aver. Υ^*
42	107	93.81	0.148	0.3285	9.74388330e-5

Figure 4.2: Example 6: accuracy (using log scale) in each iteration k, with $\texttt{tol} = 10^{-4}$ and initial points $x_0 = (0,0)$, $x_0 = (5,5)$ and $x_0 = (-2.264513, 6.911972)$.

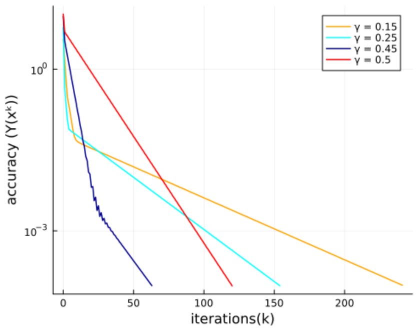
(a) $x_0 = (0,0)$.



(b) $x_0 = (5,5)$.



(c) $x_0 = (-2.264513, 6.911972)$



Chapter 5

Application to the Cournot model

A duopoly is a specific market structure where only two companies dominate the production and supply of a particular product or service. These two companies hold the majority of the market in question. In a duopoly, companies must consider their competitor's actions when making decisions, as one company's actions affect the other. Having this in mind, the Cournot model is a classical model for the analysis of a duopoly, and it is based on the fact that the two companies compete by choosing the quantity of output they will produce. The results and concepts presented along this chapter can be found in E. Junior et al. [18].

5.1 The viability condition of Cournot model as an EP

The Cournot model of duopoly has its a origin in a game theory, see Cournot [7], and it has been modeled as an equilibrium problem; see Orsborne [28]. Let us see that we can improve its presentation as an quasi-equilibrium problem through its presentation as Nash-Cournot game with a unique shared constraint, see in [9] for more details.

In a Cournot model for duopoly, it considers two producers $j = 1, 2$ an of a homogeneous good that produce the quantities $x = (x^1, x^2) \in \mathbb{R}^2_+$ of this good at a unit cost $(c^1, c^2) \in \mathbb{R}^2_+$, and in usual presentation we suppose that unit cost of production are the same, $c^1 = c^2 = c > 0$. The unit price at which this good is sold is given by

$$p(Q) = \begin{cases} a - Q \in \mathbb{R}, & \text{if } \ 0 \leqslant Q \leqslant a \\ 0, & \text{if } \ Q > a, \end{cases}$$

where $a > 0$ is the hightest price at which one unit price can be sold when Q is close to zero. Thus, we can see that this price depends on total level of production of the duopoly $Q = x^1 + x^2$. In the other words, the greater production (Q hight), the lower the price $p(Q)$ will be. The existence of this duopoly depends of the fact that the producers do not produce too much in total, because, otherwise if $Q > a$ the price go to $p(Q) = 0$, i.e, the duopoly break (because the goods sold at a negative price or zero are not rewarding).

We will see that a Cournot equilibrium x^* satisfies the conditions of QEP, i.e., $x^* \in C(x^*)$ and $f(x^*, y) \geqslant 0$, for all $y \in C(x^*)$.

5.2 Setting the fixed point condition of a QEP as a viability condition

The viability constraint of the duopoly is that the price of the final good will be non-negative, i.e., $p(Q) \in \mathbb{R}_+$, this means that, $p(x) = a - (x^1 + x^2) \geqslant 0$. Then, $x = (x^1, x^2) \in K$, where

$$K = \{x = (x^1, x^2) \in \mathbb{R}^2_+, a - (x^1 + x^2) \geqslant 0\}$$

is the two dimensional simplex if $a = 1$.

Denote by $C^j(x^{-j}) = \{x^j \in \mathbb{R}_+, 0 \leqslant x^j \leqslant a - x^j\}, j = 1, 2$ and $C(x) = C^1(x^2) \times C^2(x^1)$. It is easy to show that $x \in C(x)$ is equivalent to $x^j \in C^j(x^{-j}), j = 1, 2$, i.e., $x \in C(x)$, and it is equivalent to $x = (x^1, x^2) \in K$. The term shared constraint comes from the fact that a unique constraint K (i.e., a non-negative selling price) defines all the moving sets $C^j(x^{-j})$.

5.3 Setting the equilibrium condition of a QEP

The margins of producers $i = 1, 2$ are the difference $m^i = p - c^i, i = 1, 2$, between the price of the good and their unit cost of production and selling with $c^1 = c^2 = c > 0$. Then, at the status quo $x = (x^1, x^2) \in \mathbb{R}^2_+$ their profits are, for $i = 1, 2$,

$$\pi^i = \begin{cases} px^i - cx^i = (p - c)x^i = [(a - c - x^1 - x^2)]x^i, & \text{if } x^1 + x^2 \leqslant a, \\ -cx^i, & \text{if } x^1 + x^2 > a. \end{cases}$$

Suppose now that, with the aim to maximize their profit each producer i consider unilaterally changing his production from the status quo. This means that he hopes that the other producer will not change, producing at the same level as before. Hence, each duopolist maximizes his profit, taking into consideration that the other duopolist produces and sells at the status quo $x = (x^1, x^2) \in \mathbb{R}^2_+$. Then, if the price of the good is non-negative, i.e., if the viability constraint $p = a - (x^1 - x^2) \geqslant 0$ is satisfied, they solve the unconstrained interrelated programs

$$\max\{\pi^1(y^1, x^2) = [a - c - x^2 - y^1]y^1; \quad y^1 \in X^1 = \mathbb{R}^2\}$$

and

$$\max\{\pi^2(x^1, y^2) = [a - c - x^1 - y^2]y^2; \quad y^2 \in X^2 = \mathbb{R}^2\}.$$

Their resolution provides the best responses $x^{*,1} = \frac{1}{2}[a - c^1 - x^2]$ and $x^{*,2} = \frac{1}{2}[a - c^2 - x^1]$. Then, the Cournot equilibrium is a fixed point $x^* = (x^{*,1}, x^{*,2}) \in \mathbb{R}^2_+$. This gives us, if $a > c$,

$$x^* = (x^{*,1}, x^{*,2}) = \frac{1}{3}(a - c, a - c).$$

In turn, we can verify that $x^* \in K$. Notice that the definition of a constrained Nash equilibrium, with the viability constraint K means that the two following equilibrium conditions are satisfied:

$$\pi^1(x^{*,1}, x^{*,2}) \geqslant \pi^1(y^1, x^{*,2}), \quad \forall y^1 \in C^1(x^{*,2}) \tag{5.1}$$

$$\pi^2(x^{*,1}, x^{*,2}) \geqslant \pi^2(x^{*,1}, y^2), \quad \forall y^2 \in C^2(x^{*,1}), \tag{5.2}$$

where

$$C^1(x^{*,2}) = \{y^1 \in \mathbb{R}_+, a - c - y^1 - x^{*,2} \geqslant 0\}$$

and

$$C^2(x^{*,1}) = \{y^2 \in \mathbb{R}_+, a - c - x^{*,1} - y^2 \geqslant 0\}.$$

Then the equilibrium condition of the QEP $f(x^*, y) \geqslant 0$, for all $y \in C(x^*)$, it follows from the definition of a Nikaido-Isoda equilibrium function,

$$f(x, y) = [\pi^1(x^1, x^2) - \pi^1(y^1, x^2)] + [\pi^2(x^1, x^2) - \pi^2(x^1, y^2)]$$

with the conditions (5.1) and (5.2).

5.4 A better formulation of the Nash-Cournot model as a QEP

In this section, we reiterate the advantage of modeling the Cournot example as a quasi-equilibrium problem (QEP), rather than a traditional equilibrium problem. Furthermore, we introduce the linear and inexact regularization method. This method acts as a perturbation to the unit cost of the duopolists and can result in potentially better convergence. This is an important result, as it allows for inexact knowledge of the cost in each period, which refines over time, if the error is adequately controlled.

Considering the new presentation of the Cournot model as a QEP detailed in section 5.3 with the same unit cost of production $c^1 = c^2 = c > 0$, the new presentation the Cournot model as a QEP with different unit costs $(c^1, c^2) > 0$, $c^1 \neq c^2$ comes easily if we recall the definition of $p(Q)$,

$$p(Q) = \begin{cases} a - Q \in \mathbb{R}, & \mathtt{if} \;\; 0 \leqslant Q \leqslant a \\ 0, & \mathtt{if} \;\; Q > a. \end{cases}$$

Hence, if the unit cost of production of the producers are different the margins are not the same. And as the margin is the difference between the price and its cost of production and selling of the good, $m^i = p - c^i, i = 1, 2$. Then the profits of the duopolists are

$$\pi^i(x^i, x^{-i}) = px^i - c^i x^i = (p - c^i)x^i = m^i x^i, \quad i = 1, 2.$$

Thus, their profits are,

$$\pi^1 = \begin{cases} [(a - c^1 - x^2) - x^1]x^1, & \text{if} \;\; x^1 + x^2 \leqslant a; \\ -c^1 x^1, & \text{if} \;\; x^1 + x^2 > a. \end{cases}$$

and

$$\pi^2 = \begin{cases} [(a - c^2 - x^1) - x^2]x^2, & \text{if} \;\; x^1 + x^2 \leqslant a; \\ -c^2 x^2, & \text{if} \;\; x^1 + x^2 > a. \end{cases}$$

5.4.1 New viability conditions

The first viability condition for the existence of a duopoly is that the selling price $p = a - (x^1 + x^2)$ is non-negative. Two new viability conditions that imply this first one require that margins must be non-negatives, in other words,

$$m^1 = \alpha^1(x^2) - x^1 \geqslant 0 \;\; \text{and} \;\; m^2 = \alpha^2(x^1) - x^2 \geqslant 0,$$

with $\alpha^1(x^2) = a - c^1 - x^2$ and $\alpha^2(x^1) = a - c^2 - x^1$. Otherwise, all profits may be negative.

As seen previously, each duopolist maximizes his profit, taking into account what the other duopolist produces at the status quo $x = (x^1, x^2) \in \mathbb{R}^2_+$. Then, they solves the unconstrained interrelated programs

$$\max\{\pi^1(y^1, x^2) = [\alpha^1(x^2) - y^1]y^1, \quad y^1 \in X^1 = \mathbb{R}_+\},$$

and

$$\max\{\pi^2(x^1, y^2) = [\alpha^2(x^1) - y^2]y^2, \quad y^2 \in X^2 = \mathbb{R}_+\}.$$

Their resolutions provide the unique equilibrium

$$x^* = (x^{*,1}, x^{*,2}) = \frac{1}{3}(a + c^1 - 2c^2, a + c^2 - 2c^1),$$

with $a + c^1 - 2c^2 > 0$ and $a + c^2 - 2c^1 > 0$. If $c^1 = c^2 = c > 0$, then $x^{*,1} = x^{*,2} = \frac{1}{3}(a - c)$.

5.5 The Cournot model as a QEP

Set of profitable deviations. Two viability constraints are that the profit of each duopolist must not be negative when they try to move unilaterally from the status quo profile of production $x = (x^1, x^2)$ to another profile of production $y = (y^1, y^2) \in \mathbb{R}^2_+$. These constraints are, $y^1 \in C^1(x^2)$ and $y^2 \in C^2(x^1)$ where,

$$C^1(x^2) = \{y^1 \in \mathbb{R}_+, \pi^1(y^1, x^2) = [\alpha^1(x^2) - y^1]y^1 \geqslant 0\}$$

and

$$C^2(x^1) = \{y^2 \in \mathbb{R}_+, \pi^2(x^1, y^2) = [\alpha^2(x^1) - y^2]y^2 \geqslant 0\}.$$

These two sets represent the sets of unilateral profitable deviations (y^1, x^2) and (x^1, y^2) from the status quo x. Then, two viability (profitability) constraints can be written

$$y^1 \in C^1(x^2) = \{y^1, 0 \leqslant y^1 \leqslant \alpha^1(x^2)\},$$

and

$$y^2 \in C^2(x^1) = \{y^2, 0 \leqslant y^2 \leqslant \alpha^2(x^1)\}.$$

i.e., $y \in C(x)$ if $y = (y^1, y^2)$ and $C(x) = C^1(x^2) \times C^2(x^1)$, i.e., $y \in C(x)$ if $0 \leqslant y^1 \leqslant \alpha^1(x^2) = a - c^1 - x^2$ and $0 \leqslant y^2 \leqslant \alpha^2(x^1) = a - c^2 - x^1$. If $c^1 = c^2 = c > 0$, then $y \in C(x)$ is equivalent to $0 \leqslant y^1 \leqslant a - c - x^2$ and $0 \leqslant y^2 \leqslant a - c - x^1$.

Several viability constraints: making non-negative profits at the status quo. As a consequence the condition $x^* \in C(x^*)$ we have that is can be viable to stay at the status quo, because each duopolist makes a non-negative profit. More explicitly, $x^* \in C(x^*)$ is equivalent to $0 \leqslant x^{*,1} \leqslant \alpha^1(x^{*,2}) = a - c^1 - x^{*,2}$ and $0 \leqslant x^{*,2} \leqslant \alpha^2(x^{*,1}) = a - c^2 - x^{*,1}$, i.e., $\pi^1(x^{*,1}, x^{*,2}) \geqslant 0$ and $\pi^2(x^{*,1}, x^{*,2}) \geqslant 0$. This refers to the existence of two viability conditions instead of one when $c^1 = c^2 = c > 0$.

Equilibrium: stability of the status quo with different moving constraints. The status quo $x^* = (x^{*,1}, x^{*,2})$ is an equilibrium is it is not profitable to deviate unilaterally from it,i.e., iff

$$\pi^1(y^1, x^{*,2}) \leqslant \pi^1(x^{*,1}, x^{*,2}), \quad \forall y^1 \in C^1(x^{*,2}),$$

and

$$\pi^2(x^{*,1}, y^2) \leqslant \pi^2(x^{*,1}, x^{*,2}), \quad \forall y^2 \in C^2(x^{*,1}).$$

A better QEP formulation of the Cournot equilibrium problem. It becomes: find $x^* \in C(x^*)$ such that $f(x^*, y) \geqslant 0$, for all $y \in C(x^*)$ with a Nikaido-Isoda function

$$f(x, y) = [\pi^1(x^1, x^2) - \pi^1(y^1, x^2)] + [\pi^2(x^1, x^2) - \pi^2(x^1, y^2)].$$

5.6 The advantages of a QEP formulation

This leads to a more comprehensive definition of the status quo as an equilibrium, one that considers the fact that

(i) Viability of Staying: If producers decide to maintain the status quo, this is only viable if each duopolist makes a non-negative profit. This is a viability constraint that, interestingly, many traditional presentations of the model ignore.

(ii) Stability of Changing: If one producer decides to change its production, this change will not be profitable for that firm if the other firm does not also change its production. This is a stability constraint. Although traditional presentations use this idea, they fail to consider the local and moving constraints $y^1 \in C^1(x^{*,2})$ and $y^2 \in C^2(x^{*,1})$. This highlights a fundamental point: if you assume that other duopolist actions will constrain your own choice set, it is essential to incorporate this reality into the analysis. Traditional presentation, miss this nuance.

A better understanding of the generalized Nash equilibrium problem

Existence of QEP. Quasi-equilibrium problems ($QEPs$) are a generalization of specific cases, as generalized Nash equilibrium problems. For examples of games with shared constraints, you can see for more details Dutang [9] and Fischer et al. [12]. Our new and simplified perspective on the Cournot model allows us to better understand the intuitions behind the sufficient conditions given for the existence of equilibria in games with shared constraints.

5.6.1 A linear and inexact regularization of a QEP leads to a striking interpretation of the Cournot model

Linear regularization. As proposed in chapter 4 a linear regularization of the function $f(x, y)$. In Cournot example,

$$f(x, y) = [\pi^1(x^1, x^2) - \pi^1(y^1, x^2)] + [\pi^2(x^1, x^2) - \pi^2(x^1, y^2)],$$

and

$$\langle e, y - x \rangle = e^1(y^1 - x^1) + e^2(y^2 - x^2).$$

Then, for duopolist 1,

$$\pi^1(y^1, x^2) = (a - c^1 - x^2)y^1 - (y^1)^2$$

and

$$\pi^1(x^1, x^2) = (a - c^1 - x^2)x^1 - (x^1)^2,$$

give $\pi^1(y^1, x^2) - \pi^1(x^1, x^2) = (a - c^1 - x^2)(y^1 - x^1) - [(y^1)^2 - (x^1)^2]$. Then,

$$\pi^1(y^1, x^2) - \pi^1(x^1, x^2) - e^1(y^1 - x^1) = [a - (c^1 + e^1) - x^2](y^1 - x^1) - [(y^1)^2 - (x^1)^2].$$

And is valid the same for duopolist 2:

$$\pi^2(y^1, y^2) - \pi^2(x^1, x^2) - e^1(y^2 - x^2) = [a - (c^2 + e^2) - x^1](y^2 - x^2) - [(y^2)^2 - (x^2)^2].$$

The model QEP of the Cournot model becomes of the fact: find $x^* \in X$ such that $x^* \in C(x^*)$ with $f(x^*, y) \geqslant 0$, for all $y \in C(x^*)$.

Final interpretation: The previous calculations show that a linear and inexact perturbation represents a perturbation of the unit cost of each duopolist. This is a important result in behavioral sciences. Indeed, we show that a linear and inexact perturbation allows:

(i) to obtain a better computational performance than the exact version;

(ii) to the inexact knowledge of the unit cost which it is refines over time as a controlled error.

Chapter 6

Conclusion

In this work, we proposed a study of the proximal point method for solving quasi- equilibrium problems in its exact and inexact versions. Furthermore, we saw two varia- tions of the proximal point method used for solving the equilibrium problems, the first one using Bregman regularization and the second one using separable Bregman regularization. We have seen that the quasi-equilibrium has several applications. And here, we applied the concepts and results of quasi-equilibrium in economics to propose an interpretation of the Cournot duopoly as a quasi-equilibrium. Regarding future work, our idea is to continue the research on quasi-equilibrium problems, proposing an inexact version of the separable Bregman regularization in the context of quasi-equilibrium. Where, we consider a kind of linearized error similar to the inexact version shown in this work, and it proved that this function parametrized satisfies assumptions 2 and hence, the solution set of the equilibrium problem is nonempty, demonstrating that the algorithm is well-defined. It is worth mentioning that the condition of coercivity considered is better than the condition considered in the exact version. In convergence analysis, we prove that considering an error summable, the sequence generated by the algorithm has cluster points, and all of them belong to the solution set of QEP. Further- more, we highlight that the criterion of summability is the same considered by Eckstein and several others works.

Bibliography

[1] S.Reich, S. Sabach - *Two strong convergence theorems for Bregman strongly non-expansive operators in reflexive Banach spaces.* Nonlinear Anal., 73 (2010), 122-135.

[2] Cruz Neto, J.X., Santos, P.S.M., Silva, R.C.M. et al. *On a Bregman regularized proximal point method for solving equilibrium problems.* Optim. Letters, 13 (2019), 1143–1155.

[3] Bigi, G., and Passacantando, M. - *Gap functions for quasi-equilibria.* Journal of Global Optimization 66, 4 (2016), 791–810.

[4] Chen, G., Teboulle, M., - *Convergence analysis of proximal-like optimization algorithm using Bregman functions*, SIAM J. Optim., 3 (1993), 538-543.

[5] Burachik, R., and Kassay, G. - *On a generalized proximal point method for solving equilibrium problems in Banach spaces.* Nonlinear Analysis: Theory, Methods & Applications 75, 18 (2012), 6456–6464.

[6] Ceng, L.-C., and Yao, J.-C. *Approximate proximal methods in vector optimization.* European Journal of Operational Research 183, 1, (2007), 1–19.

[7] Cournot, A. - *Researches into the mathematical principles of the theory of wealth. In Forerunners of Realizable Values Accounting in Financial Reporting.* Routledge, (2020), 3–13.

[8] Djafari Rouhani, B., and Mohebbi, V. - *Proximal point method for quasi-equilibrium problems in Banach spaces.* Numerical Functional Analysis and Optimization 41, 9, (2020), 1007–1026.

[9] Dutang, C. - *Existence theorems for generalized Nash equilibrium problems: An analysis of assumptions.* Journal of Nonlinear Analysis and Optimization 4, 2 (2013), 115–126.

[10] Eckstein, J. - *Approximate iterations in Bregman-function-based proximal algorithms.* Mathematical programming, 83, (1998), 113–123.

[11] Facchinei, F., Kanzow, C., Karl, S., and Sagratella, S. - *The semi smooth newton method for the solution of quasi-variational inequalities.* Computational Optimization and Applications, 62, (2015), 85–109.

[12] Fischer, A., Herrich, M., and Schönefeld, K. - *Generalized Nash equilibrium problems-recent advances and challenges.* Pesquisa Operacional, 34, (2014), 521–558.

[13] Han, D., Zhang, H., Qian, G., and Xu, L.- *An improved two-step method for solving generalized Nash equilibrium problems.* European Journal of Operational Research 216, 3 (2012), 613–623.

[14] Humes Jr, C., and Silva, P. J. *Inexact proximal point algorithms and descent methods in optimization.* Optimization and Engineering 6, 2 (2005), 257–271.

[15] De Pierro, A., Iusem, A. *A relaxed version of Bregman's method for convex programming.* Journal of Optimization Theory and Applications, 51 (1986), 421-440.

[16] Iusem, A. N. *On the maximal monotonicity of diagonal subdifferential operators.* In AIP Conference Proceedings, vol. 1281, American Institute of Physics, (2010), 1705–1706.

[17] Iusem, A. N., Kassay, G., and Sosa, W. *On certain conditions for the existence of solutions of equilibrium problems.* Mathematical Programming, 116 (2009), 259-273.

[18] Júnior, E., Santos, P., Soubeyran, A., et al. *On inexact versions of a quasi-equilibrium problem: a Cournot duopoly perspective.* Journal of Global Optimization, 89 (2023) (1), 171-196.

[19] Iusem, A. N., and Nasri, M. *Inexact proximal point methods for equilibrium problems in Banach spaces.* Numerical Functional Analysis and Optimization 28, 11-12 (2007), 1279–1308.

[20] Iusem, A. N., and Sosa, W. *On the proximal point method for equilibrium problems in Hilbert spaces.* Optimization 59, 8 (2010), 1259–1274.

[21] Izmailov, A., and Solodov, M. *Otimização, volume 1: condições de otimalidade, elementos de análise convexa e de dualidade.* Impa, 2005.

[22] Kaplan, A., and Tichatschke, R. *On inexact generalized proximal methods with a weakened error tolerance criterion.* Optimization 53, 1 (2004), 3–17.

[23] Konnov, I. *Application of the proximal point method to nonmonotone equilibrium problems.* Journal of Optimization Theory and Applications 119 (2003), 317–333.

[24] Konnov, I. *Equilibrium models and variational inequalities.* Elsevier, 2007.

[25] Mohammadi, M., and Eskandani, G. Z. *Approximation solutions of quasi-equilibrium problems in Banach spaces.* Miskolc Mathematical Notes 21, 1 (2020), 261–272.

[26] Muu, L. D., and Quoc, T. D. *Regularization algorithms for solving monotone Ky-Fan inequalities with application to a Nash-Cournot equilibrium model.* Journal of optimization theory and applications 142 (2009), 185–204.

[27] Nikaidô, H., and Isoda, K. *Note on non-cooperative convex games.* Pacific Journal of Mathematics 5, 5 (1955), 807–816.

[28] Osborne, M. J., et al. *An introduction to game theory,* vol. 3. Oxford university press New York, 2004.

[29] D. Butnariu and A.N. Iusem (2000). *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization.* Kluwer, Dordrecht.

[30] Papa Quiroz, E., and Cruzado, S. *An inexact scalarization proximal point method for multiobjective quasiconvex minimization.* Annals of Operations Research 316, 2 (2022), 1445–1470.

[31] Quiroz, E. P., Ramirez, L. M., and Oliveira, P. R. *An inexact proximal method for quasiconvex minimization.* European Journal of Operational Research 246, 3 (2015), 721–729.

[32] Alber, Y. I., Iusem, A. N., and Solodov, M. V. *On the projected subgradient method for nonsmooth convex optimization in a hilbert space.* Mathematical Programming 81 (1998), 23–35.

[33] Rockafellar, R. T. *Monotone operators and the proximal point algorithm.* SIAM Journal on Control and Optimization 14, 5 (1976), 877–898.

[34] Santos, P. J. S., and Souza, J. C. O. - *A proximal point method for quasi-equilibrium problems in Hilbert spaces.* Optimization 71, 1 (2022), 55–70.

[35] Solodov, M. V., and Svaiter, B. F. *Error bounds for proximal point subproblems and associated inexact proximal point algorithms.* Mathematical programming, 88 (2000), 371–389.

[36] Solodov, M. V., and Svaiter, B. F. *An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions.* Mathematical of Operations Research 25 (2000), 214–230.

[37] Strodiot, J. J., Nguyen, T. T. V., and Nguyen, V. H. *A new class of hybrid extragradient algorithms for solving quasi-equilibrium problems.* Journal of Global Optimization 56, 2 (2013), 373–397.

[38] Svaiter, B. F. *A class of fejér convergent algorithms, approximate resolvents and the hybrid proximal-extragradient method*. Journal of Optimization Theory and Applications 162 (2014), 133–153.

[39] Censor, Y., Iusem, A. N., and Zenios, S. A. An interior point method with Bregman functions for the variational inequality problem with paramonotone operators. Mathematical Programming 81 (1998), 373–400.

[40] Van, N. T. T., Strodiot, J. J., Nguyen, V., and Vuong, P. *An extragradient type method for solving nonmonotone quasi-equilibrium problems*. Optimization 67, 5 (2018), 651–664.

[41] Zhang, J., Qu, B., and Xiu, N. *Some projection-like methods for the generalized Nash equilibria.* Computational optimization and applications 45, 1 (2010), 89–109.

[42] You, M., and Li, S. *Characterization of duality for a generalized quasi-equilibrium problem.* Applicable Analysis 97, 9 (2018), 1611–1627.

[43] Hung, P.G., Muu, L.D. *On Inexact Tikhonov and Proximal Point Regularization Methods for Pseudomonotone Equilibrium Problems.* Vietnam J. of Math. 40,(2012) 255–274.