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

Problema de Equilíbrio em espaços de Banach: um método de ponto proximal regularizado via distância de Bregman.

Elianderson Meneses Santos

Teresina - 2015

Elianderson Meneses Santos

Dissertação de Mestrado:

Problema de Equilíbrio em espaços de Banach: um método de ponto proximal regularizado via distância de Bregman.

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 Xavier da Cruz Neto

Teresina - 2015

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

S237p Santos, Elianderson Meneses.
Problema de equiblíbrio em espaços de Banach: um método de ponto proximal regularizado via distância de Bregman / Elianderson Meneses Santos. – Teresina, 2015.
79 f.il.

Dissertação (Mestrado) – Pós-Graduação em Matemática, Universidade Federal do Piauí, 2015.
Orientador: Prof. Dr. João Xavier da Cruz Neto.

1. Algoritmos – Matemática. 2. Problema de Equilíbrio. 3. Distância de Bregman. I. Título

CDD 511.8

Dedico este trabalho a todos aqueles que me apoiaram e acreditaram em mim durante esta caminhada, em especial aos meus pais Ermiro e Naide, aos meus avós Manoel e Francisca, a toda a minha família e ao meu grande amigo e orientador João Xavier. Dedico também este trabalho à minha falecida avó Maria de Jesus, a Vó Pequena, que tenho certeza que esteve olhando por mim todo esse tempo.

Agradecimentos

Agradeço primeiramente a DEUS, por ter me dado força e coragem diante de todas as dificuldades. Graças a Ele pude concluir este curso de mestrado, no qual aprendi muitas coisas e fiz novos amigos.

Agradeço também aos meus pais, Ermiro e Naide, à minha irmã Nádia, aos meus avós Manoel e Francisca, e a toda minha família pelo apoio e por tudo o que fazem por mim todos os dias da minha vida.

Agradeço imensamente ao meu amigo e orientador, Prof. Dr. João Xavier da Cruz Neto, pessoa sempre atenciosa e humilde, um verdadeiro exemplo a ser seguido. Agradeço a ele por toda a ajuda, pelos conselhos e por tudo o que fez por mim durante todos esses anos, não somente no mestrado como também na iniciação científica durante a graduação e mesmo antes da minha entrada na universidade, durante a época em que participei como aluno do PIC Jr. da Obmep. Peço a DEUS que o abençoe sempre e em todos os aspectos e que eu possa continuar contando com a confiança e o conselho desse grande amigo.

Agradeço também a todos os professores do Departamento de Matemática da UFPI, que desde a época da graduação sempre me ajudaram e apoiaram, em especial aos professores Gilvan Lima, João Benício, Mário Gomes, Paulo Alexandre, Jurandir, Paulo Sérgio e Sissy, Newton Luis, José Francisco e ao Professor Juscelino Silva, atualmente professor da Universidade Federal do Cariri, mas que à época da graduação era meu professor de Cálculo II na UFPI.

Um agradecimento especial também ao professor Luis Carlos Vieira da Silva, que foi meu professor no ensino fundamental na Escola Municipal Angelim, e que me apresentou às Olimpíadas de Matemática. Graças à oportunidade que me foi dada através dele, pude iniciar meu caminho na matemática e chegar até onde cheguei.

Agradeço aos amigos pelos conselhos, pelas broncas, pelas brincadeiras e principalmente por terem acreditado em mim. Agradeço especialmente ao Atécio, Magari, Raphaell, Suiany, João Neto, Narielle, Byanka, João Manário, Wiumar, Fernando Braz e Jailson, colegas desde a época da graduação e que sempre estiveram comigo. Não posso esquecer de agradecer também aos amigos Renan Maia, Phelps, Rômulo, Neves, Danilo, Kellwyn, Helson e Cristiane, Guilherme, Felipe, Joelson e Larissa, parceiros desde os tempos do Barão de Gurguéia. Agradeço também à Lays, Monik, Alinne, Victor, Maria Hellem, Layra e Giovanna.

Agradeço aos amigos do mestrado pelo companheirismo nos momentos de dificuldade, pelas questões resolvidas em grupo, pelas pizzas, por tudo. Obrigado ao Jeferson, Jaciel, Rafael, Pádua, Ray, Sandoel, Antonio Aguiar, Jhonata, Lucas Machado, Hércules, Tiago, Antonio Luis, Andressa, Victor, Alberone e Lucas Vidal.

Agradeço aos primos e tios do EJC (Encontro de Jovens com Cristo) pela amizade e pelo apoio, em especial à Camila Duarte, Gerson, Danilo Sousa, Paulo, Djaellyson, Lucas Moura e Caroline Santos, Clebert, Jessyca, Cleiton, Cláudia, Mariana Lopes, Marcos Paulo e Annie, Diego Leitão, Thaylanne, Suelma, Tios Chaguinha e Maria, Tios Paulo Moura e Maria Moura, Tia Márcia, Tios Gilvan e Francisca.

Agradeço aos amigos do Grupo de Teatro “Os Federais”, em especial à professora Erica Rodrigues Fontes, coordenadora do projeto. Quero deixar também um grande abraço para Débora, Melina, Daniel, Jhobert, Ana Karoline e Jainilene.

Agradeço a CAPES pelo apoio financeiro.

"*Uma criança, um professor, um livro e uma caneta podem mudar o mundo*".

Malala Yousafzai.

Resumo

Nesta dissertação apresentamos um algoritmo de ponto proximal para resolver Problemas de Equilíbrio em espaços de Banach reflexivos, o qual foi apresentado por Burachik e Kassay em [1]. Inicialmente, revisamos alguns conceitos de Análise Funcional, Análise convexa e alguns resultados preliminares e apresentamos as definições e algumas propriedades das Distâncias de Bregman e das funções totalmente convexas. Logo após, são enunciadas as hipóteses básicas do Problema de Equilíbrio, tendo como base a formulação apresentada por Blum e Oettli em [27]. Em seguida é estabelecida uma regularização do Problema de Equilíbrio envolvendo Distâncias de Bregman de Funções totalmente convexas. Por último, é apresentado o algoritmo proposto e mostra-se que, sob certas hipóteses, a sequência gerada pelas iterações está bem definida e converge assintoticamente a uma solução do Problema de Equilíbrio.

Palavras-chave: Distâncias de Bregman, Problema de Equilíbrio, Algoritmo de ponto proximal, Aplicação de dualidade, Método de regularização, Funções totalmente convexas.

Abstract

In this dissertation we present a proximal point algorithm for solving Equilibrium Problems in reflexive Banach spaces, which was presented by Burachik and Kassay in [1]. Initially, we review some concepts of Functional Analysis, Convex Analysis and some preliminary results and we present the definitions and some properties of Bregman distances and Tottaly Convex Functions. Soon after, are presented the basic assumptions of Equilibrium Problem based on the formulation presented by Blum and Oettli in [27]. It is the established a regularization for Equilibrium Problems involving Bregman distances of totally Convex Functions. Finally, the proposed algorithm is presented and it is shows that, under certain assumptions, the sequence generated by iterations is well defined and converges asymptotically to a solution of Equilibrium Problem.

Keywords: Bregman distances, Equilibrium Problem, Proximal point algorithm, Duality mapping, Regularization method, Totally convex function.

Sumário

Introdução

O objetivo deste trabalho é apresentar um algoritmo de ponto proximal para resolver o Problema de Equilíbrio em espaços de Banach reflexivos, o qual foi idealizado por Burachik e Kassay em [1] e que usa uma regularização via distâncias de Bregman e Funções Totalmente Convexas. A importância desse resultado reside em sua larga abrangência, uma vez que o Problema de Equilíbrio constitui uma classe de problema bastante ampla, visto que engloba vários outros tipos de problemas como seus casos particulares. Por outro lado, há também o fato de que a classe dos espaços de Banach reflexivos contém uma vasta gama de espaços importantes (podemos citar, por exemplo, os espaços de Hilbert e os espaços $L_p(1 < p < \infty)$).

No capítulo 1, com base nas referências [2–10, 12–14], apresentaremos uma série de definições e conceitos envolvendo Análise Funcional e Análise Convexa, que serão necessários para um melhor entendimento na leitura do presente texto.

No capítulo 2, apresentaremos a definição, algumas propriedades e exemplos de Distâncias de Bregman. Veremos que as Distâncias de Bregman foram inicialmente estudadas por Bregman [15] na década de 60, porém a nomenclatura "Distância de Bregman" foi cunhada por Censor e Lent em [16] apenas no ínicio da década de 80. Ainda nesse capítulo e com base nas referências [17, 18, 20–23] serão apresentadas a definição e algumas propriedades das Funções Totalmente Convexas. Veremos, por exemplo, que este tipo de função é mais geral que as funções estritamente convexas. Além disso, tais funções serão bastante importantes, pois ajudarão a garantir a existência e unicidade de solução para o nosso problema regularizado.

No capítulo 3, serão apresentadas as hipóteses básicas do Problema de Equilíbrio. Baseados em Blum, Oettli [27] e Iusem, Kassay, Sosa [28] abordaremos também alguns tipos de problemas que podem ser estudados como casos particulares do Problema de Equilíbrio, tais como: Problema de Otimização Convexa, Problema de Ponto Fixo, Complementaridade, Desigualdades Variacionais e o problema de equilíbrio de Nash em jogos não cooperativos. Veremos que, em cada caso, resolver o problema é equivalente a resolver um Problema de Equilíbrio para uma função auxiliar dada. Existe ainda um outro exemplo, que não se encaixa em nenhum dos casos anteriores, o qual mostra a generalidade do Problema de Equilíbrio e que pode ser visto em uma forma mais geral no trabalho de Iusem e Sosa [28].

No capítulo 4, apresentamos uma breve introdução sobre o método de ponto proximal. Em sequida, veremos a regularização para o problema de equilíbrio proposta por Burachik e Kassay em [1] e mostraremos que, sob certas hipóteses, o problema regularizado têm solução única. Na última seção apresentaremos um algoritmo tipo ponto proximal envolvendo a regularização dada e veremos que, sob certas condições, a sequência gerada converge assintoticamente a uma solução do Problema de Equilíbrio.

Capítulo 1

Noções Preliminares

1.1 Conceitos e resultados de Análise Funcional

Baseados em [2–7], apresentaremos nesta seção alguns conceitos e resultados de Análise Funcional que serão úteis para um melhor entendimento na leitura do presente trabalho.

1.1.1 Espaços de Banach

Definição 1.1.1 (Métrica). *Uma* métrica *num conjunto* M *é uma função* $d : M \times M \to \mathbb{R}$*, que associa a cada par ordenado* $(x, y) \in M \times M$ *um número real* $d(x, y)$*, de modo que sejam satisfeitas as seguintes condições para quaisquer* $x, y, z \in M$*:*

D1) $d(x, x) = 0$*;*

D2) Se $x \neq y$ *então* $d(x, y) > 0$*;*

D3) $d(x, y) = d(y, x)$*;*

D4) $d(x, z) \leqslant d(x, y) + d(y, z)$.

O par (M, d) *é chamado de* espaço métrico.

Definição 1.1.2 (Norma). *Uma* norma *em um espaço vetorial E (real ou complexo) é uma aplicação* $\|\cdot\| : E \to \mathbb{R}$ *tal que:*

N1) $\|\xi\| \geqslant 0$ *para todo* $\xi \in E$ *e,* $\|\xi\| = 0$ *se, e somente se,* $\xi = 0$.

N2) $\|\alpha\xi\| = |\alpha|\|\xi\|$*, para todo* $\xi \in E$ *e qualquer* $\alpha \in \mathbb{K}$*, onde* $\mathbb{K} = \mathbb{C}$ *ou* $\mathbb{K} = \mathbb{R}$.

N3) $\|\xi + \eta\| \leqslant \|\xi\| + \|\eta\|$*, para quaisquer* $\xi, \eta \in E$.

O par $(E, \|\cdot\|)$ *é chamado de* espaço vetorial normado *ou simplesmente* espaço normado.

A norma $\|\cdot\|: E \to \mathbb{R}$ define também uma métrica d em E, dada por

$$d(\xi,\eta) = \| \xi - \eta \|$$

Dizemos ainda que duas normas $\|\cdot\|_1$ e $\|\cdot\|_2$ em E são *equivalentes* se existirem $A, B \in \mathbb{R}_{++}$ tais que $A\|\xi\|_1 \leqslant \|\xi\|_2 \leqslant B\|\xi\|_1$ para todo $\xi \in E$.

Definição 1.1.3 (Espaço métrico completo)**.** *Dizemos que um espaço métrico* (M, d) *é* completo *se, para toda sequência de Cauchy* $\{u^k\}$ *em* E*, isto é, se* $\{u^k\} \subset E$ *é uma sequência tal que, dado* $\varepsilon > 0$*, existe* $k_0 \in \mathbb{N}$ *tal que*

$$m, n \geqslant k_0 \Rightarrow d(u^m, u^n) < \varepsilon,$$

existir $u^* \in M$ *tal que* $\lim_{k\to\infty} u^k = u^*$.

Definição 1.1.4 (Espaço de Banach)**.** *Um espaço normado* E *que é completo com a métrica induzida pela norma é chamado de* espaço de Banach.

Como exemplos de espaços de Banach temos o conjunto $\mathbb{R}$ dos números reais e conjunto $\mathbb{C}$ dos números complexos. Mais geralmente, todo espaço normado de dimensão finita E é um espaço de Banach [3, Teorema 1.1.6, pág. 6].

Nas próximas seções, abordaremos alguns importantes exemplos de espaços de Banach de dimensão infinita: os espaços $L_p(X, \Sigma, \mu)$, o espaço $L_\infty(X, \Sigma, \mu)$ e os espaços de sequências l_p e l_∞.

1.1.2 Os espaços $L_p(X, \Sigma, \mu)$

Nesta subseção abordaremos alguns conceitos e definições da Teoria da Medida e que são mais amplamente discutidos em [5].

Definição 1.1.5 (Topologia)**.** *Uma topologia em um conjunto* X *é uma coleção* τ *de subconjuntos de* X*, chamados* conjuntos abertos*, satisfazendo as seguintes propriedades:*

(T1) Qualquer união de elementos de τ *pertence a* τ*.*

(T2) Qualquer interseção finita de elementos de τ *pertence a* τ*.*

(T3) X *e* $\emptyset$ *pertencem a* τ*.*

O par (X, τ) *é chamado de* espaço topológico.

Definição 1.1.6. *Uma* $\sigma-$*álgebra no conjunto* X *é uma família* Σ *de subconjuntos de* X *que satisfaz as seguintes propriedades:*

(a) $\emptyset, X \in \Sigma$

(b) Se $A \in \Sigma$*, então* $A^C = X - A \in \Sigma$

(c) Se $A_n \in \Sigma$ *para todo* $n \in \mathbb{N}$*, então* $\bigcup_{n=1}^{\infty} A_n \in \Sigma$

Neste caso, o par (X, Σ) é chamado *espaço mensurável.* Cada elemento de Σ é chamado de *conjunto mensurável.* Dada uma coleção $\mathcal{F}$ de subconjuntos de X, a interseção de todas as $\sigma-$álgebras que contêm $\mathcal{F}$ é ainda uma $\sigma-$álgebra, chamada $\sigma-$*álgebra gerada por* $\mathcal{F}$ e denotada por $\Sigma(\mathcal{F})$. Se (X, τ) é um espaço topológico, a $\sigma-$álgebra $\Sigma(\tau)$ é chamada de $\sigma-$*álgebra de Borel de X* e denotada por $\mathcal{B} = \mathcal{B}(X)$. Os elementos de $\mathcal{B}(X)$ são chamados de *conjuntos de Borel* ou *boreliano.*

Definição 1.1.7. *A* reta estendida *é o conjunto* $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$, *onde* $-\infty < x < \infty$ *para todo* $x \in \mathbb{R}$.

Definição 1.1.8. *A topologia usual de* $\mathbb{R}$ *induz uma topologia em* $\overline{\mathbb{R}}$*, onde os abertos* $A \subset \overline{\mathbb{R}}$ *são da forma:*

(a) $A \subset \mathbb{R}$ *é aberto em* $\mathbb{R}$*, ou*

(b) $A = [-\infty, a)$ *para algum* $a \in \mathbb{R}$*, ou*

(c) $A = (a, \infty]$ *para algum* $a \in \mathbb{R}$*, ou*

(d) A *é uma união de conjuntos como os de (a), (b) ou (c).*

Consideraremos $\overline{\mathbb{R}}$ como espaço mensurável com a $\sigma-$álgebra de Borel $\mathcal{B}(\overline{\mathbb{R}})$ relativa a esta topologia.

Definição 1.1.9. *Seja* (X, Σ) *um espaço mensurável. Uma função* $f : X \longrightarrow \overline{\mathbb{R}}$ *é* mensurável *se* $f^{-1}(A) \in \Sigma$ *para todo boreliano* $A \in \mathcal{B}(\overline{\mathbb{R}})$.
O conjunto formado por essas funções será denotado por $M(X, \Sigma)$. *Consideremos ainda o subconjunto* $M^+(X, \Sigma) := \{f \in M(X, \Sigma) : f(x) \geqslant 0,\ \forall x \in X\}$

A *função característica* de A é uma função $\chi_A : X \longrightarrow \mathbb{R}$, onde $\chi_A(x) = 1$, se $x \in A$ e, $\chi_A(x) = 0$, caso contrário. É claro que χ_A é mensurável se, e somente se, $A \in \Sigma$. Uma combinção linear de funções características mensuráveis é chamada de *função simples mensurável.* Toda função simples mensurável admite uma única representação da forma

$$\varphi = \sum_{i=1}^{n} a_i \chi_{A_i},$$

onde $n \in \mathbb{N}$, $a_1, ..., a_n$ são reais não-nulos e distintos, e $A_1, ..., A_n$ são conjuntos mensuráveis não-vazios e disjuntos dois a dois. Esta é chamada a *representação canônica* da função simples mensurável φ.

Definição 1.1.10. *Dizemos que uma função* $\mu : \Sigma \to [0, \infty]$ *é uma* medida *no espaço mensurável* (X, Σ) *se:*

(a) $\mu(\emptyset) = 0$

(b) Se $(A_n)_{n=1}^{\infty}$ *é uma sequência de conjuntos disjuntos dois a dois de* Σ*, então*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

A medida μ é finita se $\mu(X) < \infty$, e é dita σ—finita se existirem conjuntos $(A_n)_{n=1}^{\infty}$ em Σ tais que $X = \bigcup\limits_{n=1}^{\infty} A_n$ e $\mu(A_n) < \infty$ para todo n. O terno (X, Σ, μ) é chamado de *espaço de medida.*

Definição 1.1.11. *Seja* (X, Σ, μ) *um espaço de medida.*

(a) A integral *da função simples* $\varphi \in M^+(X, \Sigma)$*, de representação canônica* $\varphi = \sum\limits_{i=1}^{n} a_i \chi_{A_i}$, *em relação à medida* μ *é definida por*

$$\int_X \varphi d\mu = \sum_{i=1}^{m} a_i \mu(A_i).$$

(b) A integral *da função* $f \in M^+(X, \Sigma)$ *em relação à medida* μ *é definida por*

$$\int_X f d\mu := \sup\left\{\int_X \varphi d\mu : \varphi \in M^+(X, \Sigma) \text{ é simples e } 0 \leqslant \varphi \leqslant f\right\}$$

(c) Para $f \in M^+(X, \Sigma)$ *e* $A \in \Sigma$*, define-se*

$$\int_A f d\mu = \int_X f \chi_A d\mu.$$

Sejam agora (X, Σ, μ) um espaço de medida e $1 \leqslant p < \infty$. O conjunto de todas as funções mensuráveis de X em $\mathbb{K}$ tais que

$$\|f\|_p := \left(\int_X |f|^p d\mu\right)^{1/p} < \infty$$

será denotado por $\mathcal{L}_p(X, \Sigma, \mu)$.

Observe que $\|\cdot\|_p$ satisfaz as propriedades N2 e N3 (Ver [3], Teorema 1.2.2, pág. 8) da definição de norma e que $\|f\|_p \geqslant 0$ para toda $f \in \mathcal{L}_p(X, \Sigma, \mu)$, porém, em geral, $\|\cdot\|_p$ não é uma norma em $\mathcal{L}_p(X, \Sigma, \mu)$, pois pode ocorrer $\|f\|_p = 0$ para f não identicamente nula.

De modo geral, se (X, Σ, μ) é um espaço de medida, definimos uma relação de equivalência dizendo que $f, g : X \longrightarrow \mathbb{K}$ são equivalentes se $f = g$ $\mu-$quase sempre, isto é, se existe $A \in \Sigma$ tal que $\mu(A) = 0$ e $f(x) = g(x)$ para todo $x \notin A$. A classe de equivalência de f denota-se por $[f]$. Assim, é imediato que no conjunto quociente

$$L_p(X, \Sigma, \mu) := \{[f] : f \in \mathcal{L}_p(X, \Sigma, \mu)\}$$

as operações $[f] + [g] = [f + g]$ e $c[f] = [cf]$ estão bem definidas e tornam $L_p(X, \Sigma, \mu)$ um espaço vetorial. Definido

$$\|[f]\|_p := \|f\|_p,$$

corrigimos o que faltava para $\|\cdot\|_p$ ser uma norma. Assim, $(L_p(X, \Sigma, \mu), \|\cdot\|_p)$ é um espaço vetorial normado. Além disso, $(L_p(X, \Sigma, \mu), \|\cdot\|_p)$ é um espaço de Banach com essa norma.

Teorema 1.1.1. *Se* $1 \leqslant p < \infty$*, então* $L_p(X, \Sigma, \mu)$ *é um espaço de Banach com a norma*

$$\|[f]\|_p := \left(\int_X |f|^p d\mu \right)^{1/p}$$

Demonstração. Ver [3, Teorema 1.2.3, pág. 10]. □

1.1.3 O espaço $L_\infty(X, \Sigma, \mu)$

Seja $\mathcal{L}_\infty(X, \Sigma, \mu)$ o conjunto de todas as funções mensuráveis tais que existem um conjunto $N \in \Sigma$ e um $M \in \mathbb{R}$ tais que $\mu(N) = 0$ e $|f(x)| \leqslant M$ para todo $x \notin N$. Se $f \in \mathcal{L}_\infty(X, \Sigma, \mu)$ e $N \in \Sigma$ é um conjunto de medida nula, definimos

$$S_f(N) = \sup\{|f(x)| : x \notin N\} \text{ e}$$

$$\|f\|_\infty = \inf\{S_f(N) : N \in \Sigma, \mu(N) = 0\} \tag{1.1}$$

Note que pode ocorrer $\|f\|_\infty = 0$ sem que se tenha f identicamente nula. Assim, recorremos mais uma vez às classes de equivalência e dizemos que duas funções são equivalentes se coincidem μ-quase sempre. O conjunto $L_\infty(X, \Sigma, \mu)$ é o conjunto de todas as classes de equivalência das funções mensuráveis $f : X \longrightarrow \mathbb{K}$ que são limitadas μ-quase sempre. Como anteriormente, $L_\infty(X, \Sigma, \mu)$ é um espaço vetorial com as operações $[f+g] = [f]+[g]$ e $c[f] = [cf]$. Se $[f] \in L_\infty(X, \Sigma, \mu)$, definimos

$$\|[f]\|_\infty = \|f\|_\infty.$$

Não é difícil verificar que $\|[f]\|_\infty$ é uma norma em $L_\infty(X, \Sigma, \mu)$.

Teorema 1.1.2. $(L_\infty(X,\Sigma,\mu), \|[f]\|_\infty)$ *é um espaço de Banach.*

Demonstração. Ver [3, Teorema 1.3.1, pág. 13]. □

1.1.4 Espaços de sequências

Para cada número real $p \geqslant 1$, definimos

$$\ell_p = \left\{ (a_j)_{j=1}^{\infty} : a_j \in \mathbb{K} \ \forall j \in \mathbb{N} \ e \ \sum_{j=1}^{\infty} |a_j|^p < \infty \right\}.$$

Definindo em $\mathcal{P}(\mathbb{N})$ a seguinte medida de contagem

$$\mu_c : \mathcal{P}(\mathbb{N}) \longrightarrow \mathbb{R} \cup \{\infty\}, \mu_c(X) = \begin{cases} n, \text{ se X tem } n \text{ elementos,} \\ \infty, \text{ se X tem infinitos elementos,} \end{cases}$$

temos que $\ell_p = L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_c)$ e a norma $\|\cdot\|_p$ se transforma em

$$\|(a_j)_{j=1}^{\infty}\|_p = \left(\sum_{j=1}^{\infty} |a_j|^p \right)^{\frac{1}{p}}.$$

Podemos então afirmar que ℓ_p é um espaço de Banach com as operações usuais de sequências e com a norma $\|\cdot\|_p$.

Para $p = \infty$, definimos ℓ_∞ como o espaço das sequências limitadas de escalares, ou seja:

$$\ell_\infty = \{(a_j)_{j=1}^{\infty} : a_j \in \mathbb{K}, \ \forall j \in \mathbb{N}, \text{ e} \sup_{j \in \mathbb{N}} |a_j| < \infty\}$$

Observando o fato de que $\ell_\infty = L_\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_c)$ podemos concluir que ℓ_∞ é um espaço de Banach com as operações usuais de funções e com a norma

$$\|(a_j)_{j=1}^{\infty}\| = \sup\{|a_j| : j \in \mathbb{N}\}.$$

1.1.5 Dualidade e espaços reflexivos

Os espaços normados possuem uma estrutura algébrica de espaço vetorial associada as transformações lineares. Por também serem espaços métricos, possuem ainda uma estrutura topológica associada as funções contínuas.

Definição 1.1.12. *Um operador linear contínuo do espaço normado* E *no espaço normado* F*, ambos sobre o mesmo corpo de escalares* $\mathbb{K}$*, é uma função* $T : E \to F$ *que é linear e contínua.*

O conjunto de todos os operadores lineares contínuos de E em F será denotado por $\mathcal{L}(\mathsf{E},\mathsf{F})$. É claro que $\mathcal{L}(\mathsf{E},\mathsf{F})$ é um espaço vetorial sobre $\mathbb{K}$ com as operações usuais de funções. No caso particular em que $\mathsf{F}=\mathbb{K}$, escrevemos E^* ao invés de $\mathcal{L}(\mathsf{E},\mathbb{K})$ e chamamos esse espaço de *dual topológico de E*, ou simplesmente, de *dual de E*, e dizemos que seus elementos são *funcionais lineares contínuos*.

Proposição 1.1.1. *Sejam* E *e* F *espaços normados. Então*

(a) A expressão

$$\| \mathsf{T} \| = \sup\{\| \mathsf{T}(u) \| : u \in \mathsf{E} \ e \ \| u \| \leqslant 1\}$$

define uma norma em $\mathcal{L}(\mathsf{E},\mathsf{F})$.

(b) $\| \mathsf{T}(u) \| \leqslant \| \mathsf{T} \| \cdot \| u \|$ *para quaisquer* $\mathsf{T}\in\mathcal{L}(\mathsf{E},\mathsf{F})$ *e* $u \in \mathsf{E}$.

(c) Se F *é um espaço de Banach, então* $\mathcal{L}(\mathsf{E},\mathsf{F})$ *também é um espaço de Banach.*

Demonstração. Ver [3, Proposição 2.1.4, pág. 34]. □

Corolário 1.1.1. *O dual* E^* *de qualquer espaço normado* E *é um espaço de Banach.*

De posse do corolário acima, dado um espaço normado E, sabemos que seu dual E^* é um espaço de Banach. Podemos então considerar o dual de E^*, chamado o *bidual de E* e denotado por E^{**}. Veremos a seguir que dado um espaço normado E, seu bidual E^{**} contém uma *cópia isométrica* de E, isto é, que E^{**} contém um subespaço isomorfo isometricamente a E.

Seja E um espaço normado. Definimos o *mergulho canônico de* E *em* E^{**} como

$$\mathsf{J}_\mathsf{E} : \mathsf{E} \to \mathsf{E}^{**},\ \mathsf{J}_\mathsf{E}(x)(\varphi) = \varphi(x) \text{ para todos } x \in \mathsf{E} \text{ e } \varphi \in \mathsf{E}^*,$$

onde ao escrevermos $\mathsf{J}_\mathsf{E}(x)(\varphi) = \varphi(x)$ estamos dizendo que a imagem de $x \in \mathsf{E}$ pela aplicação J_E é um funcional linear $\mathsf{T}_x \in \mathsf{E}^{**}$ tal que $\mathsf{T}_x(\varphi) = \varphi(x)$ para todo $\varphi \in \mathsf{E}^*$.

Proposição 1.1.2. *Para todo espaço normado* E, $\mathsf{J}_\mathsf{E} : \mathsf{E} \to \mathsf{E}^{**}$ *é uma* isometria linear.

Demonstração. Ver [3, Proposição 4.3.1, pág. 89]. □

Definição 1.1.13. *Um espaço normado* E *é* reflexivo *se a aplicação* $\mathsf{J}_\mathsf{E} : \mathsf{E} \to \mathsf{E}^{**}$ *for sobrejetora, ou seja,* $\mathsf{J}_\mathsf{E}(\mathsf{E}) = \mathsf{E}^{**}$.

No caso em que E é reflexivo é comum o abuso de notação $E = E^{**}$ para indicar que E e E^{**} são isomorfos isometricamente.

Todo espaço normado de dimensão finita é reflexivo [6, Teorema 4.6-5, pág. 241]. Para $p \in (1, \infty)$, os espaços $L_p(X, \Sigma, \mu)$ e ℓ_p são reflexivos [3, Proposição 4.3.12, pág. 93]. O espaço ℓ_1 não é reflexivo [3, Exemplo 4.3.8, pág. 91] e, consequentemente, seu dual ℓ_∞ também não é reflexivo [3, Exemplo 4.3.14, pág. 95].

1.1.6 Espaços de Hilbert

Definição 1.1.14 (Produto interno)**.** *Seja E um espaço vetorial sobre o corpo* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$*. Um* produto interno *em* E *é uma aplicação*

$$\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{K}, (x, y) \mapsto \langle x, y \rangle,$$

tal que, para quaisquer $x, y, z \in E$ *e* $\lambda \in \mathbb{K}$*:*

(H1) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$*;*

(H2) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$*;*

(H3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$*;*

(H4) Para todo $x \neq 0$*,* $\langle x, x \rangle \in \mathbb{R}_{++}$*.*

O par $(E, \langle \cdot \rangle)$ *é chamado* espaço com produto interno.

Não é difícil verificar que dado um produto interno $\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{K}$, a função

$$\| \cdot \| : E \longrightarrow \mathbb{R}, \|x\| = \sqrt{\langle x, x \rangle}$$

é uma norma em E, a qual é chamada de *norma induzida pelo produto interno.*

Proposição 1.1.3 (Lei do Paralelogramo)**.** *Seja* $(E, \| \cdot \|)$ *um espaço vetorial com a norma induzida pelo produto interno. Então, para quaisquer* $x, y \in E$,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Demonstração. [3, Proposição 5.1.7, pág. 107]. □

Definição 1.1.15 (Espaço de Hilbert)**.** *Um espaço de Hilbert é um espaço de Banach cuja norma é induzida de um produto interno.*

O espaço euclidiano $\mathbb{R}^n$ é um espaço de Hilbert [6, Exemplo 3.1-3, pág. 131]. São também espaços de Hilbert o espaço $\mathbb{C}^n$ [6, Exemplo 3.1-4, pág. 131] e o espaço ℓ_2 [6, Exemplo 3.1-6, pág. 133]. Além disso, se $p \neq 2$, então ℓ_p não é um espaço de Hilbert [6, Exemplo 3.1-7, pág. 133].

Ainda nesse contexto de espaços de Hilbert, temos o Teorema da Representação de Riez-Fréchet, que nos dá uma descrição de todos os funcionais lineares contínuos em um espaço de Hilbert.

Teorema 1.1.3 (Teorema da Representação de Riez-Fréchet)**.** *Sejam* H *um espaço de Hilbert e* $\varphi : H \longrightarrow \mathbb{K}$ *um funcional linear contínuo. Então existe um único* $y_0 \in H$ *tal que*

$$\varphi(x) = \langle x, y_0 \rangle, \ \forall \ x \in H,$$

além disso, $\|y_0\| = \|\varphi\|$.

Demonstração. [3, Teorema 5.5.2., pág. 126]. □

Corolário 1.1.2. *O dual de um espaço de Hilbert é um espaço de Hilbert.*

Demonstração. Seja H um espaço de Hilbert. Como H^* é espaço de Banach, por ser um dual, basta mostrar que a norma em H^* é induzida por um produto interno. Dados $\varphi_1, \varphi_2 \in H^*$, pelo Teorema 1.1.3, temos que existem e são únicos $y_1, y_2 \in H$ tais que $\varphi_1(x) = \langle x, y_1 \rangle$, $\forall \ x \in H$, e $\varphi_2(x) = \langle x, y_2 \rangle$, $\forall \ x \in H$, com $\|y_1\| = \|\varphi_1\|$ e $\|y_2\| = \|\varphi_2\|$. A expressão

$$\langle \varphi_1, \varphi_2 \rangle := \langle y_2, y_1 \rangle \tag{1.2}$$

define um produto interno em H^*. De fato, para quaisquer $\varphi_1, \varphi_2, \varphi_3 \in H^*$ e para todo $\lambda \in \mathbb{K}$, como $(\varphi_1 + \varphi_2)(x) = \langle x, y_1 + y_2 \rangle$ e $(\lambda\varphi_1)(x) = \lambda\varphi_1(x) = \lambda\langle x, y_1 \rangle = \langle x, \lambda y_1 \rangle$, temos

(H1) $\langle \varphi_1 + \varphi_2, \varphi_3 \rangle = \langle y_3, y_1 + y_2 \rangle = \langle y_3, y_1 \rangle + \langle y_3, y_2 \rangle = \langle \varphi_1, \varphi_3 \rangle + \langle \varphi_2, \varphi_3 \rangle$;

(H2) $\langle \lambda\varphi_1, \varphi_2 \rangle = \langle \lambda y_2, y_1 \rangle = \lambda\langle y_2, y_1 \rangle = \lambda\langle \varphi_1, \varphi_2 \rangle$;

(H3) $\langle \varphi_1, \varphi_2 \rangle = \langle y_2, y_1 \rangle = \overline{\langle y_1, y_2 \rangle} = \overline{\langle \varphi_2, \varphi_1 \rangle}$;

(H4) Para todo $\varphi_1 \neq 0$, temos que $y_1 \neq 0$, pois se fosse $y_1 = 0$, teríamos que $\varphi_1(x) = 0, \forall \ x \in H$, uma contradição. Logo, se $\varphi_1 \neq 0$, então $\langle \varphi_1, \varphi_1 \rangle = \langle y_1, y_1 \rangle \neq 0$. Agora, de

$$\langle \varphi_1, \varphi_1 \rangle = \langle y_1, y_1 \rangle = \|y_1\|^2 = \|\varphi_1\|^2,$$

segue que a norma $\|\cdot\| : H^* \to \mathbb{R}$, onde $\|\varphi\| = \sqrt{\langle \varphi_1, \varphi_1 \rangle}$, satisfaz a Proposição 1.1.3 e, portanto, H^* é um espaço de Hilbert.

□

Corolário 1.1.3. *Todo espaço de Hilbert é reflexivo.*

Demonstração. Seja H um espaço de Hilbert. Dado $\Phi \in H^{**}$, do corolário anterior sabemos que H^{**} é um espaço de Hilbert e assim, usando o Teorema 1.1.3 podemos tomar um $\varphi \in H^*$ tal que $\Phi(\psi) = \langle \psi, \varphi \rangle$ para todo $\psi \in H^*$. Uma segunda aplicação do Teorema 1.1.3 nos dá um vetor $y \in H$ tal que $\varphi(x) = \langle x, y \rangle$ para todo $x \in H$. Seja $\psi \in H^*$. Aplicando o Teorema 1.1.3 novamente, temos que existe $z \in H$ tal que $\psi(z) = \langle x, z \rangle$ para todo $x \in H$. De (1.2) temos que $\langle \psi, \varphi \rangle = \langle y, z \rangle$ e assim

$$J_H(y)(\psi) = \psi(y) = \langle y, z \rangle = \langle \psi, \varphi \rangle = \Phi(\psi), \forall\, \psi \in H^*.$$

A igualdade acima prova que $J_H(H) = H^{**}$, isto é, que H é reflexivo. □

1.1.7 A topologia fraca

Definição 1.1.16. *Seja* (X, τ) *um espaço topológico.*
(a) Um conjunto $F \subset X$ *é* fechado *em* X *relativamente à topologia* τ *se o seu complementar for aberto, isto é,* F *é fechado se* $(X - F) \in \tau$.
(b) Um conjunto $K \subset X$ *é* compacto *em* X *relativamente a topologia* τ *se, e somente se, para toda cobertura* $K \subset \bigcup A_\lambda$ *de* K*, por abertos* A_λ *de* X*, existe uma subcobertura finita* $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_n}$.

Todo espaço vetorial normado E é espaço métrico com a métrica d induzida por sua norma $\|\cdot\|$. Além disso, E é um espaço topológico, pois existe uma topologia natural τ_d em E induzida pela métrica d.

Quando falamos em *enfraquecer a topologia* estamos nos referindo a encontrar uma topologia em E que contenha *menos abertos* que a topologia usual τ_d, isto é, que esteja propriamente contida em τ_d, partindo do princípio de que quanto menor for a quantidade de abertos em E, maior será a chance de obtermos subconjuntos compactos em E. Porém, uma vez que consideramos topologias com menos abertos em E, diminuimos a chance das funções definidas neste espaço serem contínuas. Devemos então escolher previamente funções que devem ser (ou permanecer) contínuas e procurar a menor topologia em E que garanta essa continuidade.

Sejam X um conjunto, $(Y_i)_{i\in I}$ uma família de espaços topológicos e $(f_i)_{i\in I}$ uma família de funções $f_i : X \to Y_i$ para cada $i \in I$. Queremos definir em X a menor topologia que torna todas as funções f_i contínuas. Para cada $i \in I$ e cada aberto A_i em Y_i considere o conjunto $f_i^{-1}(A_i) = \{x \in X : f_i(x) \in A_i\}$. Chame de Φ a coleção dos subconjuntos de X que podem ser escritos como interseções finitas de conjuntos da forma $f_i^{-1}(A_i)$.

Proposição 1.1.4. *Existe uma topologia* τ *em* X *que tem* Φ *como base, isto é, os elementos de* τ *são uniões de elementos de* Φ.

Demonstração. Basta notar que para cada $i \in I$, valem as igualdades

$$\bigcup_{i\in I} f_i^{-1}(A_j) = f_i^{-1}(\bigcup_i A_i), \text{ e } \bigcap_{i\in I} f_i^{-1}(A_j) = f_i^{-1}(\bigcap_i A_i).$$

Logo, cada f_i induz uma topologia τ_i em X (Ver *Topologia induzida* em [2, pág. 66]). Como Φ é uma coleção dos subconjuntos de X que são interseções finitas de conjuntos da forma $f_i^{-1}(A_i)$, segue que Φ é *base de uma topologia* em X [2, Capítulo 3, Proposição 10, pág. 73]. □

A topologia τ da Proposição 1.1.4 acima é chamada de *topologia gerada pela família de funções* $(f_i)_{i\in I}$.

Definição 1.1.17. *A topologia fraca em um espaço normado* E*, denotada por* $\sigma(E, E^*)$*, é a topologia gerada pelos funcionais lineares contínuos* $\varphi \in E^*$*. Quando uma sequência* $(u^k)_{k=1}^{\infty}$ *em* E *converge para* $u \in E$ *na topologia fraca escrevemos* $u^k \xrightarrow{w} u$*, e dizemos que* u^k *converge fracamente a* u.

A proposição a seguir lista algumas propriedades válidas para sequências em um espaço normado E munido da topologia fraca $\sigma(E, E^*)$.

Proposição 1.1.5. *Seja* $(u^k)_{k=1}^{\infty}$ *uma sequência no espaço normado* E*. Então*
(i) $u^k \xrightarrow{w} u$ *em* $\sigma(E, E^*)$ *se, e somente se,* $\Psi(u^k) \to \Psi(u)$ *para todo* $\Psi \in E^*$.
(ii) Se $u^k \to u$ *na topologia da norma, então* $u^k \xrightarrow{w} u$ *em* $\sigma(E, E^*)$.
(iii) Se $u^k \xrightarrow{w} u$*, então a sequência* $(\|u^k\|)_{k=1}^{\infty}$ *é limitada e* $\|u\| \leqslant \liminf_k \|u^k\|$.
(iv) Se $u^k \xrightarrow{w} u$ *em* E *e* $\Psi_k \to \Psi$ *em* E^**, então* $\Psi_k(u^k) \to \Psi(u)$ *em* $\mathbb{K}$.

Demonstração. Ver [4, Proposição 3.5, pág. 58]. □

Em qualquer espaço normado E, a topologia fraca está contida na topologia da norma e elas coincidem se, e somente se, E for um espaço de dimensão finita [3, Proposição 6.2.6, pág. 146]. Para todo espaço normado E, vale também que $E^* = (E, \sigma(E, E^*))^*$, isto é, o dual de E na topologia da norma coincide com o dual de E na topologia fraca [3, Corolário 6.2.7, pág. 148]. Além disso, o Teorema de Mazur [3, Teorema 6.2.11, pág. 149] garante que um subconjunto convexo K de E é fechado na topologia da norma se, e somente se, for fechado na topologia fraca.

1.1.8 A topologia fraca-estrela

Definição 1.1.18. *A* topologia fraca-estrela *no dual* E^* *do espaço normado E, denotada por* $\sigma(E^*, E)$*, é a topologia gerada pelas funções do conjunto* $J_E(E) = \{J_E(x) : x \in E\}$*, isto é, pelas funções* $\varphi \mapsto J_E(x)(\varphi) = \varphi(x) \in \mathbb{K}$*, onde* $x \in E$*.*

Quando uma sequência $(\varphi_n)_{n=1}^{\infty}$ em E^* converge para $\varphi \in E^*$ na topologia fraca-estrela escrevemos $\varphi_n \xrightarrow{w^*} \varphi$.

Proposição 1.1.6. *Seja* E *um espaço normado.*
(a) Se $\varphi_n \xrightarrow{w} \varphi$ *em* E^**, então* $\varphi_n \xrightarrow{w^*} \varphi$*.*
(b) Se E *é espaço de Banach e* $\varphi_n \xrightarrow{w^*} \varphi$ *em* E^* *então a sequência* $(\|\varphi_n\|)_{n=1}^{\infty}$ *é limitada e* $\|\varphi\| \leqslant \liminf_{n\to\infty} \|\varphi_n\|$*.*
(c) Se E *é Banach,* $\varphi_n \xrightarrow{w^*} \varphi$ *em* E^* *e* $u_n \longrightarrow u$ *em* E*, então* $\varphi_n(u_n) \longrightarrow \varphi(u)$ *em* $\mathbb{K}$*.*

Demonstração. Ver [3, Proposição 6.3.3, pág. 153]. □

Dados um espaço normado qualquer E e seu dual E^*, a topologia fraca-estrela $\sigma(E, E^*)$ de E^* sempre está contida na topologia fraca $\sigma(E^*, E^{**})$ e elas coincidem se, e somente se, E é reflexivo [3, Proposição 6.3.8, pág. 155].

Teorema 1.1.4 (Banach-Alaoglu-Bourbaki)**.** *Para todo espaço normado* E*, a bola unitária* $B_{E^*}(0; 1)$ *é compacta na topologia fraca-estrela* $\sigma(E^*, E)$ *de* E^**.*

Demonstração. Ver [3, Teorema 6.3.9, pág. 156]. □

1.1.9 Espaços estritamente convexos e espaços uniformemente convexos

Definição 1.1.19 (Espaço estritamente convexo). *Um espaço normado* $(E, \|\cdot\|)$ *é* estritamente convexo *se, dados quaisquer* $\xi, \eta \in E$, *com* $\xi \neq \eta$ *e* $\| \xi \| = \| \eta \| = 1$, *vale a desigualdade* $\| (\xi + \eta)/2 \| < 1$.

Do ponto de vista geométrico, E é estritamente convexo se o ponto médio de qualquer segmento ligando dois pontos distintos da esfera unitária $S(0;1)$ está no interior da bola unitária fechada $\bar{B}(0;1)$.

Todo espaço de Hilbert é estritamente convexo e o espaço $C[a,b]$ das funções contínuas definidas no intervalo $[a,b] \subset \mathbb{R}$ não é estritamente convexo [6, Lema 6.2-4, pág. 333].

Definição 1.1.20 (Espaço uniformemente convexo). *Um espaço normado* $(E, \|\cdot\|)$ *é* uniformemente convexo *se, para cada* $\varepsilon \in (0,2]$, *existe* $\delta = \delta(\varepsilon) > 0$ *tal que,* $\|x+y\| \leqslant 2(1-\delta)$ *sempre que* $x, y \in B_E$ *e* $\| x - y \| \geqslant \varepsilon$.

Todo espaço de Hilbert é uniformemente convexo [3, Exemplo 6.6.2, pág. 167]. Os espaços $(\mathbb{R}^2, \|\cdot\|_\infty)$, $(\mathbb{R}^2, \|\cdot\|_1)$, ℓ_1 e ℓ_∞ não são uniformemente convexos [3, Exemplo 6.6.3, pág. 167]. Para todo $1 < p < \infty$, $L_p(X, \Sigma, \mu)$ é uniformemente convexo [3, Teorema 6.6.12, pág. 172].

Um importante resultado nesse contexto é o Teorema de Milman-Pettis [4, Teorema 3.31, pág. 77], que garante que todo espaço uniformemente convexo é reflexivo. Segue imediatamente deste resultado que todo espaço de Hilbert e os espaços L_p e ℓ_p $(1 < p < \infty)$ são sempre reflexivos.

Observe que segue imediatamente das definições acima que todo espaço uniformemente convexo é estritamente convexo.

Sejam agora $(X, \|\cdot\|)$ um espaço normado real e $(X^*, \|\cdot\|_*)$ seu dual. Denotemos ainda por X^*_w o dual de X munido da topologia fraca-estrela.

Levando em conta [1] e [13], definiremos a seguinte aplicação:

Definição 1.1.21. *Seja* X *um espaço de Banach. A* aplicação de dualidade *é dada por*

$$J : X \rightrightarrows X^*, \ J(u) = \{\psi \in X^* : \psi(u) = \|u\|^2 = \|\psi\|_*^2\}, \ \forall\, u \in X. \tag{1.3}$$

Em [13, Exemplo 1, pág. 7] mostra-se que $J = \partial\varphi$ é o subdiferencial[1] $\partial\varphi$ de φ, onde $\varphi : X \to \mathbb{R}$ é dada por $\varphi(x) = \frac{1}{2}\|x\|^2$.

Os dois próximos resultados envolvendo a aplicação $J : X \rightrightarrows X^*$, respectivamente os teoremas 1.1.5 e 1.1.6, serão utilizados ao longo da demonstração do resultado principal deste trabalho para mostrar a convergência do método.

Teorema 1.1.5. *Seja* X *um espaço de Banach. Se o espaço dual* X^* *é estritamente convexo, então* $J : X \to X^*$ *está bem definida e é contínua de* X *em* X^*_w.

Demonstração. Para cada $u \in X$, $J(u)$ é um subconjunto fechado e convexo de X^*. De fato, $J(u)$ é convexo, pois para quaisquer $\psi_1, \psi_2 \in J(u)$, temos $\psi_1(u) = \|u\|^2 = \|\psi_1\|_*^2$ e $\psi_2(u) = \|u\|^2 = \|\psi_2\|_*^2$. Assim, para todo $t \in [0, 1]$,

$$((1-t)\psi_1 + t\psi_2)(u) = (1-t)\psi_1(u) + t\psi_2(u) = (1-t)\|u\|^2 + t\|u\|^2 = \|u\|^2,$$

e portanto, $(1-t)\psi_1 + t\psi_2 \in J(u)$. Para mostrar que $J(u)$ é fechado, note que dada uma sequência $(\psi_n)_{n=1}^{\infty} \subset J(u)$, tal que $\psi_n \xrightarrow{w^*} \psi$, temos que $\psi_n(u) = \|u\|^2 = \|\psi_n\|_*^2$, para todo $n \in \mathbb{N}$, e da Proposição 1.1.6, item (c), que $\psi(u) = \|u\|^2 = \|\psi\|_*^2$. Logo, $\psi \in J(u)$ e portanto, $J(u)$ é fechado. Agora, como $J(u) \subset \partial B$, onde $B = B(0, \|u\|)$ em X^*, e por hipótese X^* é estritamente convexo, então segue que $J(u)$ consiste de um único ponto e portanto, que $J : X \longrightarrow X^*$ está bem definida.

Para mostrar a continuidade de J de X em X^*_w, seja então $(u_n)_{n=1}^{\infty} \subset X$ uma sequência tal que $u_n \longrightarrow u_0 \in X$. Seja u_0^* um ponto de acumulação fraca-estrela qualquer de $(J(u_n))_{n=1}^{\infty}$, que existe, pois a bola unitária de X^* é compacta na topologia fraca-estrela (Ver Teorema 1.1.4, acima, e [2, Capítulo 7, Proposição 12, pág. 191]). Assim, existe uma subsequência $(J(u_{n_k}))_{k=1}^{\infty} = (\varphi_k)_{k=1}^{\infty}$ de $(J(u_n))_{n=1}^{\infty}$ tal que $\varphi_k \xrightarrow{w^*} u_0^*$ $(k \to \infty)$. Isto é, dado um $\varepsilon > 0$ qualquer, existe $k_0 \in \mathbb{N}$ tal que $\|\varphi_k - u_0^*\|_* < \varepsilon$ sempre que $k \geqslant k_0$. Logo, para $k \geqslant k_0$, $u_0^*(u_{n_k}) = \varphi_k(u_{n_k}) = \|u_{n_k}\|^2 = \|\varphi_k\|_*^2$. Como $\varphi_k \xrightarrow{w^*} u_0^*$ e E é espaço de Banach, pela Proposição 1.1.6, item (b), segue que

$$\|u_0^*\| \leqslant \liminf_{k\to\infty} \|\varphi_k\|_* = \liminf_{k\to\infty} \|u_{n_k}\| \leqslant \|u_0\|.$$

Temos que $u_0^*(u_0) = \|u_0\|^2 \geqslant \|u_0^*\|_*^2$, uma vez que $\varphi_k \in B[0, \|u_0\|]$ (em X^*) e a bola fechada de raio $\|u_0\|$ em X^* é fechada na topologia fraca-estrela [3, Teorema 6.4.5, pág.

[1] Ver Definição1.2.5.

160]. Consequentemente, temos

$$
\begin{aligned}
u_0^*(u_0) &= u_0^*(\lim_{k\to\infty} u_{n_k}) = \lim_{k\to\infty} u_0^*(u_{n_k}) = \lim_{k\to\infty} \varphi_k(u_{n_k}) \\
&= \lim_{k\to\infty} \|u_{n_k}\|^2 = \lim_{k\to\infty} \|\varphi_k\|_*^2 \leqslant \|u_0^*\|_*^2,
\end{aligned}
$$

e assim,

$$\|u_0\|^2 = \|u_0^*\|_*^2 = u_0^*(u_0).$$

Em outras palavras, $u_0^* = J(u_0)$, e então

$$J(u_n) \xrightarrow{w^*} J(u_0).$$

□

Observação 1.1.1. *Quando X é um espaço de Banach reflexivo, seu dual X^* também é reflexivo [3, Proposição 4.3.13, pág. 94]. Usando isto e o fato em [13, Teorema 1.1, pág. 2], podemos dizer que o teorema acima é valido para todo espaço de Banach reflexivo, apenas tomando uma norma equivalente.*

Lema 1.1.1. *Seja X é um espaço de Banach uniformemente convexo. Suponha que $u_n \xrightarrow{w} u$ e $\limsup_{n\to\infty} \|u_n\| \leqslant \|u\|$. Então $u_n \longrightarrow u$ se $n \longrightarrow \infty$.*

Demonstração. Podemos assumir que $u \neq 0$. Por hipótese, dado $\varphi \in X^*$, temos que $\varphi(u_n) \to \varphi(u)$ para todo $u \in X$, e assim, pela semicontinuidade fraca inferior da norma em X,

$$\|u\| \leqslant \liminf_{n\to\infty} \|u_n\| \leqslant \|u\|.$$

Portanto, $\|u\| = \lim_{n\to\infty} \|u_n\|$. Agora, sejam $y_n = \frac{u_n}{\|u_n\|}$ e $y = \frac{u}{\|u\|}$. Claramente se vê que $y_n \xrightarrow{w} y$ se $n \to \infty$. Vamos agora supor que $y_n \nrightarrow y$ e argumentar por contradição. De fato, neste caso existe uma subsequência $(y_{n_k})_{k=1}^{\infty}$, tal que $\|y_{n_k} - y\| \geqslant \varepsilon$, de modo que existe $\delta > 0$ tal que $\|y_{n_k} + y\| < 2(1-\delta)$. Agora, fazendo $n_k \to \infty$, temos $\|y\| \leqslant (1-\delta)$. Contradição. □

Observação 1.1.2. *A partir de agora (salvo menção em contrário) utilizaremos o símbolo $\langle T, x\rangle$ para representar o operador T aplicado ao vetor x, isto é, $T(x) := \langle T, x\rangle$.*

Teorema 1.1.6. *Seja X um espaço de Banach. Se X^* for uniformemente convexo, então $J : X \to X^*$ é uniformemente contínua sobre todos os subconjuntos limitados de X.*

Demonstração. Suponha agora que X^* é uniformemente convexo. Dada a sequência $(u_n)_{n=1}^{\infty}$ em X, admitiremos que existem subsequências $(a_n)_{n=1}^{\infty}$ e $(b_n)_{n=1}^{\infty}$ tais que $\|a_n\| \leqslant M$, $\|b_n\| \leqslant M$, $\|a_n - b_n\| \to 0$, mas $\|J(a_n) - J(b_n)\|_* \geqslant \varepsilon$, para algum $\varepsilon > 0$ e para todo $n \in \mathbb{N}$ e argumentaremos por contradição.

Sejam $z_n = \frac{a_n}{\|a_n\|}$, $y_n = \frac{b_n}{\|b_n\|}$. Podemos supor, sem perda de generalidade, que $\|a_n\| \geqslant \alpha > 0$ e $\|b_n\| \geqslant \alpha > 0$ para todo $n \in \mathbb{N}$. Temos que

$$\|z_n - y_n\| = \|(a_n/\|a_n\|) - (b_n/\|b_n\|)\| = \frac{\|(a_n\|b_n\| - b_n\|a_n\|)\|}{\|a_n\| \cdot \|b_n\|} \leqslant$$

$$\frac{\|(a_n\|b_n\| - b_n\|a_n\|)\|}{\alpha^2} = \frac{\|((a_n - b_n) \cdot \|a_n\|) + (a_n \cdot (\|b_n\| - \|a_n\|))\|}{\alpha^2} \leqslant$$

$$\frac{\|a_n - b_n\| \cdot \|a_n\| + \|a_n\| \cdot \|(\|b_n\| - \|a_n\|)\|}{\alpha^2} \leqslant \frac{\|a_n - b_n\|2M}{\alpha^2} \underset{n\to\infty}{\longrightarrow} 0$$

e, além disso,

$$\langle J(z_n) + J(y_n), z_n \rangle = \langle J(z_n), z_n \rangle + \langle J(y_n), z_n \rangle = \|z_n\|^2 + \langle J(y_n), y_n + (z_n - y_n) \rangle$$

$$= \|z_n\|^2 + \|y_n\|^2 - \langle J(y_n), y_n - z_n \rangle \geqslant 2 - \|z_n - y_n\|,$$

pois $\|z_n\|^2 = \|y_n\|^2 = 1$ e

$$\langle J(y_n), z_n - y_n \rangle \leqslant |\langle J(y_n), z_n - y_n \rangle| \leqslant \|J(y_n)\|_* \|z_n - y_n\| = \|z_n - y_n\|.$$

E assim,

$$\frac{1}{2}\|J(y_n) + J(z_n)\|_* \geqslant 1 - \frac{1}{2}\|z_n - y_n\|, \ \forall\, n \in \mathbb{N}. \tag{1.4}$$

Visto que $\|J(y_n)\| = \|J(z_n)\| = 1$ para todo $n \in \mathbb{N}$, e X^* é uniformemente convexo, concluimos que $\lim_{n\to\infty}(J(z_n) - J(y_n)) = 0$, pois, caso contrário, existiria um $n_0 \in \mathbb{N}$ tal que a desigualdade em (1.4) geraria uma contradição se $n = n_0$. Por outro lado,

$$J(a_n) - J(b_n) = \|b_n\|(J(z_n) - J(y_n)) + (\|a_n\| - \|b_n\|)J(b_n),$$

e assim concluímos que

$$J(a_n) - J(b_n) \underset{n\to\infty}{\longrightarrow} 0.$$

□

1.2 Conceitos e resultados de Análise Convexa

Nesta seção procuramos elencar definições e conceitos básicos de Análise Convexa que serão necessários mais adiante no decorrer do presente texto. O leitor interessado pode encontrar mais conteúdo sobre esses conceitos nas referências [7], [8], [9], [10] e [12].

1.2.1 Conjuntos convexos e funções convexas

Definição 1.2.1 (Conjunto convexo). *Um subconjunto* $C \subset E$ *de um espaço vetorial* E *é dito* convexo *se, para quaisquer pontos* $a, b \in C$*, o segmento* $[a, b]$ *está contido em* C*. Onde* $[a, b]$ *é definido por*

$$y \in [a, b] \Leftrightarrow \exists\, t \in [0, 1] \text{ tal que } y = ta + (1 - t)b.$$

Abaixo temos a representação geométrica de um conjunto convexo e um conjunto não-convexo, respectivamente.

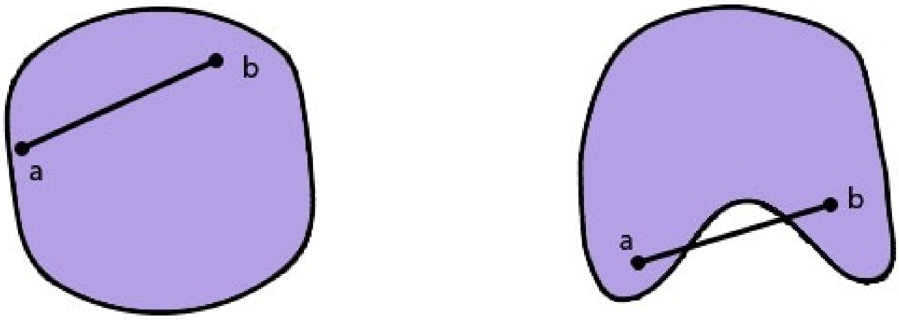

Figura 1.1: O conjunto à esquerda é convexo, enquanto que o da direita não é convexo.

Exemplo 1.2.1. *Qualquer subespaço vetorial de* E *é convexo. Com efeito, dado um subespaço vetorial* $F \subset E$*, das propriedades de subespaço, para todos* $x, y \in F$ *e* $t \in [0, 1]$*, temos que* $tx \in F$ *e* $(1 - t)y \in F$ *e, consequentemente,* $tx + (1 - t)y \in F$*. Portanto,* F *é convexo.*

Exemplo 1.2.2. *Sejam* E *um espaço normado e* $R > 0$*. Então* $C = \{x \in E : \|x\| \leqslant R\}$ *é convexo. De fato, dados quaisquer* $y, z \in C$*, então* $\|y\| \leqslant R$ *e* $\|z\| \leqslant R$*. Assim, para todo* $t \in [0, 1]$*,*

$$\|(1 - t)z + ty\| \leqslant \|(1 - t)z\| + \|ty\| = |(1 - t)|\|z\| + |t|\|y\| \leqslant (1 - t)R + tR = R,$$

donde segue que $(1 - t)z + ty \in C$*. Consequentemente,* C *é convexo.*

Exemplo 1.2.3. *Sejam* E *um espaço normado e* $R \in \mathbb{R}^+$*. Então* $C = \{x \in E : \|x\| = R\}$ *é convexo se, e somente se,* $R = 0$*. De fato, se* R *é igual a* 0*, nada há que provar, pois nesse caso* $C = \{0\}$ *e é óbvio que todo conjunto unitário é convexo. Reciprocamente, se* *supormos que* $C = \{x \in E : \|x\| = R\}$ *é convexo com* $R > 0$*, então, tomando* $x, y \in C$, $x \neq y$*, temos que* $\|x\| = \|y\| = R$ *e*

$$\left\|\frac{1}{2}(x+y)\right\| = \frac{1}{2}\|x+y\| \leqslant \frac{1}{2}\left(\|x\| + \|y\|\right) = \frac{1}{2}(R+R) = R.$$

Suponha que $\left\|\frac{1}{2}(x+y)\right\| = R$. *Segue daí que* $\|x+y\| = \|x\| + \|y\| = 2R$*, o que ocorre se, e somente se,* $x = y$. *Mas assumimos inicialmente que* $x \neq y$. *Esta contradição mostra que* C *não pode ser convexo se* $R > 0$.

Mais exemplos de conjuntos convexos podem ser encontrados na literatura em [8], [10] e [12], por exemplo.

Definição 1.2.2 (Função convexa)**.** *Se* E *é um espaço vetorial e* $C \subset E$ *é convexo, diz-se que* $\varphi : C \to \mathbb{R}$ *é uma* função convexa *se, para todos* $a, b \in C$ *e todo* $t \in [0,1]$*, vale*

$$\varphi(ta + (1-t)b) \leqslant t\varphi(a) + (1-t)\varphi(b). \tag{1.5}$$

Se a desigualdade em (1.5) *é estrita para todo* $t \in (0,1)$*, então* φ *é dita* estritamente convexa.

Embora a definição acima seja bastante boa, adotaremos a seguinte postura: se o domínio de φ não for o espaço inteiro, escreveremos $\varphi(x) = \infty$ nos pontos em que φ não está definida. Observe que aqui a nova φ será convexa se e somente a antiga também for. Indo um pouco mais longe, poderíamos admitir que φ tomasse o valor $-\infty$, mas isto não nos levaria tão longe assim. Dessa forma, de agora em diante, uma função convexa será sempre definida em todo o espaço E e tomará valores em $\mathbb{R} \cup \{\infty\}$ (salvo menção em contrário). Excluiremos também o caso $\varphi \equiv \infty$.

O *domínio efetivo* de φ é o convexo $\mathcal{D} := \mathrm{dom}(\varphi) = \{x \in E : \varphi(x) < \infty\}$.

Baseados em [8] e [10], apresentamos a seguir alguns exemplos de funções convexas.

Exemplo 1.2.4. $E = \mathbb{R}$ *e* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = e^x$ *é estritamente convexa.*

Exemplo 1.2.5. *Sejam* $E = \mathbb{R}$ *e* $1 \leqslant p < \infty$. *Então* $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ *dada por*

$$f(x) = \begin{cases} x^p, \text{ se } x \geqslant 0 \\ \infty, \text{ c.c.} \end{cases}$$

é convexa.

Exemplo 1.2.6. *Sejam* $\mathsf{E} = \mathbb{R}^n$, $a \in \mathbb{R}^n$ *e* $\alpha \in \mathbb{R}$. *Então* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *dada por* $f(x) = \langle x, a\rangle + \alpha$ *é convexa.*

Exemplo 1.2.7 (Norma euclidiana)**.** *Seja* $\mathsf{E} = \mathbb{R}^n$. *A função* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *dada por* $f(x) = \langle x, x\rangle^{1/2}$ *é convexa.*

Definição 1.2.3. *Seja* $\varphi : \mathsf{E} \to \mathbb{R} \cup \{\infty\}$. *O* epígrafo *de* φ *é o seguinte subconjunto de* $\mathsf{E} \times \mathbb{R} \cup \{\infty\}$*:*

$$\mathsf{E}_\varphi = \{(x, t) \in \mathsf{E} \times \mathbb{R} \cup \{\infty\} \mid \varphi(x) \leqslant t\}$$

Apresentaremos agora uma outra caracterização de função convexa a partir da convexidade de seu *epígrafo*.

Teorema 1.2.1. *A função* $\varphi : \mathsf{E} \to \mathbb{R} \cup \{\infty\}$ *é convexa se, e somente se, seu epígrafo é convexo.*

Demonstração. Suponhamos primeiro que E_φ é convexo. Sejam $x, y \in \mathcal{D}$ quaisquer. Temos que $(x, \varphi(x)) \in \mathsf{E}_\varphi$ e $(y, \varphi(y)) \in \mathsf{E}_\varphi$. Como E_f é convexo, para todo $t \in [0, 1]$, temos que

$$t(x, \varphi(x)) + (1 - t)(y, \varphi(y)) = (tx + (1 - t)y, t\varphi(x) + (1 - t)\varphi(y)) \in \mathsf{E}_f$$

Pela definição de epígrafo, a igualdade anterior significa dizer que

$$\varphi(tx + (1 - t)y) \leqslant t\varphi(x) + (1 - t)\varphi(y),$$

o que mostra que φ é convexa.

Reciprocamente, suponhamos que φ seja convexa. Sejam $(x, c_1) \in \mathsf{E}_\varphi$ e $(y, c_2) \in \mathsf{E}_\varphi$. Como $\varphi(x) \leqslant c_1$ e $\varphi(y) \leqslant c_2$, pela convexidade de φ, para todo $t \in [0, 1]$ tem-se

$$\begin{aligned}\varphi(tx + (1 - t)y) &\leqslant t\varphi(x) + (1 - t)\varphi(y)\\ &\leqslant tc_1 + (1 - t)c_2,\end{aligned}$$

o que significa que

$$t(x, c_1) + (1 - t)(y, c_2) = (tx + (1 - t)y, tc_1 + (1 - t)c_2) \in \mathsf{E}_\varphi.$$

Portanto, E_φ é convexo. $\square$

1.2.2 O subdiferencial de uma função convexa

Definição 1.2.4. *Sejam* X *um espaço de Banach,* $U \subset X$ *aberto e* $\varphi : U \to \mathbb{R}$*. A derivada de Gateaux de* φ *no ponto* $y \in U$ *na direção* $z \in X$ *é definida como*

$$\langle \nabla \varphi(y), z \rangle = \lim_{t \to 0^+} \frac{\varphi(y + tz) - \varphi(y)}{t}. \tag{1.6}$$

Se o limite acima existe para todo $z \in X$*, dizemos que a função* φ *é diferenciável de Gâteaux no ponto* y*.*

Se a aplicação definida por $T := \nabla\varphi(y) : X \to Y$ é linear, então $\nabla\varphi(y)$ é chamada de derivada de Gateaux (linear) de φ em y. Frequentemente, refere-se a $\nabla\varphi(y)$ também como o *gradiente* de φ em y. Observe ainda que a derivada de Gâteaux é uma generalização do conceito de derivada direcional.

Definição 1.2.5. *Sejam* E *um espaço de Banach e* $f : E \to \mathbb{R} \cup \{\infty\}$ *uma função. Dizemos que* $\xi \in E^*$ *é um* subgradiente *de* f *em* u *se*

$$f(v) \geqslant f(u) + \langle \xi, v - u \rangle, \tag{1.7}$$

para todo $v \in E$*. O conjunto de todos os subgradientes de* f *em* u *é chamado de* subdiferencial *de* f *em* u *e é denotado por* $\partial f(u)$*.*

Exemplo 1.2.8. *Seja* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = |x|$*. Então* $\partial f(0) = [-1, 1]$*. De fato, para* $u = 0$*, se* $\xi \in \partial f(0)$ *e* $v \in \mathbb{R}$*, por definição temos*

$$f(v) \geqslant f(u) + \xi \cdot (v - u),$$

isto é,

$$|v| \geqslant |0| + \xi \cdot (v - 0) = \xi \cdot v.$$

Se $v > 0$*, a desigualdade acima implica que* $\xi \leqslant 1$*.*
Se $v < 0$*, a mesma desigualdade implica que* $\xi \geqslant -1$*.*
Se $v = 0$*, então* ξ *pode ser qualquer número real.*
Logo, conclui-se que $\partial f(0) = [-1, 1]$*.*

Proposição 1.2.1. *Seja* X *um espaço de Banach e* $f : X \to \mathbb{R} \cup \{\infty\}$ *uma função convexa finita e contínua em* $u \in X$*. Então* $\partial f(u) \neq \emptyset$*.*

Demonstração. Ver [9, Proposição 10.1.19, pág. 256]. □

1.2.3 Semicontinuidade inferior e operadores monótonos

Definição 1.2.6 (Função semicontínua inferior)**.** *Seja* X *um espaço de Banach e seja* $f : X \to \mathbb{R} \cup \{\infty\}$. *Dizemos que* f *é* semicontínua inferior *se*

$$f(u) \leqslant \liminf_{k\to\infty} f(u^k),$$

onde $\{u^k\}$ *é uma sequência em* X *que converge a* u.

A partir de [7, Exemplo 3.2., pág. 40], temos o seguinte exemplo:

Exemplo 1.2.9 (Função indicadora)**.** *Sejam* X *um espaço de Banach e* $C \subset X$ *um conjunto convexo e não-vazio. A função indicadora*

$$\delta_C(x) = \begin{cases} 0, & \text{se } x \in C \\ \infty, & \text{c.c.} \end{cases}$$

é convexa e própria, e é semicontínua inferior se, e somente se, C *é fechado.*

Definição 1.2.7 (Operador monótono)**.** *Seja* X *um espaço de Banach. Dizemos que um operador* $T : X \rightrightarrows P(X^*)$ *é dito* monótono *se*

$$\langle \eta - \xi, x - y \rangle \geqslant 0, \tag{1.8}$$

para todo $\eta \in T(x)$ *e todo* $\xi \in T(y)$. *Um operador monótono é chamado de* maximal *se para qualquer operador monótono* $\tilde{T}$ *tal que* $T(x) \subseteq \tilde{T}(x)$ *para todo* $x \in X$, *então* $\tilde{T} = T$.

Como exemplos de operadores monótonos podemos citar [7, Exemplo 2.2., pág 17]:

Exemplo 1.2.10. $f : X \to \mathbb{R}\cup\{\infty\}$ *é uma função convexa e contínua, então* $T(x) = \partial f(x)$ *é um operador monótono.*

Exemplo 1.2.11. *Se* H *é um espaço de Hilbert real com produto interno* $\langle \cdot, \cdot \rangle$ *e* $T : H \to H$ *é uma aplicação linear, então* T *é monótono se, e somente se,* $\langle Tx, x \rangle \geqslant 0$ *para todo* $x \in H$.

Capítulo 2

Distâncias de Bregman e Funções Totalmente Convexas

As distâncias de Bregman foram introduzidas em [15] por Lev Bregman na década de 1960. Em seu trabalho, Bregman definiu a função $D(x, y) : S \times S \to \mathbb{R}$, onde $S \subset X$ é um subconjunto convexo de um espaço topológico X. A partir da função D, Bregman desenvolve um algoritmo que visa resolver o seguinte problema: encontrar um ponto na interseção de uma família de subcontjuntos convexos fechados do espaço topológico X, que também é conhecido como problema de viabilidade. Como aplicação, ele usa este método na resolução de alguns problemas de programação convexa.

A definição da função D de Bregman é baseada em seis axiomas, alguns dos quais atualmente fazem parte da definição de *função de Bregman*, termo este que, assim como *distância de Bregman* foi introduzido por Cersor e Lent em [16].

Neste capítulo veremos a definição de distâncias de Bregman e algumas de suas propriedades. Veremos também as noções de *módulo de convexidade* e *funções totalmente convexas*, que relacionam a distância de Bregman com a distância usual induzida pela norma do espaço X, assim como alguns resultados envolvendo esses termos que nos auxiliarão mais adiante na prova do teorema principal.

2.1 Distâncias de Bregman

Definição 2.1.1. *Sejam* X *um espaço de Banach reflexivo e* $\varphi : X \to \mathbb{R} \cup \{\infty\}$ *uma função convexa, própria e semicontínua inferior. Denote o domínio efetivo de* φ *por*

$\mathcal{D} := \mathrm{dom}(\varphi)$ *e seu interior por* $\mathcal{D}^o$. *Suponha que* $\mathcal{D}^o \neq \emptyset$ *e que* φ *é Gateaux diferenciável em* $\mathcal{D}^o$. *A* distância de Bregman *com respeito a* φ *é a função* $D_\varphi : \mathcal{D} \times \mathcal{D}^o \to \mathbb{R}$ *definida por*

$$D_\varphi(x, y) := \varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y\rangle \tag{2.1}$$

onde

$$\langle \nabla\varphi(y), z\rangle = \lim_{t\to 0^+} \frac{\varphi(y + tz) - \varphi(y)}{t} \tag{2.2}$$

é a derivada de Gateaux de φ *no ponto* y *na direção* z.

Exemplo 2.1.1. *Seja* X *um espaço de Hilbert real com produto interno* $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$, $\mathcal{D} = X$ *e* $\varphi : X \to \mathbb{R}$, $\varphi(x) = \frac{1}{2}\|x\|^2$. *A distância de Bregman associada a* φ *é a função* $D_\varphi : \mathcal{D} \times \mathcal{D}^o \to \mathbb{R}$, $D_\varphi(x, y) = \frac{1}{2}\|x - y\|^2$. *De fato, da Definição 2.1.1, temos que*

$$\begin{aligned}
D_\varphi(x, y) &= \varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y\rangle \\
&= \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 - \lim_{t\to 0^+} \frac{\|y + t(x - y)\|^2 - \|y\|^2}{2t} \\
&= \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 - \lim_{t\to 0^+} \frac{\langle y + t(x - y), y + t(x - y)\rangle - \|y\|^2}{2t} \\
&= \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 - \lim_{t\to 0^+} \frac{\|y\|^2 + 2t\langle y, x - y\rangle + t^2\|x - y\|^2 - \|y\|^2}{2t} \\
&= \frac{1}{2}\langle x, x\rangle - \frac{1}{2}\langle y, y\rangle - \langle y, x - y\rangle = \frac{1}{2}\langle x, x\rangle - \frac{1}{2}\langle y, y\rangle - \langle y, x\rangle + \langle y, y\rangle \\
&= \frac{1}{2}\langle x, x\rangle + \frac{1}{2}\langle y, y\rangle - \frac{1}{2}\langle y, x\rangle - \frac{1}{2}\langle y, x\rangle = \frac{1}{2}\langle x - y, x\rangle + \frac{1}{2}\langle y, y - x\rangle \\
&= \frac{1}{2}\langle x - y, x\rangle - \frac{1}{2}\langle x - y, y\rangle = \frac{1}{2}\langle x - y, x - y\rangle = \frac{1}{2}\|x - y\|^2.
\end{aligned}$$

Exemplo 2.1.2. $X = \mathbb{R}^n$, $\mathcal{D} = \mathbb{R}^n_{++}$, $\varphi(x) = \sum_{j=1}^n x_j \log x_j$, *estentida para a fronteira* $\partial\mathbb{R}^n_+$ *com a convenção* $0 \cdot \log 0 = 0$. *Neste caso, a distância de Bregman associada à* φ *é a função* $D_\varphi(x, y) = \sum_{j=1}^n (x_j \log \frac{x_j}{y_j} + y_j - x_j)$. *De fato, utilizando a fórmula em (2.1), temos*

$$D_\varphi(x, y) := \varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y\rangle = \sum_{j=1}^n x_j \log x_j - \sum_{j=1}^n y_j \log y_j - \langle \nabla\varphi(y), x - y\rangle,$$

onde $\langle \nabla\varphi(y), x - y\rangle$ *é a derivada de Gâteaux* φ *no ponto* y *na direção em* $(x - y)$, *a qual é*

$$\langle \nabla\varphi(y), x - y\rangle = \lim_{t\to 0^+} \frac{\varphi(y + t(x - y)) - \varphi(y)}{t} =$$

$$\lim_{t\to 0^+} \frac{\sum_{j=1}^n (y_j + t(x_j - y_j)) \log(y_j + t(x_j - y_j)) - \sum_{j=1}^n y_j \log y_j}{t} =$$

$$\lim_{t\to 0^+} \frac{t\sum_{j=1}^n (x_j - y_j)\log(y_j + t(x_j - y_j)) + \sum_{j=1}^n y_j(\log(y_j + t(x_j - y_j)) - \log y_j)}{t} =$$

$$\lim_{t\to 0^+} \frac{t\sum_{j=1}^n (x_j - y_j)\log(y_j + t(x_j - y_j))}{t} + \lim_{t\to 0^+} \frac{\sum_{j=1}^n y_j(\log(y_j + t(x_j - y_j)) - \log y_j)}{t} =$$

$$\sum_{j=1}^n (x_j - y_j)\log y_j + \lim_{t\to 0^+} \frac{\sum_{j=1}^n y_j(\log(y_j + t(x_j - y_j)) - \log y_j)}{t},$$

onde

$$\lim_{t\to 0^+} \frac{\sum_{j=1}^n y_j(\log(y_j + t(x_j - y_j)) - \log y_j)}{t} = \sum_{j=1}^n (x_j - y_j).$$

Para provar isto, basta notar que para cada $j = 1, .., n$, *o limite*

$$\lim_{t\to 0^+} \frac{y_j(\log(y_j + t(x_j - y_j)) - \log y_j)}{t} = y_j \cdot \lim_{t\to 0^+} \frac{\log((1-t)y_j + tx_j)/y_j)}{t}$$

$$= y_j \cdot \lim_{t\to 0^+} \frac{\frac{y_j}{(1-t)y_j + tx_j} \cdot \frac{(x_j - y_j)y_j}{y_j^2}}{1} = y_j \cdot \frac{(x_j - y_j)}{y_j} = x_j - y_j,$$

onde $\lim_{t\to 0^+} \frac{\log((1-t)y_j + tx_j)/y_j)}{t} = \frac{x_j - y_j}{y_j}$ *pela Regra de L'Hôspital. Logo,*

$$\langle \nabla\varphi(y), x - y\rangle = \sum_{j=1}^n (x_j - y_j)\log y_j + \sum_{j=1}^n (x_j - y_j),$$

e segue daí que

$$D_\varphi(x, y) = \sum_{j=1}^n x_j \log x_j - \sum_{j=1}^n y_j \log y_j - \sum_{j=1}^n (x_j - y_j)\log y_j - \sum_{j=1}^n (x_j - y_j) =$$
$$\sum_{j=1}^n x_j(\log x_j - \log y_j) + (y_j - x_j) = \sum_{j=1}^n (x_j \log \tfrac{x_j}{y_j} + y_j - x_j).$$

Exemplo 2.1.3. $X = \mathbb{R}^n$, $\mathcal{D} = \mathbb{R}^n_{++}$ *e* $\varphi(x) = \sum_{i=1}^n (x_i - \log x_i + 1)$. *Então*

$$D_\varphi(x, y) = \sum_{i=1}^n \log\left(\frac{y_i}{x_i}\right) + \frac{x_i}{y_i} - 1.$$

Com efeito, utilizando a fórmula da definição 2.1.1, temos

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

$$= \sum_{i=1}^n (x_i - \log x_i + 1) - \sum_{i=1}^n (y_i - \log y_i + 1) - \langle \nabla\varphi(y), x - y\rangle,$$

onde

$$\langle \nabla\varphi(y), x-y\rangle = \lim_{t\to 0^+} \frac{\varphi(y+t(x-y))-\varphi(y)}{t}$$

$$= \lim_{t\to 0^+} \frac{\sum_{i=1}^n \{y_i + t(x_i-y_i) - \log[y_i+t(x_i-y_i)]+1\} - \sum_{i=1}^n (y_i - \log y_i + 1)}{t}$$

$$= \lim_{t\to 0^+} \frac{\sum_{i=1}^n [t(x_i-y_i) - \log(y_i+t(x_i-y_i)] - \sum_{i=1}^n(-\log y_i)}{t}$$

$$= \lim_{t\to 0^+} \frac{\sum_{i=1}^n \left[t(x_i-y_i) + \log\left(\frac{y_i}{y_i+t(x_i-y_i)}\right)\right]}{t}$$

$$= \lim_{t\to 0^+} \frac{\sum_{i=1}^n t(x_i-y_i)}{t} + \lim_{t\to 0^+} \frac{\sum_{i=1}^n \left[\log\left(\frac{y_i}{y_i+t(x_i-y_i)}\right)\right]}{t}$$

$$= \lim_{t\to 0^+} \sum_{i=1}^n \frac{t(x_i-y_i)}{t} + \lim_{t\to 0^+} \sum_{i=1}^n \frac{\left[\log\left(\frac{y_i}{y_i+t(x_i-y_i)}\right)\right]}{t}$$

$$= \sum_{i=1}^n \lim_{t\to 0^+} \frac{t(x_i-y_i)}{t} + \sum_{i=1}^n \lim_{t\to 0^+} \frac{\left[\log\left(\frac{y_i}{y_i+t(x_i-y_i)}\right)\right]}{t}$$

$$= \sum_{i=1}^n \lim_{t\to 0^+} (x_i-y_i) + \sum_{i=1}^n \lim_{t\to 0^+} \frac{\left[\log\left(\frac{y_i}{y_i+t(x_i-y_i)}\right)\right]}{t}$$

Usando a Regra de L'Hôspital para o cálculo do limite no segundo somatório, obtemos

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

$$= \sum_{i=1}^n (x_i-y_i) + \sum_{i=1}^n \lim_{t\to 0^+} \frac{\left(\frac{y_i+t(x_i-y_i)}{y_i}\right)}{1} \cdot \frac{0\cdot[y_i+t(x_i-y_i)] - (x_i-y_i)\cdot y_i}{[y_i+t(x_i-y_i)]^2}$$

$$= \sum_{i=1}^n (x_i-y_i) + \sum_{i=1}^n \lim_{t\to 0^+} \frac{(y_i-x_i)}{y_i+t(x_i-y_i)} = \sum_{i=1}^n (x_i-y_i) + \sum_{i=1}^n \frac{(y_i-x_i)}{y_i}$$

Logo,

$$D_\varphi(x,y) = \sum_{i=1}^n (x_i - \log x_i) - \sum_{i=1}^n (y_i - \log y_i) - \left(\sum_{i=1}^n (x_i-y_i) + \sum_{i=1}^n \frac{(y_i-x_i)}{y_i}\right)$$

$$= \sum_{i=1}^n \log\left(\frac{y_i}{x_i}\right) + \frac{x_i}{y_i} - 1.$$

Note que segue imediatamente da convexidade de φ que, para todo $x \in \mathcal{D}$, $y \in \mathcal{D}^{\circ}$, $D_\varphi(x,y) \geqslant 0$ e $D(x,y) = 0$ se, e somente se, $x = y$. A função $D_\varphi(\cdot,\cdot)$ não é uma distância (métrica) no sentido usual do termo (em geral, não é simétrica e não satisfaz a desigualdade triangular). No entanto, existe a *propriedade dos três pontos*, a qual toma o lugar dessa desigualdade nas provas.

Proposição 2.1.1 (Propriedade dos três pontos)**.** *Seja* D_φ *a distância de Bregman associada à função* $\varphi : X \to \mathbb{R} \cup \{\infty\}$ *como na Definição 2.1.1. Se* $x \in \mathcal{D}, y, z \in \mathcal{D}^o$*, então,*

$$\langle \nabla\varphi(y) - \nabla\varphi(z), z - x\rangle = D_\varphi(x, y) - D_\varphi(x, z) - D_\varphi(z, y). \tag{2.3}$$

Demonstração. De fato, da definição 2.1.1, temos que

$$\begin{aligned}
& D_\varphi(x, y) - D_\varphi(x, z) - D_\varphi(z, y) \\
&= (\varphi(x) - \varphi(y) - \langle \nabla\varphi(y), x - y\rangle) - (\varphi(x) - \varphi(z) - \langle \nabla\varphi(z), x - z\rangle) \\
&\quad -(\varphi(z) - \varphi(y) - \langle \nabla\varphi(y), z - y\rangle) \\
&= -\langle \nabla\varphi(y), x - y\rangle + \langle \nabla\varphi(z), x - z\rangle + \langle \nabla\varphi(y), z - y\rangle \\
&= \langle \nabla\varphi(z), x - z\rangle - \langle \nabla\varphi(y), (x - y) - (z - y)\rangle \\
&= \langle \nabla\varphi(y) - \nabla\varphi(z), z - x\rangle.
\end{aligned}$$

□

Assumiremos a seguinte hipótese básica a respeito de φ.

B1: Os conjuntos de nível à direita de $D_\varphi(x, \cdot)$:

$$R^\alpha_\varphi(x) = \{z \in \mathcal{D}^o | D_\varphi(x, z) \leqslant \alpha\}$$

são limitados para todo $\alpha \geqslant 0$ e para todo $x \in \mathcal{D}$.

As condições suficientes que garantem a hipótese B1 podem ser vistas em [19, Proposição 5, pág. 1322].

Observação 2.1.1. *Uma característica inerente da técnica de Bregman para o desenvolvimento de algoritmos iterativos consiste em assegurar que as sequências* $\{x^k\}_{k\in\mathbb{N}}$ *geradas nestes algoritmos estão contidas no domínio de uma função convexa* φ *e têm a propriedade de que* $\lim_{k\to\infty} D_\varphi(x^k, x) = 0$ *para pontos* $x \in \mathcal{D} := \mathrm{dom}(\varphi)$*, os quais são soluções do problema que se supõe que o algoritmo resolve. Se a função* φ *é escolhida de modo que, para qualquer sequência* $\{x^k\}_{k\in\mathbb{N}} \subset \mathcal{D}$ *e qualquer vetor* $x \in \mathcal{D}$*, tem-se*

$$\lim_{k\to\infty} D_\varphi(x^k, x) = 0 \Rightarrow \lim_{k\to\infty} \|x^k - x\| = 0, \tag{2.4}$$

então a convergência do algoritmo para uma solução do problema é garantida. Resmerita [18] mostrou que as únicas funções convexas satisfazendo (2.4) são as funções totalmente convexas *[18, Proposição 2.2, pág. 3], as quais veremos a seguir.*

2.2 Funções totalmente convexas

Definição 2.2.1 (Função totalmente convexa). *Seja* $\mathbb{R}_+$ *o conjunto dos números reais não-negativos e considere o conjunto* $\mathbb{R}_{++} := \mathbb{R}_+ - \{0\}$. *Definimos o* módulo de convexidade local *da função* φ *no ponto* $x \in \mathcal{D}^o$, $\nu_\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, *como*

$$\nu_\varphi(x,t) := \inf\{D_\varphi(y,x) : y \in \mathcal{D},\ \|y-x\| = t\} \tag{2.5}$$

A função φ *é dita* totalmente convexa *em* x *se* $\nu_\varphi(x,t) > 0$ *sempre que* $t > 0$. *Se* φ *é totalmente convexa para todo* $x \in \mathcal{D}^o$, *então dizemos simplesmente que* φ *é totalmente convexa (em* $\mathcal{D}^o$*).*

Exemplo 2.2.1. *A função* $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *dada por*

$$\varphi(x) = \begin{cases} \sum_{i=1}^n (x_i - \log x_i + 1) & \textit{se } x \in \mathbb{R}^n_{++}, \\ \infty & \textit{c. c.} \end{cases}$$

é totalmente convexa. De fato, uma vez que a distância de Bregman associada a φ *é* $D_\varphi : \mathbb{R}^n_{++} \times \mathbb{R}^n_{++} \to \mathbb{R}$, *dada por* $D_\varphi(x,y) = \sum_{i=1}^n \log\left(\frac{y_i}{x_i}\right) + \frac{x_i}{y_i} - 1$ (Exemplo 2.1.3), *para todo* $t > 0$ *e todo* $x \in \mathbb{R}^n_{++}$ *temos*

$$\begin{aligned} \nu_\varphi(x,t) &= \inf\{D_\varphi(y,x) : y \in \mathbb{R}^n_{++},\ \|y-x\| = t\} = \\ & \inf\left\{\sum_{i=1}^n \log\left(\frac{y_i}{x_i}\right) + \frac{x_i}{y_i} - 1 : y \in \mathbb{R}^n_{++},\ \|y-x\| = t\right\} > 0, \end{aligned}$$

pois $\log\left(\frac{y_i}{x_i}\right) + \frac{x_i}{y_i} - 1 = 0$ *se, e somente se,* $x_i = y_i$ *para todo* $1 \leqslant i \leqslant n$, *o que não ocorre, pois* $\|y-x\| = t > 0$ *garante que* $x \neq y$. *Portanto,* $\nu_\varphi(x,t) > 0$ *para todo* $x \in \mathbb{R}^n_{++}$ *e todo* $t > 0$.

Vejamos agora algumas das propriedades da função módulo de convexidade local.

Proposição 2.2.1. *Se* $x \in \mathcal{D}^o$, *então*

(i) O domínio de $\nu_\varphi(x,\cdot)$ *é um intervalo* $[0,\tau_\varphi)$ *ou* $[0,\tau_\varphi]$, *com* $\tau_\varphi \in (0,\infty]$; *além disso,* τ_φ *é finito se, e somente se,* $\mathcal{D}$ *é limitado;*

(ii) Se $c \in [1,\infty)$ *e* $t \geqslant 0$, *então* $\nu_\varphi(x,ct) \geqslant c\nu_\varphi(x,t)$

(iii) Para quaisquer $s,t \in [0,\infty)$, *vale a desigualdade*

$$\nu_\varphi(x,s+t) \geqslant \nu_\varphi(x,s) + \nu_\varphi(x,t)$$

(iv) A função $\nu_\varphi(x,\cdot)$ *é não-decrescente; além disso,* $\nu_\varphi(x,\cdot)$ *é estritamente crescente se, e somente se,* φ *é totalmente convexa em* x.

Demonstração. (i) Uma vez que $x \in \mathcal{D}^{\circ}$, então existe uma bola fechada de centro x e raio $r > 0$ totalmente contida em $\mathcal{D}$. Portanto, o intervalo $[0, r]$ está contido no domínio de $\nu_{\varphi}(x, \cdot)$. Suponha agora que $\nu_{\varphi}(x, t) < \infty$. De acordo com (2.5), existe um ponto $y_t \in \mathcal{D}$ tal que $\|y_t - x\| = t$. O conjunto $\mathcal{D}$ é convexo, pois é o domínio de uma função convexa. Assim, o segmento de reta $[x, y_t]$ está contido em D. Isso implica que para qualquer $s \in [0, t]$ existe um ponto $y_s \in \mathcal{D}$ tal que $\|y_s - x\| = s$. Por consequência, temos que $[0, t]$ está contido no domínio de $\nu_{\varphi}(x, \cdot)$ sempre que $\nu_{\varphi}(x, t) < \infty$. Isto mostra que o domínio de $\nu_{\varphi}(x, \cdot)$ é convexo em $\mathbb{R}_+$, sendo portanto um intervalo do tipo $[0, \tau_{\varphi})$ ou $[0, \tau_{\varphi}]$, com $\tau_{\varphi} \in (0, \infty]$. Agora, assumindo que $\mathcal{D}$ não é limitado, então, para todo $t > 0$, existe algum ponto $y_t \in \mathcal{D}$ tal que $\|x - y_t\| = t$ e, portanto, $[0, t]$ está contido no domínio de φ, isto é, τ_{φ} não é finito. Reciprocamente, se $\mathcal{D}$ é limitado, então existe algum raio $r > 0$ tal que $\mathcal{D}$ está contido em uma bola de centro x e raio r. Logo, τ_{φ} é finito, pois $\tau_{\varphi} \leqslant r$.

(ii) Se $c = 1$ ou se $t = 0$ ou se $\nu_{\varphi}(x, t) = \infty$, então a desigualdade é óbvia. Por outro lado, seja $\varepsilon > 0$. De acordo com (2.5), existe um ponto $u \in \mathcal{D}$ tal que $\|u - x\| = ct$ e

$$\nu_{\varphi}(x, ct) + \varepsilon > D_{\varphi}(u, x) = \varphi(u) - \varphi(x) - \langle \nabla\varphi(x), u - x \rangle. \tag{2.6}$$

Para todo $\alpha \in (0, 1)$ denote por $u_{\alpha} = \alpha u + (1 - \alpha)x$. Seja $\beta = c^{-1}$ e observe que

$$\|u_{\beta} - x\| = \|\beta u + (1 - \beta)x - x\| = \|\beta u - \beta x\| = \beta\|u - x\| = c^{-1} \cdot ct = t.$$

Note também que, para todo $\alpha \in (0, 1)$,

$$\frac{\alpha}{\beta} u_{\beta} + (1 - \frac{\alpha}{\beta})x = \frac{\alpha}{\beta}[\beta u + (1 - \beta)x] + (1 - \frac{\alpha}{\beta})x = \alpha u + (1 - \alpha)x = u_{\alpha} \tag{2.7}$$

Como φ é convexa, então a função

$$a \to \frac{\varphi(x + \alpha(u - x)) - \varphi(x)}{\alpha}$$

é não decrescente em $(0, 1)$. Portanto, de acordo com (2.2) e (2.6) temos que

$$\nu_{\varphi}(x, ct) + \varepsilon > \varphi(u) - \varphi(x) - \frac{\varphi(x + \alpha(u - x)) - \varphi(x)}{\alpha}$$

para todo $\alpha \in (0, 1)$. Segue daí que

$$\nu_{\varphi}(x, ct) + \varepsilon > \frac{\alpha\varphi(u) + (1 - \alpha)\varphi(x) - \varphi(x + \alpha(u - x))}{\alpha} =$$

$$\frac{\alpha\varphi(u) + (1 - \alpha)\varphi(x) - \varphi(\alpha u + (1 - \alpha)x)}{\alpha} = \frac{\alpha\varphi(u) + (1 - \alpha)\varphi(x) - \varphi(u_{\alpha})}{\alpha} =$$

$$\frac{\alpha\varphi(u) + (1-\alpha)\varphi(x) - \frac{\alpha}{\beta}\varphi(u_\beta) - (1-\frac{\alpha}{\beta})\varphi(x)}{\alpha} + \frac{\frac{\alpha}{\beta}\varphi(u_\beta) + (1-\frac{\alpha}{\beta})\varphi(x) - \varphi(u_\alpha)}{\alpha} =$$

$$\varphi(u) - \varphi(x) + \frac{1}{\beta}(\varphi(x) - \varphi(u_\beta)) + \frac{\frac{\alpha}{\beta}\varphi(u_\beta) + (1-\frac{\alpha}{\beta})\varphi(x) - \varphi(\frac{\alpha}{\beta}u_\beta + (1-\frac{\alpha}{\beta})x)}{\alpha} =$$

$$\frac{\beta\varphi(u) + (1-\beta)\varphi(x) - \varphi(u_\beta)}{\beta} + \frac{\frac{\alpha}{\beta}\varphi(u_\beta) + (1-\frac{\alpha}{\beta})\varphi(x) - \varphi(\frac{\alpha}{\beta}u_\beta + (1-\frac{\alpha}{\beta})x)}{\alpha},$$

onde usamos a igualdade (2.7). O primeiro termo na última soma é não-negativo, pois φ é convexa. Assim,

$$\nu_\varphi(x, ct) + \varepsilon > \frac{\frac{\alpha}{\beta}\varphi(u_\beta) + (1-\frac{\alpha}{\beta})\varphi(x) - \varphi(\frac{\alpha}{\beta}u_\beta + (1-\frac{\alpha}{\beta})x)}{\alpha} =$$

$$\frac{1}{\beta}\left[\varphi(u_\beta) - \varphi(x) - \frac{\varphi(x + \frac{\alpha}{\beta}(u_\beta - x)) - \varphi(x)}{\frac{\alpha}{\beta}}\right].$$

Fazendo $\alpha \to 0^+$ e levando em conta (2.6) e (2.2) temos que

$$\nu_\varphi(x, ct) + \varepsilon > \frac{1}{\beta}D_\varphi(u_\beta, x) = cD_\varphi(u_\beta, x) \geqslant c\nu_\varphi(x, t).$$

Como $\varepsilon > 0$ é arbitrário, o item (ii) está provado.

(iii) Sejam s, t números reais positivos. Do item (ii) temos que

$$\nu_\varphi(x, s+t) = \nu_\varphi(x, \frac{s+t}{s}s) \geqslant \frac{s+t}{s}\nu_\varphi(x, s),$$

e

$$\nu_\varphi(x, s+t) = \nu_\varphi(x, \frac{s+t}{t}t) \geqslant \frac{s+t}{t}\nu_\varphi(x, t),$$

isto é, $s\nu_\varphi(x, s+t) \geqslant (s+t)\nu_\varphi(x, s)$ e $t\nu_\varphi(x, t+s) \geqslant (s+t)\nu_\varphi(x, t)$. Assim, somando ambas as desigualdades, temos $(s+t)\nu_\varphi(x, s+t) \geqslant (s+t)[\nu_\varphi(x, s) + \nu_\varphi(x, t)]$. Como $s + t > 0$, multiplicando ambos os lados por $1/(s+t)$ obtém-se a desilgualdade desejada.

(iv) Suponha que $0 < s < t$. Então,

$$\nu_\varphi(x, t) = \nu_\varphi(x, \frac{t}{s}s) \geqslant \frac{t}{s}\nu_\varphi(x, s) \geqslant \nu_\varphi(x, s), \tag{2.8}$$

onde a primeira desigualdade segue de (ii). Logo, φ é não-decrescente. Se φ é totalmente convexa em x, então a última desigualdade em (2.8) é estrita, pois $\nu_\varphi(x, s) > 0$. Segue então que $\nu_\varphi(x, \cdot)$ é crescente em $(0, \infty]$. Por outro lado, se $\nu_\varphi(x, \cdot)$ é estritamente crescente, então para todo $t > 0$, temos que $\nu_\varphi(x, t) > \nu_\varphi(x, 0) = 0$ e, por definição, φ é totalmente convexa em x. □

A respeito da continuidade da função $\nu_\varphi(x, \cdot)$, temos o seguinte.

Proposição 2.2.2. *Seja* $x \in \mathcal{D}^o$. *Então*

(i) A função $\nu_\varphi(x,\cdot)$ *é contínua à direita em* $t = 0$;

(ii) Se φ *é contínua em* $\mathcal{D}$, *então* $\nu_\varphi(x,\cdot)$ *é contínua à direita em* $[0,\tau_\varphi)$ *quando o conjunto* $\mathcal{D}$ *é aberto;*

(iii) Se X *têm dimensão finita,* $\mathcal{D}$ *é fechado e* φ *é contínua em* $\mathcal{D}$, *então* $\nu_\varphi(x,\cdot)$ *é contínua à esquerda em seu domínio;*

(iv) Se X *têm dimensão finita e* $\mathcal{D} = X$, *então* $\nu_\varphi(x,\cdot)$ *é contínua em* $(0,\tau_\varphi)$.

Demonstração. (i) Seja $\{t_k\}_{k\in\mathbb{N}}$ uma sequência em $(0,1)$ convergindo não decrescentemente a 0. Aplicando o item (ii) da Proposição 2.2.1 temos

$$\nu_\varphi(x,1) \geqslant \nu_\varphi(x,\sqrt{t_k}) = \nu_\varphi(x,\frac{t_k}{\sqrt{t_k}} \geqslant \frac{1}{\sqrt{t_k}}\nu_\varphi(x,t_k).$$

Consequentemente,

$$0 = \nu_\varphi(x,0) \leqslant \lim_{k\to\infty} \nu_\varphi(x,t_k) \leqslant \nu_\varphi(x,1)\lim_{k\to\infty}\sqrt{t_k} = 0.$$

(ii) Sejam $0 < s < t < \tau_\varphi$. Fixe $\varepsilon > 0$. De acordo com (2.5), existe um ponto $y_\varepsilon \in \mathcal{D}$ tal que $\|y_\varepsilon - x\| = s$ e

$$\nu_\varphi(x,s) + \frac{\varepsilon}{4} > D_\varphi(y_\varepsilon,x).$$

De acordo com a Proposição 2.2.1, temos

$$0 \leqslant \nu_\varphi(x,t) - \nu_\varphi(x,s) < \nu_\varphi(x,t) - D_\varphi(y_\varepsilon,s) + \frac{\varepsilon}{4}$$

A função $D_\varphi(\cdot,x)$ é contínua em $\mathcal{D}$, uma vez φ é contínua e, consequentemente, a derivada de Gâteaux $\langle\nabla\varphi(x),\cdot\rangle$ também é contínua. Portanto, existe um número $\delta(\varepsilon) > 0$ tal que, para qualquer $z \in X$ com $\|z - y_\varepsilon\| < \delta(\varepsilon)$, se tem $z \in \mathcal{D}$ e

$$|D_\varphi(z,x) - D_\varphi(y_\varepsilon,x)| < \frac{\varepsilon}{4}$$

Se $0 < t - s < \delta(\varepsilon)$, então o vetor

$$y'_\varepsilon = \frac{t}{s}y_\varepsilon + (1 - \frac{t}{s})x$$

satisfaz

$$\|y'_\varepsilon - y_\varepsilon\| = \|\frac{t}{s}y_\varepsilon + (1-\frac{t}{s})x - y_\varepsilon\| = \|(1-\frac{t}{s})(x-y_\varepsilon)\| =$$

$$|(1-\frac{t}{s})|\|x - y_\varepsilon\| = (\frac{t}{s}-1)s = (\frac{t-s}{s})\cdot s = t - s < \delta(\varepsilon)$$

e

$$\|y'_\varepsilon - x\| = \|\frac{t}{s}y_\varepsilon + (1-\frac{t}{s})x - x\| = \|\frac{t}{s}(y_\varepsilon - x)\| = \frac{t}{s}\|y_\varepsilon - x\| = \frac{t}{s}\cdot s = t$$

Portanto $y'_\varepsilon \in \mathcal{D}$ e, assim,

$$0 \leqslant \nu_\varphi(x,t) - \nu_\varphi(x,s) < D_\varphi(y'_\varepsilon, x) - D_\varphi(y_\varepsilon, x) + \frac{\varepsilon}{4} < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon$$

Isto mostra que $\nu_\varphi(x,\cdot)$ é contínua à direita.

(iii) Fixe t no domínio de $\nu_\varphi(x,\cdot)$. Seja $\{t_k\}_{k\in\mathbb{N}}$ uma sequência que converge não decrescentemente a t. Como X é de dimensão finita, os conjuntos $\{y \in \mathcal{D} : \|y - x\| = t_k\}$ são compactos, pois são limitados e fechados. A função $D_\varphi(\cdot, x)$ é contínua em $\mathcal{D}$. Logo, para cada inteiro não-negativo k, existe um vetor $y^k \in \mathcal{D}$ tal que $\|y^k - x\| = t_k$ e $\nu_\varphi(x,t_k) = D_\varphi(y^k, x)$. A sequência $\{y^k\}_{k\in\mathbb{N}}$ é limitada. Consequentemente, ela possui uma subsquência convergente $\{y^{p_k}\}_{k\in\mathbb{N}}$. Seja $y^* = \lim_{k\to\infty} y^{p_k}$. Então, de acordo com a Proposição 2.2.1, o limite a seguir existe e temos

$$\nu_\varphi(x,t) \geqslant \lim_{k\to\infty} \nu_\varphi(x,t_k) = \lim_{k\to\infty} \nu_\varphi(x,y^{p_k}) = \lim_{k\to\infty} D_\varphi(y^{p_k}, x) = D_\varphi(y^*, x) \geqslant \nu_\varphi(x,t),$$

onde a última desigualdade vale por que $y^* \in \mathcal{D}$ e

$$\|y^* - x\| = \lim_{k\to\infty} \|y^{p_k} - x\| = \lim_{k\to\infty} t_{p_k} = t.$$

Portanto, $\nu_\varphi(x,t) = \lim_{k\to\infty} \nu_\varphi(x,t_k)$.

(iv) É uma consequência de (ii) e (iii). □

Vejamos agora a relação entre funções *totalmente convexas* e funções *estritamente convexas.*

Proposição 2.2.3. *Sejam* X *um espaço de Banach,* $\varphi : X \to \mathbb{R}\cup\{\infty\}$ *uma função e considere* $\mathcal{D}$ *o domínio efetivo de* φ*. Então valem as seguintes afirmações:*

(i) Se φ *é totalmente convexa, então* φ *é estritamente convexa.*

(ii) Se X *têm dimensão finita,* $\mathcal{D}$ *é fechado,* φ *é contínua e estritamente convexa em* $\mathcal{D}$*, então* φ *é totalmente convexa.*

(iii) Se X *têm dimensão finita e* $\mathcal{D} = X$*, então* φ *é totalmente convexa se, e somente se,* φ *é estritamente convexa.*

Demonstração. (i) Lembrando que φ é estritamente convexa em $\mathcal{D}^o$ se, e somente se, para todo par $x, y \in \mathcal{D}^o$, com $x \neq y$, se tem

$$\langle \nabla\varphi(x), x - y\rangle > \langle \nabla\varphi(y), x - y\rangle. \tag{2.9}$$

Suponha, por contradição, que φ é totalmente convexa mas não é estritamente convexa em $\mathcal{D}^o$. Então, existem $x^0, y^0 \in \mathcal{D}^o$, com $x^0 \neq y^0$, tais que

$$\langle \nabla\varphi(x^0), x^0 - y^0\rangle \leqslant \langle \nabla\varphi(y^0), x^0 - y^0\rangle. \tag{2.10}$$

Como φ é convexa, também temos

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

Esta última desigualdade e (2.10) implicam que

$$\langle \nabla\varphi(x^0), x^0 - y^0\rangle = \langle \nabla\varphi(y^0), x^0 - y^0\rangle = \varphi(x^0) - \varphi(y^0). \tag{2.11}$$

Consequentemente,

$$D_\varphi(x^0, y^0) = \varphi(x^0) - \varphi(y^0) - \langle \nabla\varphi(y^0), x^0 - y^0)\rangle.$$

Segue daí que $\nu_\varphi(y^0), \|y^0 - x^0\|) = 0$ o que não acontece a menos que se tenha $\|y^0 - x^0\| = 0$, pois φ é totalmente convexa em y^0. Assim, obtemos $x^0 = y^0$, o que é uma contradição.

(ii) Seja $t > 0$ tal que $\nu_\varphi(x, t)$ é finito. Então, o conjunto $\{y \in \mathcal{D} : \|y - x\| = t\}$ é compacto e $D_\varphi(\cdot, x)$ é contínua nesse conjunto. Portanto, existe um ponto $y^* \in \mathcal{D}$ tal que $\nu_\varphi(x, t) = D_\varphi(y^*, x)$. Observe que, para qualquer $\tau \in (0, 1)$,

$$D_\varphi(y^*, x) = \varphi(y^*) - \varphi(x) - \langle x, y^* - x\rangle \geqslant \varphi(y^*) - \varphi(x) - \frac{\varphi(x + \tau(y^* - x)) - \varphi(x)}{\tau} > 0,$$

onde a primeira desigualdade segue da convexidade de φ e , como consequência, temos que a função

$$\tau \to \frac{\varphi(x + \tau(y^* - x)) - \varphi(x)}{\tau}$$

é não-crescente em $(0, 1)$. A segunda desigualdade segue da convexidade estrita de φ. Logo, $\nu_\varphi(x, t) = D_\varphi(y^*, x) > 0$ e como $t > 0$ é arbitrário, por definição, φ é totalmente convexa.

(iii) É uma consquência de (i) e (ii). $\square$

Proposição 2.2.4. *Sejam $\varphi_1, \varphi_2, ..., \varphi_m \to (-\infty, \infty]$ funções totalmente convexas, com domínios $\mathcal{D}_1, \mathcal{D}_2, ..., \mathcal{D}_m$, respectivamente. Seja $\mathcal{D} = \bigcap_{i=1}^m \mathcal{D}_i$ e denote por $\mathcal{D}^o$ seu interior. Se $\mathcal{D}^o \neq 0$ então, para quaisquer reais não-negativos $c_1, c_2, ..., c_m$ tais que $\sum_{i=1}^m c_i > 0$, a função $h := \sum_{i=1}^m c_i\varphi_i$ é totalmente convexa e, para qualquer $x \in \mathcal{D}^o$ e todo $t \in [0, \infty)$,*

$$\nu_h(x, t) \geqslant \sum_{i=1}^m c_i \nu_{\varphi_i}(x, t) \tag{2.12}$$

Demonstração. Para quaisquer $y \in \mathcal{D}$ e $x \in \mathcal{D}^o$,

$$D_h(y,x) = h(y) - h(x) - \langle \nabla h(y), y - x \rangle =$$
$$\sum_{i=1}^m c_i \varphi_i(y) - \sum_{i=1}^m c_i \varphi_i(x) - \langle \nabla (\sum_{i=1}^m c_i \varphi_i)(y), y - x \rangle =$$
$$\sum_{i=1}^m c_i \left(\varphi_i(y) - \varphi_i(x) - \langle \nabla \varphi_i(y), y - x \rangle \right) = \sum_{i=1}^m c_i D_{\varphi_i}(y,x).$$

Agora, aplicando 2.5 à função h temos

$$\nu_h(x,t) := \inf\{D_h(y,x) : y \in \mathcal{D},\ \|y - x\| = t\} =$$
$$\inf\left\{ \sum_{i=1}^m c_i D_{\varphi_i}(y,x) : y \in \mathcal{D},\ \|y-x\| = t \right\} \geqslant$$
$$\sum_{i=1}^m \inf\{c_i D_{\varphi_i}(y,x) : y \in \mathcal{D},\ \|y-x\| = t\} =$$
$$\sum_{i=1}^m c_i \inf\{D_{\varphi_i}(y,x) : y \in \mathcal{D},\ \|y-x\| = t\} =$$
$$\sum_{i=1}^m c_i \nu_{\varphi_i}(x,t).$$

Como cada uma das funções φ_i é totalmente convexa, $\nu_{\varphi_i}(x,t) > 0$ para todo $x \in \mathcal{D}^o$ e todo $t > 0$ e, portanto, $\nu_h(x,t) > 0$ e isso mostra que h é totalmente convexa. $\square$

Assumiremos também sobre φ a seguinte hipótese:

B2 (*Convexidade total em subconjuntos limitados*): A função φ é *totalmente convexa em conjuntos limitados*, isto é, para todo subconjunto limitado $C \subset \mathcal{D}^o$ e todo $t \in \mathbb{R}_{++}$

$$\inf_{x \in C} \nu_\varphi(x,t) > 0$$

Esta condição é chamada de *consistência sequencial* em [21]. Note também que se uma função φ é totalmente convexa, então ela é totalmente convexa em subconjuntos limitados.

Observação 2.2.1. *A hipótese* B2 *será utlizada mais adiante, na prova do teorema principal, como requisito alternativo na análise de convergência do algoritmo.*

Quando $X = \ell_p$ ou $X = L^p$, com $p > 1$, a família de funções $\varphi(x) = \|x\|^r$, $r > 1$, é totalmente convexa (ver [22], Proposição 3, pág. 2412). Este resultado foi melhorado em [23], onde foi mostrado que se X é um espaço de Banach uniformemente convexo, então a função $\varphi(x) = \|x\|^r$ é totalmente convexa para todo $r > 1$. Antes, porém, necessitaremos da seguinte definição.

Definição 2.2.2. *O módulo de convexidade uniforme*[1] *de* X *é a função* $\delta_X : [0,2] \to [0,1]$, *onde*

$$\delta_X(t) = \begin{cases} \inf\{1 - \frac{1}{2}\|x+y\| : \|x\| = 1 = \|y\|,\ \|x-y\| \geqslant t\}, & se\ t > 0 \\ 0, & se\ t = 0. \end{cases} \tag{2.13}$$

Observação 2.2.2. *Na demonstração do Teorema 2.2.1 a seguir, utilizamos em (2.15) a definição de Distância de Bregman dada em [23], onde* X *é um espaço de Banach uniformemente convexo,* $\varphi : X \to \mathbb{R}$ *é uma função convexa e contínua e* $D_\varphi : X \times X \to \mathbb{R}$ *é dada por*

$$D_\varphi(x,y) = \varphi(x) - \varphi(y) - \inf\{\langle \xi, x-y\rangle :\ \xi \in \partial\varphi(y)\},$$

onde $\partial\varphi(y)$ *é o subdiferencial de* φ *em* y*. Note também que, caso* φ *seja Gateaux diferenciável, esta definição coincide com a definição 2.1.1.*

Teorema 2.2.1. *Seja* X *é um espaço de Banach uniformemente convexo. Então a função* $\varphi(x) = \|x\|^r$ *é totalmente convexa para todo* $r > 1$.

Demonstração. Defina a função $\Phi : X \to P(X^*)$,

$$\Phi(x) = \frac{1}{r}\partial\varphi(x).$$

De acordo com [?, Teorema 1, pág. 195] existe um número real positivo M tal que a seguinte desigualdade é válida para todo $r > 1$:

$$\langle \xi - \eta, x - y\rangle \geqslant M(\max\{\|x\|, \|y\|\})^r \delta_X\left(\frac{\|x-y\|}{2\max\{\|x\|,\|y\|\}}\right), \tag{2.14}$$

sempre que $x, y \in X$, com $\|x\| + \|y\| \neq 0$, $\xi \in \Phi(x)$ e $\eta \in \Phi(y)$. Fixe $z \in X$, $t \in (0,\infty)$, tome $x \in U(z,t) := \{x \in X :\ \|z - x\| = t\}$ e seja $y = x - z$. Então $\|y\| = t$ e

$$D_\varphi(x,z) = D_\varphi(y+z,z) = \|y+z\|^r - \|z\|^r - r\cdot\inf\{\langle\xi, y\rangle : \xi \in \Phi(z)\} \tag{2.15}$$

Defina $\phi : [0,\infty) \to [0,\infty)$ como

$$\phi(\tau) = \frac{\|z+\tau y\|^r}{r}.$$

Então,

$$\|y+z\|^r - \|z\|^r = r\cdot(\phi(1) - \phi(0)). \tag{2.16}$$

A prova do Teorema 2.2.1 se baseará na seguinte afirmação.

[1] A função δ_X é chamada em [23] de *módulo de convexidade uniforme do espaço* X e, como o próprio nome sugere, serve para caracterizar os espaços uniformemente convexos. Um espaço vetorial X é *uniformemente convexo* se $\delta_X(t) > 0$ para todo $t > 0$.

Afirmação 2.2.1. *Se, para cada* $\tau \in [0,1]$, *podemos escolher um ponto* $\xi(\tau) \in \Phi(z+\tau y)$, *então a seguinte integral existe e*

$$\int_0^1 \langle \xi(\tau), y\rangle d\tau = \phi(1) - \phi(0) \tag{2.17}$$

Vamos agora à prova desta afirmação. Para isto, observe que, se $g \in X :\to \mathbb{R}$ é uma função convexa e contínua, então, para $u, v \in X$ arbitrários, a função $\psi : \mathbb{R} \to \mathbb{R}$ definida por

$$\psi(\tau) = g(u + \tau v)$$

também é convexa e contínua. Portanto, ψ é localmente Lipschitz [26, Proposição 2.2.6, pág. 34] e, de acordo com o Teorema de Rademacher [7, Teorema 1.18, pág. 22], ela é em quase toda parte diferenciável. Consequentemente, se ξ é um seletor da aplicação ponto-conjunto $\partial\psi$, então $\xi(\tau) = \psi'(\tau)$, para quase todo $\tau \in \mathbb{R}$. Daí, para todos $a, b \in \mathbb{R}$, com $a \leqslant b$, temos

$$\psi(b) - \psi(a) = \int_a^b \psi'(\tau)d\tau = \int_a^b \langle \xi(\tau), v\rangle d\tau, \tag{2.18}$$

para qualquer escolha de $\xi(\tau) \in \partial\psi(\tau)$. Agora, aplicamos (2.18) no caso de $\psi = \phi$, $g(x) = \|x\|^r/r$, $u = z$, $v = y$ e daí temos que vale a equação (2.17). Isto completa a prova da Afirmação 2.2.1.

Usaremos agora a Afirmação 2.2.1 para terminar a demonstração do Teorema 2.2.1. Fixe um $\tau \to \xi(\tau)$ da aplicação ponto-conjunto $\tau \to \Phi(z + \tau y)$. Aplicando a Afirmação 2.2.1 juntamente com (2.16) temos

$$\|z + y\|^r - \|z\|^r = r\int_0^1 \langle \xi(\tau), y\rangle d\tau.$$

Portanto, para qualquer $\eta \in \Phi(z)$, temos

$$\|z + y\|^r - \|z\|^r - r\langle \eta, y\rangle = r\left[\int_0^1 \langle \xi(\tau), y\rangle d\tau - \langle \eta, y\rangle\right] = r\int_0^1 \frac{1}{\tau}\langle \xi(\tau) - \eta, \tau y\rangle d\tau \tag{2.19}$$

Uma vez que, para cada $\tau \in [0,1]$, $\xi(\tau) \in \Phi(z + \tau y)$, e $\tau y = (z + \tau y) - z$, concluímos a partir de (2.19) que

$$\|z + y\|^r - \|z\|^r - r\langle \eta, y\rangle \geqslant rK\int_0^1 \frac{(\max\{\|z + \tau y\|, \|z\|\})^r}{\tau}\delta_X\left(\frac{\|\tau y\|}{2\max\{\|z + \tau y\|, \|z\|\}}\right),$$

para todo $\eta \in \Phi(z)$, onde a última integral existe por que δ_X é não-decrescente. Tomando o ínfimo sobre $\eta \in \Phi(z)$ no lado esquerdo da última desigualdade e usando (2.15), temos

$$D_\varphi(x, z) \geqslant rK\int_0^1 \frac{(\max\{\|z + \tau y\|, \|z\|\})^r}{\tau}\delta_X\left(\frac{\|\tau y\|}{2\max\{\|z + \tau y\|, \|z\|\}}\right) d\tau. \tag{2.20}$$

Claramente, $\max\{\|z+\tau y\|, \|z\|\} \leqslant \|z\| + \tau\|y\|$. Usando novamente o fato de que δ_X é monótona, segue que

$$\delta_X\left(\frac{\|\tau y\|}{2\max\{\|z+\tau y\|, \|z\|\}}\right) \geqslant \delta_X\left(\frac{\|\tau y\|}{2(\|z\|+\tau\|y\|)}\right) = \delta_X\left(\frac{\tau t}{2(\|z\|+\tau t)}\right) \tag{2.21}$$

pois $\|y\| = t$. Agora, nós afirmamos que

$$\max\{\|z+\tau y\|, \|z\|\} \geqslant \frac{\tau}{2}\|y\|.$$

Isto é claro se $\|z\| \geqslant \frac{\tau}{2}\|y\|$. Caso contrário, temos $\|z\| < \frac{\tau}{2}\|y\|$ e, portanto,

$$\|z+\tau y\| \geqslant |\tau| \cdot \|y\| - \|z\| \geqslant \tau\|y\| - \frac{\tau}{2}\|y\| = \frac{\tau}{2}\|y\| = \frac{\tau t}{2}, \tag{2.22}$$

e o resultado é válido. Substituindo (2.21) e (2.22) em (2.20) temos que, para todo $x \in U(z,t)$,

$$D_\varphi(x,z) \geqslant rK\left(\frac{t}{2}\right)^r \int_0^1 \tau^{r-1}\delta_X\left(\frac{\tau t}{2(\|z\|+\tau t)}\right) d\tau.$$

Tomando o ínfimo sobre $x \in U(z,t)$ no lado esquerdo da última desigualdade e usando (2.5) obtemos

$$\nu_\varphi(z,t) \geqslant rK\left(\frac{t}{2}\right)^r \int_0^1 \tau^{r-1}\delta_X\left(\frac{\tau t}{2(\|z\|+\tau t)}\right) d\tau > 0, \tag{2.23}$$

onde a última desigualdade é válida por que X é uniformemente convexo. Isto completa a demonstração. □

Também vamos assumir a seguinte condição sobre φ, expressa em termos de sequências, que pede a continuidade uniforme de $\nabla\varphi$ em subconjuntos limitados de $\mathcal{D}^o$.

B3: Se $\{u^k\} \subset \mathcal{D}^o$ e $\{v^k\} \subset \mathcal{D}^o$ são sequências limitadas tais que $\lim_{k\to\infty} \|u^k - v^k\| = 0$, então

$$\lim_{k\to\infty} (\nabla\varphi(u^k) - \nabla\varphi(v^k)) = 0.$$

A condição B3 foi inicialmente introduzida em [24].

Lema 2.2.1. *Suponha que* D_φ *verifica* B2. *Sejam* $\{u^k\} \subset \mathcal{D}$ *e* $\{v^k\} \subset \mathcal{D}^o$ *limitada, tais que*

$$\lim_{k\to\infty} D_\varphi(u^k, v^k) = 0.$$

Então $\lim_{k\to\infty}(u^k - v^k) = 0$, *e assim* $\{u^k\}$ *é limitada e todos os pontos de acumulação fracos de* $\{u^k\}$ *e* $\{v^k\}$ *coincidem. Além disso, se* $\{u^k\} \subset \mathcal{D}^o$ *e* D_φ *verifica B3, então*

$$\lim_{k\to\infty} \|\nabla\varphi(u^k) - \nabla\varphi(v^k)\| = 0.$$

Demonstração. Suponha, por contradição, que existam duas sequências $\{u^k\}$ e $\{v^k\}$ contidas em $\mathcal{D}^o$ e tais que $\{v^k\}$ é limitada e $\lim_{k\to\infty} D_\varphi(u^k, v^k) = 0$, mas $\{\|u^k - v^k\|\}$ não converge a zero. Então, existe um $t > 0$ e subsquências $\{u^{i_k}\}_{k\in\mathbb{N}}$ e $\{v^{i_k}\}_{k\in\mathbb{N}}$ de $\{u^k\}_{k\in\mathbb{N}}$ e $\{v^{i_k}\}_{k\in\mathbb{N}}$, respectivamente, tais que $t \leqslant \|v^{i_k} - u^{i_k}\|$ para todo $k \in \mathbb{N}$. O conjunto $E := \{v^k | k \in \mathbb{N}\}$ é limitado. Consequentemente, para todo $k \in \mathbb{N}$, temos

$$D_\varphi(u^{i_k}, v^{i_k}) \geqslant \nu_\varphi(v^{i_k}, \|u^{i_k} - v^{i_k}\|) \geqslant \nu_\varphi(v^{i_k}, t) \geqslant \inf_{x\in C} \nu_\varphi(x, t),$$

e isto implica que $\inf\limits_{x\in C} \nu_\varphi(x, t) = 0$, o que contradiz a hipótese B2. A segunda parte é óbvia a partir da hipótese B3. $\square$

A próxima proposição estabelece as propriedades da continuidade de $\nabla\varphi$.

Proposição 2.2.5. *A aplicação* $\nabla\varphi : \mathcal{D}^o \to X^*$ demicontínua. *Em outras palavras, é contínua em qualquer* $x \in \mathcal{D}^o$ *da topologia forte de* X *para a topologia fraca de* X^*. *Em particular, se* X^* *é estritamente convexo e* $\varphi = \frac{1}{2}\|\cdot\|^2$, *então a aplicação de dualidade* $J = \partial\varphi : X \to X^*$ *é demicontínua.*

Demonstração. Ver [7, Proposição 2.8, pág. 19]. $\square$

Capítulo 3

Problema de Equilíbrio

3.1 Definição do Problema de Equilíbrio

Seja X um espaço de Banach reflexivo. O *problema de equilíbrio* de f em C consiste em encontrar $x^* \in C$, tal que

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

onde $C \subset X$ é um conjunto não-vazio, convexo e fechado e $f : C \times C \to \mathbb{R}$ é uma bifunção satisfazendo às seguintes condições:

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

(P2): $f(\cdot, y) : C \to \mathbb{R}$ é hemicontínua superior para todo $y \in C$, isto é,

$$\limsup_{t \to 0^+} f(tz + (1-t)x, y) \leqslant f(x, y), \ \forall \, x, z \in C,$$

(P3): $f(x, \cdot) : C \to \mathbb{R}$ é convexa e semicontínua inferior para todo $x \in C$

Denota-se o problema de equilíbrio de f em C por $\mathsf{PE}(f, C)$. O conjunto solução do problema $\mathsf{PE}(f, C)$ será denotado por $S(f, C)$.

Em 1994, Blum e Oettli [27] mostraram que o $\mathsf{PE}(f, C)$ tem como casos particulares várias classes de problemas, tais como: *problemas de otimização convexa, problemas de ponto fixo, problema de equilíbrio de Nash em jogos não-cooperativos, problemas de desigualdade variacional e problemas de complementaridade.* Em cada caso, os problemas podem ser modelados a partir de funções auxiliares que satisfazem não somente à condição (P1), como também (P2) e (P3), caracterizando assim um caso de problema de equilíbrio. Estudaremos esses casos mais adiante.

Associado ao problema de equilíbrio $\mathsf{PE}(f, C)$ existe ainda um outro problema, chamado *problema dual* de $\mathsf{PE}(f, C)$, que consiste em encontrar $y^* \in C$ tal que $f(x, y^*) \leqslant 0$ para todo $x \in C$. O conjunto solução do problema dual será denotado por $S^d(f, C)$.

Com o objetivo de analisar a existência e unicidade dos problemas de equilíbrio, vamos considerar as seguintes propriedades:

(P4): $f(x, y) + f(y, x) \leqslant 0$, $\forall\, x, y \in C$ (monotonicidade)

(P4°): Existe $\theta \geqslant 0$ tal que $f(x, y) + f(y, x) \leqslant \theta\langle \nabla\varphi(x) - \nabla\varphi(y), x - y\rangle$, para quaisquer $x, y \in C$, onde φ é como na definição 2.1.1.

(P5): Para toda sequência $\{u^n\} \subset C$ satisfazendo $\lim_{n\to\infty} \|u^n\| = \infty$, existem $u \in C$ e $n_0 \in \mathbb{N}$ tais que $f(u^n, u) \leqslant 0$ para todo $n \geqslant n_0$.

Observação 3.1.1. *Se X^* é estritamente convexo e $\varphi = (1/2)\|\cdot\|^2$, então o lado direito da expressão na condição $(\mathsf{P4}^\circ)$ se reduz à*

$$\langle J(x) - J(y), x - y\rangle.$$

3.2 Casos particulares

Nesta seção veremos que o Problema de Equilíbrio pode ser aplicado à resolução de diversos problemas: Otimização Convexa, Problema de Ponto Fixo, Problema de Equilíbrio de Nash em jogos não cooperativos, Problemas de Desigualdade Variacional e Problema de Complementaridade. Abordaremos estes problemas como casos particulares do Problema de Equilíbrio $\mathsf{PE}(f, C)$ e, logo depois, exibiremos um exemplo que não se encaixa em nenhum dos casos anteriores, refutando assim a generalidade do Problema de Equilíbrio.

Os casos particulares apresentados a seguir foram baseados em [27]. Em geral, nesses casos consideraremos que X é um espaço topológico real, X^* seu dual topológico e que $\langle\cdot,\cdot\rangle : X^* \times X \to \mathbb{R}$ é uma forma bilinear contínua.

3.2.1 Problema de Otimização Convexa

Sejam X um espaço topológico real e $h : X \to \mathbb{R} \cup \{\infty\}$ uma função convexa, semicontínua inferior e própria. O *problema de otimização convexa* consiste em:

$$\left\| \begin{array}{l} \text{Encontrar } \ x^* \in \ C \ \text{ tal que} \\ h(x^*) \leqslant h(y) \ \text{ para todo } \ y \in C, \end{array} \right. \tag{3.1}$$

onde $C = \{u \in X : \ f(u) < \infty\}$ é convexo e fechado. Este tipo de problema possui diversas aplicações e já existe uma teoria rica e bem desenvolvida sobre ele. Podemos citar, por exemplo, as referências [8] e [10].

O resultado a seguir mostra a relação entre o Problema de Equilíbrio e o Problema de Otimização Convexa.

Teorema 3.2.1. *Seja* $f : C \to \mathbb{R} \cup \{\infty\}$*, onde* $f(x, y) = h(y) - h(x)$ *e* $h : X \to \mathbb{R} \cup \{\infty\}$ *é uma função convexa, semicontínua inferior e própria. Então* f *satisfaz às condições* $(P1) - (P3)$ *e, além disso,* $x^* \in C$ *é uma solução para o problema de equilíbrio* $PE(f, C)$ *se, e somente se,* x^* *é uma solução de (3.1).*

Demonstração.

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

P2: $f(\cdot, y) : C \to \mathbb{R}$ é hemicontínua superior para todo $y \in C$, isto é,

$$\limsup_{t \to 0^+} f(tz+(1-t)x, y) = \limsup_{t \to 0^+} [h(y)-h(tz+(1-t)x)] = h(y)-h(x) \leqslant f(x, y), \ \forall\, x, z \in C,$$

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

Provaremos primeiro a convexidade de f. Sejam $y, z \in C$ e $t \in [0, 1]$. Uma vez que h é convexa, temos

$$\begin{aligned}
f(x, ty + (1-t)z) &= h(ty + (1-t)z) - h(x) \\
&\leqslant th(y) + (1-t)h(z) - th(x) - (1-t)h(x) \\
&= t\big[h(y) - h(x)\big] + (1-t)\big[h(z) - h(x)\big] \\
&= tf(x, y) + (1-t)f(x, z).
\end{aligned}$$

e assim temos $f(x, \cdot)$ é convexa para todo $x \in C$.

A semicontinuidade inferior de $f(x, \cdot)$ segue do fato de h ser semicontínua inferior. Dada $\{y^k\}_{k \in \mathbb{N}}$ uma sequência em C tal que $y^k \to y$ tem-se

$$\begin{aligned}
\liminf_{k \to \infty} f(x, y^k) &= \liminf_{k \to \infty} \big[h(y^k) - h(x)\big] \\
&\geqslant \liminf_{k \to \infty} h(y^k) - h(x) \\
&\geqslant h(y) - h(x) \\
&= f(x, y).
\end{aligned}$$

provando, portanto que para todo $x \in C$, $f(x, \cdot)$ é semicontínua inferior.

Para a prova da segunda parte, basta notar que se $x^* \in C$ é uma solução do problema

$\mathsf{PE}(f, C)$, então

$$f(x^*, y) \geqslant 0, \text{ para todo } y \in C \Leftrightarrow h(y) - h(x^*) \geqslant 0, \text{ para todo } y \in C, \tag{3.2}$$

o que ocorre se, e somente se, $h(y) \geqslant h(x^*)$, para todo $y \in C$, donde segue que $x^* \in C$ é solução de (3.1). A recíproca é óbvia a partir de (3.2). □

3.2.2 Problema de Ponto Fixo

Seja X um espaço de Hilbert com produto interno $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$. Dado um conjunto convexo e fechado $C \subset X$ e uma aplicação contínua $T : C \to C$, o problema de ponto fixo consiste em:

$$\left\| \begin{array}{l} \text{Encontrar } x^* \in C \text{ tal que} \\ \qquad T(x^*) = x^*. \end{array} \right. \tag{3.3}$$

A seguir mostremos como (3.3) pode ser formulado no formato de um problema de equilíbrio.

Teorema 3.2.2. *Seja* $f : C \times C \to \mathbb{R}$ *dada por* $f(x, y) = \langle x - T(x), y - x \rangle$.
Então

(i) f *satisfaz* $(\mathsf{P1}) - (\mathsf{P3})$*;*

(ii) T *tem um ponto fixo se, e somente se,* $S(f, C) \neq \emptyset$*;*

(iii) f *satisfaz* (P4) *se, e somente se,* $\langle Tx - Ty, x - y \rangle \leqslant \|x - y\|^2$, $\forall\, x, y \in C$.

Demonstração. (i) Pela definição de f temos $f(x, x) = \langle x - T(x), x - x \rangle = 0$. Logo, $(\mathsf{P1})$ se verifica. Para todo $y \in C$, da continuidade de T temos

$$\begin{aligned} \limsup_{t \to 0^+} f(tz + (1-t)x, y) &= \limsup_{t \to 0^+} \langle tz + (1-t)x - T(tz + (1-t)x), y - x \rangle \\ &= \langle x - T(x), y - x \rangle \\ &\leqslant f(x, y), \end{aligned}$$

para quaisquer $x, z \in C$. E assim temos que $(\mathsf{P2})$ está verificada. Agora, dada $\{y^k\}_{k \in \mathbb{N}}$ uma sequência em C convergindo para $y \in C$, da continuidade do produto interno temos o seguinte

$$\begin{aligned} \liminf_{k \to \infty} f(x, y^k) &= \liminf_{k \to \infty} \langle x - T(x), y^k - x \rangle \\ &= \langle x - T(x), y - x \rangle \\ &= f(x, y), \end{aligned}$$

e assim $f(x,\cdot) : C \to \mathbb{R}$ é semicontínua inferior. A convexidade de $f(x,\cdot)$ segue da linearidade do produto interno. Portanto, f satisfaz (P3).

(ii) Suponha que $x^* \in C$ é tal que $T(x^*) = x^*$. Então, para todo $y \in C$, temos

$$f(x^*, y) = \langle x^* - T(x^*), y - x^* \rangle = \langle 0, y - x^* \rangle = 0.$$

Logo $x^* \in S(f, C)$. Reciprocamente, assuma que $S(f, C) \neq \emptyset$. Considere $x^* \in S(f, C)$ e escolha $y = T(x^*)$, assim segue que

$$\begin{aligned} 0 \leqslant f(x^*, y) &= \langle x^* - T(x^*), y - x^* \rangle \\ &= -\|T(x^*) - x^*\|^2. \end{aligned}$$

Implicando $\|T(x^*) - x^*\| = 0$, ou seja, $T(x^*) = x^*$.

(iii) Veja que

$$\begin{aligned} f(x,y) + f(y,x) &= \langle x - T(x), y - x \rangle + \langle y - T(y), x - y \rangle \\ &= \langle T(x) - x + y - T(y), x - y \rangle \\ &= \langle T(x) - T(y), x - y \rangle + \langle y - x, x - y \rangle \\ &= \langle T(x) - T(y), x - y \rangle - \|x - y\|^2, \quad \forall\, x, y \in C. \end{aligned} \tag{3.4}$$

Portanto, o fato de f ser monótona (isto é, $f(x,y) + f(y,x) \leqslant 0$) é suficiente e necessário para que se tenha $\langle T(x) - T(y), x - y \rangle \leqslant \|x - y\|^2 \ \ \forall\, x, y \in C$. □

Observação 3.2.1. *Em particular, se T é não expansivo, isto é, $\|T(x) - T(y)\| \leqslant \|x - y\|$, para todo $x, y \in C$, então f é monótona. De fato, usando (3.4) e a desigualdade de Cauchy-Schwarz tem-se*

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

para todo $x, y \in C$. Logo, f é monótona.

3.2.3 Equilíbrio de Nash em jogos não cooperativos

Na teoria do jogos existem os jogos cooperativos e os jogos não cooperativos. Nos jogos cooperativos é permitido que os jogadores definam entre si uma escolha de estratégias, isto é, os jogadores podem fazer acordos de modo a influenciar os resultados do jogo. No caso dos jogos não cooperativos, as decisões dos jogadores são baseadas apenas em seu próprio interesse e, portanto, cada competidor escolhe sua estratégia sozinho.

Na década de 1950, o matemático John Forbes Nash publicou importantes artigos na teoria dos jogos não cooperativos. Em [35, 36], Nash provou a existência de um equilíbrio de estratégias mistas para jogos não cooperativos, denominado *Equilíbrio de Nash*, e sugeriu uma abordagem de estudo de jogos cooperativos a partir de sua redução para a forma não cooperativa.

A seguir apresentamos o problema de equilíbrio de Nash. Consideremos um conjunto com m-jogadores que denotamos por $I = \{1, 2, \cdots, m\}$. Para cada jogador $i \in I$, associamos um conjunto de estratégias, digamos $C_i \subset \mathbb{R}^{n_i}$, o qual assumimos ser não vazio, convexo e fechado. Denotemos $C = C_1 \times C_2 \times \cdots \times C_m$, o conjunto de todas as estratégias dos m-jogadores, o qual chamamos *espaço de estratégias do jogo.* Para cada competidor $i \in I$, considere uma função contínua

$$\begin{array}{rcrcl} f_i & : & C & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & f_i(x) \end{array}$$

a qual é convexa no i-ésimo argumento e que associa o ganho (payoffs ou custo), f_i, do i-ésimo jogador a cada estratégia $x \in C$..

O Problema de Equilíbrio de Nash em Jogos Não Cooperativos associado a $\{C_i\}_{i \in I}$ e $\{f_i\}_{i \in I}$ consiste em:

$$\left\| \begin{array}{l} \text{Encontrar } \ x^* = (x_i^*)_{i \in I} \in \ C \ \text{ tal que } \ \forall\, i \in I, \\ f_i(x^*) \leqslant f_i(x_1^*, \cdots, x_{i-1}^*, y_i, x_{i+1}^*, \cdots, x_m^*) \ \ \forall\, y_i \in C_i. \end{array} \right. \tag{3.5}$$

Em outras palavras, isto nos diz que nenhum jogador tem qualquer ganho por mudar apenas suas estratégias. O ponto x^* é chamado de equilíbrio de Nash.

Podemos estabelecer a seguinte a relação entre equilíbrio de Nash em jogos não cooperativos e o problema de equilíbrio.

Teorema 3.2.3. *Seja* $f : C \times C \to \mathbb{R}$*, dada por*

$$f(x, y) = \sum_{i=1}^{m} \big(f_i(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, \cdots, x_m) - f_i(x)\big),$$

então f *satisfaz* $(P1) - (P3)$*. Além disso,* $x^* \in C$ *é um equilíbrio de Nash se, e somente se,* $x^* \in S(f, C)$*.*

Demonstração. Da definição de f, temos

$$f(x,x)=\sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_m)-f_i(x)\big)=0.$$

Portanto, f satisfaz (P1). Da continuidade das f_i temos que

$$\begin{aligned}
\liminf_{k\to\infty} f(x,y^k) &= \liminf_{k\to\infty}\sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},y_i^k,x_{i+1},\cdots,x_m)-f_i(x)\big)\\
&= \sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},y_i,x_{i+1},\cdots,x_m)-f_i(x)\big)\\
&= f(x,y)\geqslant f(x,y),
\end{aligned}$$

para toda sequência $\{y^k\}\subset C$ convergindo para $y\in C$. Isto significa dizer que a função $f(x,\cdot):C\to\mathbb{R}$ é semicontínua inferior para todo $x\in C$. Note que, para todo $y\in C$, temos

$$\begin{gathered}
\limsup_{t\to 0^+} f(tz+(1-t)x,y)=\\
\limsup_{t\to 0^+}\textstyle\sum_{i=1}^{m}\big(f_i(tz_1+(1-t)x_1,\cdots,y_i,\cdots,tz_m+(1-t)x_m)-f_i(tz+(1-t)x)\big)=\\
\textstyle\sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},y_i,x_{i+1},\cdots,x_m)-f_i(x)\big)=\\
f(x,y)\leqslant f(x,y),
\end{gathered}$$

para quaisquer $x,z\in C$. Portanto, f satisfaz (P2). Agora, sejam $y,z\in C$ e $\lambda\in[0,1]$. Pela convexidade de C e das f_i em cada argumento obtemos

$$\begin{aligned}
f(x,(1-\lambda)y+\lambda z) &= \sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},(1-\lambda)y_i+\lambda z_i,x_{i+1},\cdots,x_m)-f_i(x)\big)\\
&\leqslant (1-\lambda)\sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},y_i,x_{i+1},\cdots,x_m)-f_i(x)\big)+\\
&\quad\ \lambda\sum_{i=1}^{m}\big(f_i(x_1,\cdots,x_{i-1},z_i,x_{i+1},\cdots,x_m)-f_i(x)\big)\\
&= (1-\lambda)f(x,y)+\lambda f(x,z).
\end{aligned}$$

o que prova que $f(x,\cdot)$ é convexa. Consequentemente, a condição (P3) é válida. Suponha agora que $x^*\in C$ é um equilíbrio de Nash, isto é, que x^* satisfaz (3.5) para todo $i\in I$. Então

$$f_i(x_1^*,\cdots,x_{i-1}^*,y_i,x_{i+1}^*,\cdots,x_m^*)\geqslant f_i(x^*),\qquad \forall\, y_i\in C_i,$$

isto é,

$$f(x^*,y)=\sum_{i=1}^{m}\big(f_i(x_1^*,\cdots,x_{i-1}^*,y_i,x_{i+1}^*,\cdots,x_m^*)-f_i(x^*)\big)\geqslant 0,\qquad \forall\, y_i\in C_i.$$

Portanto, $x^* \in S(f, C)$. Reciprocamente, Se $x^* \in C$ é tal que $f(x^*, y) \geqslant 0$ para todo $y \in C$, então para qualquer $i \in I$, tome $y \in C$ tal que $y_j = x_j^*$ para qualquer $j \neq i$, obtemos

$$0 \leqslant f(x^*, y) \quad = \quad f_i(x_1^*, \cdots, x_{i-1}^*, y_i, x_{i+1}^*, \cdots, x_m^*) - f_i(x^*).$$

Logo, x^* é um equilíbrio de Nash. □

3.2.4 Desigualdade Variacional

Sejam X um espaço topológico real, $C \subset X$ um subconjunto não vazio, convexo e fechado e $T : C \to X^*$ uma aplicação contínua. O *Problema de Desigualdade Variacional*, denotado por $\mathsf{VIP}(T, C)$, consiste em

$$\mathsf{VIP}(T, C) : \begin{cases} \text{Encontrar } x^* \in C \text{ tal que} \\ \langle T(x^*), y - x^* \rangle \geqslant 0, \quad \forall\, y \in C. \end{cases}$$

As desigualdades variacionais têm oferecido uma grande possibilidade de aplicações a vários problemas fundamentais da matemática pura e econômica, da física matemática, engenharia, dentre outras. O leitor interessado pode consultar [37]. Abaixo são apresentados dois exemplos simples onde aparecem desigualdades variacionais.

Exemplo 3.2.1. *Seja* $f : [a, b] \to \mathbb{R}$ *uma função derivável. Buscamos o ponto* $x_0 \in [a, b]$ *para o qual*

$$f(x_0) = \min_{x \in [a,b]} f(x)$$

Três situações pode ocorrer:

(i) se $a < x_0 < b$, então $f'(x_0) = 0$;

(ii) se $x_0 = a$, então $f'(x_0) \geqslant 0$;

(iii) se $x_0 = b$, então $f'(x_0) \leqslant 0$.

Consequentemente, o ponto x_0 satisfaz a desigualdade variacional

$$\langle \nabla f(x_0), x - x_0 \rangle \geqslant 0 \quad \forall\, x \in [a, b].$$

Exemplo 3.2.2. *Seja* $C \subseteq \mathbb{R}^n$ *convexo, fechado e* $h : C \to \mathbb{R}$ *uma função diferenciável. Desejamos caracterizar o ponto* $x_0 \in C$ *tal que*

$$h(x_0) = \min_{x \in C} h(x).$$

Suponha que x_0 *é um ponto onde o mínimo é atingido e seja* $x \in C$. *Uma vez que* C *é convexo, o segmento* $(1-t)x_0 + tx = x_0 + t(x - x_0)$ *pertence a* C *para todo* $t \in [0, 1]$.
A função

$$\psi(t) = h(x_0 + t(x - x_0)), \;\; 0 \leqslant t \leqslant 1,$$

é tal que $\psi(0) = h(x^0)$ *e, portanto, atinge seu mínimo em* $t = 0$. *A partir do exemplo anterior, temos que*

$$\psi^{'}(0) = \langle \nabla f(x_0), x - x_0 \rangle \geqslant 0, \;\; \textit{para todo} \;\; x \in C.$$

Consequentemente, o ponto x_0 *satisfaz a desigualdade variacional*

$$\langle \nabla f(x_0), x - x_0 \rangle \geqslant 0 \quad \forall x \in C.$$

A seguir mostraremos a relação entre $VIP(T, C)$ e o problema de equilíbrio.

Teorema 3.2.4. *Seja* $f : C \times C \to \mathbb{R}$ *dada por* $f(x, y) = \big\langle T(x), y - x \big\rangle$. *Então* f *satisfaz* $(P1) - (P3)$. *Além disso,* x^* *é solução de* $VIP(T, C)$ *se, e somente se,* $x^* \in S(f, C)$.

Demonstração. Para todo $x \in C$, $f(x, x) = \big\langle T(x), x - x \big\rangle = \langle T(x), 0 \rangle = 0$. E assim, (P1) é verificada. Da continuidade de T e $\langle \cdot, \cdot \rangle$ temos que, para todo $y \in C$,

$$\begin{aligned} \limsup_{t \to 0^+} f(tz + (1-t)x, y) &= \limsup_{t \to 0^+} \langle T(tz + (1-t)x), y - (tz + (1-t)x) \rangle \\ &= \langle T(x), y - x \rangle \\ &= f(x, y) \leqslant f(x, y), \end{aligned}$$

para todos $x, z \in C$, o que mostra que a condição $(P2)$ é válida. Seja $f(x, \cdot) : C \to \mathbb{R}$ dada por $f(x, y) = \big\langle T(x), y - x \big\rangle$. Uma vez que C é um cone convexo, temos que $(1-t)y + tz \in C$, para todos $y, z \in C$ e $t \in [0, 1]$. Então,

$$\begin{aligned} f(x, (1-t)y + tz) &= \big\langle T(x), (1-t)y + tz - x \big\rangle \\ &= \big\langle T(x), (1-t)y + tz - (1-t)x - tx \big\rangle \\ &= \big\langle T(x), (1-t)(y - x) + t(z - x) \big\rangle \\ &= (1-t)\big\langle T(x), y - x \big\rangle + t\big\langle T(x), z - x \big\rangle \\ &= (1-t)f(x, y) + tf(x, z). \end{aligned}$$

Logo, f é convexa. Provemos que $f(x, \cdot)$ é semicontínua inferior para todo $x \in C$.
De fato, considere $\{y^k\}_{k \in \mathbb{N}} \subset C$ uma sequência tal que $\lim_{k \to \infty} y^k = y$. Uma vez que C é fechado, temos que $y \in C$. Pela continuidade de T, temos

$$\liminf_{k \to \infty} f(x, y^k) = \liminf_{k \to \infty} \big\langle T(x), y^k - x \big\rangle = \big\langle T(x), y - x \big\rangle = f(x, y).$$

Logo, $f(x,\cdot)$ é semicontínua inferior para todo $x \in C$. Consequentemente, (P3) é válida. Suponha agora que $x^* \in S(f,C)$. Pela definição de f, temos $\left\langle T(x^*), y - x^* \right\rangle \geqslant 0$ para todo $y \in C$. Logo, x^* é solução de $\mathsf{VIP}(T,C)$. Reciprocamente, suponha que x^* resolve $\mathsf{VIP}(T,C)$. Então, para todo $y \in C$ se tem $\left\langle T(x^*), y - x^* \right\rangle \geqslant 0$, isto é, $f(x^*, y) \geqslant 0$, para todo $y \in C$. Logo, $x^* \in S(f,C)$.

□

3.2.5 Problema de Complementaridade

Sejam X um espaço topológico real, $C \subset X$ um cone convexo e fechado e $T : C \to X^*$ uma aplicação contínua. O problema de complementaridade consiste em

$$\mathsf{PC}(T,C) : \begin{cases} \text{Encontrar } x^* \in C \text{ tal que} \\ T(x^*) \in C^* \text{ e } \left\langle T(x^*), x^* \right\rangle = 0, \end{cases}$$

onde $C^* = \{\xi \in X^* \ ; \ \langle \xi, d \rangle \geqslant 0, \ \forall \ d \in C\}$ é o *cone polar* de C.

Segundo George Isac [33], a origem do problema de complementaridade está talvez no Teorema Karush-Kuhn-Tucker para programação não-linear. A teoria de complementaridade é importante pois ela pode ser usada para modelar problemas em várias áreas do conhecimento, tais como: teoria de equilíbrio em uma economia competitiva, engenharia, teoria de elasticidade, otimização em $\mathbb{R}^n_+$, teoria dos jogos, maximização de produção de óleo, cálculo de pontos fixos e etc. Algumas dessas aplicações são encontradas nos trabalhos de George Isac [33].

A seguir apresentamos um exemplo que pode ser formulado no formato de um problema do tipo $\mathsf{PC}(T,C)$.

Exemplo 3.2.3. *Consideremos um sistema com n diferentes mercadorias (bens de consumo) e m comerciantes comprando e vendendo estas mercadorias. Sejam $u^i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1, 2, \cdots, m)$ funções de utilidade de cada comerciante. Assim, certamente cada comerciante desejará resolver o problema:*

$$\max_{s.a\ \langle p,x\rangle \leqslant \langle p,w^i\rangle} u^i(x), \tag{3.6}$$

onde:

$p = (p_1, p_2, \cdots, p_n)$ – *vetor preço, i.e, p_i é preço do i-ésima bem ($p_i \geqslant 0$);*

$w^i = (w^i_1, w^i_2, \cdots, w^i_n)$ – *vetor de mercadorias disponível ao i-ésimo comerciante;*

$x = (x_1, x_2, \cdots, x_n)$ – *vetor de bens que cada comerciante deseja obter.*

Portanto, cada comerciante deseja maximizar sua utilidade, u_i*, sujeito a uma restrição orçamentária,* $\langle p, w^i \rangle$*. Uma solução de (3.6) é denotada por* $x^i(p)$*, consideremos a função* $f : \mathbb{R}^n \to \mathbb{R}^n$ *definida por*

$$f(p) = \sum_{i=1}^{m} \big(x^i(p) - w^i\big),$$

a qual será chamada função excesso de demanda.

Neste caso, a Lei de Walras afirma que, para o preço equilíbrio p*, devemos ter* $f(p) \in \mathbb{R}^n_+$ *e* $\big\langle p, f(p) \big\rangle = 0$ *(Ver Mas-Colell [34], páginas 580-581). Isto é, temos que encontrar o equilíbrio econômico. Neste caso é necessário resolver o seguinte problema de complementaridade:*

$$PC(f, C) : \left\| \begin{array}{l} \textit{Encontrar } p \in \mathbb{R}^n_+ \textit{ tal que} \\ f(p) \in \mathbb{R}^n_+ \ \textit{ e } \ \big\langle f(p), p \big\rangle = 0, \end{array} \right.$$

onde $X = X^* = \mathbb{R}^n$ *e* $C = \mathbb{R}^n_+$*, implicando que* $C^* = \mathbb{R}^n_+$ *e* $T = f$*.*

Agora, sabendo da importância do problema de complementaridade em várias áreas do conhecimento, vamos apresentar a relação entre esse problema e o problema de equilíbrio.

Teorema 3.2.5. *Sejam* T *e* C *como em* $PC(T, C)$*. Se* $f(x, y) = \big\langle T(x), y - x \big\rangle$ *para todo* $x, y \in C$*, então* f *satisfaz* $(P1) - (P3)$*. Além disso,* x^* *é uma solução de* $PC(T, C)$ *se, e somente se,* x^* *é uma solução de* $PE(f, C)$*.*

Demonstração. Pelo Teorema 3.2.4, a função $f(x, y) = \big\langle T(x), y - x \big\rangle$ satisfaz às condições $(P1) - (P3)$. Suponha agora que x^* seja solução do problema $PC(T, C)$. Isso implica que $T(x^*) \in C^*$ (i.e., $\big\langle T(x^*), y \big\rangle \geqslant 0, \ \ \forall\, y \in C$) e $\big\langle T(x^*), x^* \big\rangle = 0$. Assim,

$$f(x^*, y) = \big\langle T(x^*), y - x^* \big\rangle = \langle T(x^*), y \rangle - \langle T(x^*), x^* \rangle = \big\langle T(x^*), y \big\rangle \geqslant 0.$$

Reciprocamente, suponha que $x^* \in S(f, C)$. Então $f(x^*, y) = \big\langle T(x^*), y - x^* \big\rangle \geqslant 0$, para todo $y \in C$. Assim, tomando $y = 2x^*$ temos

$$0 \leqslant \big\langle T(x^*), y - x^* \big\rangle = \big\langle T(x^*), x^* \big\rangle \tag{3.7}$$

e por outro lado, tomando $y = 0$, segue

$$0 \leqslant \big\langle T(x^*), y - x^* \big\rangle = \big\langle T(x^*), -x^* \big\rangle = -\big\langle T(x^*), x^* \big\rangle. \tag{3.8}$$

De (3.7) e (3.8) segue que $\big\langle T(x^*), x^* \big\rangle = 0$ e, portanto, $\big\langle T(x^*), y \big\rangle \geqslant 0$ para todo $y \in C$, isto é, $T(x^*) \in C^*$. Logo, x^* é solução de $PC(T, C)$. $\square$

Observação 3.2.2. *Observe que a prova do Teorema 3.2.5 seque quase as mesmas linhas do Teorema 3.2.4. Isto acontece por que pode-se considerar o problema de complementaridade como um caso particular de desigualdade variacional.*

3.2.6 A generalidade do Problema de Equilíbrio

Nesta seção, com o objetivo de mostrar a generalidade do Problema de Equilíbrio, apresentaremos agora o seguinte exemplo (na forma de proposição), o qual satisfaz às hipóteses básicas do $\mathsf{PE}(f, C)$ porém não se encaixa em nenhum dos casos particulares citados nas seções anteriores. Este exemplo foi baseado no trabalho de Iusem e Sosa [29].

Proposição 3.2.1. *Considere* $X = C = \mathbb{R}^2$ *e* $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \cup \{\infty\}$ *definida por*

$$f(x, y) = \|y\|^2 - \|x\|^2 + x_1 y_2 - x_2 y_1, \tag{3.9}$$

onde $x = (x_1, x_2)$, $y = (y_1, y_2)$ *e* $\|x\| \neq 0$. *Então*

(i) f *satisfaz* $(\mathsf{P}1) - (\mathsf{P}4)$.

(ii) Não existe $h : \mathbb{R}^2 \to \mathbb{R}$ *tal que* $f(x, y) = h(y) - h(x)$.

(iii) Não existe $T : \mathbb{R}^2 \to \mathbb{R}^2$ *tal que* $f(x, y) = \langle x - Tx, y - x\rangle$.

(iv) Não existe $T : \mathbb{R}^2 \to \mathbb{R}^2$ *tal que* $f(x, y) = \langle T(x), y - x\rangle$.

Demonstração. (i) De fato, considere $x, y, z \in \mathbb{R}^2$ e $t \in (0, 1)$

(P1) $f(x, x) = \|x\|^2 - \|x\|^2 + x_1 x_2 - x_2 x_1 = 0$

(P2) Para quaisquer $x, z \in \mathbb{R}^2$, temos

$$\limsup_{t \to 0^+} f(tz + (1-t)x, y)$$
$$= \limsup_{t \to 0^+} \left[\|y\|^2 - \|tz + (1-t)x\|^2 + (tz_1 + (1-t)x_1)y_2 - (tz_2 + (1-t)x_2)y_1\right]$$
$$= \|y\|^2 - \|x\|^2 + x_1 y_2 - x_2 y_1 = f(x, y) \leqslant f(x, y),$$

para todo $z \in \mathbb{R}^2$.

(P3) Mostraremos primeiro que $f(x, \cdot) : \mathbb{R}^2 \to \mathbb{R}$ é convexa para todo $x \in \mathbb{R}^2$. Temos

$$\begin{aligned}
f(x,(1-t)y+tz) &= \|(1-t)y+tz\|^2 - \|x\|^2 + x_1\big[(1-t)y_2+tz_2\big] \\
&- x_2\big[(1-t)y_1+tz_1\big] \\
&= \|(1-t)y+tz\|^2 - \|x\|^2 + (1-t)x_1y_2 + tx_1z_2 \\
&- (1-t)x_2y_1 - tx_2z_1 \\
&\leqslant (1-t)\big[\|y\|^2 - \|x\|^2 + x_1y_2 - x_2y_1\big] \\
&+ t\big[\|z\|^2 - \|x\|^2 + x_1z_2 - x_2z_1\big] \\
&= (1-t)f(x,y) + tf(x,z).
\end{aligned}$$

Agora seja $\{y^k\} = \{(y_1^k, y_2^k)\}$ uma sequência em $\mathbb{R}^2$ tal que $y^k \to y \in \mathbb{R}^2$. Assim

$$\begin{aligned}
\liminf_{k\to\infty} f(x,y^k) &= \liminf_{k\to\infty} \big[\|y^k\|^2 - \|x\|^2 + x_1y_2^k - x_2y_1^k\big] \\
&= \liminf_{k\to\infty} \|y^k\|^2 - \|x\|^2 + x_1 \liminf_{k\to\infty} y_2^k - x_2 \liminf_{k\to\infty} y_1^k \\
&= \|y\|^2 - \|x\|^2 + x_1y_2 - x_2y_1 \\
&= f(x,y).
\end{aligned}$$

Portanto, $f(x, \cdot)$ é semicontínua inferior para todo $x \in \mathbb{R}^2$.

(P4). Temos que f é monótona, i.e, $f(x,y) + f(y,x) \leqslant 0 \quad \forall\, x, y \in C$. De fato,

$$f(x,y) + f(y,x) = \big[\|y\|^2 - \|x\|^2 + x_1y_2 - x_2y_1\big] + \big[\|x\|^2 - \|y\|^2 + y_1x_2 - y_2x_1\big] = 0.$$

(ii) De fato, uma vez que $\|x\| \neq 0$, a função f dada em (3.9) não pode ser separada na forma $f(x,y) = h(y) - h(x)$.

(iii) Suponha que exista $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $f(x,y) = \langle x - Tx, y - x\rangle \quad \forall\ x, y \in \mathbb{R}^2$. Escolhendo-se um $x \in \mathbb{R}^2$, com $\|x\| \neq 0$ e tomando $y = 2x$ e $y = 3x$, respectivamente, se tem

$$\langle x - Tx, x\rangle = 4\|x\|^2 - \|x\|^2 + 2x_1x_2 - 2x_2x_1 = 3\|x\|^2 \text{ e} \tag{3.10}$$

$$\langle x - Tx, 2x\rangle = 9\|x\|^2 - \|x\|^2 + 3x_1x_2 - 3x_2x_1 = 8\|x\|^2, \tag{3.11}$$

donde segue

$$\langle x - Tx, x\rangle = 5\|x\|^2. \tag{3.12}$$

De (3.10) e (3.12) obtemos $\|x\|^2 = 0$, e daí que $\|x\| = 0$, contrariando a hipótese que assumimos originalmente de que $\|x\| \neq 0$. Portanto, não existe $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

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

(iv) Com efeito, suponha que exista $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $f(x, y) = \langle T(x), y - x \rangle$. Assumindo que $\|x\| \neq 0$, tome $y = 2x$ e $y = 3x$, respectivamete. Procedendo de maneira análoga à prova do item (iii), chega-se à mesma contradição. Consequentemente, tal aplicação T não pode existir. $\square$

Capítulo 4

Uma regularização de Bregman para a solução de Problemas de Equilíbrio em espaços de Banach

4.1 Sobre o Método de Ponto Proximal

O *Método de Ponto Proximal*, cuja formulação básica se deu com o trabalho de Rockafellar em [30], é um procedimento para encontrar zeros de um operador monótono maximal $\mathsf{T} : \mathsf{H} \to \mathsf{P}(\mathsf{H})$, onde H é um espaço de Hilbert. O algoritmo gera uma sequência $\{z^k\} \subset \mathsf{H}$, iniciando com um $z^0 \in \mathsf{H}$, e define z^{k+1} como o único zero do operador T^k dado por

$$\mathsf{T}^k(x) = \mathsf{T}(x) + \gamma_k(x - z^k), \tag{4.1}$$

ou ainda,

$$z^{k+1} = (I + \gamma_k \mathsf{T})^{-1}(z^k), \tag{4.2}$$

onde $\{\gamma_k\}$ é uma sequência limitada de números reais chamada de *coeficientes de regularização* e I é o operador identidade. Foi provado em [30] que para um operador monótono maximal T, a sequência $\{z^k\}$ é fracamente convergente a um zero de T, quando T possui zeros, e é ilimitada caso contrário. Tal convergência é global, isto é, o resultado mencionado é válido para qualquer ponto inicial $z^0 \in \mathsf{H}$.

Utilizando a notação $\bar{z}$ para a iterada atual z^k e $\bar{\mathsf{T}}$ para o operador T^k, então (4.2) se torna $\bar{\mathsf{T}}(x) = \mathsf{T}(x) + \gamma(x - \bar{z})$, com $\gamma > 0$. A partir daí pode-se definir um procedimeno de regularização para problemas de equilíbrio. Se f for uma função satisfazendo $(P1) - (P3)$, associamos a ela uma outra bifunção $\hat{f} : C \times C \to \mathbb{R}$, que será chamada de *regularização* de f e que é definida como

$$\hat{f}(x, y) = f(x, y) + \gamma\langle x - \bar{z}, y - x\rangle, \tag{4.3}$$

onde $\langle\cdot,\cdot\rangle$ é o produto interno de H e $\gamma \in \mathbb{R}_{++}$.

No caso de um espaço de Banach X, uma extensão apropriada de (4.2) pode ser obtida substituindo o operador identidade pelo gradiente $\nabla\varphi$ de uma função $\varphi \to \mathbb{R}$, estritamente convexa e diferenciável (no sentido de Gateaux), isto é,

$$z^{k+1} = [(\gamma_k T + \nabla\varphi)^{-1} \circ \nabla\varphi](z^k) \tag{4.4}$$

Em [24] foi provado que se T é monótono máximal e possui zeros, então $\{z^k\}$ é limitada e seus pontos de acumulação fracos são zeros de T, desde que φ satisfaça algumas hipóteses técnicas (B1, B2 e B3 são algumas delas).

Na seção a seguir apresentaremos a regularização proposta por Burachik e Kassay em [1], visando obter um algoritmo de ponto proximal para resolver problemas de equilíbrio em espaços de Banach reflexivos. Utilizaremos como função de regularização uma função de Bregman que, ao satisfazer certas hipóteses técnicas, garantirá a convergência do algoritmo proposto, como veremos mais adiante na Seção 4.3.

4.2 Uma regularização de Bregman para problemas de equilíbrio

Sejam X um espaço de Banach reflexivo, $C \subset X$ convexo e fechado e $f : C \times C \to \mathbb{R}$ uma bifunção tal que valem as condições P1, P2 e P3. Seja $\varphi : X \to \mathbb{R} \cup \{\infty\}$ uma função de Bregman tal como na Definição 2.1.1 e suponha $C \subset \mathcal{D}^o$.

Fixe $\gamma > 0$ e $\bar{x} \in C \subset \mathcal{D}^o$. A regularização de f será denotada por $\tilde{f} : C \times C \to \mathbb{R}$ e definida por

$$\tilde{f}(x, y) = f(x, y) + \gamma\langle\nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle. \tag{4.5}$$

Considere uma bifunção $g : C \times C \to \mathbb{R}$. De acordo com ([25], pág. 130), o *operador resolvente induzido pela função de Bregman* φ é a aplicação ponto-conjunto $\mathrm{Res}^{\varphi}_{g} : X \rightrightarrows C$

dada por

$$\text{Res}_g^{\varphi}(\bar{x}) := \{z \in C : g(z, y) + \langle \nabla\varphi(z) - \nabla\varphi(x), y - z\rangle \geqslant 0, \ \forall\, y \in C\}.$$

Esta definição e (4.5) implicam que $\hat{x}$ é solução de $\mathsf{PE}(\tilde{f}, C)$ se, e somente se, $\hat{x} \in \text{Res}^{\varphi}_{\frac{1}{\gamma}f}(\bar{x})$. De fato, para todo $y \in C$,

$$\hat{x} \in \text{Res}^{\varphi}_{\frac{1}{\gamma}f}(\bar{x}) \Leftrightarrow \frac{1}{\gamma} f(\hat{x}, y) + \langle \nabla\varphi(\hat{x}) - \nabla\varphi(\bar{x}), y - \hat{x}\rangle \geqslant 0,$$

isto é,

$$\frac{1}{\gamma}[\tilde{f}(\hat{x}, y) - \gamma\langle \nabla\varphi(\hat{x}) - \nabla\varphi(\bar{x}), y - \hat{x}\rangle] + \langle \nabla\varphi(\hat{x}) - \nabla\varphi(\bar{x}), y - \hat{x}\rangle = \frac{1}{\gamma}\tilde{f}(\hat{x}, y) \geqslant 0,$$

e como $\gamma > 0$, segue-se que $\tilde{f}(\hat{x}, y) \geqslant 0$.

Proposição 4.2.1. *Tome* $\bar{x} \in C$ *e suponha que* $C \subset \mathcal{D}^o$. *Suponha que* f *satisfaz as condições* $\mathsf{P1} - \mathsf{P3}$ *e* $\mathsf{P4}^o$ *com* $\theta < \gamma$. *Então,* $\tilde{f}$ *satisfaz as condições* $\mathsf{P1} - \mathsf{P4}$. *Além disso, se para cada sequência* $\{u^n\} \subset C$ *tal que* $\lim_{n\to\infty} \|u^n\| = \infty$, *tivermos*

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

então $\tilde{f}$ *satisfaz a condição* $\mathsf{P5}$.

Demonstração. É claro que $\tilde{f}$ satisfaz à condição (P1), pois

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

A função $y \mapsto s(y) = \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle$ é convexa e contínua. De fato, dados $u, v \in C$ e $t \in [0, 1]$, então

$$\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), (1-t)u + tv - x\rangle =$$

$$\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), (1-t)u + tv - (1 - t + t)x\rangle =$$

$$(1-t)\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), u - x\rangle + t\langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), v - x\rangle,$$

isto é,

$$s((1-t)u + tv) \leqslant (1-t)s(u) + ts(v).$$

A continuidade de s é imediata da continuidade de $\nabla\varphi$ (vide Proposição 2.2.5). Desses dois fatos segue que $\tilde{f}$ satisfaz à condição (P3). Agora, para mostrar que $\tilde{f}$ satisfaz à condição (P2), provaremos primeiro que a aplicação

$$x \mapsto \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x\rangle$$

é contínua para todo $x \in C$. Seja $\{u^n\} \subset C$ uma sequência fortemente convergente a $x \in C$. Então $\langle \nabla\varphi(\bar{x}), u^n \rangle \to \langle \nabla\varphi(\bar{x}), x \rangle$, enquanto que $\langle \nabla\varphi(u^n), y - u^n \rangle \to \langle \nabla\varphi(x), y - x \rangle$ pela Proposição 2.2.5. Assim,

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

como queríamos provar. Agora, usando o fato de que $x \mapsto \langle \nabla\varphi(x) - \nabla\varphi(\bar{x}), y - x \rangle$ é contínua e que f satisfaz $(P2)$, segue imediatamente que $\tilde{f}$ também satisfaz $(P2)$.

Afirmamos agora que $\tilde{f}$ satisfaz à condição $(P4)$. De fato, pela condição $(P4^\circ)$ temos que

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

Agora provaremos que $\tilde{f}$ satisfaz à condição $(P5)$. De (4.5), temos que

$$\begin{aligned} \tilde{f}(u^n, \bar{x}) &= f(u^n, \bar{x}) + \gamma\langle \nabla\varphi(u^n) - \nabla\varphi(\bar{x}), \bar{x} - u^n \rangle \\ &= f(u^n, \bar{x}) - \gamma\langle \nabla\varphi(\bar{x}) - \nabla\varphi(u^n), \bar{x} - u^n \rangle \\ &\leqslant -f(\bar{x}, u^n) - (\gamma - \theta)\langle \nabla\varphi(\bar{x}) - \nabla\varphi(u^n), \bar{x} - u^n \rangle \\ &= -(f(\bar{x}, u^n) + (\gamma - \theta)\langle \nabla\varphi(\bar{x}) - \nabla\varphi(u^n), \bar{x} - u^n \rangle), \end{aligned} \tag{4.6}$$

onde a desigualdade segue da condição $(P4^\circ)$ de f, pois

$$f(u^n, \bar{x}) + f(\bar{x}, u^n) \leqslant \theta\langle \nabla\varphi(\bar{x}) - \nabla\varphi(u^n), \bar{x} - u^n \rangle.$$

Por hipótese, existe $n_0 \in \mathbb{N}$ tal que a expressão entre parênteses em (4.6) é não-negativa para todo $n \geqslant n_0$. Daí segue imediatamente que $\tilde{f}$ satisfaz $(P5)$. □

Corolário 4.2.1. *Se* f *satisfaz* $(P4^\circ)$ *com* $\theta < \gamma$, *e assumindo ainda que*

(i) C *é limitado, ou*

(ii) X^* *é estritamente convexo e* $\varphi = (1/2)\|\cdot\|^2$,

então $\tilde{f}$ *satisfaz à condição* $(P5)$.

Demonstração. Usando a Proposição 4.2.1 é suficiente provar que a condição $(P6)$ vale sob (i) e (ii). No caso (i), isto é, quando C é limitado, a propriedade $(P6)$ se verifica trivialmente por vacuidade, pois não existe sequência $\{u^n\} \subset C$ tal que $\lim_{n\to\infty} \|u^n\| = \infty$. Agora devemos provar que (ii) implica a condição $(P6)$. Note que neste caso $\nabla\varphi = J$.

Pelo Teorema 1.1.5 sabemos que J está bem definida e é contínua. Tome uma sequência $\{u^n\}$ em C tal que $\|u^n\| \to \infty$ e defina $\Delta_n = \langle J(\bar{x}) - J(u^n), \bar{x} - u^n\rangle$. Nossas hipóteses implicam que

$$\Delta_n \geqslant (\|\bar{x}\| - \|u^n\|)^2 > 0, \text{ para } n \text{ suficientemente grande.} \tag{4.7}$$

Pata isto basta notar que

$$\Delta_n = \langle J(\bar{x}) - J(u^n), \bar{x} - u^n\rangle =$$

$$\langle J(\bar{x}), \bar{x}\rangle + \langle J(u^n), u^n\rangle - \langle J(\bar{x}), u^n\rangle - \langle J(u^n), \bar{x}\rangle =$$

$$\|\bar{x}\|^2 + \|u^n\|^2 - \langle J(\bar{x}), u^n\rangle - \langle J(u^n), \bar{x}\rangle \geqslant$$

$$\|\bar{x}\|^2 + \|u^n\|^2 - \|J(\bar{x})\|_*\|u^n\| - \|J(u^n)\|_*\|\bar{x}\| =$$

$$\|\bar{x}\|^2 + \|u^n\|^2 - 2\|\bar{x}\|\|u^n\| = (\|\bar{x}\| - \|u^n\|)^2 > 0.$$

Afirmamos que Dom $\partial f(\bar{x}, \cdot) \cap C \neq \emptyset$. De fato, o subdiferencial de uma função convexa, própria e semicontínua inferior é monótono maximal em qualquer espaço de Banach (ver por exemplo [14, Teorema A, pág. 210]). Se estendermos a função $f(\bar{x}, \cdot)$ a todo o espaço X, definindo-a como ∞ fora de C, então $\partial f(\bar{x}, \cdot)$ é monótono maximal. Por outro lado, o operador $T : X \rightrightarrows P(X^*)$, definido por $T(x) = \emptyset$ para todo $x \in X$ é certamente não-monótono maximal, porque seu gráfico é monótono e estritamente contido no gráfico de todo operador monótono maximal não trivial. Deste modo Dom $\partial f(\bar{x}, \cdot)$ é não-vazio. Agora, uma vez que $\partial f(\bar{x}, z) = \emptyset$ para todo $z \notin C$, segue que Dom $\partial f(\bar{x}, \cdot) \cap C \neq \emptyset$. Consequentemente, existe $v \in \partial f(\bar{x}, \hat{x})$ para algum $\hat{x} \in C$. Sendo assim, podemos escrever

$$\begin{aligned} f(\bar{x}, u^n) &\geqslant f(\bar{x}, \hat{x}) + \langle v, u^n - \hat{x}\rangle \\ &= f(\bar{x}, \hat{x}) - \langle v, \hat{x}\rangle + \langle v, u^n\rangle \\ &\geqslant A - B\|u^n\|, \end{aligned} \tag{4.8}$$

onde $A := f(\bar{x}, \hat{x}) - \langle v, \hat{x}\rangle$ e $B := \|v\|_*$. De modo geral, temos

$$\liminf_{n\to\infty}[f(\bar{x}, u^n) + (\gamma - \theta)\Delta_n] \geqslant \liminf_{n\to\infty}[A - B\|u^n\| + (\gamma - \theta)(\|\bar{x}\| - \|u^n\|)^2] = \infty$$

e então a condição (P6) é estabelecida sobre f. Pela Proposição 4.2.1 concluimos que a condição (P5) é válida para $\tilde{f}$ neste caso. □

O próximo resultado estabelece a existência e a unicidade de solução de $PE(\tilde{f}, C)$.

Corolário 4.2.2. *Considere todas as hipóteses da Proposição 4.2.1 válidas. Então valem as seguintes afirmações:*

(i) Se a condição (P6) *é válida, então* $\mathsf{PE}(\tilde{f}, C)$ *admite pelo menos uma solução.*

(ii) Se φ *é estritamente convexa, então* $\mathsf{PE}(\tilde{f}, C)$ *admite no máximo uma solução.*

De modo geral, se vale a condição (P6) *e* φ *é estritamente convexa, então* $\mathsf{PE}(\tilde{f}, C)$ *tem uma única solução.*

Demonstração. Da Proposição 4.2.1 temos que $\tilde{f}$ satisfaz a condição (P5). Usando agora [28, Teorema 4.3, pág. 270] obtemos que $\mathsf{PE}(\tilde{f}, C)$ possui uma solução. Isto prova (i). Para provar (ii), suponha que ambos x' e x'' sejam soluções de $\mathsf{PE}(\tilde{f}, C)$. Então

$$0 \leqslant \tilde{f}(x', x'') = f(x', x'') + \gamma\langle \nabla\varphi(x') - \nabla\varphi(\bar{x}), x'' - x'\rangle, \tag{4.9}$$

$$0 \leqslant \tilde{f}(x'', x') = f(x'', x') + \gamma\langle \nabla\varphi\nabla(x'') - \nabla\varphi(\bar{x}), x' - x''\rangle. \tag{4.10}$$

Usando (4.9) e (4.10) e a propriedade (P4°) de f temos

$$\begin{aligned} 0 &\leqslant \tilde{f}(x'', x') + \tilde{f}(x', x'') \\ &= f(x', x'') + \gamma\langle \nabla\varphi(x') - \nabla\varphi(\bar{x}), x'' - x'\rangle \\ &\quad + f(x'', x') + \gamma\langle \nabla\varphi(x'') - \nabla\varphi(\bar{x}), x' - x''\rangle \\ &= f(x', x'') + f(x'', x') - \gamma\langle \nabla\varphi(x'') - \nabla\varphi(x'), x'' - x'\rangle \\ &\leqslant (\theta - \gamma)\langle \nabla\varphi(x'') - \nabla\varphi(x'), x'' - x'\rangle. \end{aligned} \tag{4.11}$$

Uma vez que $\theta - \gamma < 0$ e $\nabla\varphi$ é estritamente monótono, obtemos de (4.11) que $x' = x''$. □

No resultado a seguir assumiremos que φ é *coerciva*, isto é, $\lim\limits_{\|x\|\to\infty} \frac{\varphi(x)}{\|x\|} = \infty$.

Corolário 4.2.3. *Seja* φ *uma função coerciva e Gateaux diferenciável. Se a bifunção* $g : C \times C \to \mathbb{R}$ *satisfaz às condições* (P1), (P2), (P3) *e* (P4), *então* $Res_g^{\varphi}(\bar{x}) \neq \emptyset$, *para todo* $\bar{x} \in C$.

Demonstração. Note que todas as hipóteses do Corolário 4.2.2 são válidas. Portanto, é suficiente provar que a condição (P6) é válida quando φ é coerciva. Tome uma sequência $\{u^n\} \subset C$ tal que $\lim_{n\to\infty} \|u^n\| = \infty$ e defina $\Delta_n := \langle \nabla\varphi(\bar{x}) - \nabla\varphi(u^n), \bar{x} - u^n\rangle$. Usando a convexidade de φ podemos escrever

$$\begin{aligned} \Delta_n &= \langle \nabla\varphi(\bar{x}), \bar{x} - u^n\rangle - \langle \nabla\varphi(u^n), \bar{x} - u^n\rangle \\ &\geqslant \langle \nabla\varphi(\bar{x}), \bar{x} - u^n\rangle + [\varphi(u^n) - \varphi(\bar{x})] \\ &= \|u^n\| \left[\frac{\langle \nabla\varphi(\bar{x}), \bar{x} - u^n\rangle}{\|u^n\|} + \frac{\varphi(u^n)}{\|u^n\|} - \frac{\varphi(\bar{x})}{\|u^n\|}\right] \end{aligned} \tag{4.12}$$

Por outro lado, podemos escrever

$$f(\bar{x}, u^n) \geqslant A - B\|u^n\| \tag{4.13}$$

onde $A := f(\bar{x}, \hat{x}) - \langle v, \hat{x} \rangle$ e $B := \|v\|_*$, assim como na prova do item (ii) do Corolário 4.2.1. Combinando (4.12) e (4.13) e rearranjando a expressão resultante, temos

$$\begin{aligned} &\liminf_{n\to\infty}[f(\bar{x}, u^n) + (\gamma - \theta)\Delta_n] \\ &\geqslant \liminf_{n\to\infty}\left[A - B\|u^n\| + (\gamma - \theta)\|u^n\|\left(\tfrac{\langle\nabla\varphi(\bar{x}), \bar{x}-u^n\rangle}{\|u^n\|} + \tfrac{\varphi(u^n)}{\|u^n\|} - \tfrac{\varphi(\bar{x})}{\|u^n\|}\right)\right] \\ &= \liminf_{n\to\infty}\left[A + (\gamma - \delta)\|u^n\|\left(\tfrac{\langle\nabla\varphi(\bar{x}), \bar{x}-u^n\rangle}{\|u^n\|} + \tfrac{\varphi(u^n)}{\|u^n\|} - \tfrac{\varphi(\bar{x})}{\|u^n\|} - \tfrac{B}{(\gamma-\theta)}\right)\right]. \end{aligned}$$

A hipótese de coercividade implica que a expressão no interior dos colchetes tende ao infinito e, portanto, a condição (P6) é estabelecida. □

Observação 4.2.1. *O corolário anterior prova que a condição* (P6) *é uma condição mais fraca que a hipótese de coercividade. Além disso, a condição* (P6) *é estritamente mais fraca que a coercividade, desde que existam funções de Bregman que não sejam coercivas, mas satisfaçam a condição* (P6)*. Seja* $X = \mathbb{R}$, $C \subset (0, \infty)$ *e considere* $\varphi(x) = x - \log x + 1$*. Esta função não é coerciva, uma vez que* $\lim_{x\to\infty} \varphi(x)/x < \infty$*. Entretanto, é fácil verificar a partir da definição que a condição* (P6) *vale desde que* $\gamma - \theta$ *seja suficientemente grande. De fato, seja* $t_n \to \infty$*. Com a notação usada no Corolário 4.2.3 com* $\bar{t}$ *no lugar de* $\bar{x}$ *e* t_n *ao invés de* u^n*, podemos escrever*

$$\begin{aligned} [f(\bar{t}, t_n) + (\gamma - \theta)\Delta_n] \geqslant\ & A - Bt_n + (\gamma - \theta)[(\varphi'(\bar{t}) - \varphi'(t_n))(\bar{t} - t_n)] \\ &= A - Bt_n + (\gamma - \theta)\left(t_n^{-1} - \bar{t}^{-1}\right)(\bar{t} - t_n) \\ &= A - Bt_n + (\gamma - \theta)(\bar{t} - t_n)^2(\bar{t}t_n)^{-1} \\ &= A + t_n\left[\tfrac{\gamma-\theta}{\bar{t}}\left(1 - \tfrac{\bar{t}}{t_n}\right)^2 - B\right]. \end{aligned}$$

Note que a expressão do lado direito entre colchetes é positiva para $\gamma - \theta$ *e* n *suficientemente grandes. Assim, a condição* (P6) *é válida neste caso. Quão grande deve ser* $\gamma - \theta$ *também depende da bifunção* f*, através do número* B.

Corolário 4.2.4. *Sejam todas as hipóteses da Proposição 4.2.1 válidas. Suponha que* X^* *é estritamente convexo e que* $\varphi = (1/2)\|\cdot\|^2$*. Então* $PE(\tilde{f}, C)$ *possui uma única solução.*

Demonstração. Observe que $J = \partial\varphi$ é estritamente monótona pelo Teorema de Petryshyn (Ver [32]). Consequentemente, o resultado segue do Corolário 4.2.2. □

4.3 Um Método de Ponto Proximal para a solução de Problemas de Equilíbrio

A partir de agora assumiremos que φ é estritamente convexa e que valem $(\mathsf{P1})$, $(\mathsf{P2})$, $(\mathsf{P3})$, $(\mathsf{P4}^\circ)$ e $(\mathsf{P6})$. A definição a seguir é baseada em [38, pág. 1265].

Definição 4.3.1. *Dizemos que a sequência* $\{z^k\} \subset C$ *é* assintoticamente resolvente *para o problema* $\mathsf{PE}(f, C)$ *se* $\liminf_{k\to\infty} f(z^k, y) \geqslant 0$ *para todo* $y \in C$.

Consideremos uma sequência de parâmetros de regularização $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$, para algum $\bar{\gamma} > 0$. Com base no problema regularizado (4.5) e na existência e unicidade de soluções (Corolários 4.2.2 e 4.2.4), será apresentado o seguinte algoritmo para resolver o $\mathsf{PE}(f, C)$.

Passo 1. Escolhe-se um $x^0 \in C$.

Passo 2. Constrói-se a sequência $\{x^k\} \subset C$ da seguinte forma: dado $x^k \in C$, x^{k+1} é a única solução do problema $\mathsf{PE}(f_k, C)$, onde $f_k : C \times C \to \mathbb{R}$ é dada por

$$f_k(x, y) = f(x, y) + \gamma_k \langle \nabla\varphi(x) - \nabla\varphi(x^k), y - x\rangle. \tag{4.14}$$

A formulação resolvente da iteração acima é

$$x^{k+1} \in \mathrm{Res}^{\varphi}_{1/y_k f}(x^k).$$

Teorema 4.3.1. *Seja* X *um espaço de Banach reflexivo e considere o problema* $\mathsf{PE}(f, C)$. *Para todo* $x^0 \in C$, *temos o seguinte:*

(i) A sequência $\{x^k\}$ *está bem definida.*

(ii) Se, além disso, valem B1 e B2 para D_φ *e* $S^d(f, C) \neq \emptyset$, *então a sequência* $\{x^k\}$ *é limitada e* $\lim_{k\to} \| x^{k+1} - x^k \| = 0$.

(iii) Sob as hipóteses do item (ii) e assumindo ainda que

(iiia) X^* *é uniformemente convexo e* $\varphi = (1/2) \| \cdot \|^2$, *ou*

(iiib) D_φ *satisfaz B3,*

a sequência $\{x^k\}$ *é uma sequência assintoticamente resolvente do* $\mathsf{PE}(f, C)$, *isto é,*

$$0 \leqslant \liminf_{k\to\infty} f(x^k, y), \ \forall\, y \in C \tag{4.15}$$

(iv) Se, além disso, $f(\cdot, y)$ *é fracamente semicontínua superior para todo* $y \in C$, *então todos os pontos de acumulação fracos de* $\{x^k\}$ *resolvem o* $\mathsf{PE}(f, C)$.

Demonstração. O item (i) segue imediatamente do Corolário 4.2.2 da Proposição 4.6. De fato, como cada $f_k(\cdot, y)$ ($f_k(x, y) = f(x, y) + \gamma_k \langle \nabla\varphi(x) - \nabla\varphi(x^k), y - x\rangle$) tem pelo menos uma solução (pois vale (P6)) e tem no máximo uma solução (pois φ é estritamente convexa) então, para cada k, $x^{k+1} \in S(f_k, C)$ é unicamente determinado e assim a sequência está bem definida.

Para provar (ii) tome um $x^* \in S^d(f, C)$ (pois $S^d(f, C) \neq \emptyset$) arbitrário. Como $x^{k+1} \in S(f_k, C)$ (pelo item (i)), temos por um lado que

$$0 \leqslant f_k(x^{k+1}, x^*) = f(x^{k+1}, x^*) + \gamma_k \langle \nabla\varphi(x^{k+1}) - \nabla\varphi(x^*), x^* - x^{k+1}\rangle \tag{4.16}$$

e como $f(x^{k+1}, x^*) \leqslant 0$ (pois $x^* \in S^d(f, C)$), temos que

$$\langle \nabla\varphi(x^{k+1}) - \nabla\varphi(x^*), x^* - x^{k+1}\rangle \geqslant 0. \tag{4.17}$$

Agora, usando a propriedade dos três pontos para distâncias de Bregman, temos:

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

o que nos dá

$$D_\varphi(x^{k+1}, x^k) + D_\varphi(x^*, x^{k+1}) \leqslant D_\varphi(x^*, x^k). \tag{4.18}$$

Segue daí que a sequência $\{D_\varphi(x^*, x^{k+1})\}$ é decrescente e sendo não-negativa, ela converge. Em particular, é limitada. Pela condição B1, concluímos que a sequência $\{x^k\}$ é limitada. Assim, temos que

$$0 \leqslant D_\varphi(x^{k+1}, x^k) \leqslant D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1}) \tag{4.19}$$

e como a expressão do lado direito converge a zero quando $k \to \infty$, segue que

$$\lim_{k\to\infty} D_\varphi(x^{k+1}, x^k) = 0. \tag{4.20}$$

Agora, como $\{x^k\}$ é limitada e D_φ satisfaz B2, aplicamos o Lema 2.2.1 nas sequências $\{x^k\}$ e $\{x^{k+1}\}$ para concluir que

$$\lim_{k\to\infty} \| x^{k+1} - x^k \| = 0. \tag{4.21}$$

Para a prova de (iii) sob a hipótese (iiia) observe que, sendo X^* uniformemente convexo, a aplicação J é uniformemente contínua em subconjuntos limitados de X (pelo Teorema 1.1.6) e, portanto,

$$\lim_{k\to\infty} \|J(x^{k+1}) - J(x^k)\| = 0. \tag{4.22}$$

Agora fixe qualquer $y \in C$. Uma vez que $x^{k+1} \in S(f_k, C)$ temos

$$0 \leqslant f(x^{k+1}, y)+\gamma_k\langle J(x^{k+1})-J(x^k), y-x^{k+1}\rangle \leqslant f(x^{k+1}, y)+\gamma_k\|J(x^{k+1})-J(x^k)\|\|y-x^{k+1}\|$$

Uma vez que $\{\gamma_k\}$ é limitada por $\overline{\gamma}$ e $\{x^k\}$ é limitada (pelo item (ii)), obtemos que

$$0 \leqslant \liminf_{k\to\infty} f(x^k, y), \forall y \in C \tag{4.23}$$

e assim, $\{x^k\}$ é uma sequência assintoticamente resolvente para $PE(f, C)$.

Para a prova de (iii) sob a hipótese (iiib), temos que B2 e o Lema 2.2.1 nos dão

$$\lim_{k\to\infty} \|\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)\| = 0 \tag{4.24}$$

Agora fixe $y \in C$ arbitrário. Uma vez que $x^{k+1} \in S(f_k, C)$, temos

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

Como $\{\gamma_k\}$ é limitada por $\bar{\gamma}$ e $\{x^k\}$ é limitada (pelo item (ii)), obtemos que

$$0 \leqslant \liminf_{k\to\infty} f(x^k, y), \forall y \in C \tag{4.25}$$

e assim, $\{x^k\}$ é uma sequência assintoticamente resolvente para $PE(f, C)$.

Para provar (iv) note que, pelo item (ii), a sequência $\{x^k\}$ é limitada. Agora, utilizando o Teorema de Banach-Alaoglu-Bourbaki (Teorema 1.1.4), juntamente com os resultados em [3, Teorema 6.3.9, pág. 156] e [2, Proposição 1, Cap. 7, pág. 178], concluimos que $\{x^k\}$ possui pontos de acumulação fracos, todos os quais pertencem ao fracamente fechado e convexo C. Por fim, considere $\hat{x}$ um ponto de acumulação fraco de $\{x^k\}$ e seja $\{x^{k_j}\}$ uma subsequência que converge fracamente para $\hat{x}$. Uma vez que $f(\cdot, y)$ é fracamente semicontínua superior, podemos escrever

$$f(\hat{x}, y) \geqslant \limsup_{j\to\infty} f(x^{k_j}, y) \geqslant 0, \forall\, y \in C, \tag{4.26}$$

o que equivale a dizer que $\hat{x} \in S(f, C)$.

$\square$

Proposição 4.3.1. *Suponha válidas todas as hipóteses do Teorema 4.3.1 e que* $\nabla\varphi$ *é contínuo na topologia fraca. Se* $S(f,C) = S^d(f,C)$*, então a sequência* $\{x^k\}$ *converge fracamente para uma solução do problema* $PE(f,C)$*.*

Demonstração. Pelo Teorema 4.3.1 é suficiente verificar que existe somente um ponto de acumulação fraco de $\{x^k\}$. Sejam x' e x'' dois pontos de acumulação fracos de $\{x^k\}$ e considere subsequências $\{x^{i_k}\}$ e $\{x^{j_k}\}$ convergindo fracamente para x' e x'' respectivamente. Pelo item (iv) do Teorema 4.3.1 segue que ambos x' e x'' pertencem à $S(f,C) = S^d(f,C)$. Por (4.18), temos que $D_\varphi(x', x^k)$ e $D_\varphi(x'', x^k)$ são ambas convergentes, digamos, à $\sigma \geqslant 0$ e $\mu \geqslant 0$, respectivamente. Agora, considerando a expressão

$$\langle \nabla\varphi(x^{i_k}) - \nabla\varphi(x^{j_k}), x' - x'' \rangle = [D_\varphi(x', x^{j_k}) - D_\varphi(x', x^{i_k})] - [D_\varphi(x'', x^{j_k}) - D_\varphi(x'', x^{i_k})]$$

e tomando os limites em ambos os lados quando $k \to \infty$, obtemos

$$\langle \nabla\varphi(x') - \nabla\varphi(x''), x' - x'' \rangle = [\sigma - \sigma] - [\mu - \mu] = 0,$$

e consequentemente $x' = x''$ pela monotonicidade estrita de $\nabla\varphi$. □

O resultado abaixo segue imediatamente da proposição anterior.

Corolário 4.3.1. *Suponha que* X *tem uma aplicação de dualidade normalizada* J *que é contínua na topologia fraca. Suponha também todas as hipóteses do Teorema 4.3.1 com (iiia) válida. Se* $S(f,C) = S^d(f,C)$*, então a sequência* $\{x^k\}$ *converge fracamente a uma solução do problema* $PE(f,C)$*.*

Demonstração. Basta notar que, neste caso, $J = \nabla\varphi$ e aplicar a Proposição 4.3.1. □

Referências Bibliográficas

[1] Burachik, R. S.; Kassay, G. *On a Generalized Proximal Point Method for Equilibrium Problems in Banach Spaces.* Nonlinear Analysis Series A: Theory, Methods & Applications, 75: 6456-6464, 2012.

[2] Lima, E. L. *Elementos de Topologia Geral.* Rio de Janeiro: SBM, 2009.

[3] Botelho, G., Pellegrino, D., Teixeira, E. *Fundamentos de análise funcional.* Rio de Janeiro: SBM, 2012.

[4] Brezis, Haim. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, 2010.

[5] Isnard, Carlos. *Introdução à medida e integração.* Rio de Janeiro: IMPA, 2007.

[6] Kreyszig, Erwin. *Introductory functional analysis with aplications.* Wiley, 1978.

[7] R. Phelps, *Convex Functions, Monotone Operators, and Differentiability*, second ed., in: *Lecture Notes in Mathematics*, vol. 1364, Springer, Germany, 1993.

[8] Izmailov, A., Solodov, M., *Otimização - Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*, IMPA, Rio de Janeiro. vol.1, 2005.

[9] F. Botelho, *Functional Analysis and Applied Optimization in Banach Spaces: Applications to Non-Convex Variational Models*, Springer, 2014.

[10] Rockafellar, R. T. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[11] Iusem, A. N. *Métodos de Ponto Proximal em Otimização* Coleção: 20º Colóquio Brasileiro de Matemática. Rio de Janeiro: IMPA, 1995.

[12] F. Acker, F. Dickstein. *Uma introdução à análise convexa.* Rio de Janeiro: IMPA, 1983.

[13] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, in: Monogr. Math., Springer, 2010.

[14] R.T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math. 33 (1970) 209-216.

[15] Bregman, L.M. (1967) *The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming*, USSR Computational Mathematics and Mathematical Physics, 7, 200-217.

[16] Censor, Y. e Lent, A. (1981) *An iterative row-action method for interval convex programming*, J. Optim. Theory Appl., 34, 321-353.

[17] D. Butnariu, E. Resmerita, *Bregman distances, totally convex functions, and a method for solving operator equations in Banach spaces.* Hindawi Publishing Corporation. Abstract and Aplplied Analysis. V. 2006. 1-39.

[18] Resmerita, E. *On total convexity, Bregman projections and stability in Banach spaces*, Journal of Convex Analysis, 11 (2004), 1-16.

[19] G. Kassay, S. Reich, S. Sabach, *Iterative methods for solving systems of variational inequalities in reflexive Banach spaces*, SIAM J. Optim. 21 (2011), 1319-1344.

[20] D. Butnariu, Y. Censor, S. Reich, *Iterative averaging of entropic projections for solving stochastic convex feasibility problems*, Comput. Optim. Appl. 8 (1997), 21-39.

[21] D. Butnariu, A.N. Iusem, *Totally Convex Functions for Fixed Point Computation and Infinite Dimensional Optimization*, Kluwer, Dordrecht, 2000.

[22] D. Butnariu, C.A. Isnard, A.N. Iusem, *A mixed Hölder and Minkowski inequality*, Proc. Amer. Math. Soc. 127 (1999), 2405-2415

[23] D. Butnariu, A.N. Iusem, E. Resmerita, *Total convexity of powers of the norm in uniformly convex Banach spaces*, J. Convex Anal. 7 (2000) 319-334.

[24] R.S. Burachik, S. Scheimberg, *A proximal point method for the variational inequality problem in Banach spaces*, SIAM J. Control Optim. 39 (2001) 1633-1649.

[25] S. Reich, S. Sabach, *Two strong convergence theorems for Bregman strongly nonexpansive operators in reflexive Banach spaces*, Nonlinear Anal. 73(2010) 122-135.

[26] F. H. Clarke: *Optimization and Nonsmooth Analysis*, John Willey and Sons, New York, 1983.

[27] E. Blum, W. Oettli, *From optimization and variational inequalities to equilibrium problems*, Math. Student 63 (1994) 123-145.

[28] A.N. Iusem, G. Kassay, W. Sosa, *On certain conditions for the existence of solutions of equilibrium problems*, Math. Program. 116 (2009) 259-273.

[29] A.N. Iusem, W. Sosa, *New existence results for equilibrium problems*, Nonlinear Anal. 52 (2003) 621-635.

[30] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. 14 (1976), pp. 877-898.

[31] Rolando Gárciga Otero, e Alfredo N. Iusem. *Proximal methods in reflexive Banach spaces without monotonicity.* Journal of mathematical analysis and applications 330.1 (2007): 433-450.

[32] W.V. Petryshyn, *A characterization of strict convexity of Banach spaces and other uses of duality mappings*, J. Funct. Anal. 6 (1970) 282-291

[33] Isac, G. - *Complementarity Problems.* Lecture Notes in Mathematics, Springer-Verlag, New York, 1992.

[34] Mas-colell, A., Whinston, M.D., Green, J.R. - *Microeconomic Theory.* Oxford University Press, 1995.

[35] Nash, J. -*Equilibrium points in N-person games.* Proc. Nat. Acad. Sci. USA (36) 48-49, 1950.

[36] Nash, J. - *Non-cooperative games.* Ann. Math. (54) 286-293, 1951.

[37] Kinderlehrer, D., Stampacchia, G. - *An introduction to variational inequalities and their applications.* Academic Press, New York, 1980.

[38] A.N. Iusem, W. Sosa, *On the proximal point method for equilibrium problems in Hilbert spaces*, Optimization 59 (2010) 1259-1274.