



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO EM MATEMÁTICA

**Um Método Proximal com Distâncias Proximais
Generalizadas para Problemas de Equilíbrio de Dois
Níveis**

Gean Lucas da Silva Veloso

Teresina - 2026

Gean Lucas da Silva Veloso

Dissertação de Mestrado:

**Um Método Proximal com Distâncias Proximais Generalizadas
para Problemas de Equilíbrio de Dois Níveis**

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 - 2026



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

Um método proximal com distâncias proximais generalizadas para problemas de equilíbrio de dois níveis

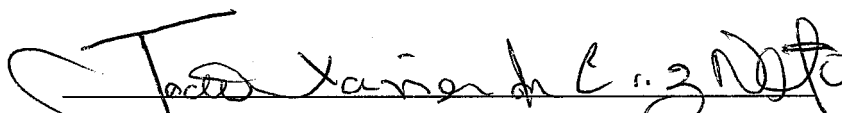
Gean Lucas da Silva Veloso

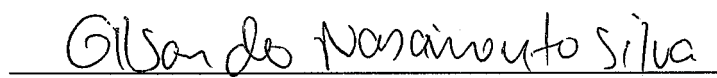
Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

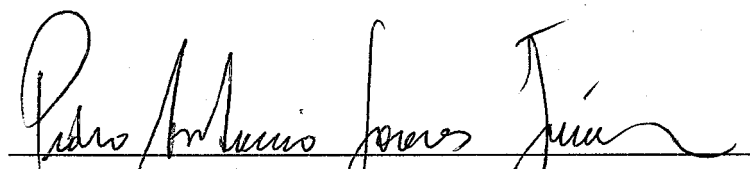
A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Dissertação aprovada em 20 de julho de 2026.

Banca Examinadora:


Prof. Dr. João Xavier da Cruz Neto – Orientador


Prof. Dr. Gilson do Nascimento Silva - UFPI


Prof. Dr. Pedro Antonio Soares Júnior - UESPI

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

V443m Veloso, Gean Lucas da Silva.
 Um método proximal com distâncias proximais
 generalizadas para problemas de equilíbrio de dois níveis /
 Gean Lucas da Silva Veloso. -- 2026.
 120 f.

 Dissertação (Mestrado) - Universidade Federal do Piauí,
 Centro de Ciências da Natureza, Programa de Pós-Graduação
 em Matemática, Teresina, 2026.
 “Orientador: Prof. Dr. João Xavier da Cruz Neto”.

 1. Otimização. 2. Método Proximal. 3. Análise Convexa.
 4. Problema do Equilíbrio. I. Cruz Neto, João Xavier da Cruz.
 II. Título.

CDD 519.3

Bibliotecária: Caryne Maria da Silva Gomes - CRB3/1461

*A todos que acreditaram em mim e fizeram parte
desta trajetória.*

Agradecimentos

Agradeço, primeiramente, a Deus, por até aqui ter me ajudado, sustentado e concedido forças e coragem para continuar, mesmo diante das dificuldades encontradas ao longo desta caminhada.

Agradeço a toda a minha família, em especial à minha mãe, Zilnete, por sempre ter me dado todo o apoio necessário; ao meu pai, Francisco (Chicão), responsável por despertar em mim, ainda muito cedo, o amor pela Matemática; à minha irmã, Kayllane, e ao meu irmão, Daniel, por todo o apoio e motivação ao longo dessa trajetória. Agradeço, sobretudo, aos meus pais, por todos os esforços e sacrifícios feitos para que a educação de seus filhos fosse sempre uma prioridade. Amo vocês.

À minha amada Letícia Carolinny, agradeço por ter estado ao meu lado em todos os momentos, pelo apoio, pelo carinho e por ter sido meu porto seguro durante toda essa caminhada. Sua presença foi fundamental para que eu chegasse até aqui. Sem você, tudo teria sido muito mais difícil. Te amo.

Agradeço ao meu amigo da graduação Thiago de Paulo, pelos conselhos, pela amizade e por estar presente em todas as horas. Aos amigos que conheci durante a pós-graduação, agradeço pelas horas de diversão, pelo companheirismo e por terem me ajudado ao longo dessa jornada. Em especial, agradeço a Elliel Matias, Francisco Júnior, Gustavo Fernandes, Hyon Ferreira, Jefferson Brito, Otávio “Almôndegas”, Pedro Henrique e Vinícius Araújo, pelos “cafezinhos” essenciais para os momentos de descontração e por terem tornado o mestrado um pouco mais leve. Agradeço também a Ana Júlia, Andreina Marcela, Antonio Victor, Eduardo Viana, Emanuely, Halan Santos, Igor Fontenele, Ismael de Carvalho, Katia de Sousa, Lucas Aurélio, Maiara Veras, Maria Angélica, Natália Brito, Raimundo Bruno, Sayd, Stefany Maria, Victor Emanuel e Ysadora Soares, pela convivência e pelos momentos compartilhados durante essa etapa. Agradeço também à tia Cruzinha, pelas muitas conversas no prédio da Pós, e à Larisse, responsável por sempre

construir um ambiente familiar, unir todos na Pós-graduação com os eventos que ela faz e que muito ajudou com seus sábios concelhos quando precisei.

Agradeço a todos os professores do Departamento de Matemática que contribuíram para a minha formação acadêmica, em especial aos professores Barnabé Pessoa Lima, Gleison Santos, Ítalo Melo, Joel Rabelo, Mykael Cardoso, Ray Serra e Sandoel Vieira. Aos professores Halyson Baltazar e Leandro Pessoa, agradeço pela excelente condução e dedicação à coordenação do programa durante esses dois anos. Agradeço também ao professor Gilson Silva, por ter aceitado fazer parte da banca de defesa e pelas valiosas contribuições dadas a este trabalho.

Ao professor Pedro Antônio Soares Jr., agradeço por ter aceitado fazer parte da banca, pela coorientação, pelas importantes contribuições para este trabalho e, sobretudo, por ter me acompanhado e orientado desde a graduação. Sou profundamente grato pelas diversas portas que me abriu ao longo dos anos, as quais foram fundamentais para a minha formação acadêmica e para a aquisição de experiência profissional, tanto na UESPI quanto na FAPEPI.

Foi também por intermédio do professor Pedro que tive a oportunidade de conhecer o professor João Xavier da Cruz Neto, a quem agradeço por ter me orientado profissionalmente durante o período em que trabalhei na FAPEPI e, posteriormente, por ter aceitado me orientar durante o mestrado. Agradeço pela confiança, pelos ensinamentos, pela dedicação e por todas as contribuições para a minha formação e para a realização deste trabalho.

Sinto-me profundamente honrado por ter sido orientado por dois grandes professores e pesquisadores que tanto contribuíram e continuam contribuindo para o desenvolvimento da Matemática e da ciência no estado do Piauí.

Agradeço à Fundação de Amparo à Pesquisa do Estado do Piauí (FAPEPI) pela oportunidade e pela experiência profissional proporcionadas durante o final da minha graduação e o início do mestrado, período que teve grande importância para a minha formação.

Por fim, agradeço à CAPES pelo apoio financeiro concedido durante o mestrado.

“O importante é não parar de questionar.”

Albert Einstein.

Resumo

Neste trabalho estudamos problemas de equilíbrio de dois níveis em $\mathbb{R}^n$ por meio de métodos proximais com distâncias proximais generalizadas. Inicialmente, apresentamos conceitos fundamentais de análise convexa, incluindo conjuntos convexos, funções convexas, subdiferenciais, funções conjugadas, cones normais e funções suporte. Em seguida, introduzimos os problemas de equilíbrio, destacando sua capacidade de unificar diversos modelos clássicos, como problemas de otimização convexa, problemas de ponto fixo e desigualdades variacionais. O foco principal da dissertação é a análise de problemas de equilíbrio de dois níveis, nos quais se busca uma solução de um problema de equilíbrio de nível superior restrita ao conjunto solução de um problema de equilíbrio de nível inferior. Para isso, estudamos a classe das distâncias proximais generalizadas e sua utilização na construção de bifunções regularizadas. Sob hipóteses adequadas de convexidade, semicontinuidade, monotonicidade e coercividade assintótica, apresentamos condições que garantem existência e unicidade de soluções dos subproblemas regularizados. Por fim, analisamos o método proximal proposto para resolver problemas de equilíbrio de dois níveis e, sob as hipóteses adotadas, demonstramos que a sequência gerada pelo algoritmo converge para uma solução do problema. Encerramos o trabalho discutindo perspectivas futuras, como versões inexatas do método e taxas de convergência.

Palavras-chave: problema de equilíbrio; problema de equilíbrio de dois níveis; método proximal; distâncias proximais generalizadas; análise convexa.

Abstract

In this work we study bilevel equilibrium problems in $\mathbb{R}^n$ by means of proximal methods with generalized proximal distances. We first present fundamental concepts of convex analysis, including convex sets, convex functions, subdifferentials, conjugate functions, normal cones and support functions. Then, we introduce equilibrium problems, emphasizing their ability to unify several classical models, such as convex optimization problems, fixed point problems and variational inequalities. The main focus of this dissertation is the analysis of bilevel equilibrium problems, in which one seeks a solution of an upper-level equilibrium problem restricted to the solution set of a lower-level equilibrium problem. To this end, we study the class of generalized proximal distances and their use in the construction of regularized bifunctions. Under suitable assumptions of convexity, semicontinuity, monotonicity and asymptotic coercivity, we present conditions ensuring existence and uniqueness of solutions of the regularized subproblems. Finally, we analyze the proximal method proposed for solving bilevel equilibrium problems and under the adopted assumptions, we demonstrate that the sequence generated by the algorithm converges to a solution of the problem. We conclude by discussing future research directions, such as inexact versions of the method and convergence rates.

Keywords: equilibrium problem; bilevel equilibrium problem; proximal method; generalized proximal distances; convex analysis.

Lista de Símbolos

$\mathbb{N}$	Conjunto dos números naturais.
$\mathbb{R}$	Conjunto dos números reais.
$\mathbb{R}^n$	Espaço euclidiano n -dimensional.
$\mathbb{R}_+$	Conjunto dos números reais não negativos.
$\langle \mathbf{x}, \mathbf{y} \rangle$	Produto interno usual entre $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
$\ \mathbf{x}\ $	Norma euclidiana de $\mathbf{x} \in \mathbb{R}^n$.
$\ \cdot\ _1$	Norma da soma em $\mathbb{R}^n$.
$B_r(\mathbf{a})$	Bola aberta de centro $\mathbf{a}$ e raio $r > 0$.
$B_r[\mathbf{a}]$	Bola fechada de centro $\mathbf{a}$ e raio $r > 0$.
$S_r(\mathbf{a})$	Esfera de centro $\mathbf{a}$ e raio $r > 0$.
$\text{int}(C)$	Interior do conjunto C .
$\text{conv}(C)$	Envoltória convexa do conjunto C .
$\text{core}_K C$	Core de C relativo ao conjunto K .
E_f	Epígrafo da função f .
$\text{dom}(f)$	Domínio efetivo da função f .
$\nabla f(\mathbf{x})$	Gradiente da função diferenciável f no ponto $\mathbf{x}$.
$\partial f(\mathbf{x})$	Subdiferencial da função f no ponto $\mathbf{x}$.
δ_Ω	Função indicadora do conjunto Ω .
$\mathcal{N}_\Omega(\mathbf{x})$	Cone normal ao conjunto Ω no ponto $\mathbf{x}$.
σ_Ω	Função suporte do conjunto Ω .
f^*	Conjugada de Fenchel da função f .

f^{**}	Biconjugada de Fenchel da função f .
$f \square g$	Inf-convolução das funções f e g .
$PE(f, K)$	Problema de equilíbrio associado à bifunção f e ao conjunto K .
$S(f, K)$	Conjunto solução do problema de equilíbrio $PE(f, K)$.
$Fix(T)$	Conjunto dos pontos fixos de T .
$VI(T, K)$	Problema de desigualdade variacional associado ao operador T e ao conjunto K .
$\mathcal{P}(\mathbb{R}^n)$	Conjunto das partes de $\mathbb{R}^n$.
$PEB(f, h, K)$	Problema de equilíbrio de dois níveis associado às bifunções f, h e ao conjunto K .
$S(h, S(f, K))$	Conjunto solução do problema de equilíbrio de nível superior sobre $S(f, K)$; corresponde ao conjunto solução do problema de dois níveis.
$\operatorname{argmin}_{x \in K} \varphi(x)$	Conjunto dos minimizadores de φ sobre K .
$o(\varepsilon_k)$	Notação de Landau para uma quantidade assintoticamente desprezível em relação a ε_k .
f_z	Função $f_z : K \rightarrow \mathbb{R}$ definida por $f_z(y) := f(z, y)$.
$\mathcal{R}(\mathcal{N}_{S(f, K)})$	Imagem da aplicação cone normal associada ao conjunto $S(f, K)$.
$(\mathcal{H})$	Hipótese geométrica de regularidade utilizada na análise de convergência do método proximal.
S	Conjunto não vazio, fechado e convexo utilizado na definição das distâncias proximais generalizadas.
d	Distância proximal associada ao conjunto S .
$\mathcal{D}(S)$	Família de todas as distâncias proximais associadas ao conjunto S .
D	Distância proximal induzida associada à distância proximal d .
$\partial_1 d(\cdot, y)$	Subdiferencial de $d(\cdot, y)$ com relação à primeira variável.

$\nabla_1 d(\mathbf{x}, \mathbf{y})$	Gradiente de $d(\cdot, \mathbf{y})$ com relação à primeira variável, avaliado em $\mathbf{x}$.
$\mathcal{F}(\mathcal{S})$	Família dos pares proximais (d, D) associados ao conjunto $\mathcal{S}$.
$\bar{f}$	Bifunção regularizada.
λ	Parâmetro positivo de regularização proximal.
ε	Parâmetro positivo de penalização.
$\{\mathbf{x}^k\}$	Sequência de iterados gerada pelo método proximal.
$\{\lambda_k\}$	Sequência de parâmetros positivos de regularização proximal.
$\{\varepsilon_k\}$	Sequência de parâmetros positivos de penalização.
f_k	Bifunção regularizada associada à k -ésima iteração do método proximal.
$\text{dist}(\mathbf{y}, \mathcal{S}(f, K))$	Distância do ponto $\mathbf{y}$ ao conjunto $\mathcal{S}(f, K)$.
$O(k^{-\alpha})$	Notação de Landau utilizada para representar uma possível taxa de convergência sublinear.
$\mathbf{x}(t)$	Trajetória contínua considerada na discussão de possíveis formulações dinâmicas do método.

Sumário

Resumo	v
Abstract	vi
Introdução	1
1 Noções Preliminares	4
1.1 Conceitos e resultados de Análise no $\mathbb{R}^n$	4
1.2 Conceitos e resultados de Análise Convexa	8
1.2.1 Conjuntos convexos	8
1.2.2 Funções convexas	10
1.2.3 Função indicadora, cone normal e função suporte	16
1.2.4 Função conjugada e desigualdade de Fenchel	18
2 Problemas de Equilíbrio	28
2.1 Problemas de Equilíbrio	28
2.2 Casos Particulares de Problemas de Equilíbrio	30
2.2.1 Problemas de Otimização Convexa	30
2.2.2 Problemas de Ponto Fixo	33
2.2.3 Problemas de Desigualdade Variacional	34
2.2.4 Equações Generalizadas	39
2.3 Monotonicidade de Bifunções	42
2.4 Resultados de Existência	44
2.5 Problemas de Equilíbrio de Dois Níveis	46
2.5.1 Definição do Problema	47
2.5.2 Interpretação Hierárquica	49

2.5.3	Problemas de Otimização de Dois Níveis como Casos Particulares .	50
2.5.4	Existência de Soluções	50
2.5.5	Relação com o Método Proximal	51
2.6	Uma hipótese geométrica de regularidade	52
3	Distâncias Proximais Generalizadas	54
3.1	Distâncias Proximais	55
3.2	Distâncias Proximais Induzidas	55
3.2.1	Propriedades Fundamentais dos Pares Proximais	63
4	Bifunções Regularizadas	65
4.1	Hipóteses para bifunções de equilíbrio	65
4.2	Bifunções regularizadas via distâncias proximais generalizadas.	68
5	Método Proximal para Problemas de Equilíbrio de Dois Níveis	74
5.1	Distância proximal generalizada com método de regularização para resolver PEBs	74
5.1.1	Algoritmo de ponto proximal	75
5.2	Convergência para um equilíbrio de dois níveis	85
5.2.1	Exemplos	87
6	Conclusões e Perspectivas Futuras	95
6.1	Método proximal inexato	96
6.2	Taxas de convergência	97
6.3	Formulações dinâmicas	98
6.4	Extensões para espaços de Hilbert e Banach	99
6.5	Considerações finais	99

Introdução

Problemas de equilíbrio constituem uma classe ampla de problemas matemáticos que permite unificar, em uma mesma formulação, diversos modelos clássicos da otimização e da análise não linear. Embora ideias relacionadas a problemas de equilíbrio já apareçam em trabalhos anteriores, como em Nikaido e Isoda [47], a formulação abstrata foi fortemente difundida a partir dos trabalhos de Ky Fan [23] e, posteriormente, de Blum e Oettli [11], que sistematizaram o problema de equilíbrio como uma estrutura unificadora. De modo geral, dado um conjunto não vazio $K \subset \mathbb{R}^n$ e uma bifunção $f : K \times K \rightarrow \mathbb{R}$, o problema de equilíbrio associado a f e K consiste em encontrar $x^* \in K$ tal que

$$f(x^*, y) \geq 0, \quad \forall y \in K.$$

O conjunto solução desse problema será denotado por

$$S(f, K) = \{x \in K; f(x, y) \geq 0, \quad \forall y \in K\}.$$

Essa formulação é importante porque inclui, como casos particulares, problemas de otimização convexa, desigualdades variacionais, problemas de ponto fixo, complementaridade e equilíbrio de Nash em jogos não cooperativos.

Nesta dissertação, o interesse principal recai sobre uma classe hierárquica desses problemas, conhecida como problema de equilíbrio de dois níveis. Dados um conjunto $K \subset \mathbb{R}^n$ e duas bifunções $f, h : K \times K \rightarrow \mathbb{R}$, o problema de equilíbrio de dois níveis consiste em encontrar $\bar{x} \in S(f, K)$ tal que

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Isto é, buscamos uma solução de um problema de equilíbrio associado à bifunção h , mas restrita ao conjunto solução do problema de equilíbrio associado à bifunção f . Assim, o problema pode ser escrito como

$$\text{encontrar } \bar{x} \in S(f, K) \text{ tal que } h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Esse tipo de formulação é particularmente relevante porque modela situações hierárquicas, nas quais uma decisão ou condição de equilíbrio de nível superior deve ser satisfeita sobre o conjunto de soluções de um problema de nível inferior. Por essa razão, problemas de equilíbrio de dois níveis estão relacionados a programas matemáticos com restrições de equilíbrio, problemas de otimização de dois níveis, modelos do tipo líder-seguidor e processos de seleção entre múltiplas soluções de um problema de equilíbrio.

O objetivo central deste trabalho é estudar o método proximal com distâncias proximais generalizadas para resolver problemas de equilíbrio de dois níveis, conforme desenvolvido em [10]. A ideia fundamental desse método consiste em substituir o problema original por uma sequência de subproblemas regularizados. Em cada iteração, considera-se uma bifunção da forma

$$f_k(x, y) = \varepsilon_k f(x, y) + h(x, y) + \frac{1}{\lambda_k} \langle \nabla_1 d(x, x^k), y - x \rangle,$$

onde d é uma distância proximal generalizada, $\{\lambda_k\}$ é uma sequência de parâmetros positivos e $\{\varepsilon_k\}$ é uma sequência de penalização.

A presença do termo proximal tem a função de regularizar o problema, estabilizando o processo iterativo e permitindo obter propriedades de existência, unicidade dos subproblemas regularizados e convergência da sequência gerada. Um aspecto importante da abordagem estudada é que a distância utilizada não precisa ser a distância euclidiana usual. Em vez disso, trabalha-se com distâncias proximais generalizadas, introduzidas em contextos de otimização por Auslender e Teboulle [7] e posteriormente desenvolvidas por outros autores, como Burachik e Dutta [14]. Essa generalidade permite incorporar informações geométricas do conjunto viável e pode tornar o método mais adequado a determinadas classes de problemas.

O trabalho está organizado da seguinte maneira. No Capítulo 1, apresentamos noções preliminares de análise em $\mathbb{R}^n$ e de análise convexa. São introduzidos conceitos como produto interno, norma, sequências, conjuntos convexos, funções convexas, funções próprias, semicontinuidade inferior, subdiferenciais, função indicadora, cone normal, função suporte, função conjugada e desigualdade de Fenchel. Esses resultados serão utilizados ao longo de toda a dissertação, especialmente nas demonstrações relacionadas à existência de soluções e à convergência dos algoritmos.

No Capítulo 2, estudamos problemas de equilíbrio. Inicialmente, apresentamos sua definição e algumas hipóteses clássicas sobre bifunções. Em seguida, discutimos casos

particulares importantes, como problemas de otimização convexa, problemas de ponto fixo e desigualdades variacionais. Também são apresentadas noções de monotonicidade para bifunções, resultados de existência de soluções e a formulação dos problemas de equilíbrio de dois níveis.

No Capítulo 3, introduzimos as distâncias proximais generalizadas. Apresentamos a definição de distância proximal associada a um conjunto S , a noção de distância proximal induzida e as propriedades fundamentais dos pares proximais. Também discutimos exemplos importantes, incluindo a distância euclidiana quadrática e distâncias do tipo Bregman. Esse capítulo fornece a estrutura geométrica necessária para a construção dos subproblemas regularizados.

No Capítulo 4, estudamos bifunções regularizadas por meio de distâncias proximais generalizadas. São apresentadas hipóteses sobre as bifunções f e h , bem como condições que garantem que as bifunções regularizadas preservam propriedades fundamentais, como convexidade, semicontinuidade e monotonicidade. Também são discutidas condições de coercividade assintótica que asseguram a existência de soluções para os problemas regularizados.

No Capítulo 5, apresentamos o método proximal para problemas de equilíbrio de dois níveis. Estudamos o algoritmo proposto em [10], sua boa definição e a análise de convergência da sequência gerada. O principal resultado mostra que, sob hipóteses adequadas, toda sequência produzida pelo algoritmo converge para uma solução do problema de equilíbrio de dois níveis. A demonstração utiliza estimativas associadas à distância proximal induzida, uma condição geométrica envolvendo função suporte e conjugada de Fenchel, e argumentos de convergência quase-Fejér.

Por fim, no Capítulo 6, apresentamos as conclusões do trabalho e algumas perspectivas futuras. Entre as direções apontadas, destacam-se o estudo de versões inexatas do método proximal, a obtenção de taxas explícitas de convergência e formulações dinâmicas.

Capítulo 1

Noções Preliminares

Neste capítulo apresentamos conceitos fundamentais de Análise Convexa e Topologia em $\mathbb{R}^n$ que servirão de base para o desenvolvimento dos resultados posteriores. Em particular, abordaremos propriedades de sequências, conjuntos convexos, funções convexas e ferramentas essenciais como subdiferenciais e conjugação de Fenchel.

1.1 Conceitos e resultados de Análise no $\mathbb{R}^n$

Nesta seção, adotamos como referências principais os trabalhos de Lima [34, 35]. Inicialmente, recordamos alguns conceitos fundamentais relacionados ao produto interno e a sequências em $\mathbb{R}^n$, os quais serão amplamente utilizados nas análises de convergência dos algoritmos estudados neste trabalho.

Definição 1.1.1. *Um produto interno no espaço euclidiano $\mathbb{R}^n$ é uma aplicação que associa a cada par de vetores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ um número real, denotado por $\langle \mathbf{x}, \mathbf{y} \rangle$, tal que, para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, são satisfeitas as seguintes propriedades:*

$$PI.1 \quad \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle;$$

$$PI.2 \quad \langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle;$$

$$PI.3 \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha \mathbf{y} \rangle;$$

$$PI.4 \quad \text{se } \mathbf{x} \neq 0, \text{ então } \langle \mathbf{x}, \mathbf{x} \rangle > 0.$$

Em outras palavras, um produto interno é uma função real definida em $\mathbb{R}^n \times \mathbb{R}^n$ que é simétrica, bilinear e definida positiva. Segundo Lima [35], o exemplo mais relevante é o

produto interno usual em $\mathbb{R}^n$, dado por

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n,$$

onde $\mathbf{x} = (x_1, \dots, x_n)$ e $\mathbf{y} = (y_1, \dots, y_n)$. Este será o produto interno adotado ao longo deste trabalho.

A partir do produto interno, define-se a norma associada. Para $\mathbf{x} \in \mathbb{R}^n$, escrevemos

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle},$$

isto é,

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

O valor $\|\mathbf{x}\|$ é denominado norma euclidiana ou, geometricamente, o comprimento do vetor $\mathbf{x} \in \mathbb{R}^n$.

A seguir, apresentamos a desigualdade de Cauchy-Schwarz, resultado fundamental que estabelece uma relação entre o produto interno e as normas dos vetores envolvidos e que será utilizado frequentemente ao longo desta dissertação.

Teorema 1 (Desigualdade de Cauchy-Schwarz). *Para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, tem-se $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$. Vale a igualdade se, e somente se, um dos vetores $\mathbf{x}, \mathbf{y}$ é um múltiplo escalar do outro.*

Demonstração. Veja Teorema 1.3, página 3 [34]. □

A norma euclidiana $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ satisfaz as seguintes propriedades, para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, onde $|\alpha|$ denota o valor absoluto de α :

N1. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|;$

N2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|;$

N3. $\mathbf{x} \neq 0 \Rightarrow \|\mathbf{x}\| > 0.$

De maneira geral, uma norma é qualquer aplicação $\|\cdot\| : E \rightarrow \mathbb{R}$ em um espaço vetorial E que satisfaça as propriedades N1, N2 e N3 acima.

Toda norma em um espaço vetorial E induz naturalmente uma noção de distância. Dados $\mathbf{x}, \mathbf{y} \in E$, define-se a distância entre $\mathbf{x}$ e $\mathbf{y}$ por

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

As propriedades da norma implicam imediatamente que essa distância satisfaz, para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$, as seguintes condições:

$$D1. \quad d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z});$$

$$D2. \quad d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x});$$

$$D3. \quad \mathbf{x} \neq \mathbf{y} \Rightarrow d(\mathbf{x}, \mathbf{y}) > 0.$$

Observa-se que uma norma arbitrária em um espaço vetorial E não necessariamente é induzida por um produto interno, isto é, pode não existir um produto interno $\langle \cdot, \cdot \rangle$ tal que $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$ para todo $\mathbf{x} \in E$.

Quando uma norma é proveniente de um produto interno, ela satisfaz a identidade do paralelogramo:

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

Exemplo 1.1.1. Considere a norma $\|\cdot\|_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$\|(\mathbf{x}_1, \dots, \mathbf{x}_n)\|_1 = |\mathbf{x}_1| + \dots + |\mathbf{x}_n| \quad \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^n.$$

Esta é chamada por Lima em [35] de norma da soma. Podemos verificar facilmente que a norma da soma satisfaz (N1), (N2) e (N3). De fato, sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_1 &= \|(\mathbf{x}_1, \dots, \mathbf{x}_n) + (\mathbf{y}_1, \dots, \mathbf{y}_n)\|_1 \\ &= \|(\mathbf{x}_1 + \mathbf{y}_1, \dots, \mathbf{x}_n + \mathbf{y}_n)\|_1 \\ &= |\mathbf{x}_1 + \mathbf{y}_1| + \dots + |\mathbf{x}_n + \mathbf{y}_n| \\ &\leq (|\mathbf{x}_1| + |\mathbf{y}_1|) + \dots + (|\mathbf{x}_n| + |\mathbf{y}_n|) \\ &= (|\mathbf{x}_1| + \dots + |\mathbf{x}_n|) + (|\mathbf{y}_1| + \dots + |\mathbf{y}_n|) \\ &= \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1, \end{aligned}$$

satisfazendo (N1). Para $\alpha \in \mathbb{R}$, temos

$$\begin{aligned} \|\alpha \mathbf{x}\|_1 &= \|\alpha(\mathbf{x}_1, \dots, \mathbf{x}_n)\|_1 = \|(\alpha \mathbf{x}_1, \dots, \alpha \mathbf{x}_n)\|_1 = |\alpha \mathbf{x}_1| + \dots + |\alpha \mathbf{x}_n| \\ &= |\alpha| |\mathbf{x}_1| + \dots + |\alpha| |\mathbf{x}_n| = |\alpha| (|\mathbf{x}_1| + \dots + |\mathbf{x}_n|) = |\alpha| \|\mathbf{x}\|_1, \end{aligned}$$

satisfazendo (N2). Por último, temos

$$\|\mathbf{x}\|_1 = \|(\mathbf{x}_1, \dots, \mathbf{x}_n)\|_1 = |\mathbf{x}_1| + \dots + |\mathbf{x}_n|,$$

se $\mathbf{x} \neq 0$, então pelo menos um $x_i \neq 0$, $i = 1, \dots, n$, de toda forma teremos pela definição de valor absoluto [33] que $|x_1| + \dots + |x_n| > 0$, satisfazendo (N3).

Para $n \geq 2$, tomando $\mathbf{x} = (1, 0, \dots, 0, 0)$ e $\mathbf{y} = (0, 0, \dots, 0, 1)$ em $\mathbb{R}^n$, temos

$$\|\mathbf{x} + \mathbf{y}\|_1^2 + \|\mathbf{x} - \mathbf{y}\|_1^2 = 4 + 4 = 8,$$

por outro lado,

$$2(\|\mathbf{x}\|_1^2 + \|\mathbf{y}\|_1^2) = 2(1 + 1) = 4.$$

Como $8 \neq 4$, a lei do paralelogramo não vale. Portanto, a norma $\|\cdot\|_1$ não pode ser induzida por nenhum produto interno.

Uma norma em $\mathbb{R}^n$ também permite introduzir noções geométricas fundamentais. Uma bola aberta de centro $\mathbf{a} \in \mathbb{R}^n$ e raio $r > 0$ é definida como o conjunto dos pontos cuja distância a $\mathbf{a}$ é menor que r , isto é,

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x} - \mathbf{a}\| < r\}.$$

De forma análoga, definem-se a bola fechada e a esfera de centro $\mathbf{a}$ e raio r por

$$B_r[\mathbf{a}] = \{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x} - \mathbf{a}\| \leq r\} \quad \text{e} \quad S_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x} - \mathbf{a}\| = r\}.$$

Definição 1.1.2. Uma sequência em $\mathbb{R}^n$ é uma aplicação $\mathbf{x} : \mathbb{N} \rightarrow \mathbb{R}^n$, definida no conjunto dos números naturais. O valor dessa aplicação no número k é denotado por $\mathbf{x}^k$ e recebe o nome de k -ésimo termo da sequência. Utilizaremos a notação $\{\mathbf{x}^k\}$ para representar a sequência cujo termo geral é $\mathbf{x}^k \in \mathbb{R}^n$.

Uma subsequência de $\{\mathbf{x}^k\}$ é obtida ao restringirmos a sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < k_2 < \dots < k_j < \dots\}$ de $\mathbb{N}$. Tal subsequência é indicada por $\{\mathbf{x}^{k_j}\}$.

Dizemos que uma sequência $\{\mathbf{x}^k\}$ é limitada quando existe uma constante $c > 0$ tal que $\|\mathbf{x}^k\| \leq c$ para todo $k \in \mathbb{N}$. Um ponto $\mathbf{a} \in \mathbb{R}^n$ é dito limite da sequência $\{\mathbf{x}^k\}$ se, para todo $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que, para todo $k > k_0$, vale $\|\mathbf{x}^k - \mathbf{a}\| < \varepsilon$. Nesse caso, dizemos que $\{\mathbf{x}^k\}$ converge para $\mathbf{a}$, e escrevemos $\lim_{k \rightarrow +\infty} \mathbf{x}^k = \mathbf{a}$ ou $\mathbf{x}^k \rightarrow \mathbf{a}$. Caso contrário, a sequência é dita divergente.

Diz-se que um ponto $\mathbf{a} \in \mathbb{R}^n$ é um valor de aderência, ou ponto de acumulação, da sequência $\{\mathbf{x}^k\}$ se existe uma subsequência de $\{\mathbf{x}^k\}$ que converge para $\mathbf{a}$.

Exemplo 1.1.2. A sequência $x^k = (-1)^k + \frac{1}{k+1}$ possui dois pontos de acumulação e por isso não converge.

De fato, temos $x^{2i} \rightarrow 1$ e $x^{2i+1} \rightarrow -1$.

1.2 Conceitos e resultados de Análise Convexa

Nesta seção, adotamos como referências principais os trabalhos de Bauschke & Combettes [9], Izmailov & Solodov [29, 30], Mordukhovich [41] e Rockafellar [51].

1.2.1 Conjuntos convexos

Um conjunto convexo se caracteriza por conter todos os segmentos cujos extremos pertencem ao conjunto.

Definição 1.2.1. Um conjunto $C \subset \mathbb{R}^n$ é convexo se, para quaisquer $x, y \in C$ e $\alpha \in [0, 1]$,

$$\alpha x + (1 - \alpha)y \in C.$$

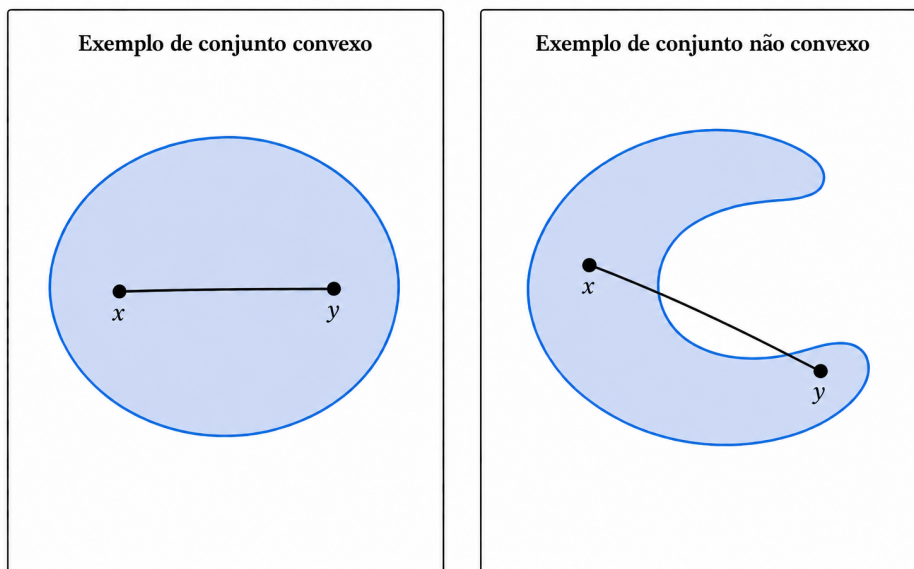


Figura 1.1: Representação geométrica de conjuntos convexo e não convexo.

Geometricamente, isso significa que o segmento de reta que liga quaisquer dois pontos do conjunto está inteiramente contido nele.

O ponto $\alpha x + (1 - \alpha)y$, onde $\alpha \in [0, 1]$, chama-se *combinação convexa* de x e y (com parâmetro α).

Exemplo 1.2.1. Os conjuntos $\mathbb{R}^n$, bolas euclidianas e semiespaços são convexos.

A próxima proposição apresenta uma importante propriedade de estabilidade dos conjuntos convexos, mostrando que a convexidade é preservada sob interseções arbitrárias.

Proposição 1.2.1. Sejam $C_i \subset \mathbb{R}^n, i \in I$, conjuntos convexos, onde I é um conjunto qualquer. Então a interseção $C = \bigcap_{i \in I} C_i$ também é um conjunto convexo.

Demonstração. Veja Proposição 3.2.1, página 70 [30]. □

Ao longo deste trabalho, será necessário considerar o interior de conjuntos convexos em diferentes momentos. Por essa razão, é importante destacar que essa operação não compromete a convexidade do conjunto. O resultado a seguir mostra que a convexidade de um conjunto é preservada quando se considera o seu interior.

Proposição 1.2.2. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Então o interior de C , denotado por $\text{int}(C)$, é um conjunto convexo.

Demonstração. Veja Proposição 3.2.4, página 71 [30]. □

Para compreender melhor a estrutura dos conjuntos convexos, é conveniente introduzir o conceito de combinação convexa. De modo intuitivo, esse tipo de combinação permite construir novos pontos a partir de pontos dados, atribuindo-lhes pesos não negativos cuja soma é igual a um. Esse conceito será utilizado frequentemente ao longo deste trabalho.

Definição 1.2.2. A combinação linear $\alpha_1 v_1 + \cdots + \alpha_k v_k$ chama-se uma combinação convexa de $v_1, \dots, v_k \in \mathbb{R}^n$ quando $\alpha_1 + \cdots + \alpha_k = 1$ e $\alpha_i \geq 0$ para $i = 1, \dots, k$.

Teorema 2. Um conjunto $C \subset \mathbb{R}^n$ é convexo se, e somente se, para quaisquer $k \in \mathbb{N}$, $x^i \in C$ e $\alpha_i \in [0, 1], i = 1, \dots, k$, tais que $\sum_{i=1}^k \alpha_i = 1$, a combinação convexa $\sum_{i=1}^k \alpha_i x^i$ pertence a C .

Demonstração. Veja Teorema 3.2.8, página 73 [30]. □

1.2.2 Funções convexas

A convexidade de funções desempenha papel fundamental na teoria de otimização, especialmente na análise de estabilidade e convergência.

Definição 1.2.3. Se $C \subset \mathbb{R}^n$ é um conjunto convexo, diz-se que uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é convexa em C quando para quaisquer $x, y \in C$ e $\alpha \in [0, 1]$, tem-se que

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

A função f diz-se estritamente convexa quando a desigualdade acima é estrita para todos $x \neq y$ e $\alpha \in (0, 1)$.

A função f diz-se fortemente convexa com módulo $\gamma > 0$, quando, para quaisquer $x, y \in C$ e $\alpha \in [0, 1]$, tem-se

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

Exemplo 1.2.2. A função $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, é convexa, mas não é estritamente convexa.

Sejam $x, y \in \mathbb{R}$ e $\alpha \in [0, 1]$. Então,

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

Por outro lado,

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

Logo,

$$f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y),$$

o que mostra que f é convexa.

Entretanto, para quaisquer $x \neq y$ e $\alpha \in (0, 1)$, a igualdade acima continua válida. Portanto, a desigualdade estrita exigida na definição de convexidade estrita não é satisfeita. Conclui-se que f não é estritamente convexa.

Exemplo 1.2.3. A função $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$, é uma função fortemente convexa com módulo $\gamma = 1$.

De fato, para quaisquer $x, y \in \mathbb{R}$ e $\alpha \in [0, 1]$,

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &= (\alpha x + (1 - \alpha)y)^2 \\ &= \alpha^2 x^2 + (1 - \alpha)^2 y^2 + 2\alpha(1 - \alpha)xy. \end{aligned}$$

Por outro lado,

$$\alpha f(x) + (1 - \alpha)f(y) = \alpha x^2 + (1 - \alpha)y^2.$$

Consequentemente,

$$\alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y) = \alpha(1 - \alpha)(x - y)^2.$$

Logo,

$$f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y) - \alpha(1 - \alpha)\|x - y\|^2,$$

isto é, f é fortemente convexa com módulo $\gamma = 1$.

Exemplo 1.2.4. A função $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = e^x$, é estritamente, mas não fortemente convexa.

De fato, para $x, y \in \mathbb{R}, x \neq y$, e $\alpha \in (0, 1)$,

$$f(\alpha x + (1 - \alpha)y) = e^{\alpha x + (1 - \alpha)y} = (e^x)^\alpha (e^y)^{1 - \alpha}.$$

Pela desigualdade entre as médias geométrica e aritmética ponderadas,

$$(e^x)^\alpha (e^y)^{1 - \alpha} < \alpha e^x + (1 - \alpha)e^y,$$

pois $e^x \neq e^y$. Logo,

$$f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y),$$

mostrando que f é estritamente convexa.

Por outro lado, suponha que f seja fortemente convexa com módulo $\gamma > 0$. Tomando $\alpha = \frac{1}{2}$, $y = x + h$ e dividindo a desigualdade de convexidade forte por e^x , obtemos

$$e^{h/2} \leq \frac{1 + e^h}{2} - \frac{\gamma}{4} e^{-x} h^2.$$

Fazendo $x \rightarrow -\infty$, com $h \neq 0$, então $e^{-x} \rightarrow +\infty$, e o lado direito tende a $-\infty$, absurdo pois o lado esquerdo é positivo. Assim, nenhum $\gamma > 0$ pode satisfazer a definição de convexidade forte para todos os $x, y \in \mathbb{R}$. Portanto, f não é fortemente convexa.

Definição 1.2.4. O epígrafo da função $f : C \rightarrow \mathbb{R}$ é o conjunto

$$E_f = \{(x, c) \in C \times \mathbb{R}; f(x) \leq c\}.$$

Exemplo 1.2.5. Considere a função $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = x^2$. O epígrafo de f é $E_f = \{(x, c) \in \mathbb{R}^2 : f(x) \leq c\}$. Substituindo $f(x) = x^2$, obtemos $E_f = \{(x, c) \in \mathbb{R}^2 : x^2 \leq c\}$.

Geometricamente, E_f corresponde à região do plano situada sobre a parábola $c = x^2$, incluindo a própria parábola.

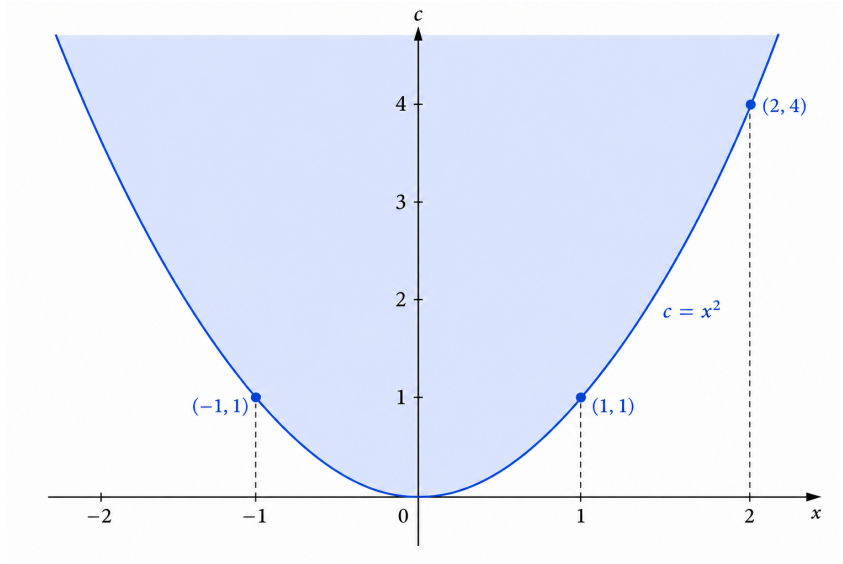


Figura 1.2: Epígrafo da função $f(x) = x^2$.

A relação entre convexidade de conjuntos e de funções é dada pelo seguinte teorema.

Teorema 3. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : C \rightarrow \mathbb{R}$ é convexa em C se, e somente se, o epígrafo de f é um conjunto convexo em $\mathbb{R}^n \times \mathbb{R}$.

Demonstração. Veja Teorema 3.1.15, página 67 [30]. □

A inclusão do valor $+\infty$ permite tratar naturalmente problemas com restrições, como será visto adiante.

Definição 1.2.5. O domínio de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é o conjunto dos pontos em que f assume valores finitos, isto é,

$$\text{dom}(f) = \{x \in \mathbb{R}^n; f(x) < +\infty\}.$$

Diz-se que f é uma função própria se seu domínio é não vazio, ou seja, $\text{dom}(f) \neq \emptyset$. Equivalentemente, f é própria se existe $x \in \mathbb{R}^n$ tal que $f(x) < +\infty$.

Exemplo 1.2.6. Considere a função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f(\mathbf{x}) = \|\mathbf{x}\|^2$. Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $\alpha \in [0, 1]$. Então

$$\begin{aligned} f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) &= \|\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}\|^2 \\ &= \alpha^2\|\mathbf{x}\|^2 + (1 - \alpha)^2\|\mathbf{y}\|^2 + 2\alpha(1 - \alpha)\langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

Pela desigualdade de Cauchy-Schwarz, $\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \|\mathbf{y}\|$ e segue que

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha\|\mathbf{x}\|^2 + (1 - \alpha)\|\mathbf{y}\|^2 = \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}).$$

Portanto, f é convexa. Além disso, $\text{dom}(f) = \{\mathbf{x} \in \mathbb{R}^n; f(\mathbf{x}) < +\infty\} = \mathbb{R}^n$, de modo que $\text{dom}(f) \neq \emptyset$. Logo, f é própria.

A semicontinuidade é uma propriedade fundamental para garantir a existência de soluções em problemas de otimização.

Definição 1.2.6. Dizemos que a função $f : C \rightarrow \mathbb{R}$ é *semicontínua inferiormente* no ponto $\mathbf{x} \in C \subset \mathbb{R}^n$, quando para qualquer sequência $\{\mathbf{x}^k\} \subset C$ tal que $\{\mathbf{x}^k\} \rightarrow \mathbf{x}$ ($k \rightarrow +\infty$), tem-se

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

Dizemos que f é *semicontínua superiormente* quando, nas mesmas condições,

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

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

Exemplo 1.2.7. Sejam $C = [\frac{1}{2}, 2] \subset \mathbb{R}$ e $f : C \rightarrow \mathbb{R}$ dada por

$$f(\mathbf{x}) = \begin{cases} \frac{1}{\mathbf{x}}, & \mathbf{x} \in [\frac{1}{2}, 1), \\ 2, & \mathbf{x} \in [1, 2], \end{cases}$$

ver Figura 1.3. O conjunto C é compacto. A função f não é *semicontínua inferiormente* em C , pois ela não é *semicontínua inferiormente* no ponto $\mathbf{x} = 1 \in C$. Tem-se que $\inf_{z \in C} f(z) = 1$, mas não existe $\mathbf{x} \in C$ tal que $f(\mathbf{x}) = 1$, isto é, não existe minimizador global.

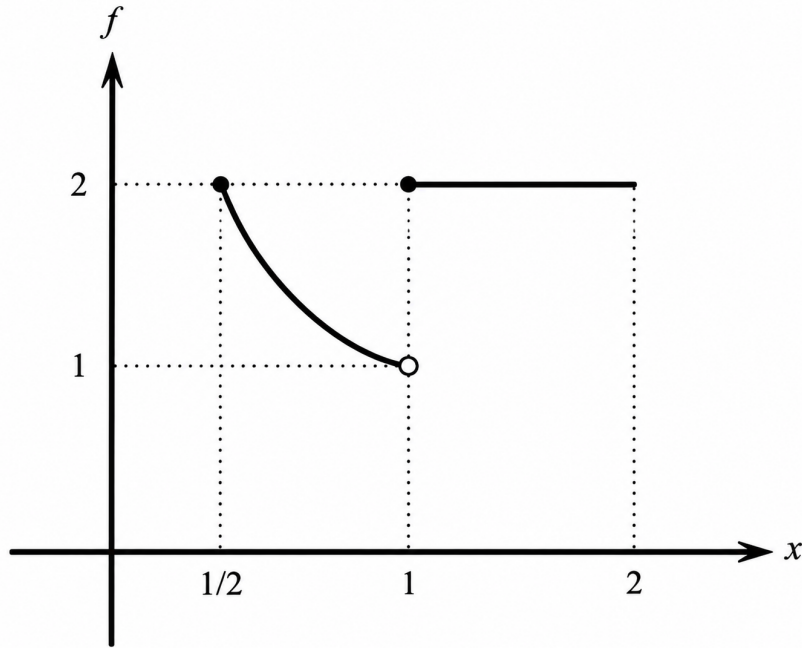


Figura 1.3: A função f definida no Exemplo 1.2.7 é semicontínua superiormente, mas não é semicontínua inferiormente em C .

Teorema 4 (Caracterização de funções convexas diferenciáveis). *Sejam $C \subset \mathbb{R}^n$ um conjunto convexo e aberto e $f : C \rightarrow \mathbb{R}$ uma função diferenciável em C . Então as propriedades seguintes são equivalentes:*

(a) *A função f é convexa em C .*

(b) *Para todo $x \in C$ e todo $y \in C$,*

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

(c) *Para todo $x \in C$ e todo $y \in C$,*

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

Demonstração. Ver Teorema 3.4.30, página 138 [30].

□

O item (b) do Teorema 4 pode ser interpretado como a propriedade de que o gráfico de f está acima de seus hiperplanos tangentes. O item (c) diz que o gradiente de uma função convexa é monótono, e será importante em demonstrações futuras.

Em muitos problemas de otimização, as funções envolvidas não são diferenciáveis. Nesse contexto, o conceito de subgradiente generaliza a noção de gradiente.

Definição 1.2.7. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa. Dizemos que $\mathbf{y} \in \mathbb{R}^n$ é um subgradiente de f no ponto $\mathbf{x} \in \mathbb{R}^n$ se*

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

O conjunto de todos os subgradientes de f em $\mathbf{x}$ se chama o subdiferencial de f em $\mathbf{x}$, denotado por $\partial f(\mathbf{x})$.

O subdiferencial de uma função convexa desempenha um papel importante no estudo de problemas de otimização e será utilizado posteriormente na verificação de propriedades fundamentais das bifunções consideradas neste trabalho. O próximo resultado apresenta algumas propriedades essenciais desse conjunto.

Teorema 5. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. Então para todo $\mathbf{x} \in \mathbb{R}^n$, o conjunto $\partial f(\mathbf{x})$ é convexo, compacto e não-vazio.*

Demonstração. Ver Teorema 3.4.52, página 154 [30]. □

Proposição 1.2.3. *Uma função convexa $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável em $\mathbf{x} \in \mathbb{R}^n$ se, e somente se, o conjunto $\partial f(\mathbf{x})$ contém um elemento só. Neste caso,*

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

Demonstração. Ver Proposição 3.4.53, página 156 [30]. □

Além disso, o subdiferencial fornece uma caracterização dos pontos de mínimo.

Proposição 1.2.4. *Se f é convexa, então $\mathbf{x}$ é minimizador de f se, e somente se,*

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

Demonstração. Ver Teorema 3.4.54, página 157 [30]. □

No estudo da análise convexa, alguns conjuntos possuem uma estrutura que permanece inalterada quando seus elementos são multiplicados por escalares não negativos. Esse tipo de conjunto, denominado cone, será importante na definição e na análise de diversos objetos utilizados ao longo deste trabalho.

Definição 1.2.8. *Um conjunto $C \subset \mathbb{R}^n$ chama-se cone quando ele contém todos os múltiplos não-negativos de seus elementos:*

$$\mathbf{d} \in C \Rightarrow t\mathbf{d} \in C \quad \forall t \in \mathbb{R}_+.$$

Por definição, se C é um cone não-vazio, então necessariamente $0 \in C$. Informalmente, um cone é um conjunto de direções. Com exceção do cone trivial $\{0\}$, um cone não-vazio sempre é ilimitado.

A partir do conceito de combinação convexa, podemos construir um conjunto convexo mesmo quando o conjunto considerado inicialmente não possui essa propriedade. Essa construção conduz à noção de envoltória convexa, que reúne todas as combinações convexas finitas de seus elementos e corresponde ao menor conjunto convexo que o contém.

Definição 1.2.9. *Seja $C \subset \mathbb{R}^n$. A envoltória convexa de C , denotada por $\text{conv}(C)$, é o conjunto de todas as combinações convexas finitas de elementos de C , isto é,*

$$\text{conv}(C) = \left\{ \sum_{i=1}^m \lambda_i x_i; x_i \in C, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1, m \in \mathbb{N} \right\}.$$

Equivalentemente, $\text{conv}(C)$ é o menor conjunto convexo que contém C .

Definição 1.2.10. *Sejam C e K conjuntos convexas com $C \subset K$. Definimos o $\text{core}_K C$, o core de C relativo a K , por*

$$\text{core}_K C = \{x \in C : C \cap (x, y) \neq \emptyset, \forall y \in K \setminus C\},$$

onde $(x, y) = \{(1-t)x + ty; t \in (0, 1)\}$.

As noções de envoltória convexa e de core relativo permitem descrever, respectivamente, o menor conjunto convexo que contém um conjunto dado e os pontos que ocupam uma posição interior em relação a um conjunto convexo maior. Esses conceitos desempenham um papel importante na formulação e no estudo dos problemas de equilíbrio, especialmente na abordagem desenvolvida por [11], e serão utilizados posteriormente na apresentação de resultados fundamentais dessa teoria.

1.2.3 Função indicadora, cone normal e função suporte

Seja $\Omega \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo. Esses objetos desempenham papel fundamental na modelagem de restrições.

Definição 1.2.11. *A função indicadora de Ω é definida por*

$$\delta_{\Omega}(x) = \begin{cases} 0, & x \in \Omega, \\ +\infty, & x \notin \Omega. \end{cases}$$

Essa função permite incorporar restrições em problemas de otimização de forma conveniente.

Definição 1.2.12. *O cone normal a Ω em $\mathbf{x} \in \Omega$ é definido por*

$$\mathcal{N}_\Omega(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n; \langle \mathbf{v}, \mathbf{y} - \mathbf{x} \rangle \leq 0, \forall \mathbf{y} \in \Omega\}.$$

O cone normal pode ser interpretado como o conjunto de direções que formam ângulo obtuso com todas as direções viáveis do conjunto [Ver [30, p. 65]].

Proposição 1.2.5. *Se Ω é convexo e fechado, então*

$$\partial\delta_\Omega(\mathbf{x}) = \mathcal{N}_\Omega(\mathbf{x}).$$

Demonstração. Pela definição de subdiferencial, tem-se

$$\partial\delta_\Omega(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n; \delta_\Omega(\mathbf{y}) \geq \delta_\Omega(\mathbf{x}) + \langle \mathbf{v}, \mathbf{y} - \mathbf{x} \rangle, \forall \mathbf{y} \in \mathbb{R}^n\}.$$

Suponha inicialmente que $\mathbf{x} \in \Omega$. Então $\delta_\Omega(\mathbf{x}) = 0$. Assim, $\mathbf{v} \in \partial\delta_\Omega(\mathbf{x})$ se, e somente se,

$$\delta_\Omega(\mathbf{y}) \geq \langle \mathbf{v}, \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{y} \in \mathbb{R}^n.$$

Se $\mathbf{y} \notin \Omega$, então $\delta_\Omega(\mathbf{y}) = +\infty$, de modo que a desigualdade é satisfeita automaticamente.

Portanto, resta considerar $\mathbf{y} \in \Omega$. Nesse caso, $\delta_\Omega(\mathbf{y}) = 0$, e obtemos

$$0 \geq \langle \mathbf{v}, \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{y} \in \Omega.$$

Logo,

$$\langle \mathbf{v}, \mathbf{y} - \mathbf{x} \rangle \leq 0, \quad \forall \mathbf{y} \in \Omega,$$

isto é, $\mathbf{v} \in \mathcal{N}_\Omega(\mathbf{x})$. Portanto,

$$\partial\delta_\Omega(\mathbf{x}) = \mathcal{N}_\Omega(\mathbf{x}).$$

Se $\mathbf{x} \notin \Omega$, então $\mathbf{x} \notin \text{dom}(\delta_\Omega)$, de modo que $\partial\delta_\Omega(\mathbf{x}) = \emptyset$.

Pela convenção usual, também se define

$$\mathcal{N}_\Omega(\mathbf{x}) = \emptyset, \quad \mathbf{x} \notin \Omega.$$

Assim, a igualdade continua válida para todo $\mathbf{x} \in \mathbb{R}^n$. □

Definição 1.2.13. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo. A função suporte de Ω é a função*

$$\sigma_\Omega : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$$

definida por

$$\sigma_\Omega(\mathbf{v}) = \sup_{\mathbf{x} \in \Omega} \langle \mathbf{v}, \mathbf{x} \rangle.$$

Segundo Rockafellar [51], ela mede o maior valor que o funcional linear $\mathbf{x} \mapsto \langle \mathbf{v}, \mathbf{x} \rangle$ assume sobre o conjunto Ω . Fixado $\mathbf{v} \in \mathbb{R}^n$, $\langle \mathbf{v}, \mathbf{x} \rangle = \beta$ representa um hiperplano ortogonal a $\mathbf{v}$. A função suporte fornece o maior valor β tal que o conjunto ainda permanece no semiespaço

$$\{\mathbf{x}; \langle \mathbf{v}, \mathbf{x} \rangle \leq \beta\}.$$

Assim,

$$\Omega \subseteq \{\mathbf{x}; \langle \mathbf{v}, \mathbf{x} \rangle \leq \beta\} \iff \beta \geq \sigma_\Omega(\mathbf{v}).$$

1.2.4 Função conjugada e desigualdade de Fenchel

Os resultados a seguir podem ser encontrados nos trabalhos de Mordukhovich [41] e Rockafellar [51].

Definição 1.2.14 (Conjugada de Fenchel). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa própria. A função $f^* : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por*

$$f^*(\mathbf{v}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{\langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x})\},$$

para todo $\mathbf{v} \in \mathbb{R}^n$, é chamada a conjugada de Fenchel da função f .

Exemplo 1.2.8. *Considere a função convexa $f(\mathbf{x}) = \|\mathbf{x}\|$ para todo $\mathbf{x} \in \mathbb{R}^n$, encontremos sua função conjugada.*

Pela definição de conjugada, para todo $\mathbf{v} \in \mathbb{R}^n$, $f^(\mathbf{v}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{\langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x})\}$. Portanto, $f^*(\mathbf{v}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{\langle \mathbf{v}, \mathbf{x} \rangle - \|\mathbf{x}\|\}$. Pela desigualdade de Cauchy-Schwarz, temos*

$$\langle \mathbf{v}, \mathbf{x} \rangle \leq \|\mathbf{v}\| \|\mathbf{x}\|$$

$$\langle \mathbf{v}, \mathbf{x} \rangle - \|\mathbf{x}\| \leq (\|\mathbf{v}\| - 1) \|\mathbf{x}\|$$

Se $\|\mathbf{v}\| \leq 1$, então

$$(\|\mathbf{v}\| - 1) \|\mathbf{x}\| \leq 0$$

de modo que

$$\langle \mathbf{v}, \mathbf{x} \rangle - \|\mathbf{x}\| \leq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Tomando $\mathbf{x} = 0$, obtemos valor 0. Portanto

$$f^*(\mathbf{v}) = 0.$$

Se $\|\mathbf{v}\| > 1$, existe $\mathbf{u} \neq 0$ tal que

$$\langle \mathbf{v}, \mathbf{u} \rangle > \|\mathbf{u}\|.$$

Tomando $\mathbf{x} = t\mathbf{u}$, com $t > 0$

$$\langle \mathbf{v}, t\mathbf{u} \rangle - \|t\mathbf{u}\| = t(\langle \mathbf{v}, \mathbf{u} \rangle - \|\mathbf{u}\|).$$

Como o termo entre parênteses é positivo, fazendo $t \rightarrow +\infty$

$$\langle \mathbf{v}, t\mathbf{u} \rangle - \|t\mathbf{u}\| \rightarrow +\infty$$

e então $f^*(\mathbf{v}) = +\infty$. Logo,

$$f^*(\mathbf{v}) = \begin{cases} 0, & \text{se } \|\mathbf{v}\| \leq 1, \\ +\infty, & \text{se } \|\mathbf{v}\| > 1. \end{cases}$$

Ou seja, a função conjugada é a função indicadora da bola unitária fechada

$$f^*(\mathbf{v}) = \delta_B(\mathbf{v})$$

onde $B = \{\mathbf{v} \in \mathbb{R}^n; \|\mathbf{v}\| \leq 1\}$.

Considere agora a função suporte de B

$$\sigma_B(\mathbf{v}) = \sup_{\mathbf{x} \in B} \langle \mathbf{v}, \mathbf{x} \rangle = \sup_{\|\mathbf{x}\| \leq 1} \langle \mathbf{v}, \mathbf{x} \rangle.$$

Pela desigualdade de Cauchy-Schwarz,

$$\langle \mathbf{v}, \mathbf{x} \rangle \leq \|\mathbf{v}\| \|\mathbf{x}\|.$$

Como $\|\mathbf{x}\| \leq 1$,

$$\langle \mathbf{v}, \mathbf{x} \rangle \leq \|\mathbf{v}\|.$$

Logo $\sigma_B(\mathbf{v}) \leq \|\mathbf{v}\|$. Por outro lado, se $\mathbf{v} \neq 0$, tome $\mathbf{x} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$. Então $\|\mathbf{x}\| = 1$, isto é, $\mathbf{x} \in B$,

e

$$\langle \mathbf{v}, \mathbf{x} \rangle = \left\langle \mathbf{v}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\rangle = \frac{\|\mathbf{v}\|^2}{\|\mathbf{v}\|} = \|\mathbf{v}\|.$$

Portanto $\sigma_B(v) = \|v\|$. Assim, a norma euclidiana pode ser escrita como função suporte da bola unitária euclidiana:

$$\sigma_B(v) = \|v\|.$$

Considere agora $f(x) = \|x\|$. Então

$$f(v) = \sigma_B(v)$$

e conseqüentemente,

$$f^*(v) = (\sigma_B(v))^* = \delta_B(v).$$

Uma consequência imediata da definição de conjugada de Fenchel é a desigualdade apresentada a seguir, que relaciona uma função à sua conjugada.

Teorema 6 (Desigualdade de Fenchel). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função com $\text{dom } f \neq \emptyset$. Então*

$$\langle v, x \rangle \leq f(x) + f^*(v)$$

para todo $x, v \in \mathbb{R}^n$.

Demonstração. Segue da definição anterior que, para cada $x \in \mathbb{R}^n$ e cada $v \in \mathbb{R}^n$, tem-se

$$f^*(v) = \sup_{x \in \mathbb{R}^n} \{\langle v, x \rangle - f(x)\} \geq \langle v, x \rangle - f(x),$$

isto é,

$$f^*(v) + f(x) \geq \langle v, x \rangle.$$

□

Essa desigualdade será utilizada posteriormente na análise de convergência dos métodos considerados.

O próximo resultado fornece uma caracterização útil dos elementos do subdiferencial em termos da conjugada de Fenchel, estabelecendo uma relação direta entre esses dois conceitos.

Teorema 7. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa própria semicontínua inferiormente, onde f é finita em $x \in \mathbb{R}^n$. Então $v \in \partial f(x)$ se, e somente se, $f^*(v) = \langle v, x \rangle - f(x)$.*

Demonstração. ($\Rightarrow$) Suponha que $v \in \partial f(x)$. Pela definição 1.2.7, para todo $z \in \mathbb{R}^n$, vale

$$\langle v, z - x \rangle \leq f(z) - f(x).$$

Rearranjando os termos, obtemos

$$\langle \mathbf{v}, \mathbf{z} \rangle - f(\mathbf{z}) \leq \langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x}).$$

Tomando o supremo com respeito a $\mathbf{z}$, segue que

$$f^*(\mathbf{v}) = \sup_{\mathbf{z} \in \mathbb{R}^n} \{\langle \mathbf{v}, \mathbf{z} \rangle - f(\mathbf{z})\} \leq \langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x}).$$

Por outro lado, escolhendo $\mathbf{z} = \mathbf{x}$ na definição de f^* , temos

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

Assim, concluímos que

$$f^*(\mathbf{v}) = \langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x}).$$

($\Leftarrow$) Reciprocamente, suponha que

$$f^*(\mathbf{v}) = \sup_{\mathbf{z} \in \mathbb{R}^n} \{\langle \mathbf{v}, \mathbf{z} \rangle - f(\mathbf{z})\} = \langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x}).$$

Da definição de supremo, segue que, para todo $\mathbf{z} \in \mathbb{R}^n$,

$$\langle \mathbf{v}, \mathbf{z} \rangle - f(\mathbf{z}) \leq \langle \mathbf{v}, \mathbf{x} \rangle - f(\mathbf{x}).$$

Reorganizando a desigualdade, obtemos

$$\langle \mathbf{v}, \mathbf{z} - \mathbf{x} \rangle \leq f(\mathbf{z}) - f(\mathbf{x}), \quad \forall \mathbf{z} \in \mathbb{R}^n.$$

Portanto, pela definição de subgradiente, segue que

$$\mathbf{v} \in \partial f(\mathbf{x}).$$

□

Este teorema garante que se $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função convexa e $\mathbf{v} \in \partial f(\mathbf{x})$ para algum $\mathbf{x} \in \mathbb{R}^n$, então $f^*(\mathbf{v}) < +\infty$.

O próximo resultado expressa uma propriedade de inversão de ordem associada à conjugação de Fenchel. Essa propriedade será utilizada posteriormente, no Capítulo 5, na análise do exemplo em que o problema de nível inferior é formulado como um problema de minimização.

Proposição 1.2.6. *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funções convexas próprias. Se $f(\mathbf{x}) \geq g(\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^n$, então $f^*(\mathbf{x}) \leq g^*(\mathbf{x})$.*

Demonstração. Pela definição 1.2.14 de conjugação de Fenchel, $f^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle y, x \rangle - f(x)\}$. Como por hipótese $f(x) \geq g(x)$, então

$$-f(x) \leq -g(x) \Rightarrow \langle y, x \rangle - f(x) \leq \langle y, x \rangle - g(x).$$

Logo

$$f^*(y) \leq g^*(y) \quad \forall y \in \mathbb{R}^n.$$

Como y é arbitrário, então segue o resultado. $\square$

O próximo resultado estabelece que, sob condições adequadas, uma função coincide com sua biconjugada.

Proposição 1.2.7. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa própria e semicontínua inferiormente. Então*

$$f = f^{**}.$$

Demonstração. Mostraremos inicialmente que $f^{**} \leq f$. De fato, por definição,

$$f^{**}(x) = \sup_{v \in \mathbb{R}^n} \{\langle v, x \rangle - f^*(v)\}.$$

Por outro lado, pela definição da conjugada de Fenchel, para todo $v \in \mathbb{R}^n$,

$$f^*(v) = \sup_{z \in \mathbb{R}^n} \{\langle v, z \rangle - f(z)\} \geq \langle v, x \rangle - f(x).$$

Logo,

$$\langle v, x \rangle - f^*(v) \leq f(x), \quad \forall v \in \mathbb{R}^n.$$

Tomando o supremo em $v \in \mathbb{R}^n$, obtemos

$$f^{**}(x) \leq f(x), \quad \forall x \in \mathbb{R}^n.$$

Reciprocamente, mostraremos que $f \leq f^{**}$. Suponha, por absurdo, que isso não ocorra. Então existe $x_0 \in \mathbb{R}^n$ tal que

$$f^{**}(x_0) < f(x_0).$$

Assim, podemos escolher $\alpha \in \mathbb{R}$ satisfazendo

$$f^{**}(x_0) < \alpha < f(x_0).$$

Em particular, $(x_0, \alpha) \notin E_f$.

Como f é convexa, própria e semicontínua inferiormente, seu epígrafo é um conjunto convexo e fechado. Portanto, pelo teorema de separação, existe uma função afim

$$h(x) = \langle v, x \rangle + \beta, \quad v \in \mathbb{R}^n, \beta \in \mathbb{R},$$

tal que

$$h(x) \leq f(x), \quad \forall x \in \mathbb{R}^n,$$

e

$$h(x_0) > \alpha.$$

Como $h \leq f$, temos, para todo $z \in \mathbb{R}^n$,

$$\langle v, z \rangle + \beta \leq f(z).$$

Logo,

$$\langle v, z \rangle - f(z) \leq -\beta, \quad \forall z \in \mathbb{R}^n.$$

Tomando o supremo em $z \in \mathbb{R}^n$, segue que

$$f^*(v) \leq -\beta.$$

Consequentemente,

$$f^{**}(x_0) = \sup_{w \in \mathbb{R}^n} \{\langle w, x_0 \rangle - f^*(w)\} \geq \langle v, x_0 \rangle - f^*(v).$$

Como $f^*(v) \leq -\beta$, obtemos

$$f^{**}(x_0) \geq \langle v, x_0 \rangle + \beta = h(x_0) > \alpha.$$

Portanto,

$$f^{**}(x_0) > \alpha,$$

o que contradiz a escolha de α , pois $\alpha > f^{**}(x_0)$.

Logo,

$$f(x) \leq f^{**}(x), \quad \forall x \in \mathbb{R}^n.$$

Como já mostramos que $f^{**} \leq f$, concluímos que

$$f = f^{**}.$$

□

As proposições a seguir estabelecem relações fundamentais entre a função indicadora de um conjunto e sua função suporte por meio da conjugação de Fenchel. Em particular, sob hipóteses adequadas sobre o conjunto, essas funções são conjugadas uma da outra.

Proposição 1.2.8. *Se Ω é um conjunto não vazio, então*

$$\delta_{\Omega}^* = \sigma_{\Omega}.$$

Demonstração. Pela definição da conjugada de Fenchel da função δ_{Ω} , tem-se

$$\delta_{\Omega}^*(s) = \sup_{x \in \mathbb{R}^n} \{\langle s, x \rangle - \delta_{\Omega}(x)\}.$$

Recordando que

$$\delta_{\Omega}(x) = \begin{cases} 0, & \text{se } x \in \Omega, \\ +\infty, & \text{se } x \notin \Omega, \end{cases}$$

segue que

$$\langle s, x \rangle - \delta_{\Omega}(x) = \begin{cases} \langle s, x \rangle, & \text{se } x \in \Omega, \\ -\infty, & \text{se } x \notin \Omega. \end{cases}$$

Portanto, $\delta_{\Omega}^*(s) = \sup_{x \in \Omega} \langle s, x \rangle$. Como, por definição, a função suporte de Ω é dada por $\sigma_{\Omega}(s) = \sup_{x \in \Omega} \langle s, x \rangle$, concluímos que

$$\delta_{\Omega}^*(s) = \sigma_{\Omega}(s), \quad \forall s \in \mathbb{R}^n.$$

Logo, $\delta_{\Omega}^* = \sigma_{\Omega}$. □

Proposição 1.2.9. *Seja $\Omega \subseteq \mathbb{R}^n$ um conjunto não vazio, convexo e fechado. Então,*

$$\sigma_{\Omega}^* = \delta_{\Omega}.$$

Demonstração. Pela definição da conjugada de Fenchel,

$$\sigma_{\Omega}^*(x) = \sup_{s \in \mathbb{R}^n} \{\langle x, s \rangle - \sigma_{\Omega}(s)\}.$$

Como

$$\sigma_{\Omega}(s) = \sup_{y \in \Omega} \langle s, y \rangle,$$

segue que

$$\sigma_{\Omega}^*(x) = \sup_{s \in \mathbb{R}^n} \left\{ \langle x, s \rangle - \sup_{y \in \Omega} \langle s, y \rangle \right\}.$$

Portanto,

$$\sigma_{\Omega}^*(x) = \sup_{s \in \mathbb{R}^n} \inf_{y \in \Omega} \langle s, x - y \rangle.$$

Analisemos dois casos.

Caso 1: $x \in \Omega$.

Neste caso, tomando $y = x$, obtemos

$$\langle s, x - y \rangle = 0, \quad \forall s \in \mathbb{R}^n.$$

Além disso,

$$\inf_{y \in \Omega} \langle s, x - y \rangle \leq 0.$$

Logo, $\sigma_{\Omega}^*(x) = 0$.

Caso 2: $x \notin \Omega$.

Como Ω é convexo e fechado, pelo teorema de separação forte [página 101 de [30]] existe $s_0 \neq 0$ e $\alpha \in \mathbb{R}$ tais que

$$\langle s_0, x \rangle > \alpha \geq \langle s_0, y \rangle, \quad \forall y \in \Omega.$$

Assim,

$$\langle s_0, x - y \rangle > 0, \quad \forall y \in \Omega.$$

Definindo $\beta := \inf_{y \in \Omega} \langle s_0, x - y \rangle$, temos $\beta > 0$. Para todo $t > 0$,

$$\inf_{y \in \Omega} \langle ts_0, x - y \rangle = t\beta.$$

Consequentemente,

$$\sigma_{\Omega}^*(x) \geq \sup_{t > 0} t\beta = +\infty.$$

Logo,

$$\sigma_{\Omega}^*(x) = +\infty.$$

Portanto,

$$\sigma_{\Omega}^*(x) = \begin{cases} 0, & \text{se } x \in \Omega, \\ +\infty, & \text{se } x \notin \Omega. \end{cases}$$

Concluimos então que $\sigma_{\Omega}^* = \delta_{\Omega}$. □

Corolário 1. Se $\Omega \subseteq \mathbb{R}^n$ é não vazio, convexo e fechado, então

$$\delta_{\Omega}^{**} = \delta_{\Omega}.$$

Demonstração. Como

$$\delta_{\Omega}^* = \sigma_{\Omega} \Rightarrow \delta_{\Omega}^{**} = \sigma_{\Omega}^*.$$

Pela proposição anterior,

$$\sigma_{\Omega}^* = \delta_{\Omega}.$$

Portanto, $\delta_{\Omega}^{**} = \delta_{\Omega}$. □

O corolário anterior mostra, portanto, que a função indicadora de um conjunto não vazio, convexo e fechado coincide com sua biconjugada.

Definição 1.2.15 (inf-convolução). *Sejam $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. A inf-convolução (ou ínfimo convolução) de f e g , $f \square g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é definida por*

$$(f \square g)(x) = \inf_{y \in \mathbb{R}^n} \{f(y) + g(x - y)\}.$$

Proposição 1.2.10. *Sejam $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. Então*

$$(f \square g)^* = f^* + g^*.$$

Demonstração. Seja $u \in \mathbb{R}^n$. Pela definição da conjugada de Fenchel, temos

$$(f \square g)^*(u) = \sup_{x \in \mathbb{R}^n} \{\langle x, u \rangle - (f \square g)(x)\}.$$

Como

$$(f \square g)(x) = \inf_{y \in \mathbb{R}^n} \{f(y) + g(x - y)\},$$

segue que

$$(f \square g)^*(u) = \sup_{x \in \mathbb{R}^n} \left\{ \langle x, u \rangle - \inf_{y \in \mathbb{R}^n} \{f(y) + g(x - y)\} \right\}.$$

Usando a identidade

$$a - \inf_y b_y = \sup_y \{a - b_y\},$$

obtemos

$$(f \square g)^*(u) = \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^n} \{\langle x, u \rangle - f(y) - g(x - y)\}.$$

Portanto,

$$(f \square g)^*(u) = \sup_{x, y \in \mathbb{R}^n} \{\langle x, u \rangle - f(y) - g(x - y)\}.$$

Fazendo a mudança de variável

$$z = x - y,$$

temos $\mathbf{x} = \mathbf{y} + \mathbf{z}$. Assim,

$$\langle \mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{y} + \mathbf{z}, \mathbf{u} \rangle = \langle \mathbf{y}, \mathbf{u} \rangle + \langle \mathbf{z}, \mathbf{u} \rangle.$$

Logo,

$$(f \square g)^*(\mathbf{u}) = \sup_{\mathbf{y}, \mathbf{z} \in \mathbb{R}^n} \{ \langle \mathbf{y}, \mathbf{u} \rangle + \langle \mathbf{z}, \mathbf{u} \rangle - f(\mathbf{y}) - g(\mathbf{z}) \}.$$

Reorganizando os termos, obtemos

$$(f \square g)^*(\mathbf{u}) = \sup_{\mathbf{y}, \mathbf{z} \in \mathbb{R}^n} \{ (\langle \mathbf{y}, \mathbf{u} \rangle - f(\mathbf{y})) + (\langle \mathbf{z}, \mathbf{u} \rangle - g(\mathbf{z})) \}.$$

Como as variáveis $\mathbf{y}$ e $\mathbf{z}$ são independentes, segue que

$$(f \square g)^*(\mathbf{u}) = \sup_{\mathbf{y} \in \mathbb{R}^n} \{ \langle \mathbf{y}, \mathbf{u} \rangle - f(\mathbf{y}) \} + \sup_{\mathbf{z} \in \mathbb{R}^n} \{ \langle \mathbf{z}, \mathbf{u} \rangle - g(\mathbf{z}) \}.$$

Portanto,

$$(f \square g)^*(\mathbf{u}) = f^*(\mathbf{u}) + g^*(\mathbf{u}).$$

Como $\mathbf{u} \in \mathbb{R}^n$ foi arbitrário, concluimos que

$$(f \square g)^* = f^* + g^*.$$

□

A definição de inf-convolução e a propriedade que relaciona sua conjugada às conjugadas das funções envolvidas serão utilizadas posteriormente na análise do exemplo de minimização associado ao problema de nível inferior no Capítulo 5.

Capítulo 2

Problemas de Equilíbrio

2.1 Problemas de Equilíbrio

O conceito de equilíbrio aparece naturalmente em diversas áreas do conhecimento, tais como Economia, Engenharia, Física e Teoria dos Jogos. Em linhas gerais, um sistema encontra-se em equilíbrio quando nenhum agente, componente ou variável envolvida possui incentivo para alterar unilateralmente o seu estado atual. A modelagem matemática desse conceito levou ao desenvolvimento de diferentes formulações de problemas de equilíbrio, muitas das quais podem ser unificadas sob uma mesma estrutura.

Historicamente, uma das primeiras formulações gerais relacionadas ao conceito moderno de problema de equilíbrio foi apresentada por Ky Fan em [23], através de sua desigualdade minimax. Posteriormente, diversos autores observaram que problemas clássicos de otimização, desigualdades variacionais e equilíbrios de Nash podiam ser interpretados dentro de uma estrutura comum.

Entretanto, foi somente com o trabalho seminal de Blum e Oettli [11] que surgiu a formulação atualmente conhecida como Problema de Equilíbrio. Os autores mostraram que uma ampla classe de problemas matemáticos pode ser modelada através de uma bifunção definida sobre um conjunto convexo, fornecendo uma estrutura unificada para o estudo de diversas aplicações.

Ao longo dos anos, essa formulação tornou-se uma ferramenta fundamental em Otimização e Análise Variacional, motivando o desenvolvimento de resultados de existência, unicidade, estabilidade e métodos numéricos para obtenção de soluções. Entre as contribuições mais relevantes destacam-se os trabalhos de Muu e Oettli [45], que estabeleceram

importantes resultados de existência para problemas de equilíbrio multifuncionais, de Iusem e Sosa [25], que obtiveram teoremas de existência sob hipóteses mais gerais, e de Combettes e Hirstoaga [19], que desenvolveram métodos proximais para problemas de equilíbrio em espaços de Hilbert.

Neste trabalho, consideraremos problemas definidos em espaços euclidianos de dimensão finita. Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo e seja

$$f : K \times K \longrightarrow \mathbb{R}$$

uma bifunção.

Definição 2.1.1. *O Problema de Equilíbrio associado à bifunção f e ao conjunto K , denotado por $PE(f, K)$, consiste em encontrar um ponto $\bar{x} \in K$ tal que*

$$f(\bar{x}, y) \geq 0, \quad \forall y \in K. \quad (2.1)$$

O conjunto de todas as soluções de (2.1) será denotado por

$$S(f, K) = \{x \in K : f(x, y) \geq 0, \forall y \in K\}.$$

Assim, resolver o problema de equilíbrio equivale a determinar os elementos do conjunto $S(f, K)$.

A condição $f(\bar{x}, y) \geq 0$ para todo $y \in K$ significa que, quando o sistema se encontra no estado $\bar{x}$, nenhuma outra decisão admissível y produz uma violação da relação de equilíbrio descrita pela bifunção f . Dessa forma, a bifunção atua como um mecanismo de comparação entre estados admissíveis do sistema.

Para garantir propriedades adequadas de existência e estabilidade das soluções, assumimos algumas hipóteses sobre a bifunção f .

$$(P1) \quad f(x, x) = 0, \quad \forall x \in K;$$

$$(P2) \quad \text{para cada } x \in K, \text{ a função } y \mapsto f(x, y) \text{ é convexa e semicontínua inferiormente;}$$

$$(P3) \quad \text{para cada } y \in K, \text{ a função } x \mapsto f(x, y) \text{ é semicontínua superiormente.}$$

Essas condições desempenham papel fundamental em diversos resultados clássicos de existência de soluções para problemas de equilíbrio.

Observação 2.1.1. *As hipóteses acima aparecem naturalmente nos trabalhos de Blum e Oettli [11] e em grande parte da literatura posterior. Em particular, elas permitem utilizar ferramentas da Análise Convexa e da teoria de pontos fixos para obter resultados de existência.*

Uma característica importante do problema de equilíbrio é seu elevado grau de generalidade. De fato, diversas classes clássicas de problemas matemáticos podem ser obtidas como casos particulares da formulação (2.1). Entre elas destacam-se: Problemas de Otimização Convexa, Problemas de Desigualdade Variacional, Problemas de Complementaridade, Problemas de Ponto Fixo, Equilíbrios de Nash em Jogos Não Cooperativos, Problemas de Inclusão Variacional, entre outros.

Essa capacidade de unificação torna o estudo dos problemas de equilíbrio particularmente relevante, pois resultados desenvolvidos para a formulação abstrata podem ser aplicados simultaneamente a diversas áreas da Matemática Aplicada.

Nosso objetivo inicial consiste em compreender a estrutura do conjunto de soluções $S(f, K)$, estabelecendo condições sob as quais ele é não vazio. Posteriormente, essa teoria será utilizada para introduzir os Problemas de Equilíbrio de Dois Níveis, que constituem o objeto central desta dissertação.

Antes disso, apresentaremos alguns dos principais exemplos de problemas que podem ser formulados como problemas de equilíbrio, evidenciando o caráter unificador da teoria desenvolvida por Blum e Oettli [11].

2.2 Casos Particulares de Problemas de Equilíbrio

Nesta seção apresentaremos alguns casos particulares, destacando a maneira pela qual eles podem ser incorporados à estrutura geral de Blum e Oettli [11].

2.2.1 Problemas de Otimização Convexa

Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo e seja $\varphi : K \rightarrow \mathbb{R}$ uma função convexa e semicontínua inferiormente.

O problema de otimização convexa consiste em determinar um ponto $\bar{x} \in K$ tal que

$$\varphi(\bar{x}) = \min_{x \in K} \varphi(x). \quad (2.2)$$

Equivalentemente,

$$\varphi(\bar{x}) \leq \varphi(y), \quad \forall y \in K.$$

A seguir mostraremos que esse problema pode ser formulado como um problema de equilíbrio.

Teorema 8. *Se $f(x, y) = \varphi(y) - \varphi(x)$ para todo $x, y \in K$, então f satisfaz P1–P3. Além disso, $\bar{x}$ é uma solução de (2.2) se, e somente se, $\bar{x}$ é uma solução de $PE(f, K)$.*

Demonstração. De fato,

P1. $f(x, x) = \varphi(x) - \varphi(x) = 0$ para todo $x \in K$.

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

Provaremos primeiro a convexidade. Sejam $y, z \in K$ e $t \in [0, 1]$, da convexidade de φ segue que

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

portanto $f(x, \cdot)$ é convexa para todo $x \in K$.

A propriedade de semicontinuidade inferior de $f(x, \cdot)$ decorre igualmente da semicontinuidade inferior de φ . Com efeito, seja $(y^k)_{k \in \mathbb{N}} \subset K$ uma sequência tal que $y^k \rightarrow y$. Então,

$$\begin{aligned} \liminf_{k \rightarrow +\infty} f(x, y^k) &= \liminf_{k \rightarrow +\infty} [\varphi(y^k) - \varphi(x)] \\ &\geq \liminf_{k \rightarrow +\infty} \varphi(y^k) - \varphi(x) \\ &\geq \varphi(y) - \varphi(x) \\ &= f(x, y). \end{aligned}$$

Portanto, $f(x, \cdot)$ é semicontínua inferiormente em K para todo $x \in K$, o que mostra que a condição P2 é satisfeita.

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

Dado $(x^k)_{k \in \mathbb{N}} \subset K$ uma sequência em K tal que $x^k \rightarrow x$, então

$$\begin{aligned} \limsup_{k \rightarrow +\infty} f(x^k, y) &= \limsup_{k \rightarrow +\infty} [\varphi(y) - \varphi(x^k)] \\ &\leq \varphi(y) + \limsup_{k \rightarrow +\infty} [-\varphi(x^k)] \\ &= \varphi(y) - \liminf_{k \rightarrow +\infty} \varphi(x^k) \\ &\leq \varphi(y) - \varphi(x) \\ &= f(x, y) \end{aligned}$$

e por definição, segue o resultado.

Suponha agora que $\bar{x}$ é uma solução de (2.2), isto é, $\varphi(\bar{x}) \leq \varphi(y)$ para todo $y \in K$. Temos então que

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

Logo, $\bar{x}$ é solução de $PE(f, K)$. Da mesma forma, seja $\bar{x}$ uma solução de $PE(f, K)$, ou seja, $f(\bar{x}, y) \geq 0$ para todo $y \in K$. Então

$$0 \leq f(\bar{x}, y) = \varphi(y) - \varphi(\bar{x}) \quad \Rightarrow \quad \varphi(\bar{x}) \leq \varphi(y) \quad \forall y \in K,$$

portanto $\bar{x}$ é solução de (2.2). □

O resultado acima mostra que o problema de minimização convexa pode ser visto como um problema de equilíbrio associado à bifunção $f(x, y) = \varphi(y) - \varphi(x)$.

Exemplo 2.2.1. Considere $K = \mathbb{R}$ e a função $\varphi(x) = x^2$. A bifunção associada é dada por $f(x, y) = y^2 - x^2$. O problema de equilíbrio consiste em encontrar $\bar{x} \in \mathbb{R}$ tal que

$$y^2 - \bar{x}^2 \geq 0, \quad \forall y \in \mathbb{R}.$$

Tomando $y = 0$, obtemos $-\bar{x}^2 \geq 0$, o que implica $\bar{x} = 0$.

Observa-se que esse é precisamente o minimizador global da função φ .

A interpretação acima evidencia que a teoria de problemas de equilíbrio constitui uma extensão natural da teoria de otimização convexa. Em particular, muitos resultados de existência e algoritmos desenvolvidos para problemas de equilíbrio generalizam técnicas clássicas da otimização.

Nos próximos tópicos veremos que essa mesma estrutura engloba também problemas de desigualdade variacional e outras formulações fundamentais da análise variacional.

2.2.2 Problemas de Ponto Fixo

A teoria de pontos fixos constitui uma das áreas mais importantes da Matemática moderna, desempenhando papel fundamental em Análise, Topologia, Economia Matemática, Teoria dos Jogos e Equações Diferenciais.

De maneira geral, dado um conjunto $K \subset \mathbb{R}^n$ e uma aplicação $T : K \rightarrow K$, o problema de ponto fixo consiste em determinar um ponto $\bar{x} \in K$ tal que

$$T(\bar{x}) = \bar{x}. \quad (2.3)$$

Um ponto satisfazendo (2.3) é denominado ponto fixo da aplicação T .

Diversos problemas matemáticos podem ser reformulados como problemas de ponto fixo. Em particular, teoremas clássicos de existência de soluções para equações diferenciais, sistemas de equações não lineares e modelos econômicos são obtidos através de resultados de ponto fixo.

Nesta subseção mostraremos que problemas de ponto fixo podem ser escritos como problemas de equilíbrio.

Teorema 9. *Seja $T : K \rightarrow K$ uma aplicação contínua e considere a bifunção*

$$f(x, y) = \langle x - T(x), y - x \rangle. \quad (2.4)$$

Então $x^ \in K$ é ponto fixo de T se, e somente se, $x^* \in S(f, K)$.*

Demonstração. Suponha inicialmente que $T(x^*) = x^*$. Então

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

Logo, $f(x^*, y) \geq 0$, para todo $y \in K$, o que implica $x^* \in S(f, K)$. Reciprocamente, suponha que $x^* \in S(f, K)$. Então

$$\langle x^* - T(x^*), y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Tomando $y = T(x^*)$, obtemos

$$\langle x^* - T(x^*), T(x^*) - x^* \rangle \geq 0.$$

Consequentemente,

$$-\|T(x^*) - x^*\|^2 \geq 0.$$

Da positividade da norma segue que

$$\|T(x^*) - x^*\| = 0,$$

ou seja, $T(x^*) = x^*$. Portanto, x^* é ponto fixo de T . $\square$

Proposição 2.2.1. *Se $T : K \rightarrow K$ é contínua, então a bifunção definida em (2.4) satisfaz as hipóteses (P1), (P2) e (P3).*

Demonstração. Temos

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

Logo, (P1) é satisfeita.

Fixado $x \in K$, a função

$$y \mapsto \langle x - T(x), y - x \rangle$$

é afim em y . Portanto é convexa e contínua. Assim, (P2) é satisfeita.

Por outro lado, fixado $y \in K$, a continuidade de T implica a continuidade da aplicação

$$x \mapsto \langle x - T(x), y - x \rangle.$$

Logo $f(\cdot, y)$ é contínua e, conseqüentemente, semicontínua superiormente. Portanto, (P3) também é satisfeita. $\square$

Observação 2.2.1. *O resultado anterior mostra que os problemas de ponto fixo constituem um caso particular da teoria dos problemas de equilíbrio. Dessa forma, técnicas desenvolvidas para problemas de equilíbrio podem ser utilizadas para estudar a existência de pontos fixos.*

2.2.3 Problemas de Desigualdade Variacional

A origem da teoria das desigualdades variacionais está associada aos trabalhos pioneiros de Stampacchia na década de 1960. De acordo com Nagurney e Zhang [46], a formulação moderna foi introduzida por Hartman e Stampacchia [24] no contexto de equações diferenciais parciais. Além disso, o próprio Stampacchia [55] atribui ao artigo [54] o primeiro resultado geral de existência para problemas de desigualdade variacional.

Definição 2.2.1. *Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo e $T : K \rightarrow \mathbb{R}^n$ uma aplicação. O problema de desigualdade variacional associado a T e K , denotado por $VI(T, K)$, consiste em determinar um ponto $\bar{x} \in K$ tal que*

$$\langle T(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K. \quad (2.5)$$

Exemplo 2.2.2. *Considere uma função diferenciável $f : [a, b] \rightarrow \mathbb{R}$ e seja $x_0 \in [a, b]$ um ponto de mínimo global de f no intervalo, isto é, $f(x_0) = \min_{x \in [a, b]} f(x)$. Dependendo da posição de x_0 no intervalo, temos três possibilidades:*

- (i) *Se $x_0 \in (a, b)$, então $f'(x_0) = 0$;*
- (ii) *Se $x_0 = a$, então $f'(x_0) \geq 0$;*
- (iii) *Se $x_0 = b$, então $f'(x_0) \leq 0$.*

Em qualquer um dos casos acima, conclui-se que x_0 satisfaz a seguinte desigualdade variacional:

$$f'(x_0)(x - x_0) \geq 0, \quad \forall x \in [a, b].$$

Exemplo 2.2.3. *Seja $K \subset \mathbb{R}^n$ um conjunto convexo e fechado, e considere uma função diferenciável $h : K \rightarrow \mathbb{R}$. Suponha que $x_0 \in K$ seja um ponto de mínimo global de h sobre K , isto é, $h(x_0) = \min_{x \in K} h(x)$. Desejamos obter uma caracterização desse ponto minimizador.*

Seja $x \in K$ arbitrário. Como K é convexo, o segmento que une x_0 a x está contido em K , ou seja,

$$(1 - t)x_0 + tx = x_0 + t(x - x_0) \in K, \quad \forall t \in [0, 1].$$

Definimos a função auxiliar $\psi : [0, 1] \rightarrow \mathbb{R}$ por

$$\psi(t) = h(x_0 + t(x - x_0)).$$

Como x_0 é ponto de mínimo de h em K , segue que ψ atinge seu mínimo em $t = 0$. Pelo resultado do exemplo anterior, obtemos

$$\psi'(0) \geq 0.$$

Aplicando a regra da cadeia, segue que

$$\psi'(0) = \langle \nabla h(x_0), x - x_0 \rangle.$$

Portanto, o ponto x_0 satisfaz a desigualdade variacional.

$$\langle \nabla h(x_0), x - x_0 \rangle \geq 0, \quad \forall x \in K.$$

Mostraremos como certos problemas de desigualdade variacional podem ser formulados como problemas de equilíbrio. Pela Definição 2.2.1, dado um conjunto não vazio, fechado e convexo $K \subset \mathbb{R}^n$ e uma aplicação $T : K \rightarrow \mathbb{R}^n$, o problema $VI(T, K)$ consiste em encontrar $\bar{x} \in K$ tal que

$$\langle T(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K.$$

Também será útil considerar uma versão multivalorada desse problema. Para isso, seja $\mathcal{P}(\mathbb{R}^n)$ o conjunto das partes de $\mathbb{R}^n$ e seja $\bar{T} : K \rightarrow \mathcal{P}(\mathbb{R}^n)$ uma aplicação ponto-conjunto tal que $\bar{T}(x)$ é não vazio, compacto e convexo para todo $x \in K$. Nesse caso, o problema de desigualdade variacional multivalorada, denotado por $VI(\bar{T}, K)$, consiste em encontrar $x^* \in K$ e $u^* \in \bar{T}(x^*)$ tais que

$$\langle u^*, y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Antes de estabelecer a equivalência entre desigualdades variacionais e problemas de equilíbrio, apresentamos o seguinte resultado auxiliar.

Lema 1. *Sejam $A \subset \mathbb{R}^n$ um conjunto compacto e convexo, $B \subset \mathbb{R}^n$ um conjunto convexo, e seja $p : A \times B \rightarrow \mathbb{R}$ uma função tal que:*

- (a) *para cada $y \in B$, a aplicação $p(\cdot, y) : A \rightarrow \mathbb{R}$ é côncava e semicontínua superiormente;*
- (b) *para cada $x \in A$, a aplicação $p(x, \cdot) : B \rightarrow \mathbb{R}$ é convexa.*

Suponha ainda que $\max_{x \in A} p(x, y) \geq 0$, para todo $y \in B$. Então existe $x^ \in A$ tal que $p(x^*, y) \geq 0$, para todo $y \in B$.*

Demonstração. Suponha, por contradição, que não exista $x \in A$ tal que $p(x, y) \geq 0$, para todo $y \in B$. Então, para cada $x \in A$, existem $y_x \in B$ e $\varepsilon_x > 0$ tais que $p(x, y_x) < -\varepsilon_x$. Para $y \in B$ e $\varepsilon > 0$, considere o conjunto $A(y, \varepsilon) := \{x \in A; p(x, y) < -\varepsilon\}$. Pela semicontinuidade superior de $p(\cdot, y)$, cada conjunto $A(y, \varepsilon)$ é aberto em A . Além disso, pela construção acima, a família desses conjuntos cobre A . Como A é compacto, existem $y^1, \dots, y^m \in B$ e $\varepsilon_1, \dots, \varepsilon_m > 0$ tais que

$$A \subset \bigcup_{i=1}^m A(y^i, \varepsilon_i).$$

Tomando $\varepsilon := \min_{1 \leq i \leq m} \varepsilon_i > 0$, obtemos, para todo $x \in A$,

$$\min_{1 \leq i \leq m} p(x, y^i) \leq -\varepsilon.$$

Por [51, Teorema 21.1, página 186] aplicado às funções convexas

$$x \mapsto -\varepsilon - p(x, y^i), \quad i = 1, \dots, m,$$

existem $\lambda_i \geq 0$, com $\sum_{i=1}^m \lambda_i = 1$, tais que

$$\sum_{i=1}^m \lambda_i (-\varepsilon - p(x, y^i)) \geq 0, \quad \forall x \in A.$$

Portanto,

$$\sum_{i=1}^m \lambda_i p(x, y^i) \leq -\varepsilon, \quad \forall x \in A.$$

Como B é convexo, o ponto

$$\hat{y} := \sum_{i=1}^m \lambda_i y^i$$

pertence a B . Pela convexidade de $p(x, \cdot)$, segue que

$$p(x, \hat{y}) = p\left(x, \sum_{i=1}^m \lambda_i y^i\right) \leq \sum_{i=1}^m \lambda_i p(x, y^i) \leq -\varepsilon, \quad \forall x \in A.$$

Logo,

$$\max_{x \in A} p(x, \hat{y}) \leq -\varepsilon < 0,$$

o que contradiz a hipótese $\max_{x \in A} p(x, y) \geq 0$, para todo $y \in B$. Portanto, existe $x^* \in A$ tal que

$$p(x^*, y) \geq 0, \quad \forall y \in B.$$

□

Teorema 10. *Temos as seguintes relações entre desigualdades variacionais e problemas de equilíbrio.*

(i) *Seja $T : K \rightarrow \mathbb{R}^n$ uma aplicação contínua e defina $f : K \times K \rightarrow \mathbb{R}$ por*

$f(x, y) = \langle T(x), y - x \rangle$. Então f satisfaz as hipóteses básicas (P1)–(P3), e x^ resolve $VI(T, K)$ se, e somente se, x^* resolve $PE(f, K)$.*

(ii) *Seja $\bar{T} : K \rightarrow \mathcal{P}(\mathbb{R}^n)$ uma aplicação ponto-conjunto, semicontínua superiormente, com valores não vazios, compactos e convexos. Defina $f : K \times K \rightarrow \mathbb{R}$ por*

$f(x, y) = \max_{u \in \bar{T}(x)} \langle u, y - x \rangle$. Então f satisfaz as hipóteses básicas (P1)–(P3), e x^ resolve $PE(f, K)$ se, e somente se, existe $u^* \in \bar{T}(x^*)$ tal que (x^*, u^*) resolve $VI(\bar{T}, K)$.*

Demonstração. Inicialmente, provemos o item (i). Seja $f(x, y) = \langle T(x), y - x \rangle$. Se x^* é solução de $VI(T, K)$, então $\langle T(x^*), y - x^* \rangle \geq 0$, para todo $y \in K$. Pela definição de f , isso equivale a $f(x^*, y) \geq 0$, para todo $y \in K$. Logo, x^* é solução de $PE(f, K)$.

Reciprocamente, se x^* é solução de $PE(f, K)$, então $f(x^*, y) \geq 0$, para todo $y \in K$. Como $f(x^*, y) = \langle T(x^*), y - x^* \rangle$, obtemos $\langle T(x^*), y - x^* \rangle \geq 0$, para todo $y \in K$. Portanto, x^* resolve $VI(T, K)$.

Além disso, verifica-se que f satisfaz (P1)–(P3). De fato,

$$f(x, x) = \langle T(x), x - x \rangle = 0, \quad \forall x \in K.$$

Como T é contínua, para cada $y \in K$, a aplicação $x \mapsto f(x, y) = \langle T(x), y - x \rangle$ é contínua e, portanto, semicontínua superiormente. Por fim, para cada $x \in K$, a aplicação $y \mapsto f(x, y) = \langle T(x), y - x \rangle$ é afim, logo convexa e semicontínua inferiormente.

Agora, provemos o item (ii). Suponha inicialmente que $x^* \in K$ e $u^* \in \bar{T}(x^*)$ resolvam $VI(\bar{T}, K)$. Então $\langle u^*, y - x^* \rangle \geq 0$, para todo $y \in K$. Como $u^* \in \bar{T}(x^*)$, temos

$$f(x^*, y) = \max_{u \in \bar{T}(x^*)} \langle u, y - x^* \rangle \geq \langle u^*, y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Logo, x^* resolve $PE(f, K)$.

Reciprocamente, suponha que x^* seja solução de $PE(f, K)$. Então $f(x^*, y) \geq 0$, para todo $y \in K$. Pela definição de f ,

$$\max_{u \in \bar{T}(x^*)} \langle u, y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Tomemos

$$A = \bar{T}(x^*) \quad \text{e} \quad p(u, y) = \langle u, y - x^* \rangle, \quad u \in A, \quad y \in K.$$

Como A é compacto e convexo, e p é linear em cada variável, as hipóteses do lema 1 são satisfeitas. Assim, existe $u^* \in A = \bar{T}(x^*)$ tal que $p(u^*, y) \geq 0$, para todo $y \in K$, isto é, $\langle u^*, y - x^* \rangle \geq 0$, para todo $y \in K$. Portanto, (x^*, u^*) resolve $VI(\bar{T}, K)$.

Por fim, as hipóteses (P1)–(P3) para a bifunção

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

seguem da compacidade e convexidade dos valores de $\bar{T}$, da semicontinuidade superior de $\bar{T}$, e do fato de que, para cada $x \in K$, a função $y \mapsto f(x, y)$ é o máximo de funções afins em y , sendo, portanto, convexa. $\square$

2.2.4 Equações Generalizadas

As equações generalizadas constituem uma importante classe de problemas da Análise Variacional, permitindo formular de maneira unificada diversos problemas de otimização, complementaridade e desigualdades variacionais. Essa formulação foi introduzida e desenvolvida por Robinson em [49]. Por outro lado, como discutido por Blum e Oettli em [11], problemas de desigualdade variacional podem ser naturalmente incorporados à estrutura dos problemas de equilíbrio. Dessa forma, uma classe de equações generalizadas também pode ser interpretada como um caso particular de problema de equilíbrio.

Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo e seja $F : K \rightarrow \mathbb{R}^n$. Consideremos a equação generalizada

$$0 \in F(x) + \mathcal{N}_K(x), \quad (2.6)$$

onde $\mathcal{N}_K(x)$ denota o cone normal a K no ponto x .

A expressão (2.6) inclui, em particular, as restrições impostas pelo conjunto K por meio do cone normal. Veremos a seguir que sua resolução é equivalente à resolução de uma desigualdade variacional e, consequentemente, de um problema de equilíbrio.

Teorema 11. *Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo e $F : K \rightarrow \mathbb{R}^n$. Defina a bifunção $f : K \times K \rightarrow \mathbb{R}$ por $f(x, y) := \langle F(x), y - x \rangle$. Então $x^* \in K$ é solução da equação generalizada $0 \in F(x) + \mathcal{N}_K(x)$ se, e somente se, $x^* \in S(f, K)$.*

Demonstração. Suponha inicialmente que x^* seja solução da equação generalizada. Então,

$$0 \in F(x^*) + \mathcal{N}_K(x^*),$$

o que equivale a

$$-F(x^*) \in \mathcal{N}_K(x^*).$$

Pela definição do cone normal,

$$\langle -F(x^*), y - x^* \rangle \leq 0, \quad \forall y \in K.$$

Multiplicando a desigualdade por -1 , obtemos

$$\langle F(x^*), y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Como

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

segue que

$$f(x^*, y) \geq 0, \quad \forall y \in K.$$

Portanto, $x^* \in S(f, K)$.

Reciprocamente, suponha que $x^* \in S(f, K)$. Pela definição do conjunto solução do problema de equilíbrio,

$$f(x^*, y) \geq 0, \quad \forall y \in K.$$

Logo,

$$\langle F(x^*), y - x^* \rangle \geq 0, \quad \forall y \in K,$$

ou, equivalentemente,

$$\langle -F(x^*), y - x^* \rangle \leq 0, \quad \forall y \in K.$$

Pela definição do cone normal, temos

$$-F(x^*) \in \mathcal{N}_K(x^*),$$

e, portanto,

$$0 \in F(x^*) + \mathcal{N}_K(x^*).$$

Assim, x^* é solução da equação generalizada. $\square$

A proposição anterior mostra que a equação generalizada $0 \in F(x) + \mathcal{N}_K(x)$ é equivalente à desigualdade variacional VI(F, K) e, por meio da bifunção $f(x, y) = \langle F(x), y - x \rangle$, também pode ser formulada como um problema de equilíbrio.

Mostraremos agora que, sob uma hipótese natural de continuidade sobre F, a bifunção associada satisfaz as condições (P1)–(P3) utilizadas neste capítulo.

Proposição 2.2.2. *Seja $F : K \rightarrow \mathbb{R}^n$ uma aplicação contínua e considere a bifunção*

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

Então f satisfaz as hipóteses (P1), (P2) e (P3).

Demonstração. Verificaremos cada uma das condições.

(P1) Para todo $x \in K$,

$$f(x, x) = \langle F(x), x - x \rangle = 0.$$

Portanto, f satisfaz (P1).

(P2) Fixado $x \in K$, considere a função $f(x, \cdot) : K \rightarrow \mathbb{R}$. Para $y, z \in K$ e $t \in [0, 1]$, temos

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

Logo, $f(x, \cdot)$ é afim e, em particular, convexa. Além disso, como $F(x)$ é um vetor fixo, a aplicação

$$y \mapsto \langle F(x), y - x \rangle$$

é contínua. Portanto, $f(x, \cdot)$ é semicontínua inferiormente. Assim, (P2) é satisfeita.

(P3) Fixado $y \in K$, considere

$$f(\cdot, y) : K \rightarrow \mathbb{R}, \quad f(x, y) = \langle F(x), y - x \rangle.$$

Seja $\{x^k\} \subset K$ uma sequência tal que $x^k \rightarrow x$. Como F é contínua,

$$F(x^k) \rightarrow F(x).$$

Além disso,

$$y - x^k \rightarrow y - x.$$

Pela continuidade do produto interno,

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

Portanto, $f(\cdot, y)$ é contínua e, consequentemente, semicontínua superiormente. Assim, f satisfaz (P3). $\square$

Portanto, as equações generalizadas da forma $0 \in F(x) + \mathcal{N}_K(x)$ podem ser incorporadas à teoria dos problemas de equilíbrio por meio da bifunção

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

Esse exemplo reforça o caráter unificador da formulação de Blum e Oettli, uma vez que uma equação generalizada pode ser interpretada sucessivamente como uma desigualdade variacional e como um problema de equilíbrio.

2.3 Monotonicidade de Bifunções

A monotonicidade desempenha papel fundamental na teoria dos problemas de equilíbrio e das desigualdades variacionais, estando diretamente relacionada a resultados de existência, unicidade e convergência de algoritmos. As noções apresentadas nesta seção são baseadas nos conceitos clássicos de monotonicidade para operadores estudados por Facchinei e Pang [21] e Bauschke e Combettes [9], bem como em suas extensões para bifunções no contexto dos problemas de equilíbrio, conforme desenvolvido por Blum e Oettli [11].

Definição 2.3.1. *Uma bifunção $f : K \times K \rightarrow \mathbb{R}$ é dita monótona quando*

$$f(x, y) + f(y, x) \leq 0, \quad \forall x, y \in K. \quad (2.7)$$

Esta definição é uma generalização natural da monotonicidade de operadores em desigualdades variacionais.

Exemplo 2.3.1. *Considere um operador monótono $T : K \rightarrow \mathbb{R}^n$ e a bifunção $f(x, y) = \langle T(x), y - x \rangle$. Então f é monótona.*

Demonstração. De fato, temos

$$\begin{aligned} f(x, y) + f(y, x) &= \langle T(x), y - x \rangle + \langle T(y), x - y \rangle \\ &= -\langle T(x) - T(y), x - y \rangle. \end{aligned}$$

Como T é monótono,

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

Portanto,

$$f(x, y) + f(y, x) \leq 0.$$

Assim, f é monótona. □

Definição 2.3.2. *A bifunção $f : K \times K \rightarrow \mathbb{R}$ com $f(x, x) = 0$ é dita estritamente monótona quando*

$$f(x, y) + f(y, x) < 0, \quad \forall x, y \in K,$$

e a igualdade ocorre somente quando $x = y$.

A monotonicidade estrita permite obter resultados de unicidade das soluções.

Proposição 2.3.1. *Seja f uma bifunção estritamente monótona, então o conjunto de soluções $S(f, K)$ possui no máximo um elemento.*

Demonstração. Suponha que existam duas soluções distintas $x^*, y^* \in S(f, K)$. Como x^* é solução, $f(x^*, y^*) \geq 0$. Como y^* também é solução, $f(y^*, x^*) \geq 0$. Somando as duas desigualdades,

$$f(x^*, y^*) + f(y^*, x^*) \geq 0.$$

Por outro lado, pela monotonicidade,

$$f(x^*, y^*) + f(y^*, x^*) \leq 0.$$

Logo,

$$f(x^*, y^*) + f(y^*, x^*) = 0.$$

Da monotonicidade estrita segue que $x^* = y^*$. Portanto $S(f, K)$ possui no máximo uma solução. $\square$

Definição 2.3.3. *A bifunção $f : K \times K \rightarrow \mathbb{R}$ é dita fortemente monótona quando existe uma constante $\mu > 0$ tal que*

$$f(x, y) + f(y, x) \leq -\mu \|x - y\|^2, \quad \forall x, y \in K. \quad (2.8)$$

É imediato observar que toda bifunção fortemente monótona é estritamente monótona.

Proposição 2.3.2. *Toda bifunção fortemente monótona é estritamente monótona.*

Demonstração. Se $x \neq y$, então $\|x - y\|^2 > 0$. Da condição (2.8),

$$f(x, y) + f(y, x) < 0.$$

Consequentemente, a igualdade

$$f(x, y) + f(y, x) = 0$$

somente pode ocorrer quando $x = y$. Portanto f é estritamente monótona. $\square$

A noção seguinte é mais fraca que a monotonicidade e aparece frequentemente em problemas de equilíbrio.

Definição 2.3.4. A bifunção $f : K \times K \rightarrow \mathbb{R}$ é dita pseudomonótona quando

$$f(x, y) \geq 0 \Rightarrow f(y, x) \leq 0,$$

para quaisquer $x, y \in K$.

Toda bifunção monótona é pseudomonótona.

Proposição 2.3.3. Se f é monótona, então f é pseudomonótona.

Demonstração. Suponha que $f(x, y) \geq 0$. Como f é monótona,

$$f(x, y) + f(y, x) \leq 0.$$

Logo,

$$f(y, x) \leq -f(x, y) \leq 0.$$

Portanto f é pseudomonótona. □

A recíproca, em geral, não é verdadeira.

A pseudomonotonicidade é suficiente para diversos resultados de existência e aparece frequentemente em modelos econômicos e problemas de equilíbrio de Nash.

Definição 2.3.5. A bifunção $f : K \times K \rightarrow \mathbb{R}$ é dita paramonótona quando

$$f(x, y) + f(y, x) = 0 \Rightarrow x \in S(f, K) \iff y \in S(f, K).$$

A paramonotonicidade desempenha papel importante na análise de convergência de métodos proximais.

2.4 Resultados de Existência

Uma das questões fundamentais associadas a um problema de equilíbrio consiste em determinar condições que garantam a não vacuidade do conjunto de soluções

$$S(f, K) = \{x \in K; f(x, y) \geq 0, \forall y \in K\}.$$

Os primeiros resultados gerais nessa direção remontam aos trabalhos de Ky Fan [23], cujas ideias serviram de base para grande parte da teoria moderna de problemas de equilíbrio.

Nesta seção apresentaremos alguns resultados clássicos de existência que serão utilizados posteriormente no estudo dos Problemas de Equilíbrio de Dois Níveis.

Começamos apresentando um resultado devido a Knaster, Kuratowski e Mazurkiewicz [32].

Lema 2 (KKM). *Seja K um subconjunto arbitrário de $\mathbb{R}^n$ e $F : K \rightarrow \mathcal{P}(\mathbb{R}^n)$ uma aplicação ponto-conjunto tal que, para cada $y \in K$, $F(y)$ é fechado.*

Suponha que

$$(i) \quad \forall y_1, y_2, \dots, y_m \in K, \text{conv}\{y_1, y_2, \dots, y_m\} \subset \bigcup_{j=1}^m F(y_j).$$

(ii) Existe $y_0 \in K$ tal que $F(y_0)$ é compacto.

Então

$$\bigcap_{y \in K} F(y) \neq \emptyset.$$

Demonstração. A demonstração deste lema é apresentada em [32]. □

O resultado a seguir desempenha papel central em toda a teoria de problemas de equilíbrio.

Teorema 12 (Ky Fan). *Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, compacto e convexo.*

Considere uma bifunção $f : K \times K \rightarrow \mathbb{R}$ satisfazendo:

(i) para cada $y \in K$, a função $x \mapsto f(x, y)$ é semicontínua superiormente;

(ii) para cada $x \in K$, a função $y \mapsto f(x, y)$ é quase-convexa.

Então existe $\bar{x} \in K$ tal que

$$\inf_{y \in K} f(\bar{x}, y) \geq \inf_{w \in K} f(w, w).$$

Demonstração. A demonstração deste teorema pode ser encontrada em [23]. □

O Teorema 12 constitui um dos resultados mais importantes da teoria minimax e pode ser interpretado como um resultado de existência para problemas de equilíbrio.

O resultado acima serviu de motivação para o desenvolvimento de diversos teoremas de existência para problemas de equilíbrio, dentre os quais destaca-se o resultado de Blum e Oettli [11] apresentado a seguir.

Teorema 13 (Blum–Oettli). *Sejam $K \subset \mathbb{R}^n$ um conjunto convexo, fechado, não-vazio e $f : K \times K \rightarrow \mathbb{R}$ uma bifunção da forma $f(x, y) = g(x, y) + h(x, y)$ com $g : K \times K \rightarrow \mathbb{R}$ e $h : K \times K \rightarrow \mathbb{R}$.*

Suponha que f satisfaça:

$$(1.1) \quad g(x, x) = 0 \text{ para todo } x \in K;$$

$$(1.2) \quad g \text{ é monótona, i.e., } g(x, y) + g(y, x) \leq 0 \quad \forall x, y \in K;$$

$$(1.3) \quad g(x, \cdot) \text{ é convexa e semicontínua inferiormente para cada } x \in K;$$

$$(1.4) \quad \text{Para todo } x, y \in K, \text{ a função } t \in [0, 1] \mapsto g((1-t)x + ty, y) \text{ é semicontínua superiormente em } t = 0.$$

$$(2.1) \quad h(x, x) = 0 \text{ para todo } x \in K;$$

$$(2.2) \quad h(x, \cdot) \text{ é convexa para cada } x \in K;$$

$$(2.3) \quad h(\cdot, y) \text{ é semicontínua superiormente para cada } y \in K.$$

$$(3) \quad (\text{Coercividade}) \text{ Existe } C \subset K \text{ compacto, convexo e não vazio tal que para todo } x \in C \setminus \text{Core}_K C \text{ existe } a \in \text{Core}_K C \text{ tal que } f(x, a) \leq 0.$$

$$\text{Então, existe } x^* \in C \text{ tal que } f(x^*, y) \geq 0, \quad \forall y \in K.$$

Demonstração. É o resultado principal de Blum e Oettli em [11]. □

2.5 Problemas de Equilíbrio de Dois Níveis

Nas seções anteriores, estudamos o Problema de Equilíbrio clássico e algumas de suas propriedades fundamentais. Em particular, vimos que, sob hipóteses adequadas, o conjunto de soluções

$$S(f, K) = \{x \in K; f(x, y) \geq 0, \quad \forall y \in K\}$$

é não vazio. Entretanto, em diversas aplicações, a determinação de uma solução arbitrária de $S(f, K)$ pode não ser suficiente. Quando o problema admite múltiplas soluções, pode ser necessário introduzir um critério adicional que permita selecionar, dentre os pontos de $S(f, K)$, aqueles que satisfazem alguma propriedade de interesse. Essa necessidade conduz

naturalmente a uma estrutura hierárquica, na qual um segundo problema de equilíbrio é considerado sobre o conjunto solução do primeiro.

Do ponto de vista histórico, uma das formulações precursoras dessa classe de problemas foi apresentada por Chadli, Chbani e Riahi em [16], no contexto da seleção de soluções de problemas de equilíbrio por meio de técnicas de penalização. Posteriormente, Moudafi, em [43], formulou explicitamente uma classe de problemas de equilíbrio de dois níveis envolvendo duas bifunções monótonas e propôs um método proximal para sua resolução. Desde então, problemas dessa natureza passaram a ser estudados sob diferentes hipóteses e por meio de distintas estratégias algorítmicas. Neste trabalho, seguiremos, em particular, a formulação considerada em [10].

Os chamados Problemas de Equilíbrio de Dois Níveis surgem, portanto, da necessidade de incorporar um critério de seleção ao problema de equilíbrio original. Esse tipo de estrutura aparece naturalmente em modelos hierárquicos de otimização e equilíbrio, nos quais as decisões admissíveis em um segundo nível são determinadas pelas soluções obtidas no primeiro. Dessa forma, o problema não consiste apenas em encontrar um ponto de equilíbrio, mas em encontrar, entre todos os equilíbrios possíveis do primeiro nível, um ponto que também satisfaça uma condição adicional de equilíbrio.

Mais precisamente, o problema possui dois níveis hierárquicos. No primeiro nível, considera-se uma bifunção f e determina-se o conjunto solução $S(f, K)$ do problema de equilíbrio associado. No segundo nível, introduz-se uma segunda bifunção h , que estabelece o critério de seleção a ser satisfeito sobre $S(f, K)$. Assim, procura-se $\bar{x} \in S(f, K)$ tal que

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Esse problema será denominado Problema de Equilíbrio de Dois Níveis e constitui um dos objetos centrais desta dissertação.

2.5.1 Definição do Problema

Sejam $f : K \times K \rightarrow \mathbb{R}$ e $h : K \times K \rightarrow \mathbb{R}$ duas bifunções satisfazendo hipóteses adequadas. Inicialmente consideramos o problema de equilíbrio de primeiro nível

$$\text{Encontrar } x \in K \text{ tal que } f(x, y) \geq 0, \quad \forall y \in K.$$

Denotamos por

$$S(f, K) = \{x \in K : f(x, y) \geq 0, \forall y \in K\}$$

o conjunto de todas as suas soluções.

Uma vez determinado o conjunto $S(f, K)$, define-se o problema de equilíbrio de segundo nível.

Definição 2.5.1. *O Problema de Equilíbrio de Dois Níveis associado às bifunções f e h consiste em encontrar um ponto $\bar{x} \in S(f, K)$ tal que*

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K). \quad (2.9)$$

Esse problema será denotado por $PEB(f, h, K)$.

Equivalentemente,

$$PEB(f, h, K) \quad \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in S(f, K) \text{ tal que} \\ h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K). \end{array} \right. \quad (2.10)$$

A formulação acima permite interpretar o problema de equilíbrio de dois níveis como um problema de equilíbrio clássico cujo conjunto viável é o conjunto solução do problema de nível inferior. Essa relação, utilizada em [10, 43], é apresentada formalmente na proposição a seguir.

Proposição 2.5.1. *Suponha que $S(f, K) \neq \emptyset$. Então, $PEB(f, h, K) = PE(h, S(f, K))$. Em particular, o conjunto solução do problema de equilíbrio de dois níveis coincide com $S(h, S(f, K))$.*

Demonstração. Pela definição do Problema de Equilíbrio de Dois Níveis, um ponto $\bar{x} \in K$ é solução de $PEB(f, h, K)$ se, e somente se, $\bar{x} \in S(f, K)$ e

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Por outro lado, considerando $S(f, K)$ como conjunto viável e a restrição de h a $S(f, K) \times S(f, K)$, o problema de equilíbrio $PE(h, S(f, K))$ consiste, por definição, em encontrar $\bar{x} \in S(f, K)$ tal que

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Portanto,

$$\bar{x} \text{ resolve } PEB(f, h, K) \iff \bar{x} \text{ resolve } PE(h, S(f, K)).$$

Consequentemente,

$$\text{PEB}(\mathbf{f}, \mathbf{h}, \mathbf{K}) = \text{PE}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})),$$

e o conjunto de soluções de ambos os problemas é dado por

$$\mathbf{S}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})) = \{\mathbf{x} \in \mathbf{S}(\mathbf{f}, \mathbf{K}); \mathbf{h}(\mathbf{x}, \mathbf{y}) \geq 0, \forall \mathbf{y} \in \mathbf{S}(\mathbf{f}, \mathbf{K})\}.$$

□

Portanto, um problema de equilíbrio de dois níveis pode ser interpretado como um problema de equilíbrio definido sobre o conjunto de soluções de outro problema de equilíbrio.

Observação 1. *Como o conjunto viável do problema de segundo nível é $\mathbf{S}(\mathbf{f}, \mathbf{K})$, segue diretamente que $\mathbf{S}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})) \subseteq \mathbf{S}(\mathbf{f}, \mathbf{K})$. Portanto, toda solução do problema de equilíbrio de dois níveis é, necessariamente, uma solução do problema de equilíbrio de primeiro nível. Mais precisamente,*

$$\mathbf{x}^* \in \mathbf{S}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})) \implies \mathbf{x}^* \in \mathbf{S}(\mathbf{f}, \mathbf{K}).$$

Em geral, a recíproca não é verdadeira.

2.5.2 Interpretação Hierárquica

A estrutura do problema pode ser representada pela seguinte hierarquia:

Nível 1: Resolver $\text{PE}(\mathbf{f}, \mathbf{K})$

↓

$\mathbf{S}(\mathbf{f}, \mathbf{K})$

e, posteriormente,

Nível 2: Resolver $\text{PE}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K}))$

↓

$\mathbf{S}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})).$

Assim, o problema de segundo nível é definido sobre o conjunto solução do problema de primeiro nível. Consequentemente,

$$\mathbf{S}(\mathbf{h}, \mathbf{S}(\mathbf{f}, \mathbf{K})) \subseteq \mathbf{S}(\mathbf{f}, \mathbf{K}) \subseteq \mathbf{K}.$$

Essa inclusão evidencia o caráter hierárquico do problema: o primeiro nível determina o conjunto de soluções admissíveis, enquanto o segundo nível atua como um critério de seleção sobre esse conjunto. Portanto, o segundo nível não introduz novas soluções fora de $S(f, K)$, mas seleciona, dentre as soluções do primeiro nível, aquelas que também satisfazem a condição de equilíbrio associada à bifunção h .

2.5.3 Problemas de Otimização de Dois Níveis como Casos Particulares

Uma importante classe de problemas que pode ser modelada através de (2.9) é a dos problemas de otimização de dois níveis.

Considere uma função convexa $\varphi : K \rightarrow \mathbb{R}$. Definimos $f(x, y) = \varphi(y) - \varphi(x)$. Então $S(f, K) = \operatorname{argmin}_{x \in K} \varphi(x)$. Seja agora uma segunda função convexa $\psi : K \rightarrow \mathbb{R}$ e definamos $h(x, y) = \psi(y) - \psi(x)$. O problema de dois níveis torna-se:

Encontrar

$$\bar{x} \in \operatorname{argmin}_{x \in K} \varphi(x)$$

tal que

$$\psi(\bar{x}) = \min \left\{ \psi(x); x \in \operatorname{argmin}_{u \in K} \varphi(u) \right\}.$$

Portanto, os problemas clássicos de otimização hierárquica constituem casos particulares de problemas de equilíbrio de dois níveis.

2.5.4 Existência de Soluções

A partir da equivalência entre o problema de equilíbrio de dois níveis e um problema de equilíbrio definido sobre $S(f, K)$, podemos obter condições de existência de soluções aplicando resultados clássicos da teoria de problemas de equilíbrio. Nesse sentido, apresentamos a seguinte proposição.

Proposição 2.5.2. *Suponhamos que:*

1. $S(f, K) \neq \emptyset$;
2. $S(f, K)$ seja compacto e convexo;

3. a bifunção h satisfaça as hipóteses do Teorema 13 de Blum–Oettli sobre $S(f, K)$.

Então existe $\bar{x} \in S(f, K)$ tal que

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Em particular, o problema de equilíbrio de dois níveis admite solução.

Demonstração. Como $S(f, K)$ é não vazio e compacto por hipótese, podemos considerá-lo como o conjunto viável do problema de segundo nível. Como a bifunção h satisfaz as hipóteses do Teorema 13 de Blum–Oettli sobre $S(f, K)$, segue diretamente desse teorema que existe um ponto $\bar{x} \in S(f, K)$ tal que

$$h(\bar{x}, y) \geq 0, \quad \forall y \in S(f, K).$$

Logo, $\bar{x}$ é solução de $PEB(f, h, K)$. □

O resultado acima mostra que muitos teoremas clássicos de existência para problemas de equilíbrio podem ser aplicados diretamente ao segundo nível.

2.5.5 Relação com o Método Proximal

Os problemas de equilíbrio de dois níveis têm sido amplamente estudados na literatura e constituem uma área ativa de pesquisa em otimização e análise variacional. Diversos autores investigaram propriedades de existência, estabilidade e algoritmos para essa classe de problemas; ver, por exemplo, Ding [20], Anh, Khanh e Van [1], Chen, Wan e Cho [18], entre outros.

Uma das principais motivações para o estudo desses problemas reside no fato de que eles abrangem importantes classes de modelos hierárquicos, incluindo problemas de otimização de dois níveis e programas matemáticos com restrições de equilíbrio (*Mathematical Programs with Equilibrium Constraints* – MPEC); ver, por exemplo, Luo, Pang e Ralph [36] e Migdalas, Pardalos e Varbrand [38]. Historicamente, problemas de dois níveis foram formalizados inicialmente no contexto de otimização por Bracken e McGill [12].

Embora resultados de existência sejam fundamentais para garantir a boa formulação do problema, tais resultados não fornecem, em geral, um procedimento efetivo para calcular soluções. Por essa razão, torna-se necessário desenvolver métodos iterativos capazes de aproximar soluções de problemas de equilíbrio de dois níveis.

Nesse contexto, diversos algoritmos foram propostos na literatura. Cabot [15], por exemplo, desenvolveu um algoritmo hierárquico para problemas de minimização sucessiva, enquanto Facchinei, Pang, Scutari e Lampariello [22] estudaram métodos numéricos para problemas variacionais hierárquicos.

Nos capítulos seguintes estudaremos um método proximal para problemas de equilíbrio de dois níveis baseado em bifunções regularizadas e distâncias proximais generalizadas. A ideia central consiste em construir uma sequência $(\mathbf{x}^k)_{k \in \mathbb{N}}$ por meio da resolução sucessiva de problemas de equilíbrio regularizados, de modo que seus pontos de acumulação pertençam ao conjunto de soluções do problema de dois níveis.

Essa abordagem foi proposta por Bento, Cruz Neto, Lopes, Soares Jr. e Soubeyran em [10] e constitui um dos principais objetos de estudo desta dissertação.

2.6 Uma hipótese geométrica de regularidade

Na análise de convergência de algoritmos proximais aplicados a problemas de equilíbrio de dois níveis, certas condições adicionais de regularidade tornam-se necessárias para garantir o comportamento assintótico da sequência iterativa gerada.

Um dos primeiros métodos proximais para problemas de equilíbrio de dois níveis foi proposto por Moudafi [43] como uma extensão do algoritmo proximal clássico para o contexto de problemas de equilíbrio também devido a Moudafi em [42]. Em sua análise de convergência, o autor assumiu a condição

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| = o(\varepsilon_k), \quad (2.11)$$

onde $(\mathbf{x}^k)$ é a sequência gerada pelo algoritmo e (ε_k) é uma sequência de parâmetros positivos.

Entretanto, a verificação da condição (2.11) é, em geral, bastante delicada, uma vez que ela depende diretamente do comportamento assintótico da própria sequência iterativa. Devido a essa dificuldade, o próprio Moudafi apontou como problema em aberto a busca por hipóteses alternativas mais naturais e geometricamente interpretáveis.

Posteriormente, Chbani e Riahi [17] propuseram substituir (2.11) por uma condição geométrica envolvendo a conjugada de Fenchel e a função suporte do conjunto solução do problema de equilíbrio de nível inferior. Condições desse tipo já haviam aparecido

anteriormente nos trabalhos de Attouch, Czarnecki e Peypouquet [5, 6], no contexto de métodos de prox-penalização para problemas variacionais restritos.

Motivados por essas ideias, Chbani e Riahi [17] introduziram a seguinte hipótese, que desempenhará papel central na análise de convergência do algoritmo proximal apresentado no Capítulo 5.

Hipótese 1. *Para todo $z \in S(f, K)$ e $p \in \mathcal{R}(N_{S(f, K)})$, vale*

$$\sum_{k=1}^{\infty} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right] < +\infty. \quad (\mathcal{H})$$

Nesta expressão, f_z^* denota a função conjugada de Fenchel associada à função f_z , definida por $f_z(y) := f(z, y)$, enquanto $\sigma_{S(f, K)}$ representa a função suporte do conjunto solução $S(f, K)$, e $\mathcal{R}(N_{S(f, K)})$ denota a imagem da aplicação cone normal de $S(f, K)$.

A hipótese $(\mathcal{H})$ pode ser interpretada como uma condição de compatibilidade entre a geometria do conjunto solução $S(f, K)$, a estrutura convexa da bifunção f e a escolha dos parâmetros do algoritmo. Em particular, ela controla assintoticamente a interação entre os termos de penalização e regularização proximal.

Embora sua formulação seja técnica, essa condição substitui uma hipótese dinâmica de difícil verificação por uma condição geométrica mais adequada ao estudo de convergência. No Capítulo 5, veremos que a hipótese $(\mathcal{H})$ será essencial para estabelecer a convergência da sequência gerada pelo algoritmo de ponto proximal.

Capítulo 3

Distâncias Proximais Generalizadas

Neste capítulo apresentamos a classe de distâncias proximais generalizadas utilizada na construção do método proximal para problemas de equilíbrio de dois níveis estudado nesta dissertação.

Os métodos proximais surgiram inicialmente nos trabalhos de Martinet [37] para desigualdades variacionais e foram posteriormente desenvolvidos por Rockafellar [50] no contexto de operadores monótonos maximais. Posteriormente, diversos autores observaram que a substituição da distância euclidiana por funções mais gerais permite incorporar informações geométricas adicionais dos problemas considerados, conduzindo frequentemente a melhores propriedades teóricas e computacionais.

No contexto de problemas de equilíbrio, as distâncias proximais desempenham papel fundamental na construção de bifunções regularizadas. A ideia central consiste em adicionar ao problema original um termo de regularização capaz de estabilizar o processo iterativo e garantir propriedades de existência e unicidade dos subproblemas resolvidos em cada etapa do algoritmo.

Neste trabalho utilizaremos a classe de distâncias proximais generalizadas introduzida por Auslender e Teboulle [7] e posteriormente estudada por Burachik e Dutta [13]. Essas distâncias abrangem como casos particulares diversas funções largamente utilizadas em Otimização Convexa, incluindo as distâncias de Bregman.

Ao longo deste capítulo apresentaremos as definições fundamentais, exemplos e propriedades que serão utilizadas na construção do método proximal desenvolvido nos capítulos posteriores.

3.1 Distâncias Proximais

Seja $S \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo. Nosso objetivo é introduzir uma classe de funções que generaliza a noção usual de distância e que será utilizada na regularização dos problemas de equilíbrio considerados neste trabalho.

Definição 3.1.1. *Uma função $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ é denominada uma **distância proximal** em relação ao conjunto S se, para cada $y \in \text{int } S$, as seguintes propriedades são satisfeitas:*

- (i) *a função $d(\cdot, y)$ é própria, semicontínua inferiormente, convexa e de classe C^1 em $\text{int } S$;*
- (ii) *$\text{dom}(d(\cdot, y)) \subset S$;*
- (iii) *$\text{dom } \partial_1 d(\cdot, y) = \text{int } S$, onde $\partial_1 d(\cdot, y)$ denota o subdiferencial de $d(\cdot, y)$ em relação à primeira variável.*

Denotaremos por $\mathcal{D}(S)$ a família de todas as distâncias proximais associadas ao conjunto S .

3.2 Distâncias Proximais Induzidas

A definição de distância proximal introduzida na seção anterior fornece apenas a estrutura básica necessária para a construção dos métodos de regularização. Entretanto, para o desenvolvimento da teoria de convergência dos algoritmos proximais, torna-se necessário associar a cada distância proximal uma segunda função que satisfaça propriedades geométricas adicionais.

A cada distância proximal $d \in \mathcal{D}(S)$ associaremos uma segunda função que desempenhará o papel de medida efetiva de afastamento entre dois pontos. Essa função será denominada distância proximal induzida.

Definição 3.2.1. *Dado $d \in \mathcal{D}(S)$. Seja $D : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ uma função tal que $\text{int } S \times \text{int } S \subseteq \text{dom } D$. Então, D é referida como a distância proximal induzida de d se as seguintes propriedades valem:*

- (H1) *para cada $x \in \text{int } S$, $D(x, \cdot)$ é contínua em $\text{int } S$;*

(H2) $D(x, x) = 0$ para todo $x \in \text{int } S$;

(H3) para todo $x \in S$ e $\alpha \in \mathbb{R}$, o conjunto $\{y \in \text{int } S; D(x, y) \leq \alpha\}$ é limitado;

(H4) para cada $x, y \in \text{int } S$, vale que

$$\langle z - x, \nabla_1 d(x, y) \rangle \leq D(z, y) - D(z, x) - \gamma D(x, y)$$

para todo $z \in S$ e algum $\gamma > 0$ fixo;

(H5) se $\{y^k\} \subset \text{int } S$ e $y^k \rightarrow y \in S$, então $D(y, y^k) \rightarrow 0$;

(H6) seja $z \in S$ e $y \in \text{int } S$, e tome $w := \alpha z + (1 - \alpha)y$ com $\alpha \in (0, 1)$. Então,

$$D(z, w) + D(w, y) \leq D(z, y);$$

(H7) se $\{x^k\}, \{y^k\} \subset \text{int } S$ são sequências tais que $\{x^k\}$ converge para x e $\{y^k\}$ converge para y , com $x \neq y$, então $\liminf_k D(x^k, y^k) > 0$.

A família de todos os pares (d, D) que satisfazem as condições acima será denotada por $\mathcal{F}(S)$. Quando $(d, D) \in \mathcal{F}(S)$, diremos que (d, D) é um **par proximal associado ao conjunto S** . As hipóteses acima possuem interpretações geométricas importantes.

A propriedade (H2) é análoga à propriedade fundamental das distâncias usuais segundo a qual a distância entre um ponto e ele próprio é nula.

A condição (H3) garante que os conjuntos de nível de D sejam limitados, propriedade essencial para a obtenção de estimativas de compacidade.

Por sua vez, a hipótese (H4) desempenha papel central na teoria desenvolvida posteriormente. Ela pode ser interpretada como uma versão generalizada da monotonicidade forte do gradiente da distância proximal.

As propriedades (H5) e (H7) estabelecem relações entre convergência topológica e convergência medida pela distância proximal induzida.

Finalmente, a condição (H6) pode ser vista como uma forma enfraquecida da desigualdade triangular, desempenhando papel fundamental nas demonstrações de convergência do método proximal.

Observação 3.2.1. A propriedade (H4) implica imediatamente que

$$\nabla_1 d(x, x) = 0, \quad \forall x \in \text{int } S.$$

De fato, tomando $z = x$ na desigualdade (H4), obtemos

$$0 \leq D(x, y) - D(x, x) - \gamma D(x, y).$$

Em particular, fazendo $y = x$, segue que

$$\langle z - x, \nabla_1 d(x, x) \rangle = 0, \quad \forall z \in S.$$

Logo, $\nabla_1 d(x, x) = 0$.

Exemplo 3.2.1. Considere $S = \mathbb{R}^n$ e a distância proximal $d(x, y) = \frac{1}{2}\|x - y\|^2$. Para essa escolha de d , a distância proximal induzida associada é dada por $D(x, y) = \frac{1}{2}\|x - y\|^2$.

De fato, como $\text{int } S = \mathbb{R}^n$, temos

$$\text{int } S \times \text{int } S = \mathbb{R}^n \times \mathbb{R}^n \subset \text{dom } D.$$

Mostremos que D satisfaz as propriedades (H1)–(H7).

(H1). Para cada $x \in \mathbb{R}^n$, a aplicação

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

é contínua em $\mathbb{R}^n$. Portanto, (H1) é satisfeita.

(H2). Para todo $x \in \mathbb{R}^n$, temos

$$D(x, x) = \frac{1}{2}\|x - x\|^2 = 0.$$

Logo, (H2) é satisfeita.

(H3). Fixe $x \in S = \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Consideremos o conjunto de nível

$$L_\alpha := \{y \in \mathbb{R}^n; D(x, y) \leq \alpha\}.$$

Se $\alpha < 0$, então $L_\alpha = \emptyset$, pois $D(x, y) \geq 0$ para todo $y \in \mathbb{R}^n$. Em particular, L_α é limitado.

Se $\alpha \geq 0$, então, para $y \in L_\alpha$,

$$\frac{1}{2}\|x - y\|^2 \leq \alpha.$$

Assim,

$$\|x - y\| \leq \sqrt{2\alpha}.$$

Portanto,

$$L_\alpha \subset B_{\sqrt{2\alpha}}[x],$$

isto é, L_α está contido em uma bola fechada de centro x e raio $\sqrt{2\alpha}$. Logo, L_α é limitado.

Portanto, (H3) é satisfeita.

(H4). Temos

$$\nabla_1 d(x, y) = x - y.$$

Assim,

$$\langle z - x, \nabla_1 d(x, y) \rangle = \langle z - x, x - y \rangle.$$

Por outro lado,

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

Como

$$z - y = (z - x) + (x - y),$$

segue que

$$\|z - y\|^2 = \|z - x\|^2 + 2\langle z - x, x - y \rangle + \|x - y\|^2.$$

Substituindo na expressão anterior, obtemos

$$D(z, y) - D(z, x) - D(x, y) = \langle z - x, x - y \rangle.$$

Logo,

$$\langle z - x, \nabla_1 d(x, y) \rangle = D(z, y) - D(z, x) - D(x, y).$$

Portanto, a desigualdade (H4) é satisfeita com $\gamma = 1$.

(H5). Seja $\{y^k\} \subset \mathbb{R}^n$ tal que $y^k \rightarrow y \in \mathbb{R}^n$. Então

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

Como $y^k \rightarrow y$, temos $\|y - y^k\| \rightarrow 0$. Consequentemente,

$$D(y, y^k) \rightarrow 0.$$

Assim, (H5) é satisfeita.

(H6). Sejam $z \in S = \mathbb{R}^n$, $y \in \text{int } S = \mathbb{R}^n$ e

$$w = \alpha z + (1 - \alpha)y, \quad \alpha \in (0, 1).$$

Então

$$z - w = z - \alpha z - (1 - \alpha)y = (1 - \alpha)(z - y)$$

e

$$w - y = \alpha z + (1 - \alpha)y - y = \alpha(z - y).$$

Logo,

$$D(z, w) + D(w, y) = \frac{1}{2}\|z - w\|^2 + \frac{1}{2}\|w - y\|^2.$$

Substituindo as expressões acima, obtemos

$$D(z, w) + D(w, y) = \frac{1}{2}(1 - \alpha)^2\|z - y\|^2 + \frac{1}{2}\alpha^2\|z - y\|^2.$$

Portanto,

$$D(z, w) + D(w, y) = \frac{1}{2}[(1 - \alpha)^2 + \alpha^2]\|z - y\|^2.$$

Como

$$(1 - \alpha)^2 + \alpha^2 = 1 - 2\alpha(1 - \alpha) \leq 1,$$

segue que

$$D(z, w) + D(w, y) \leq \frac{1}{2}\|z - y\|^2 = D(z, y).$$

Logo, (H6) é satisfeita.

(H7). Sejam $\{x^k\}, \{y^k\} \subset \mathbb{R}^n$ tais que

$$x^k \rightarrow x, \quad y^k \rightarrow y, \quad x \neq y.$$

Como D é contínua em $\mathbb{R}^n \times \mathbb{R}^n$, temos

$$D(x^k, y^k) \rightarrow D(x, y).$$

Mas

$$D(x, y) = \frac{1}{2}\|x - y\|^2 > 0,$$

pois $x \neq y$. Portanto,

$$\liminf_{k \rightarrow +\infty} D(x^k, y^k) = D(x, y) > 0.$$

Assim, (H7) é satisfeita.

Concluimos que, para $d(x, y) = \frac{1}{2}\|x - y\|^2$ e $D(x, y) = \frac{1}{2}\|x - y\|^2$, tem-se $(d, D) \in \mathcal{F}(\mathbb{R}^n)$.

Portanto, a distância euclidiana quadrática aparece como um caso particular da estrutura de distâncias proximais generalizadas.

Exemplo 3.2.2. *Seja $K = \{x \in \mathbb{R}^n; g_j(x) \leq 0, j = 1, \dots, m\}$, onde $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa e de classe C^1 para $j = 1, \dots, m$. Dado $\alpha > 0$, defina $\tilde{g}_j(x) = g_j(x) - \alpha$, $j = 1, \dots, m$. Neste caso, tome*

$$S = \{x \in \mathbb{R}^n; -\tilde{g}_j(x) > 0, j = 1, \dots, m\}.$$

Claramente, $K \subset \text{int } S = S$. Notamos que, para esta escolha de S , Auslender e Teboulle em [7] mostraram que a função

$$\varphi(x) = \begin{cases} \sum_{j=1}^m -\log(-\tilde{g}_j(x)) & \text{se } x \in \text{int } S, \\ +\infty & \text{caso contrário,} \end{cases}$$

é própria, semicontínua inferiormente, convexa e consideraram a seguinte distância proximal de Bregman associada com $\varphi_v(x) := \varphi(x) + \frac{v}{2}\|x\|^2$, $v > 0$:

$$\begin{aligned} D_{\varphi_v}(x, y) &= \varphi_v(x) - \varphi_v(y) - \langle \nabla \varphi_v(y), x - y \rangle \\ D_{\varphi_v}(x, y) &= \sum_{j=1}^m -\log(-\tilde{g}_j(x)) + \frac{v}{2}\|x\|^2 - \sum_{j=1}^m -\log(-\tilde{g}_j(y)) - \frac{v}{2}\|y\|^2 \\ &\quad - \left\langle \nabla \left(\sum_{j=1}^m -\log(-\tilde{g}_j(y)) + \frac{v}{2}\|y\|^2 \right), x - y \right\rangle \\ D_{\varphi_v}(x, y) &= \sum_{j=1}^m -\log(-\tilde{g}_j(x)) - \sum_{j=1}^m -\log(-\tilde{g}_j(y)) + \frac{v}{2}\|x\|^2 - \frac{v}{2}\|y\|^2 \\ &\quad - \left\langle \sum_{j=1}^m -\frac{\nabla \tilde{g}_j(y)}{\tilde{g}_j(y)} + vy, x - y \right\rangle \\ D_{\varphi_v}(x, y) &= \sum_{j=1}^m -\log \left(\frac{\tilde{g}_j(x)}{\tilde{g}_j(y)} \right) + \frac{v}{2}\|x\|^2 - \frac{v}{2}\|y\|^2 + \left\langle \sum_{j=1}^m \frac{\nabla \tilde{g}_j(y)}{\tilde{g}_j(y)}, x - y \right\rangle - v\langle y, x - y \rangle \\ D_{\varphi_v}(x, y) &= \sum_{j=1}^m \left(-\log \frac{\tilde{g}_j(x)}{\tilde{g}_j(y)} + \frac{\langle \nabla \tilde{g}_j(y), x - y \rangle}{\tilde{g}_j(y)} \right) + \frac{v}{2}\|x\|^2 - \frac{v}{2}\|y\|^2 - v\langle y, x - y \rangle \\ D_{\varphi_v}(x, y) &= \sum_{j=1}^m \left(-\log \frac{\tilde{g}_j(x)}{\tilde{g}_j(y)} + \frac{\langle \nabla \tilde{g}_j(y), x - y \rangle}{\tilde{g}_j(y)} \right) + \frac{v}{2}\|x - y\|^2. \end{aligned}$$

Neste caso, para nossa escolha de S associada com o conjunto dado K e definindo $d(x, y) = D(x, y) := D_{\varphi_v}(x, y)$, tem-se $(d, D) \in \mathcal{F}(S)$; veja Burachik e Dutta [13].

Verificação das propriedades:

(H1). *Para todo $x \in \text{int } S$, a função $y \mapsto D(x, y)$ é contínua em $\text{int } S$. De fato, temos que $\tilde{g}_j(y) = g_j(y) - \alpha$ é convexa e, assumindo que g_j é de classe C^1 em $\text{int } S$, segue que $\tilde{g}_j$ também é de classe C^1 e, portanto, $\nabla \tilde{g}_j$ é contínua.*

Além disso, como $\tilde{g}_j(y) < 0$ para todo $y \in \text{int } S$, as funções $y \mapsto \log(-\tilde{g}_j(y))$ e $y \mapsto \frac{1}{\tilde{g}_j(y)}$ são contínuas. Como o produto interno é contínuo e o termo quadrático $\|x - y\|^2$ também é contínuo, concluímos que $D(x, \cdot)$ é contínua.

(H2). Seja $y = x$:

$$\begin{aligned} D(x, x) &= \sum_{j=1}^m \left(-\log \frac{\tilde{g}_j(x)}{\tilde{g}_j(x)} + \frac{\langle \nabla \tilde{g}_j(x), x - x \rangle}{\tilde{g}_j(x)} \right) + \frac{\nu}{2} \|x - x\|^2 \\ D(x, x) &= \sum_{j=1}^m \left(-\log 1 + \frac{\langle \nabla \tilde{g}_j(x), 0 \rangle}{\tilde{g}_j(x)} \right) + 0 \\ D(x, x) &= \sum_{j=1}^m (-\log 1) \\ D(x, x) &= 0 \end{aligned}$$

(H3). Para $x \in S$ fixo e $\alpha \in \mathbb{R}$, considere o conjunto $L = \{y \in \text{int } S; D(x, y) \leq \alpha\}$. O termo $\frac{\nu}{2} \|x - y\|^2$ é uma parcela não negativa de $D(x, y)$. Assim,

$$\frac{\nu}{2} \|x - y\|^2 \leq D(x, y) \leq \alpha \Rightarrow \|x - y\| \leq \sqrt{\frac{2\alpha}{\nu}}.$$

Portanto, como L está contido em uma bola de raio $\sqrt{\frac{2\alpha}{\nu}}$ centrada em x , então é limitado.

(H4). Seja φ_ν a função definida anteriormente e D_{φ_ν} a distância de Bregman associada. Recordemos que

$$D_{\varphi_\nu}(x, y) = \varphi_\nu(x) - \varphi_\nu(y) - \langle \nabla \varphi_\nu(y), x - y \rangle.$$

Além disso, vale a identidade de três pontos:

$$\langle z - x, \nabla \varphi_\nu(x) - \nabla \varphi_\nu(y) \rangle = D_{\varphi_\nu}(z, y) - D_{\varphi_\nu}(z, x) - D_{\varphi_\nu}(x, y).$$

De fato,

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

Como, no presente caso, $d(x, y) = D_{\varphi_v}(x, y)$, temos

$$\nabla_1 d(x, y) = \nabla \varphi_v(x) - \nabla \varphi_v(y).$$

Logo,

$$\langle z - x, \nabla_1 d(x, y) \rangle = D(z, y) - D(z, x) - D(x, y),$$

o que implica (H4) com $\gamma = 1$.

(H5). Seja $y^k \in \text{int } S$ com $y^k \rightarrow y \in S$. Como D é contínua em $(y, \cdot)$ por (H1) e $D(y, y) = 0$ por (H2), temos $D(y, y^k) \rightarrow D(y, y) = 0$.

(H6). Sejam $z \in S$, $y \in \text{int } S$ e $w = \alpha z + (1 - \alpha)y$, com $\alpha \in (0, 1)$. Como D é a distância de Bregman associada a φ_v , temos

$$D(x, y) = \varphi_v(x) - \varphi_v(y) - \langle \nabla \varphi_v(y), x - y \rangle.$$

Somando $D(z, w)$ e $D(w, y)$ e comparando com $D(z, y)$, obtemos que a desigualdade desejada é equivalente a

$$\langle \nabla \varphi_v(w) - \nabla \varphi_v(y), z - y \rangle \geq 0.$$

Como $w = \alpha z + (1 - \alpha)y$, segue que $z - y = \frac{1}{\alpha}(w - y)$. Assim,

$$\langle \nabla \varphi_v(w) - \nabla \varphi_v(y), z - y \rangle = \frac{1}{\alpha} \langle \nabla \varphi_v(w) - \nabla \varphi_v(y), w - y \rangle \geq 0,$$

onde a última desigualdade decorre da monotonicidade do gradiente de φ_v , pois φ_v é convexa. Portanto,

$$D(z, w) + D(w, y) \leq D(z, y).$$

(H7). Sejam $x^k, y^k \in \text{int } S$ com $x^k \rightarrow x$, $y^k \rightarrow y$ e $x \neq y$. Como D é contínua (nas duas variáveis) em $\text{int } S \times \text{int } S$ (por construção), temos

$$\lim_{k \rightarrow +\infty} D(x^k, y^k) = D(x, y) > 0,$$

pois $x \neq y$ implica $D(x, y) > 0$ (a distância de Bregman é positiva para pontos distintos, devido à convexidade estrita de φ_v).

Logo, $\liminf D(x^k, y^k) = D(x, y) > 0$.

Para um melhor estudo sobre distâncias de Bregman, consulte [52] e [53].

3.2.1 Propriedades Fundamentais dos Pares Proximais

Nesta subseção estabelecemos algumas propriedades básicas dos pares proximais que serão utilizadas repetidamente nos capítulos posteriores.

Os resultados seguintes correspondem às proposições apresentadas por Burachik & Dutta [13] e constituem ferramentas fundamentais para a análise do método proximal para problemas de equilíbrio de dois níveis.

Proposição 3.2.1. *Assuma que (d, D) verifica (H6) e (H7). Se $\{x^k\} \subset S$ e $\{y^k\} \subset \text{int } S$ são sequências tais que*

$$\lim_{k \rightarrow +\infty} D(x^k, y^k) = 0$$

e uma das sequências $(\{x^k\})$ ou $(\{y^k\})$ converge, então a outra também converge para o mesmo limite.

Demonstração. Suponha por contradição, que uma das sequências, digamos, $\{y^k\}$, converge para y e que $\{x^k\}$ não converge ou converge para um limite diferente. Neste caso, existe um conjunto de índices $\{k_j\}$ e $\varepsilon > 0$ tal que $\|x^{k_j} - y^{k_j}\| > \varepsilon$. De (H6), defina $z^j := y^{k_j} + \frac{\varepsilon}{\|x^{k_j} - y^{k_j}\|}(x^{k_j} - y^{k_j})$.

z^j é um ponto no segmento entre x^{k_j} e y^{k_j} com distância ε de y^{k_j} .

$$z^j = \alpha x^{k_j} + (1 - \alpha)y^{k_j} = y^{k_j} + \alpha(x^{k_j} - y^{k_j})$$

com $\alpha = \frac{\varepsilon}{\|x^{k_j} - y^{k_j}\|} \in (0, 1)$. Então,

$$D(x^{k_j}, z^j) + D(z^j, y^{k_j}) \leq D(x^{k_j}, y^{k_j})$$

e como $D(x^{k_j}, z^j) \geq 0$

$$D(z^j, y^{k_j}) \leq D(x^{k_j}, y^{k_j}).$$

Como por hipótese, $\lim_{k \rightarrow +\infty} D(x^k, y^k) = 0$, temos que $D(z^j, y^{k_j})$ converge para 0. Notando que

$$\|z^j - y^{k_j}\| = \|y^{k_j} + \frac{\varepsilon}{\|x^{k_j} - y^{k_j}\|}(x^{k_j} - y^{k_j}) - y^{k_j}\| = \left\| \frac{\varepsilon}{\|x^{k_j} - y^{k_j}\|}(x^{k_j} - y^{k_j}) \right\| = \varepsilon \frac{\|x^{k_j} - y^{k_j}\|}{\|x^{k_j} - y^{k_j}\|} = \varepsilon$$

e que $\{y^{k_j}\}$ converge, concluímos que $\{z^j\}$ é limitada e, portanto, possui uma subsequência convergente. Chamemos de z o limite desta subsequência. Sem perda de generalidade, também denotemos a subsequência como $\{z^j\}$. No total, temos

$$\|z - y\| = \varepsilon,$$

$$\liminf_j D(z^j, y^{k_j}) = 0,$$

$$\lim_j z^j = z,$$

$$\lim_j y^{k_j} = y.$$

A lista de fatos acima contradiz (H7), que implica que se $z^j \rightarrow z$, $y^{k_j} \rightarrow y$, com $z \neq y$, então necessariamente $\liminf_j D(z^j, y^{k_j}) > 0$. Se assumirmos que $\{x^k\}$ converge e $\{y^k\}$ não converge ou converge para um limite diferente, então um argumento semelhante ao anterior leva a uma contradição. Portanto, a conclusão é verdadeira.

□

Proposição 3.2.2. *Assuma que (d, D) verifica (H6) e (H7). Se $\{x^k\} \subset S$ e $\{y^k\} \subset \text{int } S$ são sequências tais que*

$$\lim_{k \rightarrow +\infty} D(x^k, y^k) = 0,$$

e que uma das sequências $\{x^k\}$ ou $\{y^k\}$ é limitada, então se verifica o seguinte:

(a) *A outra sequência também é limitada.*

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

Demonstração. Assuma que (a) não seja verdadeiro. Por exemplo, se $\{y^k\}$ é limitada e a sequência $\{x^k\}$ é ilimitada, então existe um conjunto infinito de índices J tal que $\|x^k\| \rightarrow +\infty$ quando $k \rightarrow +\infty$ com $k \in J$. Além disso, como $\{y^k\}$ é limitada, existe um conjunto infinito de índices J' com $J' \subset J$ tal que $\{y^k\}$ converge para um limite $\bar{y}$ quando $k \rightarrow +\infty$ com $k \in J'$. Temos também $D(x^k, y^k) \rightarrow 0$ quando $k \rightarrow +\infty$ e $k \in J'$. Agora, pela proposição 3.2.1 temos que $x^k \rightarrow \bar{y}$ quando $k \rightarrow +\infty$ e $k \in J'$, contradição. Portanto se y^k é limitada, x^k também é limitada.

Para (b), assumamos que $(x^k - y^k)$ não converge para zero. Então existe um conjunto infinito de índices J e $\varepsilon > 0$ tal que

$$\|x^k - y^k\| \geq \varepsilon, \quad k \in J. \quad (3.1)$$

Agora, usando (a), temos que ambas as sequências $\{x^k\}$ e $\{y^k\}$ são limitadas, e assim podemos encontrar um conjunto infinito de índices $J' \subset J$ tal que $x^k \rightarrow x$ e $y^k \rightarrow y$ quando $k \rightarrow +\infty$ e $k \in J'$. Pela proposição 3.2.1 temos que $x = y$. Isto claramente contradiz (3.1). Portanto, temos o resultado. □

Capítulo 4

Bifunções Regularizadas

Neste capítulo, estudaremos as bifunções regularizadas obtidas a partir de distâncias proximais generalizadas. Para isso, apresentaremos as hipóteses estruturais sobre bifunções de equilíbrio e mostraremos que a regularização preserva propriedades fundamentais, como convexidade, semicontinuidade e monotonicidade. Esses resultados serão essenciais para garantir a boa definição dos subproblemas que aparecem no método proximal estudado no capítulo seguinte.

4.1 Hipóteses para bifunções de equilíbrio

Seja $K \subset \mathbb{R}^n$ um conjunto não vazio, fechado e convexo. Ao longo do trabalho, de maneira mais geral, estudaremos bifunções

$$\psi : K \times K \longrightarrow \mathbb{R} \cup \{+\infty\},$$

as quais serão utilizadas para modelar o problema de equilíbrio no nível inferior e, posteriormente, os problemas regularizados associados ao método proximal.

A introdução do valor $+\infty$ é natural no contexto da Análise Convexa, pois permite incorporar restrições e tratar, em uma mesma estrutura, funções que não são necessariamente finitas em todo o seu domínio. No caso particular em que ψ assume apenas valores reais, recuperamos a classe de bifunções considerada anteriormente.

A fim de garantir uma estrutura adequada para a análise teórica e para os resultados de existência e convergência apresentados adiante, retomamos algumas hipóteses já introduzidas no Capítulo 2. Mais precisamente, as condições de normalização, convexidade, semicontinuidade inferior e semicontinuidade superior correspondem às hipóteses

(P1) – (P3), enquanto a monotonicidade foi apresentada na Definição 2.3.1. Para efeito de notação e em consonância com a formulação utilizada nos próximos capítulos, passaremos a denotar essas condições por (L1) – (L4). Além delas, introduziremos uma condição coerciva adicional, denotada por (L5). Assim, para $x, y \in K$, consideraremos as seguintes hipóteses:

$$(L1) \quad \psi(x, x) = 0;$$

$$(L2) \quad \psi(\cdot, y) : K \longrightarrow \mathbb{R} \text{ é semicontínua superiormente};$$

$$(L3) \quad \psi(x, \cdot) : K \longrightarrow \mathbb{R} \text{ é convexa e semicontínua inferiormente.}$$

$$(L4) \quad \psi(x, y) + \psi(y, x) \leq 0;$$

$$(L5) \quad \text{para qualquer sequência } \{y^k\} \subset K \text{ com } \lim_{k \rightarrow +\infty} \|y^k\| = \infty, \text{ existe } u \in K \text{ e } k_0 \in \mathbb{N} \text{ tal que}$$

$$\psi(y^k, u) \leq 0, \quad k \geq k_0.$$

Observação 4.1.1. *Até onde sabemos, a hipótese (L5) foi considerada pela primeira vez por Iusem, Kassay e Sosa em [28]. É importante destacar que Muu e Oettli em [44] e Blum e Oettli em [11] usaram uma condição que é suficiente para (L5), a fim de estabelecer resultados de existência para o problema de equilíbrio.*

O próximo teorema foi apresentado por Ky Fan em [23].

Teorema 14. *Assuma que ψ satisfaz (L1)-(L3) e K é um conjunto compacto. Então, $S(\psi, K)$ é não vazio.*

Demonstração. Considere o problema de equilíbrio associado a ψ , isto é, encontrar $\bar{x} \in K$ tal que $\psi(\bar{x}, y) \geq 0$, para todo $y \in K$. Para cada $y \in K$, definimos

$$F(y) = \{x \in K; \psi(x, y) \geq 0\}.$$

Queremos mostrar que

$$\bigcap_{y \in K} F(y) \neq \emptyset.$$

Primeiramente, para cada $y \in K$, o conjunto $F(y)$ é fechado em K . De fato, pela hipótese (L2), a função $x \mapsto \psi(x, y)$ é semicontínua superiormente em K . Portanto, o conjunto

de nível superior $\{x \in K; \psi(x, y) \geq 0\}$ é fechado. Como K é compacto, segue que $F(y)$ é compacto.

Agora, mostraremos que a aplicação F satisfaz a propriedade KKM. Sejam $y_1, \dots, y_m \in K$. Como K é convexo, temos $\text{conv}\{y_1, \dots, y_m\} \subset K$.

Afirmamos que $\text{conv}\{y_1, \dots, y_m\} \subset \bigcup_{i=1}^m F(y_i)$.

Suponha, por contradição, que exista $z \in \text{conv}\{y_1, \dots, y_m\}$ tal que

$$z \notin \bigcup_{i=1}^m F(y_i).$$

Então

$$z \notin F(y_i), \quad i = 1, \dots, m.$$

Logo,

$$\psi(z, y_i) < 0, \quad i = 1, \dots, m.$$

Como $z \in \text{conv}\{y_1, \dots, y_m\}$, existem $\lambda_1, \dots, \lambda_m \geq 0$, com $\sum_{i=1}^m \lambda_i = 1$, tais que

$$z = \sum_{i=1}^m \lambda_i y_i.$$

Pela convexidade de $\psi(z, \cdot)$, dada em (L3), obtemos

$$\psi(z, z) = \psi\left(z, \sum_{i=1}^m \lambda_i y_i\right) \leq \sum_{i=1}^m \lambda_i \psi(z, y_i).$$

Como $\psi(z, y_i) < 0$ para todo i , segue que

$$\sum_{i=1}^m \lambda_i \psi(z, y_i) < 0.$$

Portanto,

$$\psi(z, z) < 0,$$

o que contradiz (L1), pois

$$\psi(z, z) = 0.$$

Assim,

$$\text{conv}\{y_1, \dots, y_m\} \subset \bigcup_{i=1}^m F(y_i).$$

Logo, F é uma aplicação KKM.

Pelo Lema 2 de Fan-KKM, como K é compacto e cada $F(y)$ é fechado em K , segue que

$$\bigcap_{y \in K} F(y) \neq \emptyset.$$

Portanto, existe $\bar{x} \in K$ tal que $\bar{x} \in F(y)$, para todo $y \in K$.

Pela definição de $F(y)$, isso significa que $\psi(\bar{x}, y) \geq 0$, para todo $y \in K$. Consequentemente, $\bar{x} \in S(\psi, K)$. Logo, $S(\psi, K) \neq \emptyset$.

□

Teorema 15. *Assuma que ψ satisfaz (L1)-(L5). Então, $S(\psi, K)$ é não vazio.*

Demonstração. Veja Iusem, Kassay e Sosa em [28, Thm. 4.3]

□

4.2 Bifunções regularizadas via distâncias proximais generalizadas.

De acordo com Iusem [27], a ideia de regularização surgiu em conexão com problema mal posto. Sejam $f, h : K \times K \rightarrow \mathbb{R}$ tais que as suposições (L1)-(L3) valem. Tome uma distância proximal generalizada d com respeito a S tal que $K \subset \text{int } S$. Fixe $\bar{x} \in K$, $\lambda, \varepsilon > 0$ e considere a bifunção $\bar{f} : K \times K \rightarrow \mathbb{R}$ definida por

$$\bar{f}(x, y) = g(x, y) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}), y - x \rangle, \quad (4.1)$$

onde $g := \varepsilon f + h$ e $\nabla_1 d(x, \bar{x})$ denota o gradiente de $d(\cdot, \bar{x})$ em x . Denotamos por $S(\bar{f}, K)$ o conjunto solução do problema de equilíbrio associado com $\bar{f}$.

Para garantir a boa definição dos problemas de equilíbrio regularizados, introduzimos a hipótese a seguir, que pode ser interpretada como uma condição de coercividade assintótica.

Hipótese 2. *Para cada sequência $\{y^k\} \subset K$ tal que $\lim_{k \rightarrow +\infty} \|y^k\| = +\infty$,*

$$\liminf_{k \rightarrow +\infty} [g(\bar{x}, y^k) + \frac{1}{\lambda} \langle \nabla_1 d(y^k, \bar{x}), y^k - \bar{x} \rangle] > 0.$$

Essa hipótese é inspirada em Burachik e Kassay [14, Condição P6]. No referido trabalho, os autores mostram que, sob essa condição, a bifunção regularizada satisfaz uma condição coerciva do tipo (L5), o que permite obter a existência de solução para o problema de equilíbrio regularizado.

O próximo lema mostra que a bifunção regularizada preserva as condições (L1)–(L4) e, sob uma hipótese adicional, também satisfaz (L5).

Lema 3. *Suponha que f, h satisfaçam (L1)-(L4). Então, $\bar{f}$ satisfaz (L1)-(L4). Além disso, se a Hipótese 2 vale, então $\bar{f}$ satisfaz (L5).*

Demonstração. Primeiro, verificaremos se $\bar{f}$ satisfaz (L1)-(L4).

Tome $y = x$, então

$$\bar{f}(x, x) = g(x, x) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}), x - x \rangle = g(x, x) = \varepsilon f(x, x) + h(x, x) + 0 = 0$$

logo, (L1) é satisfeita.

A função $x \mapsto \langle \nabla_1 d(x, \bar{x}), y - x \rangle$ é contínua, pois pela definição 3.1.1, $\nabla_1 d(\cdot, \bar{x})$ é de classe C^1 em $\text{int } S$, como $K \subset \text{int } S$ e o produto interno é uma bifunção contínua no $\mathbb{R}^n$, segue que função é contínua. Então $\bar{f}(\cdot, y)$ é semicontínua superiormente, satisfazendo (L2).

Como a função $y \mapsto \langle \nabla_1 d(x, \bar{x}), y - x \rangle$ é contínua e convexa, então $\bar{f}(x, \cdot)$ é convexa e semicontínua inferiormente, satisfazendo (L3).

Mostremos que $\bar{f}$ satisfaz (L4). De (4.1) obtemos

$$\begin{aligned} \bar{f}(x, y) + \bar{f}(y, x) &= g(x, y) + g(y, x) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}), y - x \rangle + \frac{1}{\lambda} \langle \nabla_1 d(y, \bar{x}), x - y \rangle \\ &= g(x, y) + g(y, x) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}), y - x \rangle - \frac{1}{\lambda} \langle \nabla_1 d(y, \bar{x}), y - x \rangle \\ &= g(x, y) + g(y, x) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}) - \nabla_1 d(y, \bar{x}), y - x \rangle \\ &= g(x, y) + g(y, x) - \frac{1}{\lambda} \langle \nabla_1 d(y, \bar{x}) - \nabla_1 d(x, \bar{x}), y - x \rangle \end{aligned}$$

A bifunção g é monótona, pois

$$g(x, y) + g(y, x) = \varepsilon (f(x, y) + f(y, x)) + h(x, y) + h(y, x) \leq 0.$$

Além disso, como $d(\cdot, \bar{x})$ é convexa e diferenciável, segue do Teorema 4 que

$$\langle \nabla_1 d(y, \bar{x}) - \nabla_1 d(x, \bar{x}), y - x \rangle \geq 0.$$

Portanto,

$$\bar{f}(x, y) + \bar{f}(y, x) = g(x, y) + g(y, x) - \frac{1}{\lambda} \langle \nabla_1 d(y, \bar{x}) - \nabla_1 d(x, \bar{x}), y - x \rangle \leq 0,$$

satisfazendo (L4).

Suponha agora que vale a Hipótese 2, mostraremos que $\bar{f}$ satisfaz (L5). Seja $\{y^k\} \subset K$ tal que $\lim_{k \rightarrow +\infty} \|y^k\| = +\infty$.

Temos

$$\bar{f}(\mathbf{y}^k, \bar{\mathbf{x}}) = g(\mathbf{y}^k, \bar{\mathbf{x}}) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \bar{\mathbf{x}} - \mathbf{y}^k \rangle.$$

Como g é monótona, então $g(\mathbf{y}^k, \bar{\mathbf{x}}) \leq -g(\bar{\mathbf{x}}, \mathbf{y}^k)$. Pela Hipótese 2, existe $k_0 \in \mathbb{N}$ tal que, para todo $k \geq k_0$,

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \bar{\mathbf{x}} \rangle > 0.$$

Portanto,

$$\bar{f}(\mathbf{y}^k, \bar{\mathbf{x}}) \leq - \left[g(\bar{\mathbf{x}}, \mathbf{y}^k) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \bar{\mathbf{x}} \rangle \right].$$

Assim,

$$\bar{f}(\mathbf{y}^k, \bar{\mathbf{x}}) \leq 0, \quad k \geq k_0.$$

Como $\bar{\mathbf{x}} \in K$, concluímos que $\bar{f}$ satisfaz (L5).

□

Corolário 2. *Se f, h satisfazem (L1)-(L4) e valem qualquer uma das duas condições seguintes,*

(i) K é limitado,

(ii) d é 1-coercivo, i.e.,

$$\lim_{\|\mathbf{x}\| \rightarrow +\infty} \frac{d(\mathbf{x}, \bar{\mathbf{x}})}{\|\mathbf{x}\|} = +\infty,$$

então a bifunção regularizada

$$\bar{f}(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y}) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{x}, \bar{\mathbf{x}}), \mathbf{y} - \mathbf{x} \rangle,$$

com $g := \varepsilon f + h$, satisfaz a condição (L5).

Demonstração. Se K é limitado, então não existem sequências em K com norma tendendo ao infinito. Logo, a Hipótese 2 é satisfeita por vacuidade. Pelo Lema 3, segue que $\bar{f}$ satisfaz (L5).

Suponha agora que d é 1-coerciva. Fixado $\bar{\mathbf{x}} \in K$, mostramos inicialmente que

$$\text{dom } \partial g(\bar{\mathbf{x}}, \cdot) \cap K \neq \emptyset.$$

De fato, como $g(\bar{\mathbf{x}}, \cdot)$ é própria, convexa e semicontínua inferiormente em $\mathbb{R}^n$, segue do Teorema 5 que seu subdiferencial é não vazio em pontos do interior relativo de seu domínio. Como K é convexo e fechado, existe $\hat{\mathbf{x}} \in K$ tal que

$$\partial g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) \neq \emptyset.$$

Seja $\mathbf{v} \in \partial g(\bar{\mathbf{x}}, \hat{\mathbf{x}})$. Então, pela definição de subgradiente,

$$g(\bar{\mathbf{x}}, \mathbf{y}) \geq g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) + \langle \mathbf{v}, \mathbf{y} - \hat{\mathbf{x}} \rangle, \quad \forall \mathbf{y} \in \mathbf{K}.$$

Seja $\{\mathbf{y}^k\} \subset \mathbf{K}$ tal que $\|\mathbf{y}^k\| \rightarrow +\infty$. Então,

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) \geq g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) + \langle \mathbf{v}, \mathbf{y}^k - \hat{\mathbf{x}} \rangle.$$

Pela desigualdade de Cauchy-Schwarz,

$$\langle \mathbf{v}, \mathbf{y}^k - \hat{\mathbf{x}} \rangle \geq -\|\mathbf{v}\| \|\mathbf{y}^k - \hat{\mathbf{x}}\|,$$

e portanto

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) \geq g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) - \|\mathbf{v}\| \|\mathbf{y}^k - \hat{\mathbf{x}}\|. \quad (4.2)$$

Por outro lado, como $\mathbf{x} \mapsto d(\mathbf{x}, \bar{\mathbf{x}})$ é convexa, temos

$$d(\hat{\mathbf{x}}, \bar{\mathbf{x}}) \geq d(\mathbf{y}^k, \bar{\mathbf{x}}) + \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \hat{\mathbf{x}} - \mathbf{y}^k \rangle,$$

isto é,

$$\langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \hat{\mathbf{x}} \rangle \geq d(\mathbf{y}^k, \bar{\mathbf{x}}) - d(\hat{\mathbf{x}}, \bar{\mathbf{x}}). \quad (4.3)$$

Por (4.2) e (4.3), obtemos

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \hat{\mathbf{x}} \rangle \geq g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) - \|\mathbf{v}\| \|\mathbf{y}^k - \hat{\mathbf{x}}\| + \frac{1}{\lambda} (d(\mathbf{y}^k, \bar{\mathbf{x}}) - d(\hat{\mathbf{x}}, \bar{\mathbf{x}})).$$

Reorganizando,

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \hat{\mathbf{x}} \rangle \geq g(\bar{\mathbf{x}}, \hat{\mathbf{x}}) + \|\mathbf{y}^k - \hat{\mathbf{x}}\| \left(-\|\mathbf{v}\| + \frac{1}{\lambda} \frac{d(\mathbf{y}^k, \bar{\mathbf{x}})}{\|\mathbf{y}^k - \hat{\mathbf{x}}\|} \right) - \frac{1}{\lambda} d(\hat{\mathbf{x}}, \bar{\mathbf{x}}).$$

Como $\|\mathbf{y}^k\| \rightarrow +\infty$, também $\|\mathbf{y}^k - \hat{\mathbf{x}}\| \rightarrow +\infty$. Pela 1-coercividade de d , segue que

$$\frac{d(\mathbf{y}^k, \bar{\mathbf{x}})}{\|\mathbf{y}^k - \hat{\mathbf{x}}\|} \rightarrow +\infty.$$

Logo,

$$-\|\mathbf{v}\| + \frac{1}{\lambda} \frac{d(\mathbf{y}^k, \bar{\mathbf{x}})}{\|\mathbf{y}^k - \hat{\mathbf{x}}\|} \rightarrow +\infty,$$

e conseqüentemente

$$g(\bar{\mathbf{x}}, \mathbf{y}^k) + \frac{1}{\lambda} \langle \nabla_1 d(\mathbf{y}^k, \bar{\mathbf{x}}), \mathbf{y}^k - \hat{\mathbf{x}} \rangle \rightarrow +\infty.$$

Isso mostra que a Hipótese 2 é satisfeita. Portanto, pelo Lema 3, $\bar{\mathbf{f}}$ satisfaz (L5). $\square$

O próximo Teorema é crucial para garantir que o problema regularizado f_k definido em (5.1) tem uma solução única em cada iteração do Algoritmo 1. Isso assegura que o algoritmo está bem definido e que cada iteração está univocamente definida, o que é essencial para a análise de convergência subsequente.

Teorema 16. *Sob as hipóteses do Lema 3, segue que $S(\bar{f}, K)$ é não vazio. Além disso, se $\nabla_1 d(\cdot, \bar{x})$ é estritamente monótona, então $S(\bar{f}, K)$ tem um único elemento.*

Demonstração. Suponha que f, h satisfaçam (L1)-(L4) e que a Hipótese 2 é válida, então pelo Lema 3 a bifunção regularizada $\bar{f}$ satisfaz as condições (L1)-(L5).

Como $\bar{f}$ satisfaz (L1)-(L5), pelo Teorema 15 segue que $S(\bar{f}, K) \neq \emptyset$.

Sejam $x_1, x_2 \in S(\bar{f}, K)$ duas soluções. Por definição, para todo $y \in K$

$$\bar{f}(x_1, y) \geq 0, \quad \bar{f}(x_2, y) \geq 0.$$

Em particular, tomando $y = x_2$ na primeira e $y = x_1$ na segunda

$$\bar{f}(x_1, x_2) \geq 0, \quad \bar{f}(x_2, x_1) \geq 0. \quad (4.4)$$

Pelo Lema 3, $\bar{f}$ satisfaz (L4), ou seja,

$$\bar{f}(x_1, x_2) + \bar{f}(x_2, x_1) \leq 0. \quad (4.5)$$

Combinando (4.4) e (4.5), obtemos

$$0 \leq \bar{f}(x_1, x_2) + \bar{f}(x_2, x_1) \leq 0 \Rightarrow \bar{f}(x_1, x_2) + \bar{f}(x_2, x_1) = 0.$$

Como ambos os termos são não negativos, isto força

$$\bar{f}(x_1, x_2) = 0 \quad \text{e} \quad \bar{f}(x_2, x_1) = 0. \quad (4.6)$$

Lembrando que $\bar{f}(x, y) = g(x, y) + \frac{1}{\lambda} \langle \nabla_1 d(x, \bar{x}), y - x \rangle$, temos

$$0 = \bar{f}(x_1, x_2) = g(x_1, x_2) + \frac{1}{\lambda} \langle \nabla_1 d(x_1, \bar{x}), x_2 - x_1 \rangle, \quad (4.7)$$

$$0 = \bar{f}(x_2, x_1) = g(x_2, x_1) + \frac{1}{\lambda} \langle \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle. \quad (4.8)$$

De (4.7) e (4.8), obtemos

$$0 = g(x_1, x_2) + g(x_2, x_1) + \frac{1}{\lambda} [\langle \nabla_1 d(x_1, \bar{x}), x_2 - x_1 \rangle + \langle \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle]$$

$$0 = g(x_1, x_2) + g(x_2, x_1) - \frac{1}{\lambda} [\langle \nabla_1 d(x_1, \bar{x}) - \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle] \quad (4.9)$$

Como f e h satisfazem (L4), temos que g também é monótona:

$$g(x_1, x_2) + g(x_2, x_1) \leq 0. \quad (4.10)$$

Substituindo (4.10) em (4.9), segue

$$0 \leq -\frac{1}{\lambda} \langle \nabla_1 d(x_1, \bar{x}) - \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle.$$

Multiplicando por $-\lambda$ (lembrando que $\lambda > 0$):

$$\langle \nabla_1 d(x_1, \bar{x}) - \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle \leq 0. \quad (4.11)$$

Por hipótese, $\nabla_1 d(\cdot, \bar{x})$ é estritamente monótona, o que significa:

$$\langle \nabla_1 d(x_1, \bar{x}) - \nabla_1 d(x_2, \bar{x}), x_1 - x_2 \rangle > 0, \quad \text{se } x_1 \neq x_2.$$

A desigualdade (4.11) contradiz essa propriedade a menos que $x_1 = x_2$.

Portanto, $x_1 = x_2$.

□

Capítulo 5

Método Proximal para Problemas de Equilíbrio de Dois Níveis

5.1 Distância proximal generalizada com método de regularização para resolver PEBs

O Problema de Equilíbrio de Dois Níveis apresenta uma estrutura hierárquica que torna sua resolução significativamente mais desafiadora do que nos problemas de equilíbrio clássicos. A principal dificuldade decorre do fato de que o conjunto viável do problema de nível superior é definido implicitamente como o conjunto solução de um problema de equilíbrio de nível inferior. Essa dependência entre os dois níveis introduz não linearidades adicionais e dificulta a aplicação direta de métodos tradicionais de otimização e de desigualdades variacionais.

Nesse contexto, métodos de regularização proximal constituem uma ferramenta poderosa para o desenvolvimento de algoritmos iterativos com boas propriedades de estabilidade e convergência. A ideia fundamental desses métodos consiste em substituir, a cada iteração, o problema original por um subproblema regularizado, cuja estrutura é mais favorável do ponto de vista analítico e computacional.

A origem dessa abordagem remonta ao método proximal introduzido por Martinet [37] no estudo de desigualdades variacionais, posteriormente generalizado por Rockafellar [50] para operadores monótonos maximais no contexto do *proximal point algorithm*. Desde então, métodos proximais têm desempenhado papel central em análise convexa, otimização

e teoria de operadores monótonos.

Em formulações mais gerais, a regularização proximal não precisa ser baseada exclusivamente na distância euclidiana. Trabalhos de Auslender e Teboulle [7] mostraram que distâncias proximais generalizadas, especialmente aquelas induzidas por distâncias de Bregman, fornecem maior flexibilidade e permitem explorar estruturas geométricas específicas do problema.

No contexto de problemas de equilíbrio de dois níveis, Bento, Cruz Neto, Lopes, Soares Jr. e Soubeyran [10] propuseram um método proximal baseado em distâncias generalizadas e em uma regularização adequada da bifunção de equilíbrio. Essa abordagem permite construir uma sequência iterativa cujos pontos são obtidos como soluções de problemas de equilíbrio regularizados, preservando propriedades fundamentais como monotonicidade, existência e unicidade das soluções auxiliares.

Neste capítulo, apresentamos o método ponto proximal considerado em [10] e estabelecemos seu principal resultado de convergência. Inicialmente, descrevemos o algoritmo iterativo e introduzimos as hipóteses necessárias para sua boa definição. Em seguida, desenvolvemos os resultados auxiliares que permitem provar que a sequência gerada converge para uma solução do problema de equilíbrio de dois níveis.

5.1.1 Algoritmo de ponto proximal

Apresentamos agora um algoritmo de ponto proximal para a resolução do Problema de Equilíbrio de Dois Níveis (2.9). A abordagem proposta por Bento, Cruz Neto, Lopes, Soares Jr. e Soubeyran [10] considera algumas ideias apresentadas por Attouch, Czarnecki e Peypouquet [6] e por Chbani e Riahi [17], no contexto de métodos de prox-penalização, para desenvolver um algoritmo de ponto proximal com distâncias generalizadas aplicado a problemas de equilíbrio de dois níveis, obtendo resultados de convergência para bifunções monótonas.

Nesse contexto, o método considerado consiste em construir, a cada iteração, um problema de equilíbrio regularizado por meio de uma distância proximal generalizada. Mais precisamente, a sequência gerada pelo algoritmo é obtida resolvendo problemas auxiliares cuja bifunção combina o problema de nível inferior, o critério de nível superior e um termo proximal. Essa formulação permite estudar a boa definição do método e estabelecer propriedades de convergência sob hipóteses adequadas de monotonicidade e

regularidade.

A seguir, descrevemos formalmente o algoritmo de ponto proximal e, posteriormente, estabelecemos seus principais resultados de convergência.

Algoritmo 1.

Tome duas seqüências de números reais positivos $\{\varepsilon_k\}$ e $\{\lambda_k\}$ tais que $\lim_{k \rightarrow +\infty} \varepsilon_k = +\infty$, $\liminf_{k \rightarrow +\infty} \lambda_k > 0$ e considere a bifunção

$$f_k(x, y) = \varepsilon_k f(x, y) + h(x, y) + \frac{1}{\lambda_k} \langle \nabla_1 d(x, x^k), y - x \rangle, \quad x, y \in K. \quad (5.1)$$

INICIALIZAÇÃO. Escolha um ponto inicial $x^0 \in K$.

PASSO ITERATIVO. Dado x^k , tome como a próxima iteração $x^{k+1} \in K$ tal que

$$x^{k+1} \in S(f_k, K). \quad (5.2)$$

CRITÉRIO DE PARADA. Se $x^{k+1} = x^k$ e $x^k \in S(f, K)$, PARE.

Em seguida, apresentamos um resultado auxiliar clássico, cuja prova pode ser encontrada em [48, Lemma 9, p. 49]. Para fins de completude do texto, incluiremos sua demonstração.

Lema 4. *Sejam $\{\xi_k\}$ e $\{\gamma_k\}$ seqüências não negativas de números reais satisfazendo*

$$(a) \quad \xi_{k+1} \leq \xi_k + \gamma_k,$$

$$(b) \quad \sum_{k=0}^{\infty} \gamma_k < +\infty.$$

Então, a seqüência $\{\xi_k\}$ converge.

Demonstração. Defina

$$u_k := \xi_k + \sum_{i=k}^{\infty} \gamma_i. \quad (5.3)$$

Calculemos $u_{k+1} - u_k$:

$$u_{k+1} - u_k = \left(\xi_{k+1} + \sum_{i=k+1}^{\infty} \gamma_i \right) - \left(\xi_k + \sum_{i=k}^{\infty} \gamma_i \right) = \xi_{k+1} - \xi_k - \gamma_k.$$

Pela hipótese (a), $\xi_{k+1} - \xi_k \leq \gamma_k$. Portanto

$$u_{k+1} - u_k \leq \gamma_k - \gamma_k = 0.$$

Logo, $u_{k+1} \leq u_k$ para todo k , ou seja, $\{u_k\}$ é uma seqüência não crescente. Como $\xi_k \geq 0$ e $\sum_{i=k}^{\infty} \gamma_i \geq 0$ (pois γ_k é uma seqüência não negativa), temos $u_k \geq 0$ para todo k . Assim,

$\{u_k\}$ é uma sequência não crescente e limitada inferiormente (por 0). Portanto ela converge para algum limite $u^* \geq 0$.

De (5.3), temos:

$$\xi_k = u_k - \sum_{i=k}^{\infty} \gamma_i \quad (5.4)$$

Como $\sum_{i=0}^{\infty} \gamma_i < +\infty$, segue que $\sum_{i=k}^{\infty} \gamma_i \rightarrow 0$ quando $k \rightarrow \infty$.

Tomando o limite em (5.4)

$$\lim_{k \rightarrow +\infty} \xi_k = \lim_{k \rightarrow +\infty} u_k - \lim_{k \rightarrow +\infty} \sum_{i=k}^{\infty} \gamma_i = u^* - 0 = u^*.$$

Portanto, $\{\xi_k\}$ converge para u^* (um número real não negativo). $\square$

No que segue, dado $z \in S(f, K)$, denotaremos por f_z e h_z as funções $f_z(y) := f(z, y)$ e $h_z(y) := h(z, y)$ para $y \in K$. Além disso, utilizaremos as notações δ_Ω , $\mathcal{N}_\Omega$ e σ_Ω para a função indicadora, o cone normal e a função suporte de um conjunto Ω , conforme definidas no Capítulo 1.

Lema 5. Tome $z \in S(f, K)$, $w \in \partial(h_z + \delta_{S(f, K)})(z)$ e $p \in \mathcal{N}_{S(f, K)}(z)$ tal que $w - p \in \partial h_z(z)$. Então,

$$\begin{aligned} D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]. \end{aligned} \quad (5.5)$$

Demonstração. Como $x^{k+1} \in S(f_k, K)$, por definição temos, para todo $y \in K$:

$$f_k(x^{k+1}, y) = \varepsilon_k f(x^{k+1}, y) + h(x^{k+1}, y) + \frac{1}{\lambda_k} \langle \nabla_1 d(x^{k+1}, x^k), y - x^{k+1} \rangle \geq 0.$$

Em particular, tomando $y = z$ (note que $z \in S(f, K) \subset K$):

$$\varepsilon_k f(x^{k+1}, z) + h(x^{k+1}, z) + \frac{1}{\lambda_k} \langle \nabla_1 d(x^{k+1}, x^k), z - x^{k+1} \rangle \geq 0,$$

Logo

$$\lambda_k \varepsilon_k f(x^{k+1}, z) + \lambda_k h(x^{k+1}, z) + \langle \nabla_1 d(x^{k+1}, x^k), z - x^{k+1} \rangle \geq 0. \quad (5.6)$$

A propriedade (H4) da distância proximal induzida nos dá, para quaisquer $z, x^{k+1}, x^k \in \text{int } S$ e $\gamma > 0$:

$$\langle z - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle \leq D(z, x^k) - D(z, x^{k+1}) - \gamma D(x^{k+1}, x^k),$$

$$-\langle z - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle \geq -D(z, x^k) + D(z, x^{k+1}) + \gamma D(x^{k+1}, x^k).$$

Substituindo em (5.6)

$$\lambda_k \varepsilon_k f(x^{k+1}, z) + \lambda_k h(x^{k+1}, z) \geq -D(z, x^k) + D(z, x^{k+1}) + \gamma D(x^{k+1}, x^k). \quad (5.7)$$

A monotonicidade implica

$$f(x^{k+1}, z) \leq -f(z, x^{k+1}) \quad \text{e} \quad h(x^{k+1}, z) \leq -h(z, x^{k+1}).$$

Substituindo em (5.7)

$$-\lambda_k \varepsilon_k f(z, x^{k+1}) - \lambda_k h(z, x^{k+1}) \geq -D(z, x^k) + D(z, x^{k+1}) + \gamma D(x^{k+1}, x^k),$$

portanto

$$D(z, x^{k+1}) + \frac{1}{2} \lambda_k \varepsilon_k f(z, x^{k+1}) \leq D(z, x^k) - \frac{1}{2} \lambda_k \varepsilon_k f(z, x^{k+1}) - \lambda_k h(z, x^{k+1}) - \gamma D(x^{k+1}, x^k). \quad (5.8)$$

Por hipótese, $w - p \in \partial h_z(z)$. Pela definição 1.2.7 de subgradiente

$$h(z, x^{k+1}) \geq h(z, z) + \langle w - p, x^{k+1} - z \rangle.$$

Como $h(z, z) = 0$ (por (L1))

$$h(z, x^{k+1}) \geq \langle w - p, x^{k+1} - z \rangle,$$

multiplicando por $-\lambda_k$

$$-\lambda_k h(z, x^{k+1}) \leq -\lambda_k \langle w - p, x^{k+1} - z \rangle = \lambda_k \langle w - p, z - x^{k+1} \rangle.$$

Substituindo em (5.8)

$$\begin{aligned} D(z, x^{k+1}) + \frac{1}{2} \lambda_k \varepsilon_k f(z, x^{k+1}) &\leq D(z, x^k) - \frac{1}{2} \lambda_k \varepsilon_k f(z, x^{k+1}) + \lambda_k \langle w - p, z - x^{k+1} \rangle - \gamma D(x^{k+1}, x^k) \\ &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle - \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) \\ &\quad - \lambda_k \langle p, z - x^{k+1} \rangle. \end{aligned} \quad (5.9)$$

Observe que

$$\lambda_k \langle p, x^{k+1} - z \rangle = \frac{\lambda_k \varepsilon_k}{2} \left[\left\langle \frac{2p}{\varepsilon_k}, x^{k+1} \right\rangle - \left\langle \frac{2p}{\varepsilon_k}, z \right\rangle - f(z, x^{k+1}) \right] + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}).$$

Substituindo em (5.9)

$$\begin{aligned} D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[\left\langle \frac{2p}{\varepsilon_k}, x^{k+1} \right\rangle - \left\langle \frac{2p}{\varepsilon_k}, z \right\rangle - f(z, x^{k+1}) \right] \end{aligned} \quad (5.10)$$

Lembremos que $f_z(x) = f(z, x)$, assim, pela definição 1.2.14 e pelo Teorema 6

$$\left\langle \frac{2p}{\varepsilon_k}, x^{k+1} \right\rangle \leq f_z(x^{k+1}) + f_z^* \left(\frac{2p}{\varepsilon_k} \right). \quad (5.11)$$

Se $p \in \mathcal{N}_{S(f, K)}(z)$, então pela definição 1.2.12

$$\langle p, y - z \rangle \leq 0 \quad \forall y \in S(f, K).$$

Logo, $\langle p, z \rangle = \sup_{y \in S(f, K)} \langle p, y \rangle$. Porém, $\sigma_{S(f, K)}(p) = \sup_{y \in S(f, K)} \langle p, y \rangle$ é a função suporte do conjunto, portanto

$$\left\langle \frac{2p}{\varepsilon_k}, z \right\rangle = \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right). \quad (5.12)$$

Substituindo (5.11) e (5.12) em (5.10), obtemos

$$\begin{aligned} D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[\left\langle \frac{2p}{\varepsilon_k}, x^{k+1} \right\rangle - \left\langle \frac{2p}{\varepsilon_k}, z \right\rangle - f(z, x^{k+1}) \right] \\ &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[f_z(x^{k+1}) + f_z^* \left(\frac{2p}{\varepsilon_k} \right) - f_z(x^{k+1}) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]. \end{aligned}$$

E por fim

$$\begin{aligned} D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]. \end{aligned} \quad (5.13)$$

□

O próximo teorema reúne propriedades fundamentais da sequência gerada pelo algoritmo, estabelecendo sua boa definição e resultados importantes para a análise de sua convergência.

Teorema 17. *Assuma que $S(h, S(f, K)) \neq \emptyset$. Para todo $x^0 \in K$, temos o seguinte:*

- (i) *A sequência $\{x^k\}$ gerada pelo Algoritmo 1 está bem definida.*
- (ii) *Para todo $z \in S(h, S(f, K))$,*

$$D(z, x^{k+1}) \leq D(z, x^k) + \frac{1}{2} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]. \quad (5.14)$$

- (iii) *Se $(\mathcal{H})$ vale, então para todo $z \in S(h, S(f, K))$*

- (a) $\lim_{k \rightarrow +\infty} D(z, x^k)$ existe;
- (b) a sequência $\{x^k\}$ é limitada;
- (c) $\lim_{k \rightarrow +\infty} D(x^{k+1}, x^k) = 0$ e $\lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0$;
- (d) $\sum_{k=1}^{\infty} \varepsilon_k f(z, x^{k+1}) < +\infty$.

Demonstração. Item (i)

Pelo Algoritmo 1, dado $x^k \in K$, define-se a bifunção

$$f_k(x, y) = \varepsilon_k f(x, y) + h(x, y) + \frac{1}{\lambda_k} \langle \nabla_1 d(x, x^k), y - x \rangle, \quad x, y \in K.$$

Pelo Teorema 16, o problema de equilíbrio associado a f_k admite solução, isto é, $S(f_k, K) \neq \emptyset$. Assim, é possível escolher $x^{k+1} \in S(f_k, K)$. Como $S(f_k, K) \subset K$, segue que $x^{k+1} \in K$. Portanto, por indução, a sequência $\{x^k\}$ gerada pelo Algoritmo 1 está bem definida.

Item (ii)

Seja $z \in S(h, S(f, K))$. Por definição, $h(z, y) \geq 0 \quad \forall y \in S(f, K)$. Isso significa que z minimiza a função $h_z(y) = h(z, y)$ sobre $S(f, K)$. Portanto

$$z \in \operatorname{argmin}_{y \in S(f, K)} h_z(y) \iff 0 \in \partial(h_z + \delta_{S(f, K)})(z).$$

Do Lema 5, temos

$$\begin{aligned} D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) &\leq D(z, x^k) - \gamma D(x^{k+1}, x^k) + \lambda_k \langle w, z - x^{k+1} \rangle \\ &\quad + \frac{\lambda_k \varepsilon_k}{2} \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]. \end{aligned}$$

Usando $0 \in \partial(h_z + \delta_{S(f, K)})(z)$ podemos escolher $w = 0$. Observando que $f(z, x^{k+1}) \geq 0$ (pois $z \in S(f, K)$ e $x^{k+1} \in K$) e $-\gamma D(x^{k+1}, x^k) \leq 0$, obtemos

$$D(z, x^{k+1}) \leq D(z, x^k) + \frac{\lambda_k \varepsilon_k}{2} \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]$$

que é exatamente (5.14).

Item (iii)-(a)

Assuma agora a condição $(\mathcal{H})$:

$$\sum_{k=1}^{\infty} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right] < +\infty.$$

Da desigualdade (5.14):

$$D(z, x^{k+1}) \leq D(z, x^k) + \eta_k$$

onde $\eta_k = \frac{1}{2}\lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f,K)} \left(\frac{2p}{\varepsilon_k} \right) \right]$. Como $\sum \eta_k < +\infty$, o Lema 4 implica que $\lim_{k \rightarrow \infty} D(z, x^k)$ existe.

Item (iii)-(b)

Como a sequência $\{D(z, x^k)\}$ converge, ela é limitada. Assim, existe $c > 0$ tal que $D(z, x^k) \leq c$, $\forall k \in \mathbb{N}$. Portanto, $x^k \in \{x \in \text{int } S; D(z, x) \leq c\}$, $\forall k \in \mathbb{N}$. Pela propriedade (H3) da distância proximal induzida, esse conjunto de nível é limitado. Consequentemente, a sequência $\{x^k\}$ é limitada.

Item (iii)-(c)

Usando novamente o Lema 5 com $w = 0$, obtemos

$$\gamma D(x^{k+1}, x^k) \leq D(z, x^k) - D(z, x^{k+1}) + \eta_k.$$

Como a sequência $\{D(z, x^k)\}$ converge, segue que

$$D(z, x^k) - D(z, x^{k+1}) \rightarrow 0.$$

Além disso, como $\sum_{k=1}^{\infty} \eta_k < +\infty$ e $\eta_k \geq 0$, temos $\eta_k \rightarrow 0$. Portanto,

$$0 \leq \gamma D(x^{k+1}, x^k) \leq D(z, x^k) - D(z, x^{k+1}) + \eta_k \rightarrow 0.$$

Como $\gamma > 0$, concluímos que

$$D(x^{k+1}, x^k) \rightarrow 0.$$

Pela Proposição 3.2.2, segue que

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

Item (iii)-(d)

Da desigualdade do Lema 5 com $w = 0$, obtemos

$$D(z, x^{k+1}) + \frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) \leq D(z, x^k) + \eta_k.$$

Logo,

$$\frac{\lambda_k \varepsilon_k}{2} f(z, x^{k+1}) \leq D(z, x^k) - D(z, x^{k+1}) + \eta_k.$$

Como $\liminf_{k \rightarrow +\infty} \lambda_k > 0$, existe $\theta > 0$ tal que

$$\frac{\lambda_k}{2} \geq \theta, \quad \forall k \in \mathbb{N}.$$

Assim,

$$\theta \varepsilon_k f(z, x^{k+1}) \leq D(z, x^k) - D(z, x^{k+1}) + \eta_k.$$

Somando de $k = 1$ até N , obtemos

$$\theta \sum_{k=1}^N \varepsilon_k f(z, x^{k+1}) \leq \sum_{k=1}^N (D(z, x^k) - D(z, x^{k+1})) + \sum_{k=1}^N \eta_k.$$

Como a soma envolvendo D é telescópica, segue que

$$\theta \sum_{k=1}^N \varepsilon_k f(z, x^{k+1}) \leq D(z, x^1) - D(z, x^{N+1}) + \sum_{k=1}^N \eta_k.$$

Como $D(z, x^{N+1}) \geq 0$, temos

$$\theta \sum_{k=1}^N \varepsilon_k f(z, x^{k+1}) \leq D(z, x^1) + \sum_{k=1}^N \eta_k.$$

Portanto,

$$\sum_{k=1}^N \varepsilon_k f(z, x^{k+1}) \leq \frac{1}{\theta} \left(D(z, x^1) + \sum_{k=1}^N \eta_k \right).$$

Como $\sum_{k=1}^{\infty} \eta_k < +\infty$, fazendo $N \rightarrow \infty$, concluímos que

$$\sum_{k=1}^{\infty} \varepsilon_k f(z, x^{k+1}) < +\infty.$$

□

Agora introduziremos um lema fundamental para a análise dos pontos de acumulação da sequência gerada pelo algoritmo, mostrando que tais pontos pertencem ao conjunto solução do problema de equilíbrio de dois níveis.

Lema 6. *Assuma que $(\mathcal{H})$ vale, $S(h, S(f, K)) \neq \emptyset$, e seja $\{x^k\}$ a sequência gerada pelo*

Algoritmo 1. Se $\lim_{j \rightarrow +\infty} x^{k_j} = x^$, então*

(a) $\lim_{j \rightarrow +\infty} x^{k_j+1} = x^*$,

(b) $x^* \in S(f, K)$,

(c) $x^* \in S(h, S(f, K))$.

Demonstração. (a)

Da desigualdade triangular, temos

$$\|x^{k_j+1} - x^*\| = \|x^{k_j+1} - x^{k_j} + x^{k_j} - x^*\| \leq \|x^{k_j+1} - x^{k_j}\| + \|x^{k_j} - x^*\|, \quad j = 0, 1, \dots,$$

pelo Teorema 17 (iii)-(c), $\lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0$ e da hipótese da subsequência $x^{k_j} \rightarrow x^*$, segue que $\|x^{k_j+1} - x^{k_j}\| \rightarrow 0$, $\|x^{k_j} - x^*\| \rightarrow 0$ e portanto $\|x^{k_j+1} - x^*\| \rightarrow 0$ quando $j \rightarrow +\infty$.

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

(b)

Pelo algoritmo 1, $x^{k+1} \in S(f_k, K)$, isto significa que $f_k(x^{k+1}, y) \geq 0 \quad \forall y \in K$. Substituindo em (5.1)

$$\frac{1}{\varepsilon_{k_j}} f_k(x^{k_j+1}, y) = f(x^{k_j+1}, y) + \frac{1}{\varepsilon_{k_j}} h(x^{k_j+1}, y) + \frac{1}{\varepsilon_{k_j} \lambda_{k_j}} \langle \nabla_1 d(x^{k_j+1}, x^{k_j}), y - x^{k_j+1} \rangle \geq 0.$$

Usando (H4), obtemos

$$\langle \nabla_1 d(x^{k_j+1}, x^{k_j}), y - x^{k_j+1} \rangle \leq D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j}),$$

logo,

$$f(x^{k_j+1}, y) + \frac{1}{\varepsilon_{k_j}} h(x^{k_j+1}, y) + \frac{1}{\varepsilon_{k_j} \lambda_{k_j}} [D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j})] \geq 0. \quad (5.15)$$

Pelo Teorema 17 (iii)-(c), temos $D(x^{k_j+1}, x^{k_j}) \rightarrow 0$. Além disso, pelo item anterior, $x^{k_j+1} \rightarrow x^*$, e, como $x^{k_j} \rightarrow x^*$, segue da propriedade (H1) que

$$D(y, x^{k_j}) \rightarrow D(y, x^*) \quad \text{e} \quad D(y, x^{k_j+1}) \rightarrow D(y, x^*).$$

Consequentemente,

$$D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j}) \rightarrow 0. \quad (5.16)$$

Como $\varepsilon_{k_j} \rightarrow \infty$ e $\liminf_{j \rightarrow +\infty} \lambda_{k_j} > 0$, obtemos

$$\frac{1}{\lambda_{k_j} \varepsilon_{k_j}} [D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j})] \rightarrow 0.$$

Além disso, pela semicontinuidade superior de $h(\cdot, y)$ e pelo fato de $x^{k_j+1} \rightarrow x^*$, temos

$$\limsup_{j \rightarrow +\infty} h(x^{k_j+1}, y) \leq h(x^*, y) < +\infty.$$

Logo,

$$\limsup_{j \rightarrow +\infty} \frac{1}{\varepsilon_{k_j}} h(x^{k_j+1}, y) \leq 0.$$

Tomando $\limsup$ em (5.15) e usando a semicontinuidade superior de $f(\cdot, y)$, dada pela hipótese (L2), segue que

$$0 \leq \limsup_{j \rightarrow +\infty} f(x^{k_j+1}, y) \leq f(x^*, y).$$

Portanto,

$$f(x^*, y) \geq 0, \quad \forall y \in K.$$

Assim, $x^* \in S(f, K)$.

(c)

Sabemos agora que $x^* \in S(f, K)$. Resta provar que

$$h(x^*, y) \geq 0, \quad \forall y \in S(f, K).$$

Para isso, tomemos $y \in S(f, K)$. Como $y \in S(f, K)$, temos $f(y, z) \geq 0$, para todo $z \in K$. Em particular, como $x^{k_j+1} \in K$, segue que $f(y, x^{k_j+1}) \geq 0$. Pela monotonicidade de f , $-f(x^{k_j+1}, y) \geq 0$. Usando novamente (5.15), obtemos

$$\frac{1}{\varepsilon_{k_j}} h(x^{k_j+1}, y) + \frac{1}{\varepsilon_{k_j} \lambda_{k_j}} [D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j})] \geq -f(x^{k_j+1}, y) \geq 0.$$

Multiplicando por $\varepsilon_{k_j} > 0$, segue que

$$h(x^{k_j+1}, y) + \frac{1}{\lambda_{k_j}} [D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j})] \geq 0. \quad (5.17)$$

Pelo Teorema 17 (iii)-(c), temos $D(x^{k_j+1}, x^{k_j}) \rightarrow 0$. Além disso, como $x^{k_j} \rightarrow x^*$ e $x^{k_j+1} \rightarrow x^*$, segue de (H1) que

$$D(y, x^{k_j}) \rightarrow D(y, x^*) \quad \text{e} \quad D(y, x^{k_j+1}) \rightarrow D(y, x^*).$$

Portanto,

$$D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j}) \rightarrow 0.$$

Como $\liminf_{k \rightarrow +\infty} \lambda_k > 0$, obtemos

$$\frac{1}{\lambda_{k_j}} [D(y, x^{k_j}) - D(y, x^{k_j+1}) - \gamma D(x^{k_j+1}, x^{k_j})] \rightarrow 0.$$

Assim, tomando o limite superior em (5.17), obtemos

$$\limsup_{j \rightarrow +\infty} h(x^{k_j+1}, y) \geq 0.$$

Por outro lado, como $h(\cdot, y)$ é semicontínua superiormente e $x^{k_j+1} \rightarrow x^*$, temos

$$\limsup_{j \rightarrow +\infty} h(x^{k_j+1}, y) \leq h(x^*, y).$$

Consequentemente,

$$h(x^*, y) \geq 0.$$

Como $y \in S(f, K)$ foi tomado arbitrariamente, concluímos que

$$h(x^*, y) \geq 0, \quad \forall y \in S(f, K).$$

Portanto,

$$x^* \in S(h, S(f, K)).$$

□

5.2 Convergência para um equilíbrio de dois níveis

Definição 5.2.1. *Uma sequência $\{z^k\}$ é dita quase-Fejer convergente para um conjunto $U \neq \emptyset$ com respeito à distância proximal generalizada (d, D) , se existe uma sequência não negativa $\{\gamma_k\}$ com $\sum_{k=0}^{\infty} \gamma_k < \infty$ tal que, para cada $k \in \mathbb{N}$,*

$$D(u, z^{k+1}) \leq D(u, z^k) + \gamma_k, \quad u \in U.$$

O próximo passo é importante para estabelecer a convergência da sequência gerada pelo Algoritmo 1. Tanto a demonstração do Lema quanto a definição anterior podem ser encontradas em [26].

Lema 7. *Seja $U \subset \mathbb{R}^n$ não vazio. Suponha que a sequência $\{z^k\} \subset \text{int } S$ seja quase-Fejér convergente para U com respeito à distância proximal generalizada (d, D) , isto é, para todo $u \in U$, existe uma sequência $\{\gamma_k\} \subset \mathbb{R}_+$ tal que $\sum_{k=0}^{\infty} \gamma_k < +\infty$ e*

$$D(u, z^{k+1}) \leq D(u, z^k) + \gamma_k, \quad k = 0, 1, 2, \dots$$

Então $\{z^k\}$ é limitada. Além disso, se um ponto de acumulação z de $\{z^k\}$ pertence a U , então

$$\lim_{k \rightarrow \infty} z^k = z.$$

Demonstração. Fixe $u \in U$. Pela quase-Fejér convergência,

$$D(u, z^{k+1}) - D(u, z^k) \leq \gamma_k.$$

Somando de $i = 0$ até k , obtemos

$$D(u, z^{k+1}) - D(u, z^0) \leq \sum_{i=0}^k \gamma_i.$$

Logo,

$$D(u, z^{k+1}) \leq D(u, z^0) + \sum_{i=0}^{\infty} \gamma_i.$$

Portanto, existe $c > 0$ tal que

$$D(u, z^k) \leq c, \quad \forall k \in \mathbb{N}.$$

Assim,

$$\{z^k\} \subset \{y \in \text{int } S; D(u, y) \leq c\}.$$

Pela propriedade (H3), esse conjunto é limitado. Consequentemente, $\{z^k\}$ é limitada.

Agora, seja $z \in U$ um ponto de acumulação de $\{z^k\}$. Então existe uma subsequência $\{z^{k_j}\}$ tal que $z^{k_j} \rightarrow z$.

Como $z \in U$, pela quase-Fejér convergência, temos

$$D(z, z^{k+1}) \leq D(z, z^k) + \gamma_k.$$

Somando de $i = k_j$ até k , com $k \geq k_j$, segue que

$$D(z, z^{k+1}) \leq D(z, z^{k_j}) + \sum_{i=k_j}^k \gamma_i.$$

Pela propriedade (H5), como $z^{k_j} \rightarrow z$, temos $D(z, z^{k_j}) \rightarrow 0$.

Além disso, como $\sum_{k=0}^{\infty} \gamma_k < +\infty$,

$$\sum_{i=k_j}^{\infty} \gamma_i \rightarrow 0.$$

Portanto,

$$\lim_{k \rightarrow +\infty} D(z, z^k) = 0.$$

Mostremos que $z^k \rightarrow z$. Suponha, por contradição, que z^k não converge para z . Como $\{z^k\}$ é limitada, existe uma subsequência $\{z^{k_j}\}$ convergindo para algum $\bar{z} \neq z$. Pela propriedade (H7), $\liminf_{j \rightarrow +\infty} D(z, z^{k_j}) > 0$, o que contradiz $D(z, z^k) \rightarrow 0$. Logo, $z^k \rightarrow z$. $\square$

Teorema 18 (Convergência do Algoritmo 1). *Sob as hipóteses do Teorema 17, toda sequência $\{x^k\}$, gerada pelo Algoritmo 1, converge para uma solução de (2.10).*

Demonstração. Combinemos a desigualdade

$$D(z, x^{k+1}) \leq D(z, x^k) + \frac{1}{2} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right] \quad (5.14)$$

com

$$\sum_{k=1}^{\infty} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right] < +\infty. \quad (\mathcal{H})$$

Definimos $\gamma_k := \frac{1}{2} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2p}{\varepsilon_k} \right) - \sigma_{S(f, K)} \left(\frac{2p}{\varepsilon_k} \right) \right]$, por $(\mathcal{H})$ temos que $\sum_{k=1}^{\infty} \gamma_k < +\infty$ e portanto

$$D(z, x^{k+1}) \leq D(z, x^k) + \gamma_k$$

satisfaz a definição de $\{x^k\}$ ser quase-Fejér convergente para o conjunto $S(h, S(f, K))$ com respeito à distância proximal generalizada (d, D) . Logo, pelo Lema 7, $\{x^k\}$ é limitada. Seja x^* um ponto de acumulação de $\{x^k\}$. O item (c) do Lema 6 implica que $x^* \in S(h, S(f, K))$. Como todo ponto de acumulação pertence a $S(h, S(f, K))$, o Lema 7 implica que a sequência inteira converge para esse ponto.

□

5.2.1 Exemplos

Nesta subseção, apresentam-se duas classes de problemas para as quais as hipóteses dos principais resultados de convergência são verificadas. O objetivo é mostrar que a teoria desenvolvida para o problema de equilíbrio de dois níveis não se restringe a uma formulação abstrata, mas engloba situações clássicas da análise variacional e da otimização.

Mais precisamente, pretende-se identificar escolhas particulares da bifunção f para as quais o problema de nível inferior assume interpretações bem conhecidas. Em cada caso, o conjunto solução

$$S(f, K) = \{u \in K; f(u, z) \geq 0, \forall z \in K\}$$

pode ser reescrito em termos de objetos familiares, como conjuntos de pontos fixos ou conjuntos de minimizadores de funções convexas.

Para a primeira classe de exemplos, será utilizada a noção de solução trivial de um problema de equilíbrio. Seja $\psi : K \times K \rightarrow \mathbb{R}$ uma bifunção e considere o problema de equilíbrio associado à ψ :

$$\text{Encontrar } x^* \in K \text{ tal que } \psi(x^*, y) \geq 0, \quad \forall y \in K.$$

Diz-se que $\hat{x}$ é uma solução trivial do problema de equilíbrio associado a ψ se

$$\psi(\hat{x}, y) = 0 \quad \forall y \in K.$$

Nesse caso, em particular, o problema de equilíbrio apresenta uma estrutura simplificada, o que permite verificar de maneira mais direta a hipótese $(\mathcal{H})$, utilizada na análise de convergência do algoritmo proximal.

A seguir, serão discutidos dois casos particulares relevantes:

- o caso em que o problema de nível inferior pode ser interpretado como um problema de ponto fixo;
- o caso em que o problema de nível inferior se reduz a um problema de minimização convexa.

Esses exemplos evidenciam que a formulação estudada neste trabalho é suficientemente ampla para abranger classes importantes de problemas hierárquicos.

Ponto fixo no problema de nível inferior

Neste exemplo, consideramos uma classe de problemas em que o nível inferior pode ser descrito em termos de pontos fixos de um operador. Seja $T : K \rightarrow K$ um operador e defina a bifunção $f : K \times K \rightarrow \mathbb{R}$ por

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

Nesse caso, o problema de equilíbrio associado a f consiste em encontrar $x \in K$ tal que

$$\langle x - Tx, y - x \rangle \geq 0, \quad \forall y \in K.$$

Pelo Teorema 9 o conjunto solução desse problema coincide com o conjunto de pontos fixos de T , dado por

$$\text{Fix}(T) := \{x \in K; Tx = x\}.$$

Portanto,

$$S(f, K) = \text{Fix}(T).$$

Dessa forma, o problema de equilíbrio de dois níveis, neste caso, pode ser escrito como:

$$\text{encontrar } \bar{x} \in \text{Fix}(T) \text{ tal que } h(\bar{x}, y) \geq 0, \quad \forall y \in \text{Fix}(T).$$

Assim, o problema de nível inferior determina o conjunto de pontos fixos de T , enquanto o nível superior seleciona, nesse conjunto, um ponto que satisfaz o critério definido pela bifunção h .

Além disso, pela proposição 2.2.1 a bifunção $f(x, y) = \langle x - Tx, y - x \rangle$ satisfaz as hipóteses (L1)–(L3) sob hipóteses usuais de regularidade sobre T .

A seguir, mostramos que, se T é não expansivo, então f satisfaz também a hipótese de monotonicidade (L4).

Proposição 5.2.1. *Seja $T : K \rightarrow K$ um operador não expansivo, isto é,*

$$\|Tx - Ty\| \leq \|x - y\|, \quad \forall x, y \in K.$$

Então a bifunção $f(x, y) = \langle x - Tx, y - x \rangle$ é monótona.

Demonstração. Tomemos $x, y \in K$. Temos

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

Como $x - y = -(y - x)$, segue que

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

Equivalentemente,

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

Logo,

$$f(x, y) + f(y, x) = -\|x - y\|^2 + \langle Tx - Ty, x - y \rangle.$$

Pela desigualdade de Cauchy–Schwarz,

$$\langle Tx - Ty, x - y \rangle \leq \|Tx - Ty\| \|x - y\|.$$

Como T é não expansivo,

$$\|Tx - Ty\| \leq \|x - y\|.$$

Portanto,

$$\langle Tx - Ty, x - y \rangle \leq \|x - y\|^2.$$

Substituindo na expressão anterior, obtemos

$$f(x, y) + f(y, x) \leq 0.$$

Assim, f satisfaz (L4). □

Consequentemente, quando T é não expansivo, a bifunção definida acima se enquadra nas hipóteses estruturais consideradas neste trabalho. Além disso, para $z \in S(f, K) = \text{Fix}(T)$, tem-se

$$f(z, y) = 0, \quad \forall y \in K,$$

isto é, todo ponto de $S(f, K)$ é uma solução trivial do problema de equilíbrio de nível inferior. Sob as condições adicionais indicadas em [17], a hipótese $(\mathcal{H})$ é satisfeita nesse caso, permitindo aplicar os resultados de convergência do algoritmo proximal a essa classe de problemas.

Minimização no problema de nível inferior

Um caso particular importante do problema de equilíbrio de dois níveis ocorre quando o problema de nível inferior pode ser interpretado como um problema de minimização convexa. Suponha que exista uma função $\Psi : K \rightarrow \mathbb{R}$ convexa e semicontínua inferiormente tal que

$$f(x, y) = \Psi(y) - \Psi(x), \quad x, y \in K.$$

Como visto no capítulo 2, a condição de equilíbrio $f(x, y) \geq 0$, para todo $y \in K$, é equivalente a $\Psi(x) \leq \Psi(y)$, para todo $y \in K$, ou seja,

$$S(f, K) = \operatorname{argmin}_{x \in K} \Psi(x).$$

Portanto, o problema de dois níveis consiste em encontrar

$$\bar{x} \in \operatorname{argmin}_{x \in K} \Psi(x) \quad \text{tal que} \quad h(\bar{x}, y) \geq 0, \quad y \in \operatorname{argmin}_{x \in K} \Psi(x).$$

Pelo Teorema 8 f satisfaz (L1)–(L3). (L4) é satisfeito pois para quaisquer $x, y \in K$

$$f(x, y) + f(y, x) = \psi(y) - \psi(x) + \psi(x) - \psi(y) = 0.$$

Definimos a função estendida $\Psi_0 : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ por

$$\Psi_0(y) = \begin{cases} \Psi(y), & \text{se } y \in K, \\ +\infty, & \text{caso contrário.} \end{cases}$$

Suponha que $\operatorname{argmin}_{x \in K} \Psi(x) \neq \emptyset$, que

$$\Psi(x) = 0, \quad \forall x \in S(f, K),$$

e que existam constantes $\theta > 0$ e $m > 1$ tais que

$$\Psi_0(y) \geq \frac{\theta}{m} \text{dist}(y, S(f, K))^m, \quad \forall y \in \mathbb{R}^n. \quad (5.18)$$

Essa hipótese estabelece um controle do crescimento da função objetivo em termos da distância ao conjunto de minimizadores.

Utilizando a definição de distância ao conjunto,

$$\text{dist}(y, S(f, K)) = \inf_{x \in S(f, K)} \|x - y\|,$$

segue que

$$\text{dist}(y, S(f, K))^m = \inf_{x \in S(f, K)} \|x - y\|^m.$$

Assim, de (5.18), obtemos

$$\Psi_0(y) \geq \inf_{x \in S(f, K)} \left(\frac{\theta}{m} \|x - y\|^m \right).$$

Usando a função indicadora $\delta_{S(f, K)}$ e como para qualquer função g ,

$$\inf_{x \in S(f, K)} g(x) = \inf_{x \in \mathbb{R}^n} (g(x) + \delta_{S(f, K)}),$$

tomando $g(x) = \frac{\theta}{m} \|x - y\|^m$, podemos reescrever a expressão acima como

$$\Psi_0(y) \geq \inf_{x \in \mathbb{R}^n} \left(\frac{\theta}{m} \|x - y\|^m + \delta_{S(f, K)}(x) \right). \quad (5.19)$$

Reconhecendo a expressão acima como uma inf-convolução, lembrando que a inf-convolução de duas funções u, v é definida por

$$(u \square v)(y) := \inf_{x \in \mathbb{R}^n} \{u(y - x) + v(x)\}.$$

Tome

$$u(z) = \frac{\theta}{m} \|z\|^m, \quad v(x) = \delta_{S(f, K)}(x).$$

Então

$$\inf_{x \in \mathbb{R}^n} \left(\frac{\theta}{m} \|x - y\|^m + \delta_{S(f, K)}(x) \right) = \left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f, K)} \right)(y). \quad (5.20)$$

Juntando (5.19) e (5.20) obtemos

$$\Psi_0(y) \geq \inf_{x \in \mathbb{R}^n} \left(\frac{\theta}{m} \|x - y\|^m + \delta_{S(f, K)}(x) \right) = \left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f, K)} \right)(y), \quad \forall y \in \mathbb{R}^n. \quad (5.21)$$

A partir de (5.21) e tomando a proposição 1.2.6, obtemos

$$\Psi_0^*(\mathbf{y}) \leq \left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f,K)} \right)^* (\mathbf{y}).$$

Pela proposição 1.2.10, segue que

$$\left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f,K)} \right)^* (\mathbf{y}) = \left(\frac{\theta}{m} \|\cdot\|^m \right)^* (\mathbf{y}) + (\delta_{S(f,K)})^* (\mathbf{y}).$$

Porém, a conjugada da indicadora é a função suporte:

$$(\delta_{S(f,K)})^* (\mathbf{y}) = \sigma_{S(f,K)} (\mathbf{y}),$$

Logo,

$$\left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f,K)} \right)^* (\mathbf{y}) = \left(\frac{\theta}{m} \|\cdot\|^m \right)^* (\mathbf{y}) + \sigma_{S(f,K)} (\mathbf{y}). \quad (5.22)$$

$$\left(\frac{\theta}{m} \|\cdot\|^m \right)^* (\mathbf{y}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{ \langle \mathbf{y}, \mathbf{x} \rangle - \frac{\theta}{m} \|\mathbf{x}\|^m \},$$

escolhemos $\mathbf{x}$ na direção de $\mathbf{y}$ para maximizar $\langle \mathbf{y}, \mathbf{x} \rangle - \frac{\theta}{m} \|\mathbf{x}\|^m$, então escrevemos

$$\mathbf{x} = t \frac{\mathbf{y}}{\|\mathbf{y}\|}, \quad t \geq 0,$$

quando $\mathbf{y} \neq 0$. Assim

$$\langle \mathbf{y}, \mathbf{x} \rangle = t \|\mathbf{y}\|, \quad \|\mathbf{x}\| = t.$$

Logo o problema vira

$$\sup_{t \geq 0} \{ t \|\mathbf{y}\| - \frac{\theta}{m} t^m \}.$$

Definimos

$$\begin{aligned} \varphi(t) &= t \|\mathbf{y}\| - \frac{\theta}{m} t^m \\ \Rightarrow \varphi'(t) &= \|\mathbf{y}\| - \theta t^{m-1}. \end{aligned}$$

Então o ponto crítico satisfaz

$$\theta t^{m-1} = \|\mathbf{y}\|$$

isto é,

$$t = \left(\frac{\|\mathbf{y}\|}{\theta} \right)^{\frac{1}{m-1}}.$$

Tomando $s > 0$ tal que

$$\frac{1}{m} + \frac{1}{s} = 1,$$

então

$$\frac{1}{m-1} = s-1.$$

Logo

$$t = \left(\frac{\|y\|}{\theta} \right)^{s-1}.$$

Substituindo em φ :

$$\varphi(t) = t\|y\| - \frac{\theta}{m} t^m$$

temos $t\|y\| = \|y\| \left(\frac{\|y\|}{\theta} \right)^{s-1} = \theta^{1-s} \|y\|^s$. Além disso,

$$t^m = \left(\frac{\|y\|}{\theta} \right)^{m(s-1)},$$

como

$$s - 1 = \frac{1}{m - 1},$$

segue que

$$m(s - 1) = \frac{m}{m - 1} = s.$$

Portanto,

$$t^m = \left(\frac{\|y\|}{\theta} \right)^s = \theta^{-s} \|y\|^s.$$

Substituindo tudo

$$\varphi(t) = \theta^{1-s} \|y\|^s - \frac{\theta^{1-s}}{m} \|y\|^s = \theta^{1-s} \left(1 - \frac{1}{m} \right) \|y\|^s.$$

Como $1 - \frac{1}{m} = \frac{1}{s}$, obtemos

$$\varphi(t) = \frac{\theta^{1-s}}{s} \|y\|^s. \quad (5.23)$$

De (5.22) e (5.23), tem-se

$$\left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f,K)} \right)^* (y) = \frac{\theta^{1-s}}{s} \|y\|^s + \sigma_{S(f,K)}(y).$$

Portanto

$$\Psi_0^*(y) \leq \left(\frac{\theta}{m} \|\cdot\|^m \square \delta_{S(f,K)} \right)^* (y) = \frac{\theta^{1-s}}{s} \|y\|^s + \sigma_{S(f,K)}(y), \quad y \in \mathbb{R}^n. \quad (5.24)$$

A desigualdade (5.24) fornece um controle explícito da conjugada da função objetivo estendida em termos de um termo polinomial e da função suporte do conjunto solução.

Como $z \in S(f, K)$ e, por hipótese, $\Psi(z) = 0$, temos

$$f_z(y) = f(z, y) = \Psi(y) - \Psi(z) = \Psi(y), \quad y \in K.$$

Considerando a extensão Ψ_0 de Ψ a $\mathbb{R}^n$, segue que $f_z^* = \Psi_0^*$. Portanto, de (5.24), obtemos

$$f_z^*(q) - \sigma_{S(f,K)}(q) \leq \frac{\theta^{1-s}}{s} \|q\|^s, \quad q \in \mathbb{R}^n.$$

Tomando $\mathbf{q} = \frac{2\mathbf{p}}{\varepsilon_k}$, temos

$$f_z^* \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) - \sigma_{S(f,K)} \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) \leq \frac{\theta^{1-s}}{s} \left\| \frac{2\mathbf{p}}{\varepsilon_k} \right\|^s.$$

Assim,

$$\lambda_k \varepsilon_k \left[f_z^* \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) - \sigma_{S(f,K)} \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) \right] \leq C_p \frac{\lambda_k}{\varepsilon_k^{s-1}},$$

onde

$$C_p := \frac{\theta^{1-s} 2^s}{s} \|\mathbf{p}\|^s.$$

Consequentemente, se

$$\sum_{k=1}^{\infty} \frac{\lambda_k}{\varepsilon_k^{s-1}} < +\infty,$$

então

$$\sum_{k=1}^{\infty} \lambda_k \varepsilon_k \left[f_z^* \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) - \sigma_{S(f,K)} \left(\frac{2\mathbf{p}}{\varepsilon_k} \right) \right] < +\infty.$$

Logo, a hipótese $(\mathcal{H})$ é satisfeita neste caso particular.

Capítulo 6

Conclusões e Perspectivas Futuras

O estudo desenvolvido nesta dissertação teve como eixo central o método proximal com distâncias proximais generalizadas para problemas de equilíbrio de dois níveis, conforme proposto em [10]. A relevância dessa abordagem está no fato de que ela combina, em uma mesma estrutura, problemas de equilíbrio, formulações hierárquicas e regularização proximal por meio de distâncias não necessariamente euclidianas.

O problema de equilíbrio de dois níveis considerado pode ser escrito na forma

$$\text{encontrar } \bar{\mathbf{x}} \in S(\mathbf{f}, \mathbf{K}) \text{ tal que } h(\bar{\mathbf{x}}, \mathbf{y}) \geq 0, \quad \forall \mathbf{y} \in S(\mathbf{f}, \mathbf{K}),$$

onde

$$S(\mathbf{f}, \mathbf{K}) = \{\mathbf{u} \in \mathbf{K}; f(\mathbf{u}, \mathbf{z}) \geq 0, \forall \mathbf{z} \in \mathbf{K}\}.$$

Essa formulação generaliza diversos modelos clássicos de otimização, desigualdades variacionais e problemas com restrições de equilíbrio. Além disso, permite tratar situações hierárquicas nas quais uma condição de equilíbrio deve ser satisfeita sobre o conjunto solução de outro problema de equilíbrio.

O artigo [10] representa uma contribuição importante nessa linha, pois mostra que esse tipo de problema pode ser tratado por um método proximal construído a partir de distâncias proximais generalizadas. A substituição da distância euclidiana por uma distância proximal mais geral permite incorporar informações geométricas do problema, o que torna o método mais flexível e adequado para conjuntos viáveis com estruturas específicas.

A literatura posterior confirma que problemas de equilíbrio de dois níveis continuam sendo um tema ativo de pesquisa. Além dos métodos proximais clássicos, como os estuda-

dos por Moudafi em [43], desenvolveram-se técnicas de penalização, splitting e convergência fraca e forte em espaços de Hilbert, como em [17]. Mais recentemente, surgiram abordagens baseadas em métodos extragradientes, subgradientes, algoritmos inerciais, versões auto-adaptativas, problemas não convexos e sistemas dinâmicos associados a problemas de equilíbrio de dois níveis; ver, por exemplo, [57, 2, 4, 3, 31, 56, 40, 8].

Nesse contexto, a dissertação não encerra o tema, mas evidencia várias direções naturais para investigações futuras. A seguir, destacamos algumas delas.

6.1 Método proximal inexato

Uma primeira direção de pesquisa consiste em desenvolver uma versão inexata do algoritmo proximal estudado em [10]. No método original, cada iteração exige resolver exatamente o subproblema regularizado $\mathbf{x}^{k+1} \in S(\mathbf{f}_k, \mathbf{K})$, isto é,

$$\mathbf{f}_k(\mathbf{x}^{k+1}, \mathbf{y}) \geq 0, \quad \forall \mathbf{y} \in \mathbf{K},$$

onde

$$\mathbf{f}_k(\mathbf{x}, \mathbf{y}) = \varepsilon_k f(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y}) + \frac{1}{\lambda_k} \langle \nabla_1 d(\mathbf{x}, \mathbf{x}^k), \mathbf{y} - \mathbf{x} \rangle.$$

Do ponto de vista teórico, essa condição é conveniente, pois permite estabelecer diretamente as desigualdades fundamentais usadas na análise de convergência. Entretanto, do ponto de vista computacional, resolver exatamente cada subproblema pode ser uma exigência forte. Em aplicações numéricas, é comum que cada subproblema seja resolvido apenas aproximadamente, seja por limitações computacionais, seja porque o subproblema não admite solução explícita.

Assim, uma extensão natural seria substituir a condição exata por uma condição inexata da forma

$$\mathbf{f}_k(\mathbf{x}^{k+1}, \mathbf{y}) \geq -\eta_k, \quad \forall \mathbf{y} \in \mathbf{K},$$

onde $\{\eta_k\} \subset \mathbb{R}_+$ representa uma sequência de erros permitidos.

A questão central seria determinar hipóteses sobre $\{\eta_k\}$ que preservem a convergência do método. Condições naturais seriam $\eta_k \rightarrow 0$ e $\sum_{k=0}^{\infty} \lambda_k \eta_k < +\infty$. Além disso, como algumas estimativas do algoritmo envolvem divisão por ε_k , poderia ser necessário assumir também $\frac{\eta_k}{\varepsilon_k} \rightarrow 0$.

A expectativa é que a desigualdade fundamental do método exato seja substituída por uma desigualdade quase-Fejér perturbada do tipo

$$D(z, x^{k+1}) \leq D(z, x^k) + \gamma_k + \lambda_k \eta_k = D(z, x^k) + \rho_k,$$

com $\rho_k := \gamma_k + \lambda_k \eta_k$ e $\sum_{k=0}^{\infty} \rho_k < +\infty$. Nesse caso, a convergência poderia ser obtida pelos argumentos quase-Fejér utilizados em [10].

Entretanto, precisamos ter cuidado pois não basta introduzir o erro na definição do algoritmo. Também é necessário adaptar os lemas que garantem que todo ponto de acumulação pertence ao conjunto solução $S(h, S(f, K))$. Esse é o ponto principal da versão inexata, pois os erros precisam desaparecer de modo compatível com os parâmetros λ_k e ε_k .

Essa direção é particularmente promissora porque aproximaria o método de situações computacionais reais. Além disso, dialoga com uma tendência recente da literatura, na qual métodos subgradientes, extragradientes e auto-adaptativos são desenvolvidos para problemas de equilíbrio de dois níveis em espaços de Hilbert; ver, por exemplo, [31].

6.2 Taxas de convergência

Outra linha relevante consiste em investigar taxas explícitas de convergência. O resultado principal de [10] garante a convergência da sequência gerada pelo algoritmo, mas não fornece estimativas quantitativas para a velocidade dessa convergência.

Em aplicações, uma informação qualitativa de convergência pode ser insuficiente. É importante saber se a sequência converge linearmente, sublinearmente ou se não há uma taxa explícita sob as hipóteses consideradas.

A obtenção de taxas exigiria hipóteses adicionais. Uma possibilidade seria assumir monotonicidade forte da bifunção de nível superior sobre o conjunto solução do problema inferior. Por exemplo, supondo que exista $\mu > 0$ tal que

$$h(x, y) + h(y, x) \leq -\mu \|x - y\|^2, \quad \forall x, y \in S(f, K),$$

poderia ser possível obter estimativas quantitativas para a aproximação da solução de dois níveis.

Outra possibilidade seria impor equivalência entre a distância proximal generalizada

e a norma euclidiana, isto é, assumir que existem constantes $0 < m \leq M$ tais que

$$m\|x - y\|^2 \leq D(x, y) \leq M\|x - y\|^2.$$

Sob hipóteses desse tipo, estimativas em termos de D poderiam ser convertidas em estimativas em norma.

Uma meta futura seria obter resultados do tipo

$$\|x^k - x^*\| \leq C\rho^k, \quad 0 < \rho < 1,$$

ou, ao menos, taxas sublineares da forma

$$\|x^k - x^*\| = O\left(\frac{1}{k^\alpha}\right),$$

para algum $\alpha > 0$.

Essa perspectiva também se relaciona com trabalhos recentes que combinam técnicas inerciais, extragradientes e escolhas auto-adaptativas de parâmetros, buscando melhorar o desempenho computacional dos métodos; ver [2, 31, 58].

6.3 Formulações dinâmicas

Outra linha recente envolve a análise de sistemas dinâmicos associados a problemas de equilíbrio de dois níveis. Em [40], por exemplo, é estudada uma dinâmica do tipo forward-backward-forward para problemas de equilíbrio de dois níveis em espaços de Hilbert.

Essa abordagem é interessante porque conecta métodos iterativos discretos com sistemas dinâmicos contínuos. Em vez de estudar apenas uma sequência $\{x^k\}$, investiga-se a trajetória $x(t)$ de um sistema diferencial ou inclusão diferencial associado ao problema de dois níveis.

Uma perspectiva futura seria investigar se o método proximal com distâncias proximais generalizadas estudado em [10] admite uma formulação dinâmica correspondente. Tal abordagem poderia fornecer novas ferramentas para analisar estabilidade, dependência em relação a perturbações e comportamento assintótico.

Trabalhos recentes também consideram aplicações de abordagens dinâmicas a problemas de controle e estabilidade quantitativa, como em [39]. Isso sugere que a conexão entre equilíbrio de dois níveis, sistemas dinâmicos e controle pode ser uma direção fértil para investigações posteriores.

6.4 Extensões para espaços de Hilbert e Banach

O estudo desta dissertação foi conduzido em dimensão finita. Uma continuação natural seria investigar versões do método em espaços de Hilbert ou Banach.

Em espaços de Hilbert, parte da estrutura geométrica euclidiana é preservada por meio do produto interno. Por isso, várias técnicas de análise convexa e métodos proximais podem ser adaptadas. Já em espaços de Banach surgem dificuldades adicionais, pois não há, em geral, produto interno, e a dualidade precisa ser tratada de forma mais cuidadosa.

Nesse contexto, seria necessário adaptar a definição de distância proximal generalizada, bem como as hipóteses de semicontinuidade, convexidade e coercividade. Também seria necessário distinguir convergência fraca e convergência forte, aspecto que não desempenha papel essencial em $\mathbb{R}^n$, mas que se torna central em dimensão infinita.

Os trabalhos [17, 31, 40, 8] mostram que problemas de equilíbrio de dois níveis em espaços de Hilbert continuam sendo objeto de investigação atual.

6.5 Considerações finais

As perspectivas apresentadas mostram que o trabalho de [10] permanece conectado a linhas recentes de pesquisa em otimização e problemas de equilíbrio. Entre as direções mais promissoras estão a construção de versões inexatas do algoritmo, a obtenção de taxas de convergência, o estudo de unicidade, o desenvolvimento de métodos inerciais e auto-adaptativos, a análise de problemas não convexos, a formulação dinâmica e a implementação computacional.

Dentre essas direções, a versão inexata do método proximal parece particularmente adequada como continuação direta desta dissertação. Ela surge naturalmente da limitação computacional do método exato e pode ser abordada por meio de desigualdades quase-Fejér perturbadas, combinadas com hipóteses de controle sobre a sequência de erros.

Assim, embora esta dissertação tenha como foco principal a análise do método proximal com distâncias proximais generalizadas para problemas de equilíbrio de dois níveis, o estudo realizado aponta para um conjunto amplo de problemas em aberto. Essas questões indicam que a teoria de métodos proximais para problemas de equilíbrio de dois níveis permanece uma área fértil tanto do ponto de vista teórico quanto computacional.

Referências Bibliográficas

- [1] L. Q. Anh, P. Q. Khanh, and N. D. T. M. Van. Well-posedness under relaxed semicontinuity for bilevel equilibrium and optimization problems with equilibrium constraints. *Journal of Optimization Theory and Applications*, 153(1):42–59, 2012.
- [2] P. N. Anh and L. T. H. An. New subgradient extragradient methods for solving monotone bilevel equilibrium problems. *Optimization*, 68(11):2099–2124, 2019.
- [3] P. N. Anh, D. D. Thanh, N. K. Linh, and H. P. Tu. New explicit extragradient methods for solving a class of bilevel equilibrium problems. *Bulletin of the Malaysian Mathematical Sciences Society*, 44:3285–3305, 2021.
- [4] P. N. Anh and H. P. Tu. Subgradient projection methods extended to monotone bilevel equilibrium problems in hilbert spaces. *Numerical Algorithms*, 86:55–74, 2021.
- [5] H. Attouch, M.-O. Czarnecki, and J. Peypouquet. Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities. *SIAM Journal on Optimization*, 21(4):1251–1274, 2011.
- [6] H. Attouch, M.-O. Czarnecki, and J. Peypouquet. Prox-penalization and splitting methods for constrained variational problems. *SIAM Journal on Optimization*, 21(1):149–173, 2011.
- [7] A. Auslender and M. Teboulle. Interior gradient and proximal methods for convex and conic optimization. *SIAM Journal on Optimization*, 16:697–725, 2006.
- [8] A. Balhag, Z. Mazgouri, and M. Théra. Convergence of inertial prox-penalization and inertial forward–backward algorithms for solving monotone bilevel equilibrium problems. *Optimization*, 74(12):2885–2929, 2025.

- [9] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2 edition, 2017.
- [10] G. C. Bento, J. X. Cruz Neto, J. O. Lopes, P. A. Soares Jr., and A. Soubeyran. Generalized proximal distances for bilevel equilibrium problems. *SIAM Journal on Optimization*, 26(1):810–830, 2016.
- [11] E. Blum and W. Oettli. From optimization and variational inequality to equilibrium problems. *Math. Stud.*, 63:127–149, 1994.
- [12] J. Bracken and J. T. McGill. Mathematical programs with optimization problems in the constraints. *Operations Research*, 21(1):37–44, 1973.
- [13] R. Burachik and J. Dutta. Inexact proximal point methods for variational inequality problems. *SIAM Journal on Optimization*, 20:2653–2678, 2010.
- [14] R. Burachik and G. Kassay. On a generalized proximal point method for solving equilibrium problems in banach spaces. *Nonlinear Anal. Theory*, 75:6456–6464, 2012.
- [15] A. Cabot. Proximal point algorithm controlled by a slowly vanishing term: Applications to hierarchical minimization. *SIAM Journal on Optimization*, 15(2):555–572, 2005.
- [16] O. Chadli, Z. Chbani, and H. Riahi. Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities. *Journal of Optimization Theory and Applications*, 105(2):299–323, 2000.
- [17] Z. Chbani and H. Riahi. Weak and strong convergence of prox-penalization and splitting algorithms for bilevel equilibrium problems. *Numerical Algebra, Control and Optimization*, 3(2):353–366, 2013.
- [18] J. Chen, Z. Wan, and Y. J. Cho. The existence of solutions and well-posedness for bilevel mixed equilibrium problems in banach spaces. *Taiwanese Journal of Mathematics*, 17:725–748, 2013.
- [19] P. L. Combettes and S. A. Hirstoaga. Equilibrium programming in hilbert spaces. *Journal of Nonlinear and Convex Analysis*, 6(1):117–136, 2005.

- [20] X. P. Ding. Auxiliary principle and algorithm for mixed equilibrium problems and bilevel mixed equilibrium problems in banach spaces. *Journal of Optimization Theory and Applications*, 146:347–357, 2010.
- [21] F. Facchinei and J. S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*, volume I. Springer, 2003.
- [22] F. Facchinei, J. S. Pang, G. Scutari, and L. Lampariello. Vi-constrained hemivariational inequalities: Distributed algorithms and power control in ad-hoc networks. *Mathematical Programming*, 145(1):59–96, 2014.
- [23] Ky Fan. A minimax inequality and applications. *Inequalities III*, pages 103–113, 1972.
- [24] P. Hartman and G. Stampacchia. On some nonlinear elliptic differential functional equations. *Acta Mathematica*, 115:271–310, 1966.
- [25] A. N. Iusem and W. Sosa. New existence results for equilibrium problems. *Nonlinear Analysis*, 52(2):621–635, 2003.
- [26] A. N. Iusem, B. F. Svaiter, and M. Teboulle. Entropy-like proximal methods in convex programming. *Mathematics of Operations Research*, 19:790–814, 1994.
- [27] Alfredo N. Iusem. *Métodos de ponto proximal em otimização*. IMPA, Rio de Janeiro, 1995. 20^o Colóquio Brasileiro de Matemática.
- [28] A.N. Iusem, G. Kassay, and W. Sosa. Certain conditions for the existence of solutions of equilibrium problems. *Math. Program.*, 116:259–273, 2009.
- [29] A. Izmailov and M. Solodov. *Otimização, vol. 2: métodos computacionais*. IMPA, 3 edition, 2018.
- [30] A. Izmailov and M. Solodov. *Otimização, vol. 1: condições de otimalidade, elementos de análise convexa e de dualidade*. IMPA, 4 edition, 2020.
- [31] L. O. Jolaoso, K. O. Aremu, and O. K. Oyewole. A self-adaptive inertial subgradient extragradient algorithm for solving bilevel equilibrium problems. *Rendiconti del Circolo Matematico di Palermo Series 2*, 72:3637–3658, 2023.

- [32] B. Knaster, K. Kuratowski, and S. Mazurkiewicz. Ein beweis des fixpunktsatzes für n -dimensionale simplexe. *Fundamenta Mathematicae*, 14(1):132–137, 1929.
- [33] E. L. Lima. *Curso de Análise, vol. 1*. IMPA, 2014.
- [34] E. L. Lima. *Análise Real, vol. 2 - Funções de n Variáveis*. IMPA, 6 edition, 2016.
- [35] E. L. Lima. *Curso de Análise, vol. 2*. IMPA, 12 edition, 2020.
- [36] Z.Q. Luo, J.S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- [37] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Revue Française d'Informatique et de Recherche Opérationnelle*, 4(Ser. R-3):154–158, 1970.
- [38] A. Migdalas, P. Pardalos, and P. Varbrand. *Multilevel Optimization: Algorithms and Applications*. Kluwer, 1997.
- [39] K. Mittal, O. Chadli, X. Li, and V. Vetrivel. A dynamical approach for bilevel equilibrium problems and its applications to control problems. *Optimization*, 2025. Published online.
- [40] K. Mittal, P. Gautam, and V. Vetrivel. Forward–backward–forward dynamics for bilevel equilibrium problem. *Annals of Operations Research*, 2024. Published online.
- [41] B. S. Mordukhovich and N. M. Nam. *An Easy Path to Convex Analysis and Applications*. Williston: Morgan & Claypool, 2014.
- [42] A. Moudafi. Proximal point algorithm extended for equilibrium problems. *Journal of Natural Geometry*, 15:91–100, 1999.
- [43] A. Moudafi. Proximal methods for a class of bilevel monotone equilibrium problems. *Journal of Global Optimization*, 47:287–292, 2010.
- [44] L. D. Muu and W. Oettli. Convergence of an adaptive penalty scheme for finding constrained equilibria. *Nonlinear Anal. Theory*, 18:1159–1161, 1992.
- [45] L. D. Muu and W. Oettli. On equilibrium problems for multifunctions. *Numerical Functional Analysis and Optimization*, 13(1-2):43–57, 1992.

- [46] A. Nagurney and D. Zhang. *Projected Dynamical Systems and Variational Inequalities with Applications*. Kluwer Academic Publishers, Boston, 1996.
- [47] H. Nikaido and K. Isoda. Note on noncooperative convex games. *Pacific Journal of Mathematics*, 5:807–815, 1955.
- [48] B. T. Polyak. *Introduction to optimization*. New York, Optimization Software,, 1987.
- [49] Stephen M. Robinson. Generalized equations and their solutions, part i: Basic theory. In *Point-to-Set Maps and Mathematical Programming*, volume 10 of *Mathematical Programming Studies*, pages 128–141. Springer, 1979.
- [50] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.
- [51] R. T. Rockafellar. *Convex analysis*. Princeton university press, 1997.
- [52] M. V. Solodov and B. F. Svaiter. An inexact hybrid generalized proximal point algorithm and some new results on the theory of bregman functions. *Mathematics of Operations Research*, 25(2):214–230, 2000.
- [53] S. S. Souza, P. R. Oliveira, J. X. Cruz Neto, and A. Soubeyran. A proximal method with separable bregman distances for quasiconvex minimization over the nonnegative orthant. *European Journal of Operational Research*, 201(2):365–376, 2010.
- [54] G. Stampacchia. Formes bilinéaires coercitives sur les ensembles convexes. *Comptes Rendus de l'Académie des Sciences de Paris*, 258:4413–4416, 1964.
- [55] G. Stampacchia. Variational inequalities. In *Proceedings of the International Congress of Mathematicians*, pages 877–883, Nice, France, 1970.
- [56] T. V. Thang. Proximal subgradient algorithm for a class of nonconvex bilevel equilibrium problems. *Bulletin of the Malaysian Mathematical Sciences Society*, 47(2):65, 2024.
- [57] T. Yuying, B. V. Dinh, D. S. Kim, and S. Plubtieng. Extragradient subgradient methods for solving bilevel equilibrium problems. *Journal of Inequalities and Applications*, 2018(327), 2018.

-
- [58] T. Yuying and S. Plubtieng. Accelerated subgradient extragradient algorithm for solving bilevel system of equilibrium problems. *Symmetry*, 15(9):1681, 2023.