

Universidade Federal do Amazonas – UFAM
Universidade Federal do Pará – UFPA
Programa de Doutorado em Matemática – PDM

Lauriano de Souza e Souza

Método extragradiente subgradiente com a técnica da proporção
áurea para resolver problemas de equilíbrio em espaços de Hilbert

Manaus–AM
2025

Lauriano de Souza e Souza

Método extragradiante subgradiente com a técnica da proporção áurea para resolver problemas de equilíbrio em espaços de Hilbert

Tese apresentada ao Programa de Doutorado em Matemática em Associação Ampla UFPA–UFAM, como requisito parcial para obtenção do grau de Doutor em Matemática.

Área de concentração: Matemática Aplicada.

Orientador: Prof. Dr. Roberto Cristóvão Mesquita Silva
Coorientadora: Profa. Dra. Flávia Morgana de Oliveira Jacinto

Manaus–AM
2025

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S729m Souza, Lauriano de Souza e
 Método extragradiante subgradiente com a técnica da proporção áurea
 para resolver problemas de equilíbrio em espaços de Hilbert / Lauriano de
 Souza e Souza. - 2025.
 85 f. : il., color. ; 31 cm.

 Orientador(a): Roberto Cristóvão Mesquita Silva.
 Coorientador(a): Flávia Morgana de Oliveira Jacinto.
 Tese (doutorado) - Universidade Federal do Amazonas, Programa de
 Pós-Graduação em Matemática, Manaus, 2025.

 1. Problemas de equilíbrio. 2. Pseudomonotonicidade. 3. Condição do
 tipo-Lipschitz. 4. Método extragradiante subgradiente. 5. Proporção áurea.
 I. Silva, Roberto Cristóvão Mesquita. II. Jacinto, Flávia Morgana de
 Oliveira. III. Universidade Federal do Amazonas. Programa de Pós-
 Graduação em Matemática. IV. Título

Lauriano de Souza e Souza

Método extragradiante subgradiente com a técnica da proporção áurea para resolver problemas de equilíbrio em espaços de Hilbert

Tese apresentada ao Programa de Doutorado em Matemática em Associação Ampla UFPA–UFAM, como requisito parcial para obtenção do grau de Doutor em Matemática.

Área de concentração: Matemática Aplicada.

Aprovada em 17 de setembro de 2025.

BANCA EXAMINADORA

Documento assinado digitalmente
 **ROBERTO CRISTOVAO MESQUITA SILVA**
Data: 22/10/2025 14:47:14-0300
Verifique em <https://validar.iti.gov.br>

.....
Prof. Dr. Roberto Cristóvão Mesquita Silva (orientador – presidente)
Universidade Federal do Amazonas – UFAM

Documento assinado digitalmente
 **ORIZON PEREIRA FERREIRA**
Data: 20/10/2025 20:18:10-0300
Verifique em <https://validar.iti.gov.br>

.....
Prof. Dr. Orizon Pereira Ferreira (membro externo)
Universidade Federal de Goiás – UFG

Documento assinado digitalmente
 **PAULO SERGIO MARQUES DOS SANTOS**
Data: 21/10/2025 10:01:26-0300
Verifique em <https://validar.iti.gov.br>

.....
Prof. Dr. Paulo Sérgio Marques dos Santos (membro externo)
Universidade Federal do Delta do Parnaíba – UFDPAr

Documento assinado digitalmente
 **PEDRO JORGE SOUSA DOS SANTOS**
Data: 21/10/2025 10:15:03-0300
Verifique em <https://validar.iti.gov.br>

.....
Prof. Dr. Pedro Jorge Sousa dos Santos (membro externo)
Universidade Federal do Delta do Parnaíba – UFDPAr

Documento assinado digitalmente
 **DAIANA OLIVEIRA DOS SANTOS**
Data: 22/10/2025 11:49:36-0300
Verifique em <https://validar.iti.gov.br>

.....
Profa. Dra. Daiana Oliveira dos Santos (membro externo)
Universidade Federal de São Paulo – UNIFESP

Dedico esta tese de doutorado à minha família e à memória de meu pai, cuja ausência desde meus 12 anos jamais apagou sua presença em minhas conquistas.

Agradecimentos

A Deus, por permitir esta conquista, concedendo saúde, proteção, sabedoria e perseverança ao longo de todo o processo.

À minha esposa, Regina Carla, cujo apoio, compreensão e carinho tornaram possível esta jornada.

Ao meu filho, Lohan Pietro, que mesmo tão novo compreendeu minhas ausências e me presenteou com apoio silencioso, sorrisos e afeto.

Aos meus pais, José Roberto Andrade (in memoriam) e Erlita de Souza, que, mesmo com todas as dificuldades surgidas no caminho, sempre priorizaram os estudos dos filhos.

Às minhas irmãs, Elda Jane e Joerlen Souza, e aos demais familiares, por sempre acreditarem em mim.

Ao meu orientador, professor Roberto Cristóvão, pela orientação competente e rigorosa, e pela confiança depositada.

À minha coorientadora, professora Flávia Morgana, pelas valiosas recomendações e pelo incentivo contínuo.

Ao professor Pedro Jorge dos Santos, pelo suporte nos experimentos computacionais.

Ao colega de doutorado e de trabalho, Manoel Ricardo, pelas discussões técnicas e pela amizade.

Aos professores da banca examinadora, pelas leituras atentas e pelas contribuições que aprimoraram substancialmente esta tese.

Ao Programa de Doutorado em Matemática em Associação Ampla UFPA–UFAM e à Universidade do Estado do Amazonas (UEA), que viabilizaram este estudo.

Aos colegas do Ciclo Básico de Engenharia da Escola Superior de Tecnologia da UEA, pela amizade e apoio em todos os sentidos.

A todos que, de alguma forma, contribuíram para a realização desta tese, o meu sincero muito obrigado.

“A Geometria tem dois grandes tesouros: um é o Teorema de Pitágoras; o outro, a divisão de um segmento em média e extrema razão. O primeiro pode ser comparado a uma medida de ouro; o segundo podemos chamar de joia preciosa.”

(Johannes Kepler)

Resumo

Nesta tese, propomos um método extragradiente subgradiente acelerado e explicitamente implementável, baseado na técnica da proporção áurea, para resolver problemas de equilíbrio pseudomonótonos em espaços de Hilbert reais. O algoritmo assume que a bifunção associada satisfaz uma condição do tipo-Lipschitz, mas incorpora uma regra autoadaptativa de atualização de passo que dispensa o conhecimento prévio das constantes do tipo-Lipschitz, ampliando sua aplicabilidade prática. A convergência forte da sequência gerada é garantida com o auxílio da técnica de aproximação por viscosidade. Adicionalmente, demonstramos que, sob a hipótese de forte pseudomonotonicidade, o método apresenta taxa de convergência R-linear. Para potencializar seu desempenho, introduzimos um segundo passo inercial, cuja eficácia é validada por meio de experimentos numéricos comparativos. Os resultados indicam que a abordagem proposta é teoricamente consistente e computacionalmente eficiente.

Palavras-chave: Problemas de equilíbrio, pseudomonotonicidade, condição do tipo-Lipschitz, método extragradiente subgradiente, passo inercial, proporção áurea.

Abstract

In this thesis, we propose an accelerated and explicitly implementable subgradient extragradient method, based on the golden ratio technique, to solve pseudomonotone equilibrium problems in real Hilbert spaces. The algorithm assumes that the associated bifunction satisfies a Lipschitz-type condition, while incorporating a self-adaptive stepsize update rule that dispenses with prior knowledge of the Lipschitz-type constants, thus enhancing its practical applicability. Strong convergence of the generated sequence is ensured through the viscosity approximation technique. Moreover, under the additional assumption of strong pseudomonotonicity, we establish an R-linear convergence rate. To further improve performance, we introduce a second inertial step, whose effectiveness is validated by comparative numerical experiments. The results indicate that the proposed approach is theoretically consistent and computationally efficient.

Keywords: Equilibrium problems, pseudomonotonicity, Lipschitz-type condition, subgradient extragradient method, inertial step, golden ratio.

Sumário

Introdução	8
1 Fundamentos Teóricos	14
1.1 Tópicos de Análise Convexa em Espaços de Hilbert	14
1.1.1 Definições e Resultados	14
1.2 Problemas de Equilíbrio em Espaços de Hilbert	20
1.2.1 Formulação do Problema (PE)	20
1.3 O Método Extragradiante e sua Extensão Subgradiente	33
1.3.1 Surgimento e Motivação	33
1.3.2 O Método Extragradiante Subgradiente	35
2 Um Método Extragradiante Subgradiente Acelerado para PE	40
2.1 Descrição do Algoritmo 1 e sua Boa Definição	40
2.2 Análise de Convergência	46
2.2.1 Resultados Auxiliares	46
2.2.2 Resultado Principal	55
2.3 Taxa de Convergência	59
3 Método Extragradiante Subgradiente Duplamente Acelerado para PE	63
3.1 Descrição do Algoritmo 2	63
3.2 Análise de Convergência	64
4 Experimentos Numéricos	71
4.1 Ilustrações Numéricas do Algoritmo 1	71
4.2 Experimentos Numéricos Comparativos	77
4.2.1 Avaliação do impacto do duplo passo inercial	77
4.2.2 Influência do Parâmetro μ no Desempenho dos Algoritmos	78
5 Considerações Finais	81

Introdução

O Problema de Equilíbrio (PE) é uma formulação matemática abrangente que generaliza diversas classes fundamentais de problemas da Análise Variacional e da Otimização. Destaca-se por sua capacidade de modelar, de forma unificada, situações em que se busca um ponto que satisfaça simultaneamente condições de viabilidade e de estabilidade, expressas por meio de uma bifunção associada. Essa versatilidade faz com que o (PE) surja naturalmente em contextos aplicados e teóricos, como Teoria dos Jogos, Economia Matemática, Engenharia, Física, Química, dentre outros. Graças a isso, muitos resultados relativos à existência de soluções para problemas de equilíbrio foram estabelecidos e generalizados por vários autores (veja, por exemplo, [6, 7, 39] e suas referências).

Nesta tese, estudamos o problema de equilíbrio em espaços de Hilbert reais, com o objetivo de propor e analisar algoritmos iterativos que assegurem a convergência forte da sequência gerada para uma solução de equilíbrio sob hipóteses moderadas sobre a bifunção. O foco está em bifunções pseudomonótonas e que satisfazem uma condição do tipo-Lipschitz, estruturas frequentemente encontradas em aplicações práticas e teóricas. O método proposto combina técnicas recentes de atualização autoadaptativa de passos (para a sequência $\{\lambda_n\}$), aproximação por viscosidade e uso de passo inercial.

Seja H um espaço de Hilbert real com produto interno $\langle \cdot, \cdot \rangle$ e norma induzida $\| \cdot \|$. Seja $C \subset H$ um subconjunto convexo, fechado e não vazio, e $f : H \times H \rightarrow \mathbb{R}$ uma bifunção tal que $f(x, x) = 0$ para todo $x \in C$. O problema de equilíbrio, associado a f e C , consiste em:

$$(PE) : \text{ Encontrar } x^* \in C \text{ tal que } f(x^*, y) \geq 0, \forall y \in C.$$

Esse modelo abrange como casos particulares importantes problemas clássicos, como problemas de otimização, pontos fixos, desigualdades variacionais, complementariedade, equilíbrios de Nash, dentre outros. Ou seja, trata-se de uma formulação unificadora, cuja especialização depende de escolhas apropriadas para a bifunção f e o conjunto viável C .

Por exemplo, quando $f(x, y) = \langle F(x), y - x \rangle$, $\forall x, y \in C$, onde $F : C \rightarrow H$ é um operador não linear, o problema (PE) se reduz ao problema de desigualdade variacional (PDV): Encontrar $x^* \in C$ tal que

$$\langle F(x^*), y - x^* \rangle \geq 0, \forall y \in C. \quad (1)$$

Para mais detalhes veja a Seção 1.2 deste trabalho; recomendamos também as referências [7, 16, 30, 31].

A formulação (PE) foi proposta de forma explícita por Blum e Oettli [7] e por Muu e Oettli [38], seguindo os trabalhos pioneiros de Ky Fan sobre desigualdades do tipo minimax [17]. Por esse motivo, o (PE) também é conhecido na literatura como desigualdade de Ky Fan.

Nos últimos anos, muitos pesquisadores se interessaram pelo problema (PE), veja [2, 13, 14, 47, 55] e suas referências, e vários métodos foram propostos para resolvê-lo, um dos quais é o chamado método do tipo-proximal [18] ou método extragradiente (MEG) [41], que é uma

extensão do método extragradiente de Korpelevich (MEGK) introduzido em [32] para resolver problemas de ponto de sela e, em particular, problemas de desigualdade variacional em espaços euclidianos. Lembramos que o (MEGK) para (PDV) é descrito da seguinte forma:

$$\begin{cases} x_0 \in C, \\ y_n = P_C(x_n - \lambda F(x_n)), \\ x_{n+1} = P_C(x_n - \lambda F(y_n)), \end{cases} \quad (2)$$

onde F é um operador monótono e L -Lipschitz-contínuo, $\lambda \in \left(0, \frac{1}{L}\right)$ e P_C a projeção euclidiana sobre C .

Quando a bifunção f no (PE) for $f(x, y) = \langle F(x), y - x \rangle$, então as projeções ortogonais no (MEGK) (2) podem ser reescritas como dois problemas de minimização fortemente convexos:

$$\begin{cases} y_n = P_C(x_n - \lambda F(x_n)) = \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \\ x_{n+1} = P_C(x_n - \lambda F(y_n)) = \operatorname{argmin}_{y \in C} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \end{cases}$$

veja uma demonstração em (1.11). Essa é a maneira de estender o método extragradiente de desigualdades variacionais para problemas de equilíbrio.

Em 2008, Quoc et al. [41] estenderam o método extragradiente de (PDV) para (PE) em espaços euclidianos. Anos depois, Vinh e Muu [60] estenderam este resultado para um espaço de Hilbert real de dimensão infinita, obtendo um resultado de convergência fraca sob as hipóteses de pseudomonotonicidade e condição do tipo-Lipschitz impostas à bifunção de equilíbrio.

Em [41], foi proposta a seguinte versão do método extragradiente para (PE):

$$\begin{cases} x_0 \in C, \\ y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \\ x_{n+1} = \operatorname{argmin}_{y \in C} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \end{cases}$$

onde $\lambda > 0$ é um parâmetro adequado e $f(x, \cdot)$ é convexa, semicontínua inferior e subdiferenciável em C , para todo $x \in C$.

Neste caso, é necessário resolver dois problemas de otimização fortemente convexos no conjunto de restrições C por iteração. Esta é uma grande desvantagem do método, pois pode se tornar computacionalmente caro, se o conjunto C não tiver uma estrutura simples, ou seja, não for explícito, como mencionado em [11, 12].

Para contornar essa dificuldade, métodos do tipo extragradiente subgradiente (MEGS) foram propostos; veja (1.12) para o caso de (PDV). Esses métodos substituem a segunda projeção sobre C por uma projeção sobre um semiespaço fechado construído a partir de um subgradiente, reduzindo o custo computacional sem comprometer a convergência. Abaixo apresentamos um algoritmo deste tipo para (PE), proposto por Hieu [55], seguindo os passos do método extragradiente subgradiente para (PDV) de Censor et al. [11] e a iteração de Halpern em [19].

Algoritmo (Hieu [55])

Passo 0: Escolha $x_0 \in H$, $\lambda > 0$ e $\{\alpha_n\} \subset (0, 1)$ tal que

$$(i) 0 < \lambda < \min \left\{ \frac{1}{2c_1}, \frac{1}{2c_2} \right\}, \quad (ii) \lim_{n \rightarrow \infty} \alpha_n = 0, \quad (iii) \sum_{n=1}^{\infty} \alpha_n = +\infty.$$

Passo 1: Resolva os dois problemas de otimização fortemente convexos:

$$\begin{cases} y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \\ z_n = \operatorname{argmin}_{y \in T_n} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \end{cases}$$

onde $T_n = \{v \in H : \langle (x_n - \lambda w_n) - y_n, v - y_n \rangle \leq 0\}$ e $w_n \in \partial_2 f(x_n, y_n)$.

Passo 2: Calcule $x_{n+1} = \alpha_n x_0 + (1 - \alpha_n) z_n$. Defina $n := n + 1$ e volte ao **Passo 1**.

O Algoritmo acima é uma extensão do método extragradiente subgradiente de Halpern para resolver desigualdades variacionais em espaços de Hilbert, estudado em [33].

Um outro algoritmo que resolve o problema (PE) é o chamado algoritmo da proporção áurea, que utiliza apenas um problema de otimização fortemente convexo sobre C por iteração. Tal algoritmo foi proposto por Vinh [59], em 2018, motivado pelos trabalhos de Malitsky em [35] para resolver desigualdades variacionais. Nesse artigo, Vinh introduziu e analisou os dois algoritmos de proporção áurea para (PE) a seguir:

$$\begin{cases} x_0, y_1 \in C, \\ x_n = \frac{(\varphi - 1)y_n + x_{n-1}}{\varphi}, \\ y_{n+1} = \operatorname{argmin}_{y \in C} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \end{cases} \quad (3)$$

onde $0 < \lambda \leq \min \left\{ \frac{\varphi}{4c_1}, \frac{\varphi}{4c_2} \right\}$ e $\varphi = \frac{1 + \sqrt{5}}{2}$ é a proporção áurea.

E

$$\begin{cases} x_0, y_1 \in C, \\ x_n = \frac{(\varphi - 1)y_n + x_{n-1}}{\varphi}, \\ y_{n+1} = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \end{cases} \quad (4)$$

onde $\{\lambda_n\} \subset (0, +\infty)$ é uma sequência satisfazendo as hipóteses

$$\lim_{n \rightarrow \infty} \lambda_n = 0, \quad \sum_{n=0}^{\infty} \lambda_n = +\infty.$$

Enfatizamos que, no Algoritmo apresentado em (3), é utilizado um tamanho de passo fixo λ que depende do conhecimento prévio das constantes bifuncionais do tipo-Lipschitz c_1 e c_2 , o que pode limitar a aplicabilidade do algoritmo, visto que determinar ou estimar os valores dessas constantes, em geral, é muito difícil. Já no Algoritmo em (4), usa-se uma sequência decrescente e não somável de tamanho de passos $\{\lambda_n\}$. Segundo [49], o uso de ambos os tipos de tamanhos de passos afeta a aplicabilidade e eficiência computacional dos dois algoritmos.

Em 2019, Hieu et al. [22] propuseram uma modificação do algoritmo da proporção áurea de Vinh [59], melhorando sua eficiência computacional. Para isso, eles utilizaram uma sequência monótona de tamanho de passos λ_n que são gerados em cada iteração e são baseados em algumas iterações anteriores, os chamados passos autoadaptativos. Com a escolha dessa sequência, o algoritmo pode ser implementado sem o conhecimento prévio das constantes do tipo-Lipschitz. Abaixo sua descrição:

Algoritmo (Hieu et al. [22])

Inicialização: Seja $\varphi = \frac{1 + \sqrt{5}}{2}$. Escolha $\bar{x}_{-1} \in H, x_{-1}, x_0 \in C, \lambda_0 > 0$ e $\mu \in \left(0, \frac{\varphi}{2}\right)$.

Passos iterativos: Suponha que $\bar{x}_{n-1} \in H, x_{n-1}, x_n \in C$ e λ_n são conhecidos. Calcule x_{n+1} e λ_{n+1} como segue:

$$\begin{cases} \bar{x}_n = \frac{(\varphi - 1)x_n + \bar{x}_{n-1}}{\varphi}, \\ x_{n+1} = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(x_n, y) + \frac{1}{2} \|\bar{x}_n - y\|^2 \right\}, \\ \lambda_{n+1} = \min \left\{ \lambda_n, \frac{\mu (\|x_{n-1} - x_n\|^2 + \|x_n - x_{n+1}\|^2)}{2 [f(x_{n-1}, x_{n+1}) - f(x_{n-1}, x_n) - f(x_n, x_{n+1})]_+} \right\}, \end{cases}$$

onde $[t]_+ := \max\{0, t\}$.

Mais recentemente, Long et al. [34] propuseram um método extragradiente subgradiente acelerado com um critério de tamanho de passo dado por uma busca linear do tipo-Armijo, baseada na proporção áurea, para resolver problemas de desigualdades variacionais pseudomonótonas, como segue:

Algoritmo (Long et al. [34])

Inicialização: Dados $\lambda_1 > 0, \gamma > 0, v, u_n \in (0, 1), \varphi = \frac{1 + \sqrt{5}}{2}, \mu \in \left(0, \frac{2}{\varphi}\right), \theta = \varphi - 1$. Sejam $x_0, x_1 \in H$.

Iterativo. Calcule x_{n+1} como segue:

Passo 1. Defina

$$w_n = x_n + u_n(x_n - x_{n-1})$$

e calcule

$$y_n = P_C(w_n - \lambda_n F(w_n)).$$

Se $w_n = y_n$ ou $F(y_n) = 0$, então pare e y_n é uma solução do problema (1.6). Caso contrário:

Passo 2. Calcule

$$z_n = P_{T_n}(w_n - \lambda_n F(y_n)),$$

onde $\lambda_n := \gamma v^{m_n}$, m_n é o menor número inteiro não-negativo m satisfazendo

$$\gamma v^m \langle F(w_n) - F(y_n), z_n - y_n \rangle \leq \frac{\mu}{4\varphi} [\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2], \quad (5)$$

e $T_n := \{w \in H : \langle w_n - \lambda_n F(w_n) - y_n, w - y_n \rangle \leq 0\}$.

Passo 3. Calcule $e_n = (1 - \theta)w_n + \theta x_n$ e

$$x_{n+1} = \alpha_n g(e_n) + (1 - \alpha_n)z_n,$$

onde $g : H \rightarrow H$ é uma aplicação contração e $\alpha_n \in (0, 1)$ tal que

$$\lim_{n \rightarrow \infty} \alpha_n = 0, \quad \sum_{n=1}^{+\infty} \alpha_n = +\infty.$$

No algoritmo anterior, o vetor definido w_n é conhecido na literatura como método inercial, que é um esquema iterativo que tem como principal característica ser construído a partir do conhecimento das iterações anteriores x_{n-1} e x_n . O fator inercial $u_n(x_n - x_{n-1})$ acelera as propriedades de convergência da sequência $\{x_n\}$, no sentido de aumentar o desempenho computacional do algoritmo (veja [1]). Por esse motivo, algoritmos do tipo inercial têm sido amplamente estudados para casos específicos do problema (PE), nos últimos anos; alguns dos quais podemos encontrar em [5, 8, 9, 21, 53, 57].

Já o esquema iterativo $x_{n+1} = \alpha_n g(e_n) + (1 - \alpha_n)z_n$, no Passo 3, é conhecido como método de aproximação por viscosidade (veja [47, 61]), que aplicado ao problema de desigualdade variacional, gera uma sequência que converge fortemente para uma solução do problema (PDV), mesmo em espaços de dimensão infinita.

Ressaltamos também que o uso do critério de tamanho de passo λ_n , no Algoritmo de [34], permite que ele funcione de forma autoadaptativa, porém com carga computacional mais elevada, visto que uma busca linear geralmente requer muitos cálculos em cada iteração (veja [58]).

Inspirados e motivados por avanços recentes da literatura, em especial por [22, 34, 51], propomos um método extragradiante subgradiente acelerado, explicitamente implementável, para resolver problemas de equilíbrio com bifunções pseudomonótonas do tipo-Lipschitz em espaços de Hilbert reais. O algoritmo utiliza uma regra de passo autoadaptativa monótona (para $\{\lambda_n\}$), inspirada em [22], que evita procedimentos de busca linear, como os de [34], reduzindo assim o custo computacional por iteração. A convergência forte é assegurada pela inserção do método de aproximação por viscosidade, enquanto o uso de termos inerciais visa acelerar o processo iterativo. No caso particular em que o problema de equilíbrio se reduz a uma desigualdade variacional, o algoritmo proposto pode ser visto como uma modificação do algoritmo de [34].

Destacamos ainda que a técnica da proporção áurea exerce papel central na formulação do nosso algoritmo, viabilizando uma atualização explícita e estável dos tamanhos de passo λ_n , mesmo sem o conhecimento prévio de constantes do tipo-Lipschitz, o que amplia sua aplicabilidade prática em problemas nos quais tais constantes são de difícil ou impossível estimativa. Essa estratégia é controlada por um parâmetro μ , cuja escolha adequada — inclusive com $\mu \geq 1$ — preserva a estabilidade computacional observada nos experimentos. Assim, o método proposto oferece ganhos significativos tanto do ponto de vista teórico, ao estender resultados clássicos da literatura, quanto do ponto de vista prático, ao manter um desempenho robusto e eficiente.

A tese está organizada da seguinte forma: No Capítulo 1, reunimos definições e resultados preliminares necessários para a fundamentação teórica dos capítulos seguintes. Abordamos tópicos de análise convexa, propriedades de bifunções e aspectos relevantes de espaços de Hilbert reais. No Capítulo 2, propomos o algoritmo principal da tese, com análise detalhada de sua boa definição, convergência forte e, em um caso particular, taxa de convergência R-linear. No

Capítulo 3, propomos uma extensão do algoritmo original, que incorpora um segundo passo inercial. No Capítulo 4, apresentamos experimentos numéricos que avaliam a eficácia dos algoritmos propostos sob diferentes instâncias e parâmetros. Finalmente, no Capítulo 5, reunimos as considerações finais e sugerimos possíveis direções para trabalhos futuros.

Capítulo 1

Fundamentos Teóricos

Este capítulo apresenta os conceitos e ferramentas fundamentais que embasarão os métodos e algoritmos desenvolvidos ao longo desta tese. Iniciamos com tópicos de análise convexa em espaços de Hilbert, seguido pela formulação geral dos problemas de equilíbrio nesses espaços, e por fim, discutimos os métodos extragradiente e extragradiente subgradiente aplicados a desigualdades variacionais e sua extensão para problemas de equilíbrio.

1.1 Tópicos de Análise Convexa em Espaços de Hilbert

A análise convexa exerce um papel fundamental no estudo de problemas de equilíbrio, especialmente quando o conjunto viável é convexo e a bifunção associada apresenta propriedades de regularidade, como convexidade e semicontinuidade. Nesta seção, reunimos os principais conceitos e resultados de análise convexa em espaços de Hilbert que servirão de base para o desenvolvimento teórico e para a análise dos métodos apresentados nos capítulos seguintes.

1.1.1 Definições e Resultados

Seja H um espaço de Hilbert real, isto é, um espaço vetorial real completo com um produto interno $\langle \cdot, \cdot \rangle$, cuja norma associada é dada por $\|x\| = \sqrt{\langle x, x \rangle}$.

Começamos por recordar noções de convergência fundamentais no contexto de espaços de Hilbert, que serão amplamente utilizadas nas demonstrações.

Definição 1.1.1 (Convergência forte). *Dizemos que uma sequência $\{x_n\} \subset H$ converge fortemente para $x \in H$, denotada por $x_n \rightarrow x$, se a norma da diferença tende a zero, ou seja, $\|x_n - x\| \rightarrow 0$ quando $n \rightarrow \infty$.*

Definição 1.1.2 (Convergência fraca). *Dizemos que uma sequência $\{x_n\} \subset H$ converge fracamente para $x \in H$, denotada por $x_n \rightharpoonup x$, se para todo $y \in H$, o produto interno converge, ou seja, $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$ quando $n \rightarrow \infty$.*

Como o conjunto viável $C \subset H$ será, em geral, convexo e fechado, apresentamos a seguir as definições formais dessas propriedades.

Definição 1.1.3. [26, Definição 3.1.1, p.60] *Um conjunto $C \subset H$ é convexo se, para quaisquer $x, y \in C$ e $t \in [0, 1]$, tem-se $tx + (1 - t)y \in C$. Noutras palavras, um conjunto convexo contém todos os segmentos de reta cujos extremos pertencem ao conjunto.*

Observação 1.1.1. O ponto $tx + (1 - t)y$, onde $t \in [0, 1]$, é chamado *combinação convexa* de x e y , com parâmetro t .

Definição 1.1.4. Um conjunto $C \subset H$ é *fechado* (ou *fortemente fechado*) se, para toda sequência $\{x_n\} \subset C$ que converge fortemente para um ponto $x \in H$, tem-se $x \in C$.

A seguir um exemplo de conjunto convexo e fechado que utilizaremos na construção dos algoritmos propostos.

Exemplo 1.1.1. Seja H um espaço de Hilbert real. O conjunto

$$S := \{x \in H : \langle x, u \rangle \leq \alpha\},$$

com $u \in H$ e $\alpha \in \mathbb{R}$ fixos, é convexo e fechado, sendo denominado **semiespaço fechado** de H .

Demonstração. Convexidade. Se $x, y \in S$ e $t \in [0, 1]$, então, pela linearidade do funcional $x \mapsto \langle x, u \rangle$,

$$\langle tx + (1 - t)y, u \rangle = t\langle x, u \rangle + (1 - t)\langle y, u \rangle \leq t\alpha + (1 - t)\alpha = \alpha,$$

logo $tx + (1 - t)y \in S$. Portanto, S é convexo.

Fechamento. Seja $\{x_n\} \subset S$ tal que $x_n \rightarrow x$ em H . Sendo o funcional linear

$$L : H \rightarrow \mathbb{R}, \quad L(z) = \langle z, u \rangle,$$

contínuo, segue que $\langle x_n, u \rangle = L(x_n) \rightarrow L(x) = \langle x, u \rangle$. Como $\langle x_n, u \rangle \leq \alpha$ para todo n , passando ao limite obtemos $\langle x, u \rangle \leq \alpha$. Logo $x \in S$, e S é fechado. $\square$

A projeção métrica sobre conjuntos convexos e fechados é uma ferramenta central em métodos iterativos. A seguir, apresentamos sua definição.

Definição 1.1.5. [23, Capítulo 3, p. 116] Seja $C \subset H$ um subconjunto não vazio, convexo e fechado. A projeção métrica de H sobre C é a aplicação $P_C : H \rightarrow C$ que associa a cada $x \in H$ o único ponto $P_C(x) \in C$ tal que $\|P_C(x) - x\| = \min\{\|x - y\| : y \in C\}$. Equivalentemente, a projeção métrica é caracterizada como a solução do problema de otimização

$$P_C(x) = \arg \min \left\{ \frac{1}{2} \|x - y\|^2 : y \in C \right\}.$$

Além disso, a projeção satisfaz a seguinte condição:

$$\langle x - P_C(x), y - P_C(x) \rangle \leq 0, \forall y \in C.$$

Em seguida, introduzimos o conceito de função convexa, que desempenha papel essencial na formulação de problemas de otimização e de equilíbrio. Intuitivamente, uma função é convexa quando seu gráfico permanece abaixo do segmento de reta que liga quaisquer dois de seus pontos.

Definição 1.1.6 (Função convexa). Seja $C \subset H$ um conjunto convexo. Dizemos que uma função $f : C \rightarrow \mathbb{R}$ é *convexa* se, para quaisquer $x, y \in C$ e $t \in [0, 1]$, tem-se:

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y). \quad (1.1)$$

Dizemos que f é *estritamente convexa* em C se a desigualdade (1.1) é estrita sempre que $x \neq y$.

Uma noção ainda mais forte é a de convexidade forte:

Definição 1.1.7. *Seja $C \subset H$ um conjunto convexo. Dizemos que uma função $f : C \rightarrow \mathbb{R}$ é fortemente convexa com módulo $\lambda > 0$ se, para quaisquer $x, y \in C$ e $t \in [0, 1]$, vale a seguinte desigualdade:*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\lambda}{2}t(1-t)\|x - y\|^2.$$

Exemplo 1.1.2. *Para $w \in H$ fixo, a função $\psi : H \rightarrow \mathbb{R}$ dada por $\psi(y) = \frac{1}{2}\|y - w\|^2$ é 1-fortemente convexa em H (e, portanto, em qualquer subconjunto convexo $C \subset H$).*

Demonstração. Para quaisquer $x, y \in H$ e $t \in [0, 1]$, aplicando o Lema 1.1.4(a) aos vetores $x - w$ e $y - w$, obtemos

$$\|t(x - w) + (1-t)(y - w)\|^2 = t\|x - w\|^2 + (1-t)\|y - w\|^2 - t(1-t)\|x - y\|^2.$$

Multiplicando por $\frac{1}{2}$ resulta

$$\psi(tx + (1-t)y) = t\psi(x) + (1-t)\psi(y) - \frac{1}{2}t(1-t)\|x - y\|^2,$$

logo ψ é 1-fortemente convexa. □

Lema 1.1.1. *Seja $C \subset H$ convexo. Se $g : C \rightarrow \mathbb{R}$ é convexa e $h : C \rightarrow \mathbb{R}$ é λ -fortemente convexa em C ($\lambda > 0$), então $g + h$ é λ -fortemente convexa em C .*

Demonstração. Para quaisquer $x, y \in C$ e $t \in [0, 1]$, pela convexidade de g ,

$$g(tx + (1-t)y) \leq tg(x) + (1-t)g(y),$$

e pela λ -convexidade forte de h ,

$$h(tx + (1-t)y) \leq th(x) + (1-t)h(y) - \frac{\lambda}{2}t(1-t)\|x - y\|^2.$$

Somando as duas desigualdades, obtemos a λ -convexidade forte de $g + h$. □

Em muitos contextos, é conveniente considerar funções que podem assumir o valor $+\infty$, por exemplo, para representar restrições via penalização. Para isso, utilizamos o formalismo das funções estendidas, isto é, funções com valores em $\mathbb{R} \cup \{+\infty\}$.

Um exemplo clássico é a função indicadora de um conjunto $C \subset H$, definida por:

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

Com isso, problemas com restrições explícitas podem ser reformulados como problemas irrestritos:

$$\min_{x \in C} \psi(x) \Leftrightarrow \min_{x \in H} [\psi(x) + \delta_C(x)].$$

Para esse tipo de função, introduzimos os conceitos de domínio efetivo e função própria.

Definição 1.1.8. *Seja $f : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função real estendida. O domínio efetivo (ou, simplesmente, o domínio) de f é o conjunto de pontos onde ela é finita. Em outras palavras,*

$$\text{dom}(f) := \{x \in H : f(x) < +\infty\}.$$

A função f é dita própria, se $\text{dom}(f) \neq \emptyset$.

Também podemos estender naturalmente a noção de convexidade ao caso de funções estendidas.

Definição 1.1.9 (Função Convexa Estendida). *Seja $f : H \longrightarrow \mathbb{R} \cup \{+\infty\}$. Dizemos que f é convexa se, para quaisquer $x, y \in \text{dom}(f)$ e $t \in [0, 1]$, vale:*

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Observação 1.1.2. *Se $\text{dom}(f) = H$, então f é uma função $f : H \longrightarrow \mathbb{R}$ usual, e neste caso dizemos que f é convexa em H .*

Uma vez estabelecido o conceito de função convexa estendida, é natural introduzir a noção de subgradiente, que generaliza a derivada para funções convexas que não são necessariamente diferenciáveis. Esse conceito será fundamental para a construção dos algoritmos baseados em direções subgradientes.

Definição 1.1.10 (Subgradiente e Subdiferencial). *Seja $f : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria e convexa. Um vetor $w \in H$ é chamado subgradiente de f no ponto $x \in \text{dom}(f)$ se*

$$f(y) \geq f(x) + \langle w, y - x \rangle, \forall y \in H.$$

O conjunto de todos os subgradientes de f em x , chama-se subdiferencial de f em x , e é denotado por $\partial f(x) := \{w \in H : f(y) \geq f(x) + \langle w, y - x \rangle, \forall y \in H\}$. Se $\partial f(x) \neq \emptyset$, dizemos que f é subdiferenciável em x .

Nos problemas de equilíbrio, a bifunção $f : H \times H \rightarrow \mathbb{R}$ desempenha papel central. Assim, estendemos os conceitos anteriores (convexidade e subgradiente) à análise da bifunção em sua segunda variável, uma vez que muitas das hipóteses e construções dos métodos dependerão dessas propriedades.

Definição 1.1.11 (Convexidade na segunda variável). *Uma bifunção $f : H \times H \longrightarrow \mathbb{R}$ é dita convexa em H com relação à segunda variável se, para todo $x \in H$, a função $f(x, \cdot) : H \longrightarrow \mathbb{R}$ é convexa em H .*

Definição 1.1.12. [40, Seção 3.3, p.41] *Seja $f(x, \cdot) : H \rightarrow \mathbb{R}$ uma função convexa em H , $\forall x \in H$. Dizemos que $w \in H$ é um subgradiente de $f(x, \cdot)$ em $y \in H$, se*

$$f(x, z) \geq f(x, y) + \langle w, z - y \rangle, \forall z \in H.$$

O conjunto de todos os subgradientes de $f(x, \cdot)$ em y , chama-se o subdiferencial de $f(x, \cdot)$ em y , e denotamos por $\partial_2 f(x, y)$. Se $\partial_2 f(x, y) \neq \emptyset$, dizemos que $f(x, \cdot)$ é subdiferenciável em y .

Observação 1.1.3. *A função $f(x, \cdot)$ é subdiferenciável em H , se o for em todos os pontos de H .*

Além da convexidade e subdiferenciabilidade, também são relevantes noções fracas de continuidade via sequências.

Definição 1.1.13. [40, Seção 2.2, p.29] Uma função $f(x, \cdot) : H \rightarrow \mathbb{R}$ é dita *semicontínua inferior (s.c.i.)* em um ponto $y \in H$, $\forall x \in H$, quando para qualquer sequência $\{y_n\} \subset H$ tal que $y_n \rightarrow y$ ($n \rightarrow \infty$), tem-se

$$\liminf_{n \rightarrow \infty} f(x, y_n) \geq f(x, y).$$

Observação 1.1.4. A função $f(x, \cdot)$ é s.c.i. em H , $\forall x \in H$, quando for s.c.i. em todos os pontos de H .

A função Φ , no próximo exemplo, aparece na formulação dos nossos algoritmos.

Exemplo 1.1.3. Assuma que $f(x, \cdot)$ é convexa e semicontínua inferior em C , para todo $x \in C$. Fixe $\lambda > 0$ e $w \in H$. Então a função

$$\Phi(y) := \lambda f(x, y) + \frac{1}{2} \|w - y\|^2, \quad y \in C,$$

é 1-fortemente convexa em C . Em particular, o problema

$$\min_{y \in C} \Phi(y)$$

admite solução única.

Demonstração. Como $f(x, \cdot)$ é convexa em C e $\lambda > 0$, a função $y \mapsto \lambda f(x, y)$ é convexa em C . Pelo Exemplo 1.1.2, $y \mapsto \frac{1}{2} \|w - y\|^2$ é 1-fortemente convexa em C . Logo, pelo Lema 1.1.1, Φ é 1-fortemente convexa em C .

Além disso, como $\frac{1}{2} \|w - y\|^2 \rightarrow +\infty$ quando $\|y\| \rightarrow \infty$, segue que Φ é coerciva em C . Como $f(x, \cdot)$ é semicontínua inferior em C e o termo quadrático é contínuo, concluímos que Φ é semicontínua inferior em C . Portanto, Φ atinge seu mínimo em C (que é fechado e não vazio), isto é, existe minimizador.

Por fim, a forte convexidade implica que esse minimizador é único. $\square$

Definição 1.1.14. Seja $y \in C \subset H$ qualquer. Uma função $f(\cdot, y) : C \rightarrow \mathbb{R}$ é dita *fracamente semicontínua superior* em um ponto $x \in C$, quando para qualquer sequência $\{x_n\} \subset C$ tal que $x_n \rightarrow x$ ($n \rightarrow \infty$), tem-se

$$\limsup_{n \rightarrow \infty} f(x_n, y) \leq f(x, y).$$

Observação 1.1.5. A função $f(\cdot, y)$ é fracamente semicontínua superior em C , $\forall y \in C$, quando for fracamente semicontínua superior em todos os pontos de C .

Para tratar de problemas de minimização convexa com restrições, é fundamental introduzir o conceito de cone normal a conjuntos convexos e fechados. Esse objeto será central na formulação das condições de otimalidade e será essencial na demonstração de diversos resultados dos Capítulos 2 e 3.

Definição 1.1.15. Seja $C \subset H$ um conjunto convexo e fechado. O cone normal a C no ponto $x \in C$ é o conjunto

$$N_C(x) := \{w \in H : \langle w, y - x \rangle \leq 0, \forall y \in C\}.$$

A seguir, apresentamos o lema que fornece um critério necessário e suficiente de otimalidade para problemas convexos com restrições.

Lema 1.1.2. [40, Proposição 3.61, p.59] *Seja C um subconjunto convexo, fechado e não vazio de H e $\psi : H \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria, convexa e semicontínua inferior em H . Suponha que ψ é contínua em algum ponto de C ou que existe um ponto interior de C , onde ψ é finita. Então $\bar{x}$ é uma solução para o seguinte problema convexo $\min\{\psi(x) : x \in C\}$ se, e somente se, $0 \in \partial\psi(\bar{x}) + N_C(\bar{x})$, onde $\partial\psi(\cdot)$ denota o subdiferencial de ψ e $N_C(\bar{x})$ é o cone normal de C em $\bar{x}$.*

Alguns resultados técnicos adicionais serão utilizados nos capítulos seguintes. O próximo lema fornece uma ferramenta importante para concluir a convergência de seqüências, especialmente no contexto de aproximação por viscosidade.

Lema 1.1.3. [42, Lema 2.6] *Seja $\{a_n\}$ uma seqüência de números reais não negativos, $\{\gamma_n\}$ uma seqüência de números reais em $(0, 1)$ com $\sum_{n=1}^{+\infty} \gamma_n = +\infty$ e $\{b_n\}$ uma seqüência de números reais. Suponha que*

$$a_{n+1} \leq (1 - \gamma_n)a_n + \gamma_n b_n, \forall n \geq 1.$$

Se $\limsup_{k \rightarrow +\infty} b_{n_k} \leq 0$ para toda subsequência $\{a_{n_k}\}$ de $\{a_n\}$ satisfazendo

$$\liminf_{k \rightarrow +\infty} (a_{n_k+1} - a_{n_k}) \geq 0,$$

então $\lim_{n \rightarrow +\infty} a_n = 0$.

Além disso, as identidades a seguir, típicas de espaços de Hilbert, são recorrentes nas manipulações algébricas das demonstrações, sobretudo nas estimativas envolvendo normas e produtos internos.

Lema 1.1.4. [4, Corolário 2.15, p. 31] *Seja H um espaço de Hilbert real, então a seguintes identidades são válidas:*

- a) $\|\alpha x + (1 - \alpha)y\|^2 = \alpha\|x\|^2 + (1 - \alpha)\|y\|^2 - \alpha(1 - \alpha)\|x - y\|^2, \forall x, y \in H, \alpha \in \mathbb{R};$
- b) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle, \forall x, y \in H;$
- c) $\|x \pm y\|^2 = \|x\|^2 + \|y\|^2 \pm 2\langle x, y \rangle, \forall x, y \in H.$

Outro conceito essencial para o método de aproximação por viscosidade, utilizado nesta tese, é o de aplicação contração.

Definição 1.1.16. *Uma aplicação $g : H \rightarrow H$ é uma aplicação contração, se existe uma constante $\rho \in [0, 1)$ tal que*

$$\|g(x) - g(y)\| \leq \rho\|x - y\|, \forall x, y \in H.$$

Por fim, apresentamos a definição de convergência R -linear, que será empregada na Seção 2.3 para descrever a taxa de convergência do algoritmo proposto sob hipóteses adicionais.

Definição 1.1.17. [51, Definição 2.1] *Seja $\{x_n\}$ uma seqüência em H .*

- i) *Diz-se que $\{x_n\}$ converge R -linearmente para x^* se existem constantes $c > 0, r \in (0, 1)$ e um índice $n_0 \in \mathbb{N}$ tais que*

$$\|x_n - x^*\| \leq c r^n, \forall n \geq n_0;$$

ii) Diz que $\{x_n\}$ converge Q -linearmente para x^* se existe uma constante $r \in (0, 1)$ e um índice $n_0 \in \mathbb{N}$ tais que

$$\|x_{n+1} - x^*\| \leq r \|x_n - x^*\|, \forall n \geq n_0.$$

Observação 1.1.6. [37, p.25] A convergência Q -linear implica convergência R -linear, mas a recíproca não é verdadeira. Assim, a convergência R -linear é uma noção mais fraca que a Q -linear, sendo mais geral e frequentemente suficiente em contextos de análise assintótica.

1.2 Problemas de Equilíbrio em Espaços de Hilbert

Nesta seção, apresentamos a formulação precisa do *Problema de Equilíbrio* (PE) em espaços de Hilbert reais, bem como complementamos as propriedades exigidas da bifunção associada — algumas das quais, como convexidade, subdiferenciabilidade e semicontinuidade em sua segunda variável, já foram discutidas na seção anterior. Esses conceitos serão fundamentais para a análise teórica e para a construção dos algoritmos desenvolvidos nos capítulos seguintes.

1.2.1 Formulação do Problema (PE)

A noção de equilíbrio está presente em diversas áreas da matemática e das ciências aplicadas, como Teoria dos Jogos, Economia, Física, Engenharia e Otimização. Geralmente se refere às condições ou estados de um sistema nos quais todas as influências concorrentes estão balanceadas. Muitos desses "problemas de equilíbrio" podem ser modelados por diferentes modelos matemáticos, como: desigualdades variacionais, otimização convexa, equilíbrio de Nash em jogos não cooperativos, dentre outros. Todos esses modelos matemáticos podem ser reformulados como casos particulares de uma desigualdade de Ky Fan, que hoje é amplamente conhecida como Problema de Equilíbrio; veja [6, 7, 28] para mais detalhes. A seguir relembramos essa definição.

Definição do Problema (PE)

Seja H um espaço de Hilbert real, $C \subset H$ um conjunto não vazio, convexo e fechado, e $f : H \times H \rightarrow \mathbb{R}$ uma bifunção satisfazendo $f(x, x) = 0$, para todo $x \in C$. O problema de equilíbrio consiste em:

$$\text{Encontrar } x^* \in C \text{ tal que } f(x^*, y) \geq 0, \forall y \in C. \quad (1.2)$$

Observação 1.2.1. Um ponto x^* que satisfaz (1.2) é denominado ponto de equilíbrio, e a bifunção f , quando satisfaz $f(x, x) = 0$ para todo $x \in C$, é chamada bifunção de equilíbrio em C .

Adotaremos a seguinte notação para o conjunto-solução do problema:

$$S_{PE}(f, C) := \{x \in C : f(x, y) \geq 0, \forall y \in C\},$$

o qual assumiremos sempre não vazio ao longo da tese.

Para garantir propriedades de existência e desenvolver métodos numéricos eficientes, é necessário impor certas condições à bifunção f . A seguir, apresentamos as principais delas, extraídas da literatura especializada [7, 54].

Definição 1.2.1. Seja H um espaço de Hilbert real, $C \subset H$ um subconjunto não vazio, convexo e fechado, e $f : H \times H \rightarrow \mathbb{R}$ uma bifunção. Dizemos que f é:

(1) fortemente monótona em C , se existe uma constante $\gamma > 0$ tal que

$$f(x, y) + f(y, x) \leq -\gamma \|x - y\|^2, \forall x, y \in C;$$

(2) monótona em C , se $f(x, y) + f(y, x) \leq 0, \forall x, y \in C$;

(3) pseudomonótona em C , se $f(x, y) \geq 0 \Rightarrow f(y, x) \leq 0, \forall x, y \in C$;

(4) fortemente pseudomonótona em C , se existe uma constante $\gamma > 0$ tal que

$$f(x, y) \geq 0 \Rightarrow f(y, x) \leq -\gamma \|x - y\|^2, \forall x, y \in C,$$

onde $\|\cdot\|$ denota a norma induzida pelo produto interno $\langle \cdot, \cdot \rangle$ de H .

Proposição 1.2.1. Seja $f : H \times H \rightarrow \mathbb{R}$ uma bifunção sobre C . Valem as seguintes implicações entre as propriedades acima.

$$(1) \Rightarrow (2) \Rightarrow (3) \quad e \quad (1) \Rightarrow (4) \Rightarrow (3).$$

Demonstração. As implicações seguem diretamente das definições, por substituição e uso de desigualdades elementares.

((1) $\Rightarrow$ (2)) Suponha que f satisfaz a condição (1). Então existe $\gamma > 0$ tal que

$$f(x, y) + f(y, x) \leq -\gamma \|x - y\|^2, \forall x, y \in C.$$

Como $\|x - y\|^2 \geq 0, \forall x, y \in C$, segue que

$$f(x, y) + f(y, x) \leq 0, \forall x, y \in C,$$

isto é, f satisfaz a condição (2).

((2) $\Rightarrow$ (3)) Suponha que f satisfaz a condição (2), ou seja,

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

Se $f(x, y) \geq 0, \forall x, y \in C$, então

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

isto é, f satisfaz a condição (3).

((1) $\Rightarrow$ (4)) Suponha que f satisfaz a condição (1), ou seja, existe $\gamma > 0$ tal que

$$f(x, y) + f(y, x) \leq -\gamma \|x - y\|^2, \forall x, y \in C.$$

Se $f(x, y) \geq 0, \forall x, y \in C$, tem-se $-f(x, y) \leq 0$. Logo,

$$f(y, x) \leq -\gamma \|x - y\|^2 - f(x, y) \leq -\gamma \|x - y\|^2,$$

isto é, f satisfaz a condição (4).

((4) $\Rightarrow$ (3)) Suponha que f satisfaz a condição (4), ou seja, existe $\gamma > 0$ tal que

$$f(x, y) \geq 0 \Rightarrow f(y, x) \leq -\gamma \|x - y\|^2, \forall x, y \in C.$$

Logo,

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

isto é, f satisfaz a condição (3). □

O exemplo a seguir mostra como a noção de monotonicidade se aplica a bifunções geradas por operadores monótonos.

Exemplo 1.2.1 (Bifunção monótona). *Seja $F : C \rightarrow H$ um operador monótono, ou seja, $\langle F(x) - F(y), x - y \rangle \geq 0$, para todo $x, y \in C$. Defina a bifunção $f(x, y) = \langle F(x), y - x \rangle$, então f é monótona em C .*

Demonstração. De fato, para $x, y \in C$, temos:

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

pela monotonicidade de F . □

Outra propriedade fundamental imposta à bifunção para garantir a convergência dos métodos propostos é a condição do tipo-Lipschitz de Mastroeni, descrita a seguir.

Definição 1.2.2. [36, Teorema 3.1, p.294] *Dizemos que uma bifunção f satisfaz uma condição do tipo-Lipschitz em C , se existem constantes $c_1 > 0$, $c_2 > 0$ tais que*

$$f(x, y) + f(y, z) \geq f(x, z) - c_1 \|x - y\|^2 - c_2 \|y - z\|^2, \forall x, y, z \in C.$$

A condição do tipo-Lipschitz será central na análise de convergência dos algoritmos propostos nesta tese. O próximo exemplo mostra uma classe importante de bifunções que satisfazem essa propriedade.

Exemplo 1.2.2 (Bifunção do tipo-Lipschitz). *Seja $F : C \rightarrow H$ um operador Lipschitz-contínuo com constante $L > 0$, isto é, $\|F(x) - F(y)\| \leq L\|x - y\|$, $\forall x, y \in C$. A bifunção definida por*

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

satisfaz a condição do tipo-Lipschitz em C .

Demonstração. Sejam $x, y, z \in C$. Temos que:

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

Pela desigualdade de Cauchy-Schwarz e L -Lipschitz-continuidade de F , obtemos

$$|\langle F(x) - F(y), y - z \rangle| \leq \|F(x) - F(y)\| \|y - z\| \leq L\|x - y\| \|y - z\|.$$

Mas sabemos que $ab \leq \frac{1}{2}(a^2 + b^2)$, $\forall a, b \in \mathbb{R}$. Então,

$$L\|x - y\| \|y - z\| \leq \frac{L}{2} (\|x - y\|^2 + \|y - z\|^2).$$

Logo,

$$\langle F(x) - F(y), y - z \rangle \geq -\frac{L}{2} (\|x - y\|^2 + \|y - z\|^2),$$

e conseqüentemente,

$$f(x, y) + f(y, z) \geq f(x, z) - \frac{L}{2} (\|x - y\|^2 + \|y - z\|^2),$$

ou seja, f satisfaz a condição do tipo-Lipschitz com constantes $c_1 = c_2 = \frac{L}{2}$. □

Apresentamos a seguir uma bifunção concreta, inspirada em [34], que satisfaz simultaneamente as propriedades de monotonicidade e do tipo-Lipschitz no espaço de Hilbert finito-dimensional $\mathbb{R}^2$.

Exemplo 1.2.3. A aplicação não-linear $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$F(a, b) = (a + b + \text{sen } a, -a + b + \text{sen } b)$$

é monótona e Lipschitz-contínua em $C = [-2, 5] \times [-2, 5]$. Consequentemente, a bifunção $f : C \times C \rightarrow \mathbb{R}$ dada por $f((a, b), (c, d)) = \langle F(a, b), (c - a, d - b) \rangle$ é monótona e do tipo-Lipschitz, pelos Exemplos 1.2.1 e 1.2.2.

Demonstração. Primeiramente, mostraremos que F é monótona em C .

Sejam $(a_1, b_1), (a_2, b_2) \in C$. Temos:

$$\begin{aligned} F(a_1, b_1) - F(a_2, b_2) = & \left((a_1 - a_2) + (b_1 - b_2) + (\text{sen } a_1 - \text{sen } a_2), \right. \\ & \left. - (a_1 - a_2) + (b_1 - b_2) + (\text{sen } b_1 - \text{sen } b_2) \right). \end{aligned} \quad (1.3)$$

e

$$(a_1, b_1) - (a_2, b_2) = (a_1 - a_2, b_1 - b_2).$$

Assim, resolvendo o produto interno, obtemos:

$$\begin{aligned} \langle F(a_1, b_1) - F(a_2, b_2), (a_1 - a_2, b_1 - b_2) \rangle = & (a_1 - a_2)^2 + (b_1 - b_2)^2 \\ & + (\text{sen } a_1 - \text{sen } a_2)(a_1 - a_2) \\ & + (\text{sen } b_1 - \text{sen } b_2)(b_1 - b_2). \end{aligned} \quad (1.4)$$

Mas a função seno é Lipschitz-contínua em $\mathbb{R}$, com constante 1, isto é,

$$|\text{sen } a - \text{sen } b| \leq |a - b|, \forall a, b \in \mathbb{R}.$$

De fato, pelo Teorema do Valor Médio, para a função seno, existe $t \in (a, b)$ tal que

$$\frac{\text{sen } a - \text{sen } b}{a - b} = \cos t.$$

Aplicando o módulo e usando o fato que $|\cos t| \leq 1, \forall t \in \mathbb{R}$, chegamos em:

$$|\text{sen } a - \text{sen } b| \leq |a - b|, \forall a, b \in \mathbb{R}.$$

Logo,

$$|(\text{sen } a - \text{sen } b)(a - b)| \leq |a - b|^2,$$

donde

$$(\text{sen } a - \text{sen } b)(a - b) \geq -(a - b)^2, \forall a, b \in \mathbb{R}. \quad (1.5)$$

Aplicando (1.5) em (1.4), concluímos:

$$\langle F(a_1, b_1) - F(a_2, b_2), (a_1 - a_2, b_1 - b_2) \rangle \geq 0, \forall (a_1, b_1), (a_2, b_2) \in C.$$

Portanto, F é monótona em C .

Agora, mostraremos que F é Lipschitz-contínua, com constante $L = 3$.

Sejam $(a_1, b_1), (a_2, b_2) \in C$. Utilizando a norma euclidiana em (1.3), obtemos:

$$\begin{aligned} \|F(a_1, b_1) - F(a_2, b_2)\|^2 &= ((a_1 - a_2) + (b_1 - b_2) + (\sin a_1 - \sin a_2))^2 \\ &\quad + (-(a_1 - a_2) + (b_1 - b_2) + (\sin b_1 - \sin b_2))^2. \end{aligned}$$

Aplicando a desigualdade triangular e o fato de que $|\sin a - \sin b| \leq |a - b|, \forall a, b \in \mathbb{R}$, segue que:

$$|(a_1 - a_2) + (b_1 - b_2) + (\sin a_1 - \sin a_2)| \leq 2|a_1 - a_2| + |b_1 - b_2|$$

e

$$|-(a_1 - a_2) + (b_1 - b_2) + (\sin b_1 - \sin b_2)| \leq |a_1 - a_2| + 2|b_1 - b_2|.$$

Consequentemente,

$$\|F(a_1, b_1) - F(a_2, b_2)\|^2 \leq (2|a_1 - a_2| + |b_1 - b_2|)^2 + (|a_1 - a_2| + 2|b_1 - b_2|)^2.$$

Mas

$$(2|a_1 - a_2| + |b_1 - b_2|)^2 = 4(a_1 - a_2)^2 + 4|a_1 - a_2||b_1 - b_2| + (b_1 - b_2)^2$$

e

$$(|a_1 - a_2| + 2|b_1 - b_2|)^2 = (a_1 - a_2)^2 + 4|a_1 - a_2||b_1 - b_2| + 4(b_1 - b_2)^2.$$

Somando:

$$\|F(a_1, b_1) - F(a_2, b_2)\|^2 \leq 5(a_1 - a_2)^2 + 5(b_1 - b_2)^2 + 8|a_1 - a_2||b_1 - b_2|.$$

Aplicando a desigualdade $2ab \leq a^2 + b^2$ (válida $\forall a, b \in \mathbb{R}$) ao termo misto, temos:

$$8|a_1 - a_2||b_1 - b_2| \leq 4(a_1 - a_2)^2 + 4(b_1 - b_2)^2.$$

Assim,

$$\|F(a_1, b_1) - F(a_2, b_2)\|^2 \leq 9(a_1 - a_2)^2 + 9(b_1 - b_2)^2 = 9\|(a_1 - a_2, b_1 - b_2)\|^2.$$

Tomando a raiz quadrada em ambos os lados, concluimos:

$$\|F(a_1, b_1) - F(a_2, b_2)\| \leq 3\|(a_1, b_1) - (a_2, b_2)\|.$$

Portanto, F é Lipschitz-contínua em C , com constante de Lipschitz $L = 3$. □

Hipóteses de trabalho sobre a bifunção f

Ao longo desta tese, consideraremos o problema (PE) associado a uma bifunção $f : H \times H \rightarrow \mathbb{R}$ que satisfaz:

(A1) $f(x, x) = 0$, para todo $x \in C$;

(A2) f é pseudomonótona em C ;

(A3) f verifica a condição do tipo-Lipschitz em H , com constantes c_1 e c_2 ;

(A4) $f(x, \cdot)$ é convexa e semicontínua inferior em H , para todo $x \in H$;

(A5) $f(\cdot, y)$ é fracamente semicontínua superior em C , para todo $y \in C$.

Vejamos um exemplo de uma f , definida em um espaço de Hilbert de dimensão infinita, que verifica as hipóteses acima.

Exemplo 1.2.4 (Operador integral de Hilbert–Schmidt em $L^2(0, 1)$). *Seja $H := L^2(0, 1)$, com produto interno $\langle u, v \rangle = \int_0^1 u(t)v(t) dt$ e norma $\|u\| = \langle u, u \rangle^{1/2}$. Considere o operador integral (limitado) $K : H \rightarrow H$ definido por*

$$(Kx)(t) = \int_0^1 k(t, s) x(s) ds,$$

onde o núcleo $k \in L^2((0, 1) \times (0, 1))$ é simétrico ($k(t, s) = k(s, t)$) e positivo semidefinido, de modo que K é autoadjunto e positivo semidefinido ($K \succeq 0$).

Defina a bifunção $f : H \times H \rightarrow \mathbb{R}$ por

$$f(x, y) := \langle Kx, y - x \rangle, \quad x, y \in H,$$

e (se desejar impor restrição de viabilidade) tome o conjunto não vazio, convexo e fechado $C := \{x \in H : a \leq x(t) \leq b \text{ q.t.p. em } (0, 1)\}$, com $a < b$. Então f satisfaz as hipóteses (A1)–(A5).

Demonstração. (A1) f é de equilíbrio. Para todo $x \in H$,

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

(A2) **(Pseudo)monotonicidade.** Para $x, y \in H$, temos:

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

Como $K \succeq 0$ (isto é, $\langle K(u - v), u - v \rangle \geq 0$ para todos u, v), obtemos

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

isto é, f é monótona em H e, consequentemente, pseudomonótona em H .

(A3) **Condição do tipo-Lipschitz em H .** Para $x, y, z \in H$,

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

Pela desigualdade de Cauchy–Schwarz e pela continuidade de K ,

$$|\langle Kx - Ky, y - z \rangle| \leq \|K\| \|x - y\| \|y - z\| \leq \frac{\|K\|}{2} (\|x - y\|^2 + \|y - z\|^2),$$

onde usamos $ab \leq \frac{1}{2}(a^2 + b^2)$. Assim,

$$f(x, y) + f(y, z) \geq f(x, z) - \frac{\|K\|}{2} \|x - y\|^2 - \frac{\|K\|}{2} \|y - z\|^2,$$

isto é, (A3) vale em todo H com $c_1 = c_2 = \frac{\|K\|}{2}$.

(A4) **Convexidade e semicontinuidade inferior em y .** Fixe $x \in H$. Como

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

temos que $f(x, \cdot)$ é *afim* em y , logo convexa e contínua (portanto s.c.i.) em H .

(A5) **Fraca semicontinuidade superior de $f(\cdot, y)$ em C .** Fixe $y \in C$ e tome uma sequência $x_n \rightarrow x$ em C . Escrevemos

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

(i) Como K é linear e contínuo, de $x_n \rightarrow x$ em C segue que $Kx_n \rightarrow Kx$. Assim, para todo $y \in C$ fixo, obtemos

$$\langle Kx_n, y \rangle \rightarrow \langle Kx, y \rangle,$$

o que mostra que o funcional $x \mapsto \langle Kx, y \rangle$ é fracamente contínuo.

(ii) Como K é autoadjunto e positivo ($K \geq 0$), existe um operador $S := K^{1/2}$ limitado, autoadjunto e positivo, tal que $K = S^2$. Logo,

$$\langle Kx, x \rangle = \|Sx\|^2, \quad \forall x \in H.$$

Como a norma é fracamente semicontínua inferior em espaços de Hilbert, temos

$$\|Sx\|^2 \leq \liminf_{n \rightarrow \infty} \|Sx_n\|^2.$$

Portanto, a forma quadrática $q(x) := \langle Kx, x \rangle$ é fracamente semicontínua inferior, e consequentemente $-q$ é fracamente semicontínua superior.

Assim, $f(x, y)$ é a soma de uma função fracamente contínua com outra fracamente semicontínua superior, logo $f(\cdot, y)$ é fracamente semicontínua superior em C , como desejado. $\square$

Problemas clássicos no formato PE

Nesta subseção, apresentamos classes tradicionais de problemas que se enquadram na formulação (1.2). Damos um tratamento especial ao *Problema de Desigualdade Variacional* (PDV): enunciando hipóteses suficientes que garantem (A1)–(A5) e demonstrando a equivalência $\text{PDV} \Leftrightarrow \text{PE}$.

Para os demais modelos — *Otimização Convexa*, *Ponto Fixo*, *Complementariedade* e *Equilíbrio de Nash* — registramos suas instâncias como casos particulares de (1.2) e indicamos, quando pertinente, condições adicionais sob as quais (A1)–(A5) também são satisfeitas (por exemplo, potencial convexo no caso de jogos, ou monotonicidade/Lipschitz do pseudogradiante na redução a PDV).

Por fim, exibimos um *exemplo genuíno* de bifunção que satisfaz (A1)–(A5), mas que não se reduz a PDV nem admite representação separável do tipo $\phi(y) - \phi(x)$, tampouco a forma agregada de *Equilíbrio de Nash*. Esse exemplo evidencia que o PE é estritamente mais geral do que os casos (i)–(v) a seguir.

(i) **Desigualdade Variacional.** Seja $F : C \rightarrow H$ um operador satisfazendo as condições:

(C1) F é pseudomonótono em C , isto é,

$$\langle F(x), y - x \rangle \geq 0 \Rightarrow \langle F(y), x - y \rangle \leq 0, \forall x, y \in C;$$

(C2) Para algum $L > 0$, F é L -Lipschitz contínuo em H , isto é,

$$\|F(x) - F(y)\| \leq L\|x - y\|, \forall x, y \in H;$$

(C3) $\limsup_{n \rightarrow \infty} \langle F(x_n), y - x_n \rangle \leq \langle F(x), y - x \rangle$, para todo $y \in C$ e toda sequência $\{x_n\} \subset C$ tal que $x_n \rightarrow x$ em C .

Lembramos que o PDV associado a (F, C) é definido por:

$$\text{Encontrar } x^* \in C \text{ tal que } \langle F(x^*), y - x^* \rangle \geq 0, \forall y \in C. \quad (1.6)$$

Se tomarmos

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

então f satisfaz as condições (A1)–(A5), e a solução do PDV (1.6) coincide com a solução do PE (1.2).

Demonstração. Faremos a prova apenas da equivalência PE–PDV, pois a verificação de que f satisfaz (A1)–(A5) decorre diretamente das hipóteses assumidas para F .

Se x^* resolve o PE, então

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

isto é, x^* também resolve o PDV.

Reciprocamente, se x^* resolve o PDV, segue da definição de f que $f(x^*, y) \geq 0$ para todo $y \in C$, ou seja, x^* resolve o PE. Portanto, PE e PDV são equivalentes.

□

(ii) **Otimização convexa.** Considere o problema

$$\min\{\phi(x) : x \in C\},$$

onde $\phi : C \rightarrow \mathbb{R} \cup \{+\infty\}$ é própria, convexa e s.c.i., e $C \subset H$ é convexo e fechado. Definindo

$$f(x, y) := \phi(y) - \phi(x), \quad x, y \in C,$$

obtem-se que

$$x^* \text{ resolve PE}(f, C) \iff x^* \text{ é minimizador global de } \phi \text{ em } C.$$

Além disso, f satisfaz (A1)–(A5).

Se, adicionalmente, ϕ é diferenciável em C e definimos

$$F(x) := \nabla \phi(x), \quad x \in C,$$

então as condições de otimalidade de primeira ordem implicam

$$x^* \in \arg \min_{x \in C} \phi(x) \iff \langle F(x^*), y - x^* \rangle \geq 0, \quad \forall y \in C,$$

isto é, x^* resolve também o PDV associado ($\text{PDV}(F, C)$).

Em particular, sob diferenciabilidade, o problema de otimização convexa, o PE em (ii) e o PDV associado são três formulações equivalentes do mesmo problema (ver, por exemplo, [16, Seção 1.3.1, p. 13]). No caso geral em que ϕ é apenas própria, convexa e s.c.i., a formulação natural é via PE; a passagem a uma desigualdade variacional requer considerar operadores multivalorados do tipo

$$T(x) := \partial\phi(x) + N_C(x).$$

Essa abordagem conduz a desigualdades variacionais generalizadas, coerentes com a caracterização subdiferencial de soluções de problemas de otimização convexa restritos, mas fogem ao escopo da definição de PDV (monovalorado) adotada neste texto.

- (iii) **Ponto Fixo.** Dado um operador $T : C \rightarrow C$, o Problema de Ponto Fixo consiste em encontrar $x^* \in C$ tal que

$$x^* = T(x^*).$$

Defina a bifunção

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

Então x^* resolve o PE associado a $f \Leftrightarrow x^*$ é ponto fixo de T , isto é, $x^* = T(x^*)$, o que também pode ser visto como um PDV com $F(x) = x - T(x)$.

Além disso, se T é não-expansivo (veja [4]), então $F := I - T$ é monótono e L -Lipschitz ($L \leq 2$), logo f satisfaz (A1)–(A5) pelas condições suficientes já listadas para o caso $f(x, y) = \langle F(x), y - x \rangle$.

- (iv) **Complementariedade.** Seja $K \subset H$ um cone convexo fechado, $K^* = \{u \in H : \langle u, v \rangle \geq 0, \forall v \in K\}$ o seu dual, e $F : K \rightarrow H$ um operador. Considere o Problema de Complementariedade (PC): encontrar $x^* \in K$ tal que $F(x^*) \in K^*$ e $\langle F(x^*), x^* \rangle = 0$. Definindo

$$C := K, \quad f(x, y) = \langle F(x), y - x \rangle,$$

temos que a solução do PE coincide com a do PC, portanto também com a do PDV em K . Se F é (pseudo)monótono em K e L -Lipschitz (em H), então f satisfaz (A1)–(A5) pelas mesmas condições do caso PDV.

- (v) **Equilíbrio de Nash (m jogadores).** Considere um jogo não cooperativo com m jogadores, em que cada jogador i escolhe uma estratégia em um conjunto convexo e fechado $C_i \subset \mathbb{R}^{n_i}$, e seja

$$C = C_1 \times \cdots \times C_m.$$

Para cada $i \in \{1, \dots, m\}$, o custo do jogador i é dado por uma função $\phi_i : C \rightarrow \mathbb{R}$.

Um vetor $x^* = (x_1^*, \dots, x_m^*) \in C$ é chamado *equilíbrio de Nash* (se todos os jogadores minimizam custos) se, para todo $i \in \{1, \dots, m\}$,

$$\phi_i(x_i^*, x_{-i}^*) \leq \phi_i(y_i, x_{-i}^*), \quad \forall y_i \in C_i,$$

onde x_{-i}^* denota o vetor das estratégias dos demais jogadores.

Defina a bifunção

$$f(x, y) = \sum_{i=1}^m (\phi_i(y_i, x_{-i}) - \phi_i(x_i, x_{-i})), \quad x, y \in C.$$

Então $x^* \in C$ é equilíbrio de Nash $\Leftrightarrow f(x^*, y) \geq 0$ para todo $y \in C$.

Além disso, f satisfaz (A1)–(A5) se, por exemplo:

- (a) o jogo for *potencial convexo*, caso em que $f(x, y) = \Phi(y) - \Phi(x)$ (ver (i) **Otimização Convexa**);
- (b) sob *convexidade por bloco e diferenciabilidade por bloco*, isto é: para cada $i \in \{1, \dots, m\}$ a função $x_i \mapsto \phi_i(x_i, x_{-i})$ é convexa e continuamente diferenciável (C^1), com x_{-i} fixo, obtém-se uma redução a PDV, onde

$$F(x) = (\nabla_{x_1} \phi_1(x), \dots, \nabla_{x_m} \phi_m(x))^T,$$

isto é, x é Nash $\Leftrightarrow x$ é solução do PDV, com $C = \prod_{i=1}^m C_i$ ([16, Prop. 1.4.2, pp. 24–25]).

Em contraste, há bifunções que satisfazem (A1)–(A5) e *não* se enquadram em (i)–(v), fornecendo exemplos genuínos de PE que não reduzem a PDV nem a formulações separáveis do tipo $h(y) - h(x)$. Vejamos um exemplo.

Exemplo 1.2.5. [25, Proposição 2.6, p.627] Considere $H = \mathbb{R}^n$ e $C = H$. Defina $f : C \times C \rightarrow \mathbb{R}$ por

$$f(x, y) := y^T A y + x^T B y - x^T A x,$$

onde $A \in \mathbb{R}^{n \times n}$ é semidefinida positiva ($A \succeq 0$) e $B \in \mathbb{R}^{n \times n}$ é antissimétrica ($B^T = -B$). Suponha que exista $x_0 \in \mathbb{R}^n$ tal que $x_0^T A x_0 > 0$. Então:

- (a) f satisfaz (A1)–(A5);
- (b) f **não** é do tipo (i) (PDV), i.e., não existe operador F tal que

$$f(x, y) = \langle F(x), y - x \rangle \quad \text{para todos } x, y \in C;$$

- (c) f **não** é do tipo (ii) Otimização convexa, i.e., não existe ϕ convexa e s.c.i. com

$$f(x, y) = \phi(y) - \phi(x);$$

- (d) f **não** é do tipo (v) Equilíbrio de Nash, i.e., não existem funções $\{\phi_i\}_{i=1}^m$ e uma decomposição $x = (x_1, \dots, x_m)$ tal que

$$f(x, y) = \sum_{i=1}^m (\phi_i(y_i, x_{-i}) - \phi_i(x_i, x_{-i})) \quad \text{para todos } x, y.$$

Demonstração. (a) (A1)–(A5).

(A1) **f é de equilíbrio.**

Temos

$$f(x, x) = x^\top Ax + x^\top Bx - x^\top Ax = x^\top Bx.$$

Como B é antissimétrica ($B^\top = -B$), segue que

$$x^\top Bx = (x^\top Bx)^\top = x^\top B^\top x = x^\top (-B)x = -x^\top Bx,$$

logo, $x^\top Bx = 0$. Assim, $f(x, x) = 0$ para todo $x \in C$.

(A2) **f é pseudomonótona em C .**

Para quaisquer $x, y \in C$,

$$f(x, y) + f(y, x) = x^\top By + y^\top Bx = x^\top By + x^\top B^\top y = x^\top (B + B^\top)y = 0,$$

pela antissimetria de B .

Logo $f(y, x) = -f(x, y)$, e em particular $f(x, y) \geq 0 \Rightarrow f(y, x) \leq 0, \forall x, y \in C$.

(A3) **f é tipo-Lipschitz.**

Para quaisquer $x, y, z \in H$,

$$\begin{aligned} f(x, y) + f(y, z) - f(x, z) &= (y^\top Ay + x^\top By - x^\top Ax) + (z^\top Az + y^\top Bz - y^\top Ay) \\ &\quad - (z^\top Az + x^\top Bz - x^\top Ax) \\ &= \underbrace{y^\top Ay - y^\top Ay}_0 + \underbrace{z^\top Az - z^\top Az}_0 + \underbrace{(-x^\top Ax) - (-x^\top Ax)}_0 \\ &\quad + x^\top B(y - z) + y^\top Bz \\ &= (x - y)^\top B(y - z), \end{aligned}$$

onde usamos $y^\top By = 0$, pois $B^\top = -B$.

Pela desigualdade de Cauchy–Schwarz e norma induzida de B , obtemos

$$|f(x, y) + f(y, z) - f(x, z)| = |(x - y)^\top B(y - z)| \leq \|B\| \|x - y\| \|y - z\|.$$

Aplicando $ab \leq \frac{1}{2}(a^2 + b^2)$,

$$|f(x, y) + f(y, z) - f(x, z)| \leq \frac{\|B\|}{2} (\|x - y\|^2 + \|y - z\|^2).$$

Logo, para todos $x, y, z \in H$,

$$f(x, y) + f(y, z) \geq f(x, z) - c_1 \|x - y\|^2 - c_2 \|y - z\|^2, \quad c_1 = c_2 = \frac{\|B\|}{2}.$$

(A4) **Convexidade e semicontinuidade inferior de $f(x, \cdot)$ em $H = \mathbb{R}^n$.**

Passo 1 (Convexidade de $f(x, \cdot)$). Fixe $x \in H = \mathbb{R}^n$ e escreva

$$f(x, y) = \underbrace{y^\top Ay}_{=: q_A(y)} + \underbrace{x^\top By}_{\text{linear em } y} - x^\top Ax.$$

A função quadrática $q_A(y) = y^\top Ay$ é C^2 com

$$\nabla q_A(y) = (A + A^\top)y \quad \text{e} \quad \nabla^2 q_A(y) = A + A^\top.$$

Como $A \succeq 0$, isto é, $u^\top Au \geq 0$ para todo u , então

$$u^\top (A + A^\top)u = u^\top Au + u^\top A^\top u = u^\top Au + (u^\top Au) = 2u^\top Au \geq 0,$$

logo $A + A^\top \succeq 0$. Portanto, pela caracterização clássica via Hessiana semidefinida positiva, q_A é convexa.

Além disso, o termo $y \mapsto x^\top By$ é linear em y e $-x^\top Ax$ é constante em y , logo também são convexas. Assim, $f(x, \cdot)$ é a soma de funções convexas, portanto é convexa.

Passo 2 (semicontinuidade inferior de $f(x, \cdot)$). Como q_A é polinomial e $x^\top By$ é linear, $f(x, \cdot)$ é contínua em $\mathbb{R}^n$; em particular, é semicontínua inferior.

(A5) **Semicontinuidade superior de $f(\cdot, y)$.**

Como estamos em dimensão finita ($H = \mathbb{R}^n$), a topologia fraca coincide com a topologia forte. Assim, basta provar que $f(\cdot, y)$ é semicontínua superior em C para a topologia da norma.

Fixe $y \in C$ e escreva

$$f(x, y) = x^\top By - x^\top Ax + y^\top Ay.$$

A função $x \mapsto x^\top By$ é linear e contínua; $x \mapsto x^\top Ax$ é polinomial (grau 2) e, portanto, contínuo; e $y^\top Ay$ é constante em x . Logo $x \mapsto f(x, y)$ é contínua em $C = \mathbb{R}^n$ e, em particular, semicontínua superior em C .

(b) **f não é do tipo (i)(PDV).** Suponha, por contradição, que exista $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que

$$f(x, y) = \langle F(x), y - x \rangle \quad (\forall x, y \in C).$$

Pela hipótese, existe $x \in \mathbb{R}^n$ com $a := x^\top Ax > 0$. Usando $x^\top Bx = 0$ (pois $B^\top = -B$), obtemos

$$f(x, \lambda x) = (\lambda x)^\top A(\lambda x) + x^\top B(\lambda x) - x^\top Ax = (\lambda^2 - 1)a \quad (\forall \lambda \in \mathbb{R}).$$

Por outro lado, pela forma de PDV,

$$f(x, \lambda x) = \langle F(x), (\lambda - 1)x \rangle = (\lambda - 1) \langle F(x), x \rangle.$$

Igualando as expressões para $\lambda = 2$ e $\lambda = 3$,

$$\begin{cases} 3a = \langle F(x), x \rangle, \\ 8a = 2 \langle F(x), x \rangle, \end{cases} \implies 3a = 4a \implies a = 0,$$

o que contradiz $a > 0$. Logo, tal operador F não existe e f não é do tipo PDV.

(c) **f não é do tipo (ii) Otimização convexa (não é separável).** Suponha, por contradição, que exista ϕ convexa e s.c.i. tal que $f(x, y) = \phi(y) - \phi(x)$ para todos x, y . Então f é *separável* em (x, y) , logo todas as derivadas mistas de segunda ordem se anulam:

$$\frac{\partial^2 f}{\partial x \partial y} \equiv 0.$$

Por outro lado, para

$$f(x, y) = y^\top Ay + x^\top By - x^\top Ax,$$

temos

$$\frac{\partial f}{\partial x}(x, y) = By - (A + A^\top)x \implies \frac{\partial^2 f}{\partial x \partial y}(x, y) = B.$$

Se $B \neq 0$, segue que $\frac{\partial^2 f}{\partial x \partial y} \equiv B \neq 0$, o que contradiz a separabilidade. Logo tal ϕ não existe e f não é do tipo $\phi(y) - \phi(x)$.

(d) **f não é do tipo (v) Equilíbrio de Nash (caso $m = 2$).**

Escreva a decomposição em dois jogadores

$$x = (x_1, x_2), \quad y = (y_1, y_2), \quad x_i, y_i \in \mathbb{R}^{n_i}, \quad n_1 + n_2 = n,$$

e particione as matrizes em blocos compatíveis:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Suponha, por contradição, que existam funções ϕ_1, ϕ_2 tais que, para todos x, y ,

$$f(x, y) = (\phi_1(y_1, x_2) - \phi_1(x_1, x_2)) + (\phi_2(y_2, x_1) - \phi_2(x_2, x_1)). \quad (1.7)$$

A forma (1.7) impõe as seguintes *condições necessárias* de nulidade de derivadas mistas:

$$\frac{\partial^2 f}{\partial x_1 \partial y_1} \equiv 0, \quad \frac{\partial^2 f}{\partial x_2 \partial y_2} \equiv 0, \quad \frac{\partial^2 f}{\partial y_1 \partial y_2} \equiv 0, \quad (1.8)$$

pois, em (1.7), y_i não aparece acoplado com x_i no mesmo termo e y_1 e y_2 pertencem a termos distintos.

Por outro lado, para a nossa bifunção

$$f(x, y) = y^\top A y + x^\top B y - x^\top A x,$$

temos (com a convenção euclidiana $\langle u, v \rangle = v^\top u$):

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= B y - (A + A^\top)x \implies \frac{\partial^2 f}{\partial x \partial y}(x, y) = B, \\ \frac{\partial f}{\partial y}(x, y) &= (A + A^\top)y + B^\top x \implies \frac{\partial^2 f}{\partial y \partial y}(x, y) = A + A^\top. \end{aligned}$$

Logo, em blocos,

$$\frac{\partial^2 f}{\partial x_1 \partial y_1} = B_{11}, \quad \frac{\partial^2 f}{\partial x_2 \partial y_2} = B_{22}, \quad \frac{\partial^2 f}{\partial y_1 \partial y_2} = A_{12} + A_{21}^\top. \quad (1.9)$$

Comparando (1.8) com (1.9), a existência da representação de Nash (1.7) exigiria, necessariamente,

$$B_{11} = 0, \quad B_{22} = 0, \quad A_{12} + A_{21}^\top = 0. \quad (1.10)$$

Entretanto, no contexto do exemplo, $B^\top = -B$ é antissimétrica arbitrária (em particular, não se impõe que os blocos diagonais B_{11}, B_{22} sejam nulos), e $A \succeq 0$ é apenas semidefinida positiva no sentido quadrático (não se assume que seja bloco-diagonal). Assim, para uma partição qualquer $n = n_1 + n_2$, é típico que pelo menos uma das igualdades em (1.10) *falhe*:

$$\text{ou } B_{11} \neq 0, \quad \text{ou } B_{22} \neq 0, \quad \text{ou } A_{12} + A_{21}^\top \neq 0.$$

Nesses casos, as condições necessárias para a forma de Nash são violadas, e (1.7) é impossível. Conclui-se que f **não** é do tipo (v) Equilíbrio de Nash (no caso genuíno $B \neq 0$, ou sempre que $A_{12} + A_{21}^\top \neq 0$ para a partição considerada).

□

A Figura 1.1, abaixo, resume as relações de representabilidade (inclusão) e de equivalência de soluções, indicando também quando a redução ao PDV exige condições adicionais.

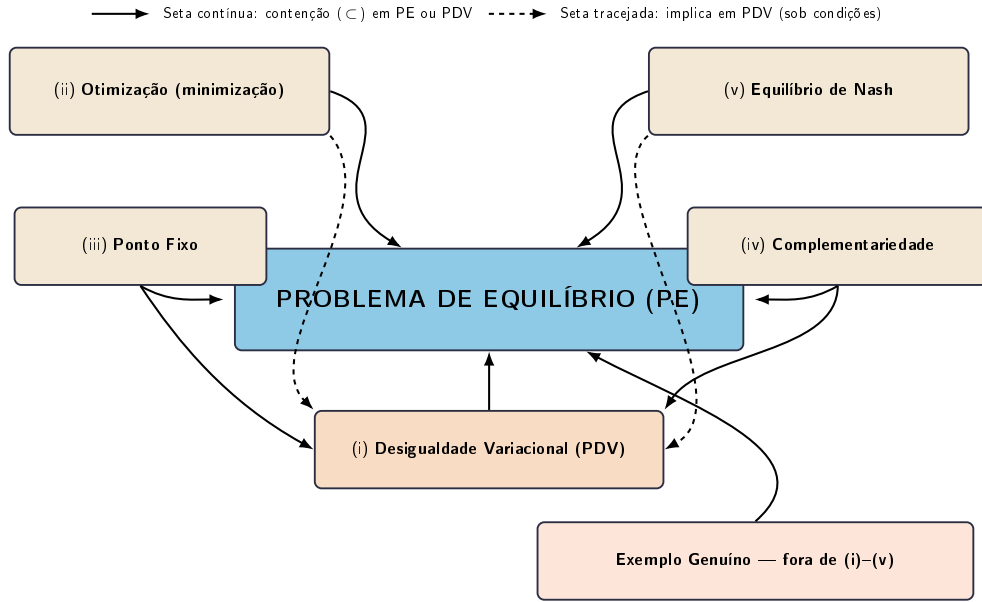


Figura 1.1: **Diagrama:** PE, casos (i)–(v) e relações com PDV.

1.3 O Método Extragradiante e sua Extensão Subgradiente

Os métodos extragradiante ocupam papel central na resolução de desigualdades variacionais e, mais amplamente, de problemas de equilíbrio, especialmente quando a monotonicidade do operador ou da bifunção é apenas fraca (por exemplo, pseudomonotonicidade). Nesta seção, abordamos duas variantes fundamentais: o Método Extragradiante clássico e sua extensão conhecida como Método Extragradiante Subgradiente.

1.3.1 Surgimento e Motivação

O Método Extragradiante foi proposto por Korpelevich [32], em 1976, com o objetivo de resolver problemas de ponto de sela associados a operadores monótonos definidos sobre subconjuntos convexos e fechados de espaços euclidianos. O método surgiu como uma alternativa aos métodos de projeção direta (como o gradiente projetado), que podem falhar em garantir convergência ou apresentar desempenho insatisfatório quando o operador não é fortemente monótono (ver exemplo em [24]). Desde então, o método tem sido adaptado e estendido para tratar uma ampla variedade de problemas, destacando-se, nesta tese, os problemas de equilíbrio com bifunções pseudomonótonas em espaços de Hilbert reais e, em particular, os problemas de desigualdade variacional. A literatura especializada sobre o tema é vasta, e mencionamos, a título de exemplo, os trabalhos [11, 29, 41, 45, 50, 62] e suas referências.

Apresentamos a seguir a formulação original do Método Extragradiante de Korpelevich (MEGK) para resolver um problema de desigualdade variacional em espaços de Hilbert.

O algoritmo extragradiante (também chamado de método da dupla projeção) para resolver (PDV) é descrito por:

$$\begin{cases} y_n = P_C(x_n - \lambda F(x_n)), \\ x_{n+1} = P_C(x_n - \lambda F(y_n)), \end{cases}$$

em que $x_0 \in C$ é dado, $\lambda > 0$ é o tamanho de passo e P_C é a projeção métrica sobre C .

Quando $H = \mathbb{R}^n$, a convergência do método para uma solução da desigualdade variacional foi obtida por Korpelevich [32] sob as hipóteses de que o conjunto-solução do (PDV) é não vazio, F é monótono e L -Lipschitz-contínuo e λ satisfazendo a condição $0 < \lambda < \frac{1}{L}$.

Em espaços de Hilbert, uma demonstração da convergência fraca das sequências $\{x_n\}$ e $\{y_n\}$ para uma mesma solução do (PDV), sob as mesmas hipóteses de Korpelevich, bem como uma ilustração do processo iterativo correspondente, encontram-se em Censor et al. [11].

A figura abaixo reproduz essa ilustração.

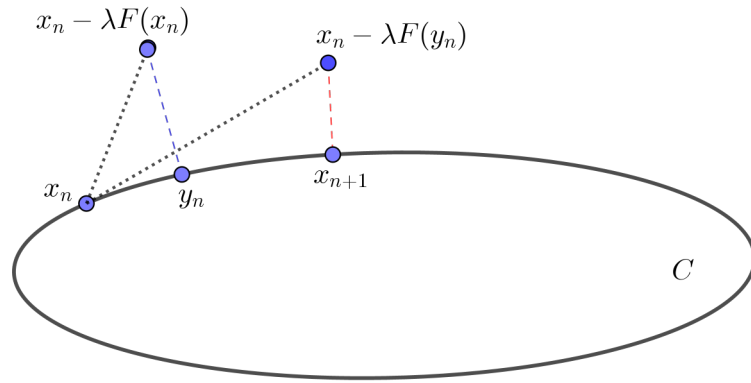


Figura 1.2: Representação geométrica de uma iteração do (MEGK). A partir de x_n , realiza-se uma projeção sobre C após um deslocamento na direção oposta ao vetor $F(x_n)$, obtendo-se o ponto intermediário y_n e realiza-se uma nova projeção para determinar o próximo iterado x_{n+1} .

A extensão do método extragradiente para problemas de equilíbrio é natural quando se observa que, ao definir $f(x, y) = \langle F(x), y - x \rangle$, o problema (PE) se reduz ao problema (PDV), conforme visto na seção anterior (**Problemas clássicos no formato PE**).

Assim, o algoritmo extragradiente para (PE) pode ser reescrito como dois problemas de otimização fortemente convexos:

$$\begin{cases} y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}, \\ x_{n+1} = \operatorname{argmin}_{y \in C} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\}. \end{cases} \quad (1.11)$$

Esse formato segue da equivalência entre projeção métrica e problema de minimização quadrática.

Demonstração da equivalência.

Da Definição 1.1.5, segue que

$$\begin{aligned} y_n &= P_C(x_n - \lambda F(x_n)) \\ &= \operatorname{argmin}_{y \in C} \left\{ \frac{1}{2} \|(x_n - \lambda F(x_n)) - y\|^2 \right\}. \end{aligned}$$

E do Lema 1.1.4(c), obtemos

$$\begin{aligned}
 y_n &= \operatorname{argmin}_{y \in C} \left\{ \frac{1}{2} \left(\|x_n - y\|^2 + \lambda^2 \|F(x_n)\|^2 - 2\lambda \langle x_n - y, F(x_n) \rangle \right) \right\} \\
 &= \operatorname{argmin}_{y \in C} \left\{ \lambda \langle F(x_n), y - x_n \rangle + \frac{1}{2} \|x_n - y\|^2 + \frac{\lambda^2}{2} \|F(x_n)\|^2 \right\} \\
 &= \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\},
 \end{aligned}$$

onde a última igualdade segue do fato que $\operatorname{argmin}_{y \in C} \{g(y) + a\} = \operatorname{argmin}_{y \in C} \{g(y)\}$.

Analogamente, obtemos o argumento mínimo x_{n+1} .

Variações deste método incorporam passos variáveis ou adaptativos, regularizações e estratégias inerciais, com o objetivo de melhorar a convergência e reduzir o custo computacional; veja, por exemplo, [10, 20, 21, 56].

Em síntese, o método extragradiente representa uma das ferramentas fundamentais na análise e resolução de problemas de equilíbrio e, em particular, de desigualdade variacional, tanto na teoria quanto em aplicações numéricas.

1.3.2 O Método Extragradiente Subgradiente

Como discutido anteriormente, o método extragradiente de Korpelevich exige a realização de duas projeções no conjunto viável C por iteração. Quando C possui uma estrutura simples — por exemplo, é uma bola, uma caixa ou um subespaço fechado — essas projeções são computacionalmente eficientes. No entanto, em muitos problemas aplicados, C é definido por restrições não explícitas, tornando a dupla projeção um passo oneroso do ponto de vista computacional.

Para superar essa dificuldade, Censor, Gibali e Reich [11, 12] propuseram o Método Extragradiente Subgradiente (MEGS), no qual a segunda projeção sobre C é substituída por uma projeção sobre um semiespaço fechado que contém C . Essa modificação preserva as propriedades de convergência do método, ao mesmo tempo em que reduz significativamente o custo computacional por iteração.

Antes de formalizar o MEGS, recordamos alguns resultados fundamentais da teoria convexa.

Proposição 1.3.1. *Seja $\psi : H \rightarrow \mathbb{R}$ uma função convexa e semicontínua inferior. Para todo $\alpha \in \mathbb{R}$, o conjunto de subnível*

$$C_\alpha := \{x \in H : \psi(x) \leq \alpha\}$$

é convexo e fechado.

Demonstração. Primeiramente, provaremos a convexidade de C_α . Sejam $x, y \in C_\alpha$ e $t \in [0, 1]$. Sendo ψ convexa, vale:

$$\psi(tx + (1-t)y) \leq t\psi(x) + (1-t)\psi(y).$$

Como $\psi(x) \leq \alpha$ e $\psi(y) \leq \alpha$, temos:

$$t\psi(x) + (1-t)\psi(y) \leq t\alpha + (1-t)\alpha = \alpha.$$

Logo,

$$\psi(tx + (1 - t)y) \leq \alpha,$$

isto é, $tx + (1 - t)y \in C_\alpha$. Portanto, C_α é convexo.

Agora, provaremos o fechamento de C_α .

Seja $\{x_n\} \subset C_\alpha$ uma sequência fortemente convergente para algum ponto $x \in H$, isto é, $x_n \rightarrow x$. Como $x_n \in C_\alpha$, temos $\psi(x_n) \leq \alpha$, para todo $n \in \mathbb{N}$. Assim, pela s.c.i. de ψ , obtemos:

$$\psi(x) \leq \liminf_{n \rightarrow \infty} \psi(x_n) \leq \alpha.$$

Portanto, $x \in C_\alpha$, o que mostra que C_α é fechado. $\square$

Observação 1.3.1. *Em particular, o conjunto $C := \{x \in H : \psi(x) \leq 0\}$ é convexo e fechado, sempre que ψ for convexa e semicontínua inferior.*

Para a próxima proposição, relembremos os tipos de separação entre conjuntos convexos.

Definição 1.3.1. [52, Capítulo 3, p.34] *Sejam $A, B \subset H$ dois subconjuntos convexos e disjuntos de um espaço de Hilbert real H . Um hiperplano $H_{w,\alpha} := \{x \in H : \langle w, x \rangle = \alpha\}$, onde $w \in H \setminus \{0\}$ e $\alpha \in \mathbb{R}$ é dito que:*

- (i) *separa A e B , se $\langle w, a \rangle \leq \alpha \leq \langle w, b \rangle$, $\forall a \in A, \forall b \in B$;*
- (ii) *separa A e B propriamente, se o hiperplano separa A e B e pelo menos um dos conjuntos não está contido em $H_{w,\alpha}$, isto é, existe $x \in A \cup B$ tal que $\langle w, x \rangle \neq \alpha$;*
- (iii) *separa A e B estritamente, se $\langle w, a \rangle < \alpha < \langle w, b \rangle$, $\forall a \in A, \forall b \in B$.*

A construção do MEGS pode ser interpretada geometricamente como a substituição de C por um semiespaço que o contém. Quando C é descrito como um conjunto de subnível $C = \{x \in H : \psi(x) \leq 0\}$, esse semiespaço pode ser obtido via subgradiente, conforme a proposição a seguir.

Proposição 1.3.2. [3, Lema 7.3, p.410] *Seja $\psi : H \rightarrow \mathbb{R}$ uma função convexa, $z \in H$, e suponha que ψ é subdiferenciável em z . Para qualquer $w \in \partial\psi(z)$, defina o conjunto convexo fechado $T(z) := \{x \in H : \psi(z) + \langle w, x - z \rangle \leq 0\}$. Se $C = \{x \in H : \psi(x) \leq 0\}$ é não vazio, então valem as seguintes propriedades:*

- (i) $C \subset T(z)$;
- (ii) *Se $\psi(z) > 0$ e $w \neq 0$, então o hiperplano que delimita $T(z)$, $H_z := \{x \in H : \langle w, x \rangle = \langle w, z \rangle - \psi(z)\}$, separa z de C , isto é,*

$$\langle w, y \rangle \leq \langle w, z \rangle - \psi(z) < \langle w, z \rangle, \forall y \in C;$$

- (iii) *Se $w = 0$, então $T(z) = H$;*

- (iv) *Se $\psi(z) = 0$ e $w \neq 0$, então o hiperplano $H_z = \{x \in H : \langle w, x \rangle = \langle w, z \rangle\}$ atua como superfície de suporte de C no ponto z .*

Demonstração. (i) Se $x \in C$, então $\psi(x) \leq 0$. Como $w \in \partial\psi(z)$, segue da definição de subgradiente que

$$\psi(z) + \langle w, x - z \rangle \leq \psi(x) \leq 0.$$

Logo, $x \in T(z)$, o que mostra que $C \subset T(z)$.

(ii) Reescrevendo $T(z)$, obtemos

$$T(z) = \{x \in H : \langle w, x \rangle \leq \langle w, z \rangle - \psi(z)\}.$$

Do item (i), temos $C \subset T(z)$, isto é,

$$\langle w, y \rangle \leq \langle w, z \rangle - \psi(z), \quad \forall y \in C.$$

Se $\psi(z) > 0$, então $z \notin C$. Além disso,

$$\psi(z) + \langle w, z - z \rangle = \psi(z) > 0 \quad \Rightarrow \quad z \notin T(z).$$

Portanto,

$$\langle w, y \rangle \leq \langle w, z \rangle - \psi(z) < \langle w, z \rangle, \quad \forall y \in C,$$

isto é, o hiperplano

$$H_z := \{x \in H : \langle w, x \rangle = \langle w, z \rangle - \psi(z)\}$$

separa z de C .

(iii) Se $w = 0$, então

$$T(z) = \{x \in H : \psi(z) + \langle 0, x - z \rangle \leq 0\} = \{x \in H : \psi(z) \leq 0\}.$$

Suponha, por absurdo, que $\psi(z) > 0$. Como $0 \in \partial\psi(z)$, segue da definição de subgradiente que

$$\psi(x) \geq \psi(z) > 0, \quad \forall x \in H,$$

o que implica $C = \emptyset$, em contradição com a hipótese. Logo, $\psi(z) \leq 0$ e, portanto, $T(z) = H$.

(iv) Do item (i), temos $C \subset T(z)$, isto é,

$$\langle w, x \rangle \leq \langle w, z \rangle - \psi(z), \quad \forall x \in C.$$

Se $\psi(z) = 0$, então

$$\langle w, x \rangle \leq \langle w, z \rangle, \quad \forall x \in C,$$

de modo que C está contido no semiespaço fechado

$$\{x \in H : \langle w, x \rangle \leq \langle w, z \rangle\}.$$

Além disso, $\psi(z) = 0$ implica $z \in C$, e claramente $\langle w, z \rangle = \langle w, z \rangle$, logo $z \in H_z$, onde

$$H_z := \{x \in H : \langle w, x \rangle = \langle w, z \rangle\}.$$

Portanto, o hiperplano H_z contém z e suporta C em z , como desejado. $\square$

Outro resultado importante para a construção do MEGS é a seguinte proposição.

Proposição 1.3.3. [3, Fatos 7.2, p.410] *Seja $\psi : H \rightarrow \mathbb{R}$ uma função convexa e $x_0 \in H$. Então:*

(i) ψ é contínua em x_0 e $\partial\psi(x_0)$ é um conjunto unitário se, e somente se, ψ é s.c.i. e Gâteaux diferenciável em x_0 . Neste caso, o único subgradiente de ψ em x_0 coincide com a derivada Gâteaux de ψ em x_0 , isto é, $\partial\psi(x_0) = \{\nabla\psi(x_0)\}$;

(ii) Se ψ é contínua em x_0 , então ψ é subdiferenciável em x_0 ;

(iii) Se H tem dimensão finita, então ψ é contínua e subdiferenciável em toda parte.

Com essas ferramentas, formalizamos o MEGS para resolver (PDV). Dado $x_0 \in H$, considere a iteração:

$$\begin{cases} y_n = P_C(x_n - \lambda F(x_n)), \\ x_{n+1} = P_{T_n}(x_n - \lambda F(y_n)), \end{cases} \quad (1.12)$$

onde

$$T_n = \left\{ x \in H : \langle (x_n - \lambda F(x_n)) - y_n, x - y_n \rangle \leq 0 \right\}$$

é um semiespaço fechado que contém C .

Observação 1.3.2. Defina $a_n := x_n - \lambda F(x_n)$ e note que $y_n = P_C(a_n)$. Pela caracterização da projeção métrica (Definição 1.1.5),

$$\langle a_n - y_n, y - y_n \rangle \leq 0, \quad \forall y \in C,$$

o que implica $C \subset T_n$. Além disso, $a_n - y_n \in N_C(y_n)$, isto é, o vetor $(x_n - \lambda F(x_n)) - y_n$ pode ser interpretado como um vetor normal a C no ponto y_n .

Observação 1.3.3. Quando C é descrito como um conjunto de subnível $C = \{x \in H : \psi(x) \leq 0\}$, a Proposição 1.3.2 mostra que semiespaços do tipo $T(z) = \{x \in H : \psi(z) + \langle w, x - z \rangle \leq 0\}$, com $w \in \partial\psi(z)$, contêm C e podem ser usados como aproximações externas de C . Essa interpretação geométrica motiva a denominação extragradiente subgradiente, pois o semiespaço auxiliar pode ser construído a partir de informações subgradientes associadas à descrição de C .

Vejamos uma interpretação geométrica:

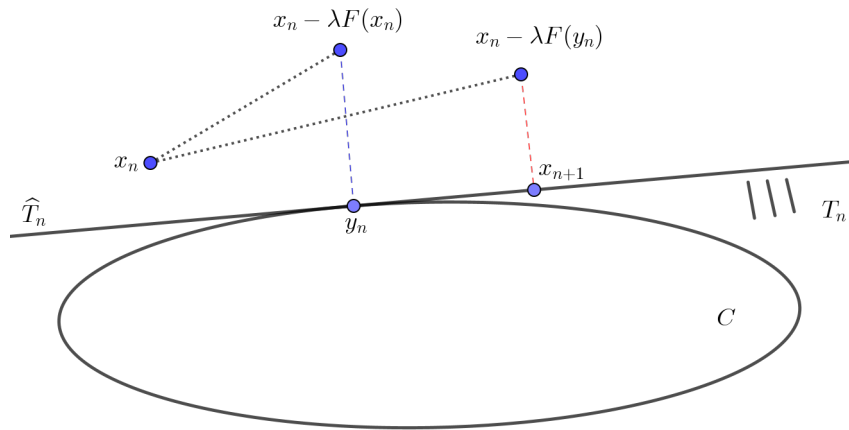


Figura 1.3: x_{n+1} é a projeção do ponto $x_n - \lambda F(y_n)$ sobre o semiespaço T_n .

Sob as mesmas hipóteses de Korpelevich, porém com F Lipschitz-contínuo em H , os autores em [11] demonstraram a convergência fraca do MEGS, o que valida o método como uma alternativa eficiente ao extragradiente clássico, sobretudo em contextos onde a projeção sobre C é custosa ou difícil de implementar.

Extensão do MEGS ao caso de Problemas de Equilíbrio

A formulação do MEGS para PE fica assim:

$$\begin{cases} y_n = \arg \min_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|y - x_n\|^2 \right\}, \\ x_{n+1} = \arg \min_{y \in T_n} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|y - x_n\|^2 \right\}, \end{cases} \quad (1.13)$$

onde

$$T_n := \{x \in H : \langle (x_n - \lambda v_n) - y_n, x - y_n \rangle \leq 0\} \quad \text{e} \quad v_n \in \partial_2 f(x_n, y_n).$$

Note que, no caso particular em que

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

com $F : C \rightarrow H$, a função $y \mapsto f(x, y)$ é afim, logo convexa e Fréchet-diferenciável em H , para todo $x \in C$. Assim, pela caracterização do subdiferencial de funções convexas diferenciáveis (ver, por exemplo, Proposição 1.3.3(i)), temos

$$\partial_2 f(x, y) = \{F(x)\}, \quad \forall x, y \in C.$$

Em particular, para os índices do algoritmo,

$$v_n \in \partial_2 f(x_n, y_n) \implies v_n = F(x_n),$$

e o semiespaço T_n assume a forma

$$T_n = \{x \in H : \langle (x_n - \lambda F(x_n)) - y_n, x - y_n \rangle \leq 0\},$$

que coincide exatamente com o semiespaço utilizado no MEGS para (PDV) em (1.12). Dessa forma, o esquema MEGS-PE recupera, como caso particular, o método extragradiente subgradiente de Censor, Gibali e Reich para problemas de desigualdade variacional.

Capítulo 2

Um Método Extragradiante Subgradiente Acelerado para PE

Neste capítulo, propomos e analisamos um método extragradiante subgradiente para a resolução de problemas de equilíbrio em espaços de Hilbert reais. Inicialmente, estabelecemos um conjunto de hipóteses que asseguram a convergência do algoritmo proposto. Em seguida, enunciamos o algoritmo e discutimos sua boa definição, destacando, em particular, as propriedades fundamentais da sequência de tamanhos de passo autoadaptativos incorporada em sua estrutura. Posteriormente, enunciamos e demonstramos o teorema principal que garante a convergência forte do método para uma solução de equilíbrio. Como caso particular, apresentamos uma versão adaptada do algoritmo voltada à resolução de problemas de desigualdades variacionais. Por fim, identificamos condições adicionais sob as quais é possível estabelecer uma taxa de convergência linear para o método.

2.1 Descrição do Algoritmo 1 e sua Boa Definição

Nesta seção, propomos uma extensão modificada do algoritmo encontrado em [34] para o contexto de Problema de Equilíbrio, o qual incorpora três estratégias distintas e complementares: um termo inercial, uma atualização explícita de tamanho de passo baseada na proporção áurea e uma etapa de aproximação por viscosidade. Por isso, chamamos esse procedimento de Algoritmo Extragradiante Subgradiente Acelerado Explícito por Viscosidade (AESAEV).

Para a análise de convergência da sequência gerada pelo método, assumimos que a bifunção $f : H \times H \rightarrow \mathbb{R}$ satisfaz as seguintes condições:

Condição A

- (A1) $f(x, x) = 0$, para todo $x \in C$;
- (A2) f é pseudomonótona em C ;
- (A3) f satisfaz a condição do tipo-Lipschitz em H , com constantes c_1 e c_2 ;
- (A4) $f(x, \cdot)$ é convexa e semicontínua inferior em H , para todo $x \in H$;
- (A5) $f(\cdot, y)$ é fracamente semicontínua superior em C , para todo $y \in C$;
- (A6) $\text{int}C \neq \emptyset$ ou $f(x, \cdot)$ é contínua em algum ponto de C , para todo $x \in H$.

Observação 2.1.1. Se f satisfaz (A1), (A2), (A4) e (A5), então o conjunto solução $S_{PE}(f, C)$ é convexo e fechado. Para uma demonstração ver, por exemplo, [41].

Também consideraremos uma segunda condição para a análise de convergência, como segue:

Condição B

Seja $\{\alpha_n\} \subset (0, 1)$ uma sequência satisfazendo

$$\lim_{n \rightarrow \infty} \alpha_n = 0 \text{ e } \sum_{n=1}^{+\infty} \alpha_n = +\infty.$$

Escolha $u_n \in [0, 1)$ tal que

$$\lim_{n \rightarrow +\infty} \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| = 0.$$

Observação 2.1.2. A Condição B é facilmente implementada em experimentos numéricos, uma vez que usaremos o efeito inercial no método proposto e o valor de $\|x_n - x_{n-1}\|$ é conhecido antes da escolha de u_n , ver [34, 46] para mais detalhes. Podemos escolher, por exemplo,

$$u_n = \begin{cases} \min \left\{ \tau, \frac{\alpha_n^2}{\|x_n - x_{n-1}\|} \right\} & \text{se } x_n \neq x_{n-1} \\ \tau, & \text{caso contrário,} \end{cases}$$

onde $\tau \in [0, 1)$.

A seguir, formalizamos o algoritmo proposto.

Algoritmo 1 : AESAEV para Problema de Equilíbrio

Passo 0 (Inicialização):

Escolha parâmetros exógenos $u_1 \in [0, 1)$, $\lambda_1 > 0$, $\varphi = \frac{1 + \sqrt{5}}{2}$, $\theta = \varphi - 1$, $\mu \in (0, \frac{2}{\varphi})$.

Escolha pontos iniciais $x_0, x_1 \in H$ e defina $n = 1$.

Passo 1:

Assuma que $x_{n-1}, x_n \in H$, $u_n \in [0, 1)$ e $\lambda_n > 0$ são conhecidos.

Calcule o vetor inercial

$$w_n = x_n + u_n(x_n - x_{n-1}),$$

e resolva o problema de minimização

$$y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(w_n, y) + \frac{1}{2} \|w_n - y\|^2 \right\}.$$

Se $y_n = w_n$, então **PARE**: y_n é solução do (PE). Caso contrário, continue.

Algoritmo 1 : AESAEV para Problema de Equilíbrio (continuação)

Passo 2:

Escolha $v_n \in \partial_2 f(w_n, y_n)$ e $q_n \in N_C(y_n)$ tais que $q_n = w_n - \lambda_n v_n - y_n$ e construa o semiespaço

$$T_n = \{x \in H : \langle w_n - \lambda_n v_n - y_n, x - y_n \rangle \leq 0\}.$$

Em seguida, calcule

$$z_n = \operatorname{argmin}_{y \in T_n} \left\{ \lambda_n f(y_n, y) + \frac{1}{2} \|w_n - y\|^2 \right\}.$$

Passo 3 (Atualização com viscosidade):

Defina o ponto auxiliar $e_n = (1 - \theta)w_n + \theta x_n$ e atualize

$$x_{n+1} = \alpha_n g(e_n) + (1 - \alpha_n) z_n,$$

onde g é uma ρ -contração.

Atualize também o tamanho de passo:

$$\lambda_{n+1} = \begin{cases} \min \left\{ \lambda_n, \frac{\mu(\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2)}{4\varphi d_n} \right\}, & \text{se } d_n > 0, \\ \lambda_n, & \text{caso contrário,} \end{cases}$$

com $d_n := f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)$.

Defina $n := n + 1$ e volte ao **Passo 1**.

Comentários sobre o Algoritmo 1

- No **Passo 0**, os números dados $\varphi = \frac{1 + \sqrt{5}}{2}$ e $\theta = \varphi - 1$ são conhecidos na literatura como proporção áurea (ou número de ouro), usados, por exemplo, em [15, 22, 34, 35, 59]. Lembramos que o número φ satisfaz a equação $\varphi^2 = \varphi + 1$, o qual será utilizada nas demonstrações dos lemas 2.1.3 e 2.2.1.
- Em comparação com algoritmos extragradiente subgradiente que não utilizam a proporção áurea (veja, por exemplo, o Algoritmo 3.2 em [51]), o parâmetro μ é expandido de $(0, 1)$ para $\left(0, \frac{2}{\varphi}\right)$. Noutras palavras, nosso algoritmo é uma extensão dessa classe. Este relaxamento também foi obtido em [34, Algoritmo 3.1], mas usando uma busca linear do tipo Armijo, baseada na proporção áurea, para determinar o tamanho de passo autoadaptativo λ_n . Ressaltamos que, por meio de experimentos numéricos, os autores observaram que o aumento do parâmetro μ melhora o desempenho desse algoritmo e quando $\mu = 1$, o algoritmo tem o melhor desempenho, independente da dimensão de R^m . No Capítulo 4, desta tese, avaliamos o desempenho numérico do parâmetro μ do nosso algoritmo.
- No **Passo 1**, acrescentamos um fator inercial, representado pelo vetor w_n , que melhora o

desempenho computacional do algoritmo.

- No **Passo 2**, a projeção no semiespaço T_n reduz o custo computacional em relação à projeção sobre C , como em métodos extragradiente clássicos.
- No **Passo 3**, a contração g garante que a sequência gerada convirja fortemente, mesmo em espaços de dimensão infinita — este é o princípio da aproximação por viscosidade.
- Caso $d_n \rightarrow 0$, evitamos problemas de instabilidade numérica — incluindo indeterminações do tipo $\frac{0}{0}$ ou explosões do tipo $\frac{a}{0}$ com $a > 0$ — adotando a convenção de que, nesses casos, o quociente é tomado como $+\infty$.

Observação 2.1.3. *Embora o método tenha sido formulado assumindo que f satisfaz a condição do tipo-Lipschitz, a regra autoadaptativa para λ_n permite sua aplicação mesmo quando as constantes c_1, c_2 são desconhecidas.*

O resultado a seguir assegura que todos os passos do algoritmo estão bem definidos.

Lema 2.1.1. *O Algoritmo 1 está bem definido.*

Demonstração. De fato:

Os dados, no **Passo 0**, são todos conhecidos.

O vetor w_n , no **Passo 1**, está bem definido, pois x_{n-1} e x_n são conhecidos. Além disso, y_n existe e é único, pelo Lema 1.1.2 (é solução de um problema de otimização fortemente convexo).

No **Passo 2**, a existência de $v_n \in \partial_2 f(w_n, y_n)$ e $q_n \in N_C(y_n)$ é garantida pelo Lema 1.1.2, pois, no passo anterior,

$$y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(w_n, y) + \frac{1}{2} \|w_n - y\|^2 \right\}.$$

Assim, podemos construir o semiespaço T_n , e novamente pelo Lema 1.1.2, garantir a existência e unicidade de z_n .

Os vetores e_n e x_{n+1} , na última etapa, bem como a atualização do tamanho de passo λ_{n+1} , estão bem definidos, pois dependem de informações já obtidas nos passos anteriores e de parâmetros exógenos ao algoritmo. Portanto, o Algoritmo 1 está bem definido. $\square$

Agora, validamos o critério de parada no **Passo 1** do Algoritmo 1. O lema a seguir, mostra que se o algoritmo termina após um número finito de iterações, então y_n é uma solução de equilíbrio.

Lema 2.1.2. *Sob as condições (A1), (A4) e (A6), se $y_n = w_n$ então $y_n \in S_{PE}(f, C)$.*

Demonstração. Se $y_n = w_n$, então

$$y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(y_n, y) + \frac{1}{2} \|y_n - y\|^2 \right\}.$$

Assim, do Lema 1.1.2, temos

$$0 \in \partial_2 \left[\lambda_n f(y_n, y) + \frac{1}{2} \|y_n - y\|^2 \right] (y_n) + N_C(y_n),$$

isto é, existe $v_n \in \partial_2 f(y_n, y_n)$ tal que

$$0 \in \lambda_n v_n + N_C(y_n),$$

donde

$$-\lambda_n v_n \in N_C(y_n).$$

Pela definição de cone normal, segue que

$$\langle -\lambda_n v_n, y - y_n \rangle \leq 0, \forall y \in C.$$

Como $\lambda_n > 0, \forall n$, multiplicando a desigualdade anterior por $-\frac{1}{\lambda_n}$, chegamos a

$$\langle v_n, y - y_n \rangle \geq 0, \forall y \in C. \quad (2.1)$$

Por outro lado, como $v_n \in \partial_2 f(y_n, y_n)$ e $f(y_n, \cdot)$ é convexa, por (A4), então

$$f(y_n, y) \geq f(y_n, y_n) + \langle v_n, y - y_n \rangle = \langle v_n, y - y_n \rangle \geq 0, \forall y \in C,$$

onde a igualdade segue de (A1) e a última desigualdade é devido a (2.1).

Portanto, $y_n \in S_{PE}(f, C)$. □

A seguinte observação será crucial nas demonstrações de alguns lemas da próxima seção, bem como os fatos, logo em seguida, referentes a sequência $\{\lambda_n\}$.

Observação 2.1.4. *Pelo Passo 2, temos que $C \subset T_n$.*

De fato, sendo

$$y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(w_n, y) + \frac{1}{2} \|w_n - y\|^2 \right\},$$

então, pelo Lema 1.1.2,

$$0 \in \partial_2(\lambda_n f(w_n, y) + \frac{1}{2} \|w_n - y\|^2)(y_n) + N_C(y_n).$$

Assim, existem $v_n \in \partial_2 f(w_n, y_n)$ e $q_n \in N_C(y_n)$ tal que $0 = \lambda_n v_n + y_n - w_n + q_n$, isto é,

$$q_n = w_n - \lambda_n v_n - y_n.$$

Agora, seja $y \in C$. Segue de $q_n \in N_C(y_n)$ que

$$\langle w_n - \lambda_n v_n - y_n, y - y_n \rangle = \langle q_n, y - y_n \rangle \leq 0.$$

Isso implica que $y \in T_n$, donde $C \subset T_n$.

Lema 2.1.3. *A sequência $\{\lambda_n\}$ gerada pelo Algoritmo 1 é uma sequência monótona não-crescente e limitada inferiormente por $\min \left\{ \lambda_1, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\}$.*

Demonstração. Seja $n \in \mathbb{N}$ fixo. Como f satisfaz a condição do tipo-Lipschitz, segue que, $\forall n$,

$$\begin{aligned} f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n) &\leq c_1 \|w_n - y_n\|^2 + c_2 \|y_n - z_n\|^2 \\ &\leq \max\{c_1, c_2\} (\|w_n - y_n\|^2 + \|y_n - z_n\|^2) \\ &\leq \max\{c_1, c_2\} (\|w_n - y_n\|^2 + (1 + \varphi) \|y_n - z_n\|^2). \end{aligned}$$

Logo, quando $f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n) > 0$, obtemos:

$$\begin{aligned} \frac{\mu(\|w_n - y_n\|^2 + (1 + \varphi) \|y_n - z_n\|^2)}{4\varphi(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n))} &\geq \frac{\mu(\|w_n - y_n\|^2 + (1 + \varphi) \|y_n - z_n\|^2)}{4\varphi \max\{c_1, c_2\} (\|w_n - y_n\|^2 + (1 + \varphi) \|y_n - z_n\|^2)} \\ &= \frac{\mu}{4\varphi \max\{c_1, c_2\}}. \end{aligned}$$

Segue da desigualdade acima e da definição de λ_{n+1} que

$$\begin{aligned} \min \left\{ \lambda_n, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} &\leq \min \left\{ \lambda_n, \frac{\mu(\|w_n - y_n\|^2 + (1 + \varphi) \|y_n - z_n\|^2)}{4\varphi(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n))} \right\} \\ &= \lambda_{n+1} \leq \lambda_n \end{aligned} \quad (2.2)$$

Por outro lado, quando $f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n) \leq 0$, por convenção, $\lambda_{n+1} = \lambda_n$. Disto e de (2.2), concluímos que $\{\lambda_n\}$ é monótona não-crescente.

Agora, segue de (2.2) que, $\forall n \geq 1$,

$$\begin{aligned} \min \left\{ \lambda_1, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} &\leq \lambda_2 \\ \min \left\{ \lambda_2, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} &\leq \lambda_3 \\ &\vdots \\ \min \left\{ \lambda_n, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} &\leq \lambda_{n+1}, \end{aligned}$$

donde obtemos,

$$\begin{aligned} \min \left\{ \lambda_1, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} &\leq \min \left\{ \lambda_2, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} \\ &\leq \min \left\{ \lambda_3, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} \\ &\vdots \\ &\leq \min \left\{ \lambda_n, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\} \\ &\leq \lambda_{n+1} \leq \lambda_n, \forall n \geq 1. \end{aligned}$$

Noutras palavras, a sequência $\{\lambda_n\}$ é limitada inferiormente por $\min \left\{ \lambda_1, \frac{\mu}{4\varphi \max\{c_1, c_2\}} \right\}$. $\square$

Observação 2.1.5. Como $\{\lambda_n\}$ é uma sequência de números reais positivos monótona e limitada inferiormente, então $\lambda_n \rightarrow \lambda > 0$.

Os resultados obtidos garantem a boa definição do algoritmo proposto e asseguram propriedades essenciais para sua análise. Com isso, estamos aptos a iniciar a investigação da convergência da sequência gerada, tema da próxima seção.

2.2 Análise de Convergência

Nesta seção, desenvolvemos a análise de convergência do algoritmo proposto. A abordagem é dividida em duas partes: primeiramente, apresentamos uma sequência de lemas auxiliares que estabelecem propriedades fundamentais das iteradas geradas; em seguida, utilizamos esses resultados para demonstrar a convergência forte da sequência $\{x_n\}$ para uma solução do problema de equilíbrio.

2.2.1 Resultados Auxiliares

O primeiro lema fornece uma estimativa entre a distância de z_n até uma solução x^* e a distância de w_n até x^* . Essa desigualdade será usada para controlar o comportamento da sequência e medir o avanço do algoritmo em direção à solução.

Lema 2.2.1. Suponha que a bifunção f satisfaz a Condição A e $\{z_n\}$ é uma sequência gerada pelo Algoritmo 1. Então, para todo $x^* \in S_{PE}(f, C)$, a seguinte desigualdade é válida:

$$\|z_n - x^*\|^2 \leq \|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2, \forall n \geq 1. \quad (2.3)$$

Demonstração. De $z_n = \operatorname{argmin}_{y \in T_n} \left\{ \lambda_n f(y_n, y) + \frac{1}{2} \|w_n - y\|^2 \right\}$ e Lema 1.1.2, segue que

$$0 \in \partial_2 \left[\lambda_n f(y_n, y) + \frac{1}{2} \|w_n - y\|^2 \right] (z_n) + N_{T_n}(z_n).$$

Daí, existe $\eta_n \in \partial_2 f(y_n, z_n)$ tal que $0 \in \lambda_n \eta_n + z_n - w_n + N_{T_n}(z_n)$, donde

$$-(\lambda_n \eta_n + z_n - w_n) \in N_{T_n}(z_n),$$

e pela definição de cone normal, $\langle \lambda_n \eta_n + z_n - w_n, y - z_n \rangle \geq 0, \forall y \in T_n$. Isso implica que

$$\lambda_n \langle \eta_n, y - z_n \rangle \geq \langle w_n - z_n, y - z_n \rangle, \forall y \in T_n. \quad (2.4)$$

Como $f(y_n, \cdot)$ é convexa e $\eta_n \in \partial_2 f(y_n, z_n)$, pela definição de subgradiente,

$$f(y_n, y) \geq f(y_n, z_n) + \langle \eta_n, y - z_n \rangle, \forall y \in H. \quad (2.5)$$

Consequentemente, de (2.4) e (2.5), obtemos

$$\lambda_n (f(y_n, y) - f(y_n, z_n)) \geq \lambda_n \langle \eta_n, y - z_n \rangle \geq \langle w_n - z_n, y - z_n \rangle, \forall y \in T_n. \quad (2.6)$$

Uma vez que $x^* \in S_{PE}(f, C) \subset C \subset T_n$, podemos tomar $y = x^*$ e obter de (2.6) que

$$\lambda_n (f(y_n, x^*) - f(y_n, z_n)) \geq \langle w_n - z_n, x^* - z_n \rangle. \quad (2.7)$$

Além disso, de $x^*, y_n \in C$, temos $f(x^*, y_n) \geq 0$, e pela pseudomonotonicidade de f , $f(y_n, x^*) \leq 0$, o que de (2.7) implica que

$$-\lambda_n f(y_n, z_n) \geq \langle w_n - z_n, x^* - z_n \rangle. \quad (2.8)$$

Por outro lado, como $v_n \in \partial_2 f(w_n, y_n)$, segue da definição de subgradiente que

$$f(w_n, y) - f(w_n, y_n) \geq \langle v_n, y - y_n \rangle, \forall y \in H.$$

Em particular, a desigualdade acima vale para $y = z_n$, isto é,

$$f(w_n, z_n) - f(w_n, y_n) \geq \langle v_n, z_n - y_n \rangle,$$

donde

$$\lambda_n (f(w_n, z_n) - f(w_n, y_n)) \geq \lambda_n \langle v_n, z_n - y_n \rangle. \quad (2.9)$$

Sendo $z_n \in T_n$, então, pela definição deste semiespaço,

$$\langle w_n - \lambda_n v_n - y_n, z_n - y_n \rangle \leq 0$$

ou

$$\lambda_n \langle v_n, z_n - y_n \rangle \geq \langle w_n - y_n, z_n - y_n \rangle. \quad (2.10)$$

Combinando (2.9) e (2.10), teremos

$$\lambda_n (f(w_n, z_n) - f(w_n, y_n)) \geq \langle w_n - y_n, z_n - y_n \rangle. \quad (2.11)$$

Somando (2.8) e (2.11), chegamos em

$$\lambda_n (f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \geq \langle w_n - y_n, z_n - y_n \rangle + \langle w_n - z_n, x^* - z_n \rangle.$$

Daí,

$$\begin{aligned} 2\lambda_n (f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) &\geq 2 \langle w_n - y_n, z_n - y_n \rangle + 2 \langle w_n - z_n, x^* - z_n \rangle \\ &= \|w_n - y_n\|^2 + \|z_n - y_n\|^2 - \|z_n - w_n\|^2 + \|w_n - z_n\|^2 \\ &\quad + \|z_n - x^*\|^2 - \|w_n - x^*\|^2 \\ &= \|w_n - y_n\|^2 + \|z_n - y_n\|^2 + \|z_n - x^*\|^2 - \|w_n - x^*\|^2, \end{aligned}$$

onde na primeira igualdade usamos o item c) do Lema 1.1.4.

Da desigualdade anterior,

$$\|z_n - x^*\|^2 \leq \|w_n - x^*\|^2 - \|w_n - y_n\|^2 - \|z_n - y_n\|^2 + 2\lambda_n (f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)). \quad (2.12)$$

Defina

$$d_n := f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n).$$

Pela regra de atualização do algoritmo, quando $d_n > 0$ tem-se

$$\lambda_{n+1} \leq \frac{\mu (\|w_n - y_n\|^2 + (1 + \varphi) \|z_n - y_n\|^2)}{4\varphi d_n}.$$

Se, por outro lado, $d_n \leq 0$, então $\lambda_{n+1} = \lambda_n$ e, além disso, $2d_n \leq 0$, de modo que a desigualdade abaixo permanece válida (pois o lado direito é não negativo):

$$2d_n \leq \frac{\mu}{2\varphi \lambda_{n+1}} \left(\|w_n - y_n\|^2 + (1 + \varphi) \|z_n - y_n\|^2 \right). \quad (2.13)$$

Por fim, substituindo (2.13) em (2.12), obtemos:

$$\begin{aligned}
\|z_n - x^*\|^2 &\leq \|w_n - x^*\|^2 - \|w_n - y_n\|^2 - \|z_n - y_n\|^2 + \\
&\quad \frac{\mu\lambda_n}{2\varphi\lambda_{n+1}} (\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2) \\
&= \|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\
&\quad - \left(1 - \frac{\mu(1 + \varphi)}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 \\
&= \|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\
&\quad - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2, \quad \forall n \geq 1,
\end{aligned}$$

onde, na última igualdade, usamos o fato que $1 + \varphi = \varphi^2$. $\square$

O próximo lema estabelece uma condição suficiente para que um ponto de acumulação fraco da sequência $\{w_n\}$ seja solução do problema de equilíbrio. Veremos, na demonstração do teorema principal, que as hipóteses deste lema surgem naturalmente no processo de análise.

Lema 2.2.2. *Suponha que a Condição A seja válida e que $\{w_n\}$ é uma sequência gerada pelo Algoritmo 1. Se existe uma subsequência $\{w_{n_k}\}$ convergindo fracamente para $z \in H$ e $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$, $\lim_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\| = 0$, $\lim_{k \rightarrow \infty} \|z_{n_k} - y_{n_k}\| = 0$, então $z \in S_{PE}(f, C)$.*

Demonstração. De $w_{n_k} \rightharpoonup z$ e $\|w_{n_k} - y_{n_k}\| \rightarrow 0$, obtemos $y_{n_k} \rightharpoonup z$. Como $\{y_n\} \subset C$ e C é convexo e (fortemente) fechado, então C é fracamente fechado; logo $z \in C$.

Mostremos agora que $z \in S_{PE}(f, C)$.

De fato, sendo $z_n = \operatorname{argmin}_{y \in T_n} \{\lambda_n f(y_n, y) + \frac{1}{2} \|w_n - y\|^2\}$, pela estimativa (2.13) obtemos

$$2f(y_n, z_n) \geq 2(f(w_n, z_n) - f(w_n, y_n)) - \frac{\mu}{2\varphi\lambda_{n+1}} (\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2). \quad (2.14)$$

Mas, por (2.11),

$$(f(w_n, z_n) - f(w_n, y_n)) \geq \frac{1}{\lambda_n} \langle w_n - y_n, z_n - y_n \rangle. \quad (2.15)$$

Assim, combinando (2.14) e (2.15), obtemos

$$2f(y_n, z_n) \geq \frac{2}{\lambda_n} \langle w_n - y_n, z_n - y_n \rangle - \frac{\mu}{2\varphi\lambda_{n+1}} (\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2). \quad (2.16)$$

Por outro lado, como $C \subset T_n$, segue de (2.6) que

$$f(y_n, y) \geq f(y_n, z_n) + \frac{1}{\lambda_n} \langle w_n - z_n, y - z_n \rangle, \quad \forall y \in C. \quad (2.17)$$

Substituindo (2.16) em (2.17), ganhamos:

$$f(y_n, y) \geq \frac{1}{\lambda_n} \langle w_n - y_n, z_n - y_n \rangle - \frac{\mu}{4\varphi\lambda_{n+1}} (\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2)$$

$$+\frac{1}{\lambda_n} \langle w_n - z_n, y - z_n \rangle, \forall y \in C. \quad (2.18)$$

Lembramos que $y_{n_k} \rightarrow z \in C$, $\lambda_n \rightarrow \lambda > 0$, $f(\cdot, y)$ é fracamente semicontínua superior,

$$\|w_{n_k} - y_{n_k}\| \rightarrow 0, \quad \|w_{n_k} - z_{n_k}\| \rightarrow 0, \quad \|z_{n_k} - y_{n_k}\| \rightarrow 0.$$

Portanto, fazendo $n = n_k$ em (2.18) e tomando o limite superior quando $k \rightarrow \infty$, obtemos

$$f(z, y) \geq \limsup_{k \rightarrow \infty} f(y_{n_k}, y) \geq 0, \quad \forall y \in C,$$

isto é, $z \in S_{PE}(f, C)$. □

No lema a seguir, obtemos a limitação de todas as sequências envolvidas no algoritmo.

Lema 2.2.3. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 1 sob as condições A e B. Então, para qualquer $x^* \in S_{PE}(f, C)$, as sequências $\{x_n\}$, $\{w_n\}$, $\{z_n\}$, $\{e_n\}$ e $\{g(e_n)\}$ são limitadas.*

Demonstração. Seja $x^* \in S_{PE}(f, C)$, pelo Lema 2.2.1, vimos que

$$\|z_n - x^*\|^2 \leq \|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2. \quad (2.19)$$

Além disso,

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) = 1 - \frac{\mu}{2\varphi} \text{ e } \lim_{n \rightarrow \infty} \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) = 1 - \frac{\mu\varphi}{2}.$$

Por hipótese, $0 < \mu < \frac{2}{\varphi}$, que multiplicando por $-\frac{1}{2\varphi}$, resulta em

$$0 > -\frac{\mu}{2\varphi} > -\frac{1}{\varphi^2}. \quad (2.20)$$

Somando 1 em cada membro de (2.20) e usando a identidade $1 + \varphi = \varphi^2$, chegamos em

$$1 - \frac{\mu}{2\varphi} > 1 - \frac{1}{\varphi^2} = \frac{1}{\varphi}. \quad (2.21)$$

Agora, se multiplicarmos $0 < \mu < \frac{2}{\varphi}$ por $-\frac{\varphi}{2}$ e somarmos 1 em cada membro da desigualdade resultante, obteremos

$$1 - \frac{\mu\varphi}{2} > 0. \quad (2.22)$$

Logo, de (2.21) e (2.22),

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) > \frac{1}{\varphi}$$

e

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) > 0.$$

Da permanência do sinal do limite, existe $n_1 \in \mathbb{N}$ tal que

$$1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} > \frac{1}{\varphi} \text{ e } 1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} > 0, \forall n \geq n_1.$$

Assim, de (2.19),

$$\|z_n - x^*\| \leq \|w_n - x^*\|, \forall n \geq n_1. \quad (2.23)$$

Agora, representando por *I*, *II*, *III*, *IV*, a definição de $\{x_n\}$, a desigualdade triangular, a ρ -contração g e a definição de $\{e_n\}$, respectivamente, obtemos

$$\begin{aligned} \|x_{n+1} - x^*\| &\stackrel{I}{=} \|\alpha_n g(e_n) + (1 - \alpha_n)z_n - x^*\| \\ &= \|\alpha_n g(e_n) + (1 - \alpha_n)z_n - [\alpha_n x^* + (1 - \alpha_n)x^*]\| \\ &= \|\alpha_n(g(e_n) - x^*) + (1 - \alpha_n)(z_n - x^*)\| \\ &\stackrel{II}{\leq} \alpha_n \|g(e_n) - x^*\| + (1 - \alpha_n) \|z_n - x^*\| \\ &= \alpha_n \|g(e_n) - g(x^*) + g(x^*) - x^*\| + (1 - \alpha_n) \|z_n - x^*\| \\ &\stackrel{II}{\leq} \alpha_n (\|g(e_n) - g(x^*)\| + \|g(x^*) - x^*\|) + (1 - \alpha_n) \|z_n - x^*\| \\ &\stackrel{III}{\leq} \rho \alpha_n \|e_n - x^*\| + \alpha_n \|g(x^*) - x^*\| + (1 - \alpha_n) \|z_n - x^*\| \\ &\stackrel{IV}{=} \rho \alpha_n \|(1 - \theta)(w_n - x^*) + \theta(x_n - x^*)\| + \alpha_n \|g(x^*) - x^*\| \\ &\quad + (1 - \alpha_n) \|z_n - x^*\| \\ &\stackrel{II}{\leq} \rho(1 - \theta) \alpha_n \|w_n - x^*\| + \rho \theta \alpha_n \|x_n - x^*\| + \alpha_n \|g(x^*) - x^*\| \\ &\quad + (1 - \alpha_n) \|z_n - x^*\|. \end{aligned} \quad (2.24)$$

Da definição de $\{w_n\}$ e desigualdade triangular, segue que

$$\begin{aligned} \|w_n - x^*\| &= \|x_n + u_n(x_n - x_{n-1}) - x^*\| \\ &\leq \|x_n - x^*\| + u_n \|x_n - x_{n-1}\| \\ &= \|x_n - x^*\| + \alpha_n \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \end{aligned} \quad (2.25)$$

Como $\lim_{n \rightarrow \infty} \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| = 0$, existe uma constante $H_1 > 0$ tal que

$$\frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \leq H_1, \forall n \geq 1. \quad (2.26)$$

Assim, de (2.23), (2.25) e (2.26)

$$\|z_n - x^*\| \leq \|w_n - x^*\| \leq \|x_n - x^*\| + \alpha_n H_1, \forall n \geq n_1. \quad (2.27)$$

Substituindo (2.27) em (2.24), produzimos, $\forall n \geq n_1$,

$$\begin{aligned} \|x_{n+1} - x^*\| &\stackrel{(2.27)}{\leq} \rho(1 - \theta) \alpha_n (\|x_n - x^*\| + \alpha_n H_1) + \rho \theta \alpha_n \|x_n - x^*\| + \\ &\quad \alpha_n \|g(x^*) - x^*\| + (1 - \alpha_n) (\|x_n - x^*\| + \alpha_n H_1) \\ &= [\rho(1 - \theta) \alpha_n + \rho \theta \alpha_n + 1 - \alpha_n] \|x_n - x^*\| \\ &\quad + \rho(1 - \theta) \alpha_n^2 H_1 + (1 - \alpha_n) \alpha_n H_1 + \alpha_n \|g(x^*) - x^*\| \\ &= [1 - \alpha_n(1 - \rho)] \|x_n - x^*\| + [\rho(1 - \theta) - 1] \alpha_n^2 H_1 \\ &\quad + \alpha_n H_1 + \alpha_n \|g(x^*) - x^*\| \\ &\triangleq [1 - \alpha_n(1 - \rho)] \|x_n - x^*\| + \alpha_n(1 - \rho) \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \\ &\leq \max \left\{ \|x_n - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}, \end{aligned} \quad (2.28)$$

onde $\Delta := \rho(1 - \theta) - 1 < 0$, pois $\theta \cong 0,62$ e $0 \leq \rho < 1$.

Por indução matemática (em $p \in \mathbb{N}$ abaixo),

$$\|x_{n+1} - x^*\| \leq \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}, \forall n \geq n_1.$$

De fato, para $n = n_1$,

$$\|x_{n_1+1} - x^*\| \stackrel{(2.28)}{\leq} \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}. \quad (2.29)$$

Para $n = n_1 + 1$,

$$\begin{aligned} \|x_{n_1+2} - x^*\| &\stackrel{(2.28)}{\leq} \max \left\{ \|x_{n_1+1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\} \\ &\stackrel{(2.29)}{\leq} \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}. \end{aligned}$$

Suponhamos que a desigualdade seja válida para $n = n_1 + p$, com $p \geq 0$ natural, isto é,

$$\|x_{n_1+p} - x^*\| \leq \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}. \quad (2.30)$$

Assim, para $n = n_1 + (p + 1)$,

$$\begin{aligned} \|x_{n_1+(p+1)} - x^*\| &= \|x_{(n_1+p)+1} - x^*\| \\ &\stackrel{(2.28)}{\leq} \max \left\{ \|x_{n_1+p} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\} \\ &\stackrel{(2.30)}{\leq} \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}. \end{aligned}$$

Ou, de forma equivalente,

$$\|x_{n+1} - x^*\| \leq \max \left\{ \|x_{n_1} - x^*\|, \frac{\|g(x^*) - x^*\| + H_1}{1 - \rho} \right\}, \forall n \geq n_1, \quad (2.31)$$

onde $n = n_1 + p$.

Segue de (2.31) que $\{x_n\}$ é limitada.

Além disso, por (2.27), $\{w_n\}$ e $\{z_n\}$ são limitadas. Sendo $e_n = (1 - \theta)w_n + \theta x_n$, ganhamos a limitação de $\{e_n\}$, e como consequência disso, $\{g(e_n)\}$ também é limitada, visto que $g : H \rightarrow H$ é contínua, por ser uma aplicação contração. $\square$

O próximo resultado fornece uma estimativa da diferença entre $\|x_n - x^*\|^2$ e $\|x_{n+1} - x^*\|^2$, ajudando a mostrar que a sequência se aproxima cada vez mais da solução.

Lema 2.2.4. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 1 sob as condições A e B. Então, para todo $x^* \in S_{PE}(f, C)$, vale a seguinte desigualdade:*

$$\begin{aligned} (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|w_n - y_n\|^2 + (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \\ \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\alpha_n H_3, \end{aligned}$$

onde

$$H_3 := \sup_{n \in \mathbb{N}} \{H_2 + \|g(e_n) - x^*\| \|g(x^*) - x^*\|\} \text{ e } H_2 := \sup_{n \in \mathbb{N}} \left\{ \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \right\}.$$

Demonstração. De fato, por *I*, *III*, *IV* (apresentados no lema anterior), desigualdade de Cauchy-Schwarz (DCS) e Lema 1.1.4, temos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\stackrel{I}{=} \|\alpha_n g(e_n) + (1 - \alpha_n)z_n - x^*\|^2 \\
&= \|\alpha_n(g(e_n) - x^*) + (1 - \alpha_n)(z_n - x^*)\|^2 \\
&\stackrel{L.1.1.4(a)}{=} \alpha_n \|g(e_n) - x^*\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 - \alpha_n(1 - \alpha_n) \|g(e_n) - z_n\|^2 \\
&\leq \alpha_n \|g(e_n) - g(x^*) + g(x^*) - x^*\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 \\
&\stackrel{L.1.1.4(b)}{\leq} \alpha_n (\|g(e_n) - g(x^*)\|^2 + 2 \langle g(e_n) - x^*, g(x^*) - x^* \rangle) \\
&\quad + (1 - \alpha_n) \|z_n - x^*\|^2 \\
&\stackrel{III;DCS}{\leq} \alpha_n \rho^2 \|e_n - x^*\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\
&\stackrel{IV}{=} \alpha_n \rho^2 \|(1 - \theta)w_n + \theta x_n - x^*\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\
&= \alpha_n \rho^2 \|(1 - \theta)(w_n - x^*) + \theta(x_n - x^*)\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\
&\stackrel{L.1.1.4(a)}{\leq} \alpha_n \rho^2 [(1 - \theta) \|w_n - x^*\|^2 + \theta \|x_n - x^*\|^2] + (1 - \alpha_n) \|z_n - x^*\|^2 \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \tag{2.32}
\end{aligned}$$

Segue de (2.32) e Lema 2.2.1 que

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\stackrel{L.2.2.1}{\leq} \alpha_n \rho^2 \left[(1 - \theta) \|w_n - x^*\|^2 + \theta \|x_n - x^*\|^2 \right] \\
&\quad + (1 - \alpha_n) \left[\|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 \right] \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\
&= [\alpha_n \rho^2 (1 - \theta) + 1 - \alpha_n] \|w_n - x^*\|^2 + \theta \alpha_n \rho^2 \|x_n - x^*\|^2 \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\
&\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2.
\end{aligned}$$

Usando o fato que $\alpha_n \rho^2 (1 - \theta) + 1 - \alpha_n \leq 1 - \theta \alpha_n \rho^2$, teremos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq (1 - \theta \alpha_n \rho^2) \|w_n - x^*\|^2 + \theta \alpha_n \rho^2 \|x_n - x^*\|^2 \\
&\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\
&\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2. \tag{2.33}
\end{aligned}$$

Agora, da definição de $\{w_n\}$, Lema 1.1.4 (b) e desigualdade de Cauchy-Schwarz, segue que:

$$\begin{aligned}
\|w_n - x^*\|^2 &= \|x_n + u_n(x_n - x_{n-1}) - x^*\|^2 \\
&\leq \|x_n - x^*\|^2 + 2 \langle u_n(x_n - x_{n-1}), w_n - x^* \rangle \\
&\leq \|x_n - x^*\|^2 + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|. \tag{2.34}
\end{aligned}$$

Observemos que

$$\lim_{n \rightarrow \infty} \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| = 0,$$

pela Condição B e pela limitação de $\{\|w_n - x^*\|\}$. Assim, a sequência

$$\left\{ \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \right\}$$

é limitada.

Como $H_2 := \sup_{n \in \mathbb{N}} \left\{ \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \right\}$, segue de (2.34) que

$$\begin{aligned} \|w_n - x^*\|^2 &\leq \|x_n - x^*\|^2 + 2\alpha_n \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \\ &\leq \|x_n - x^*\|^2 + 2\alpha_n H_2. \end{aligned}$$

Substituindo a desigualdade anterior em (2.33), obtemos

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\leq (1 - \theta\alpha_n\rho^2)(\|x_n - x^*\|^2 + 2\alpha_n H_2) + \theta\alpha_n\rho^2\|x_n - x^*\|^2 \\ &\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2. \end{aligned}$$

Do primeiro termo da direita, temos

$$\begin{aligned} (1 - \theta\alpha_n\rho^2)(\|x_n - x^*\|^2 + 2\alpha_n H_2) + \theta\alpha_n\rho^2\|x_n - x^*\|^2 \\ = \|x_n - x^*\|^2 + 2\alpha_n H_2 - 2\theta\alpha_n^2\rho^2 H_2. \end{aligned}$$

Logo,

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\leq \|x_n - x^*\|^2 + 2\alpha_n H_2 - 2\theta\alpha_n^2\rho^2 H_2 + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2. \end{aligned}$$

Como o termo $-2\theta\alpha_n^2\rho^2 H_2 \leq 0$, podemos descartá-lo para obter

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\leq \|x_n - x^*\|^2 - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 \\ &\quad + 2\alpha_n \left(H_2 + \|g(e_n) - x^*\| \|g(x^*) - x^*\|\right). \end{aligned}$$

Sendo

$$H_3 := \sup_{n \in \mathbb{N}} \{H_2 + \|g(e_n) - x^*\| \|g(x^*) - x^*\|\},$$

obtemos finalmente

$$\begin{aligned} (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|w_n - y_n\|^2 + (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \\ \leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\alpha_n H_3, \end{aligned}$$

como queríamos. $\square$

No último lema, obtemos uma relação recorrente que descreve como a distância até a solução evolui em cada passo. Essa estimativa será usada para provar que a sequência converge fortemente para um ponto de equilíbrio, via Lema 1.1.3.

Lema 2.2.5. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 1 sob as condições A e B. Então, para todo $x^* \in S_{PE}(f, C)$, vale a seguinte estimativa:*

$$\|x_{n+1} - x^*\|^2 \leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 + \alpha_n(1 - \rho)\xi_n,$$

onde

$$\xi_n = \frac{2 \langle x_{n+1} - x^*, g(x^*) - x^* \rangle}{1 - \rho} + \frac{2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|}{\alpha_n(1 - \rho)}.$$

Demonstração. De fato, de I, III, IV, (2.27) e Lema 1.1.4, temos:

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\stackrel{I}{=} \|\alpha_n g(e_n) + (1 - \alpha_n)z_n - x^*\|^2 \\ &= \|\alpha_n(g(e_n) - g(x^*)) + (1 - \alpha_n)(z_n - x^*) + \alpha_n(g(x^*) - x^*)\|^2 \\ &\stackrel{L.1.1.4(b)}{\leq} \|\alpha_n(g(e_n) - g(x^*)) + (1 - \alpha_n)(z_n - x^*)\|^2 \\ &\quad + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle \\ &\stackrel{L.1.1.4(a)}{\leq} \alpha_n \|g(e_n) - g(x^*)\|^2 + (1 - \alpha_n) \|z_n - x^*\|^2 \\ &\quad + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle \\ &\stackrel{III; (2.23)}{\leq} \alpha_n \rho^2 \|e_n - x^*\|^2 + (1 - \alpha_n) \|w_n - x^*\|^2 + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle \\ &\stackrel{IV}{=} \alpha_n \rho^2 \|(1 - \theta)w_n + \theta x_n - x^*\|^2 + (1 - \alpha_n) \|w_n - x^*\|^2 \\ &\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle \\ &= \alpha_n \rho^2 \|(1 - \theta)(w_n - x^*) + \theta(x_n - x^*)\|^2 + (1 - \alpha_n) \|w_n - x^*\|^2 \\ &\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle \\ &\stackrel{L.1.1.4(a)}{\leq} \alpha_n \rho^2 [(1 - \theta) \|w_n - x^*\|^2 + \theta \|x_n - x^*\|^2] + (1 - \alpha_n) \|w_n - x^*\|^2 \\ &\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle. \end{aligned}$$

Por (2.34),

$$\|w_n - x^*\|^2 \leq \|x_n - x^*\|^2 + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|,$$

de modo que, substituindo esta estimativa na desigualdade anterior, obtemos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq \alpha_n \rho^2 \left[(1 - \theta) (\|x_n - x^*\|^2 + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|) + \theta \|x_n - x^*\|^2 \right] \\
&\quad + (1 - \alpha_n) \left(\|x_n - x^*\|^2 + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\| \right) \\
&\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle \\
&\leq (\alpha_n \rho^2 + 1 - \alpha_n) \|x_n - x^*\|^2 \\
&\quad + \alpha_n \rho^2 (1 - \theta) 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\| \\
&\quad + (1 - \alpha_n) 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\| \\
&\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle.
\end{aligned}$$

Como $0 \leq \rho < 1$, temos $\rho^2 \leq \rho$, e portanto

$$\alpha_n \rho^2 + 1 - \alpha_n \leq \alpha_n \rho + 1 - \alpha_n = 1 - \alpha_n(1 - \rho).$$

Usando também $\rho^2 \leq \rho$ nos termos com u_n , obtemos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 \\
&\quad + (\alpha_n \rho(1 - \theta) + (1 - \alpha_n)) 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\| \\
&\quad + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle.
\end{aligned}$$

Além disso,

$$\alpha_n \rho(1 - \theta) + 1 - \alpha_n = 1 - \alpha_n(1 - \rho(1 - \theta)) \leq 1,$$

pois $\rho(1 - \theta) \leq 1$ e $\alpha_n > 0$. Assim,

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 \\
&\quad + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\| + 2\alpha_n \langle x_{n+1} - x^*, g(x^*) - x^* \rangle.
\end{aligned}$$

Por fim, escrevemos o lado direito como

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 \\
&\quad + \alpha_n(1 - \rho) \left[\frac{2\langle x_{n+1} - x^*, g(x^*) - x^* \rangle}{1 - \rho} + \frac{2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|}{\alpha_n(1 - \rho)} \right] \\
&= (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 + \alpha_n(1 - \rho) \xi_n,
\end{aligned}$$

onde

$$\xi_n := \frac{2\langle x_{n+1} - x^*, g(x^*) - x^* \rangle}{1 - \rho} + \frac{2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|}{\alpha_n(1 - \rho)}.$$

Isto conclui a demonstração. $\square$

2.2.2 Resultado Principal

Com os lemas anteriores, temos as estimativas necessárias para demonstrar que a sequência gerada pelo algoritmo converge fortemente para uma solução do problema de equilíbrio.

Teorema 2.2.1. *Seja C um subconjunto não vazio, convexo e fechado de um espaço de Hilbert H , e seja $f : H \times H \rightarrow \mathbb{R}$ uma bifunção satisfazendo a Condição A. Escolha $u_n \in [0, 1)$ verificando a Condição B e suponha que o conjunto solução $S_{PE}(f, C)$ é não vazio. Então, a sequência $\{x_n\}$ gerada pelo Algoritmo 1 converge fortemente para um elemento $x^* \in S_{PE}(f, C)$, onde $x^* = P_{S_{PE}(f, C)}(g(x^*))$.*

Demonstração. Afirmamos que a sequência $\{\|x_n - x^*\|^2\}$ converge para zero. De fato, vimos no Lema 2.2.5 que

$$a_{n+1} \leq (1 - \gamma_n)a_n + \gamma_n b_n, \forall n \geq 1,$$

onde

$$a_n = \|x_n - x^*\|^2, \gamma_n = (1 - \rho)\alpha_n \text{ e } b_n = \xi_n.$$

Logo, pelo Lema 1.1.3, é suficiente mostrar que $\limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_{k+1}} - x^* \rangle \leq 0$, para toda subsequência $\{\|x_{n_k} - x^*\|\}$ de $\{\|x_n - x^*\|\}$ satisfazendo

$$\liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^*\| - \|x_{n_k} - x^*\|) \geq 0.$$

Seja, então, $\{\|x_{n_k} - x^*\|\}$ uma subsequência de $\{\|x_n - x^*\|\}$ tal que

$$\liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^*\| - \|x_{n_k} - x^*\|) \geq 0,$$

segue que

$$\begin{aligned} \liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^*\|^2 - \|x_{n_k} - x^*\|^2) \\ = \liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^*\| - \|x_{n_k} - x^*\|)(\|x_{n_{k+1}} - x^*\| + \|x_{n_k} - x^*\|), \end{aligned}$$

donde

$$\liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^*\|^2 - \|x_{n_k} - x^*\|^2) \geq 0. \quad (2.35)$$

Pelo Lema 2.2.4, obtemos

$$\begin{aligned} \limsup_{k \rightarrow \infty} \left[(1 - \alpha_{n_k}) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_{n_k}}{\lambda_{n_k+1}} \right) \|w_{n_k} - y_{n_k}\|^2 \right. \\ \left. + (1 - \alpha_{n_k}) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_{n_k}}{\lambda_{n_k+1}} \right) \|z_{n_k} - y_{n_k}\|^2 \right] \leq \limsup_{k \rightarrow \infty} [\|x_{n_k} - x^*\|^2 - \|x_{n_{k+1}} - x^*\|^2] \\ + 2H_3 \limsup_{k \rightarrow \infty} \alpha_{n_k} \\ = - \liminf_{k \rightarrow \infty} [\|x_{n_{k+1}} - x^*\|^2 - \|x_{n_k} - x^*\|^2] \\ \leq 0. \end{aligned}$$

Isso implica que

$$\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0 \quad (2.36)$$

e

$$\lim_{k \rightarrow \infty} \|z_{n_k} - y_{n_k}\| = 0. \quad (2.37)$$

Combinando (2.36) e (2.37), teremos

$$\lim_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\| = 0, \quad (2.38)$$

pois $0 \leq \|w_{n_k} - z_{n_k}\| \leq \|w_{n_k} - y_{n_k}\| + \|y_{n_k} - z_{n_k}\|$.

Além disso,

$$\begin{aligned} \|w_{n_k} - x_{n_k}\| &= \|x_{n_k} + u_{n_k}(x_{n_k} - x_{n_{k-1}}) - x_{n_k}\| = u_{n_k} \|x_{n_k} - x_{n_{k-1}}\| \\ &= \alpha_{n_k} \frac{u_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_{k-1}}\| \rightarrow 0, \end{aligned} \quad (2.39)$$

pois $\lim_{k \rightarrow \infty} \frac{u_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_{k-1}}\| = 0$ e $\{\alpha_{n_k}\}$ é limitada.

Da definição de $\{x_n\}$ e desigualdade triangular, segue que

$$\begin{aligned} \|x_{n_k+1} - x_{n_k}\| &= \|\alpha_{n_k}g(e_{n_k}) + (1 - \alpha_{n_k})z_{n_k} - x_{n_k}\| \\ &= \|\alpha_{n_k}(g(e_{n_k}) - x_{n_k}) + (1 - \alpha_{n_k})(z_{n_k} - x_{n_k})\| \\ &\leq \alpha_{n_k} \|g(e_{n_k}) - x_{n_k}\| + (1 - \alpha_{n_k}) \|z_{n_k} - x_{n_k}\| \\ &\leq \alpha_{n_k} \|g(e_{n_k}) - x^*\| + \alpha_{n_k} \|x_{n_k} - x^*\| + \|z_{n_k} - w_{n_k}\| \\ &\quad + \|w_{n_k} - x_{n_k}\|, \end{aligned} \quad (2.40)$$

onde, na última desigualdade, usamos o fato que $1 - \alpha_{n_k} \leq 1$.

Aplicando o limite em (2.40) quando $k \rightarrow \infty$, obtemos

$$\lim_{k \rightarrow \infty} \|x_{n_k+1} - x_{n_k}\| = 0. \quad (2.41)$$

Sendo $\{x_{n_k}\}$ limitada, existe uma subsequência $\{x_{n_{k_j}}\}$ de $\{x_{n_k}\}$ tal que $x_{n_{k_j}} \rightarrow q \in H$.

Dai,

$$\begin{aligned} \limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k} - x^* \rangle &= \lim_{j \rightarrow \infty} \langle g(x^*) - x^*, x_{n_{k_j}} - x^* \rangle \\ &= \langle g(x^*) - x^*, q - x^* \rangle. \end{aligned}$$

Por outro lado, de $x_{n_{k_j}} \rightarrow q$ e $\|w_{n_{k_j}} - x_{n_{k_j}}\| \rightarrow 0$, temos $w_{n_{k_j}} \rightarrow q$. Assim, pelo Lema 2.2.2, $q \in S_{PE}(f, C)$. Como $x^* = P_{S_{PE}(f, C)}(g(x^*))$, isto é, $\langle g(x^*) - x^*, x - x^* \rangle \leq 0, \forall x \in S_{PE}(f, C)$, segue que

$$\limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k} - x^* \rangle = \langle g(x^*) - x^*, q - x^* \rangle \leq 0.$$

Por (2.41), $\|x_{n_k+1} - x_{n_k}\| \rightarrow 0$, então

$$\limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k+1} - x^* \rangle \leq 0. \quad (2.42)$$

Logo, da definição de ξ_n , relação (2.42), Condição B e limitação da sequência $\{\|w_{n_k} - x^*\|\}$, temos que

$$\begin{aligned} \limsup_{k \rightarrow \infty} b_{n_k} &= \limsup_{k \rightarrow \infty} \xi_{n_k} \leq \frac{2}{1 - \rho} \limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k+1} - x^* \rangle \\ &\quad + \frac{2}{1 - \rho} \limsup_{k \rightarrow \infty} \frac{u_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_{k-1}}\| \|w_{n_k} - x^*\| \\ &\leq 0. \end{aligned}$$

Portanto, pelo Lema 1.1.3, $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \|x_n - x^*\|^2 = 0$, donde $\{x_n\}$ converge fortemente para $x^* \in S_{PE}(f, C)$. $\square$

A seguir, mostramos como esse resultado se aplica, de forma direta e natural, ao caso particular das desigualdades variacionais.

Aplicação a Problemas de Desigualdades Variacionais

Vimos que, quando a bifunção é dada por $f(x, y) = \langle F(x), y - x \rangle$, $\forall x, y \in C$, com $F : C \rightarrow H$ um operador, o problema de equilíbrio (1.2) se reduz ao problema de desigualdade variacional (1.6), e o esquema iterativo correspondente é descrito por

$$\begin{cases} y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda f(x_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\} = P_C(x_n - \lambda F(x_n)), \\ x_{n+1} = \operatorname{argmin}_{y \in T_n} \left\{ \lambda f(y_n, y) + \frac{1}{2} \|x_n - y\|^2 \right\} = P_{T_n}(x_n - \lambda F(y_n)). \end{cases}$$

Além disso, se o operador F for pseudomonótono em C e L -Lipschitz contínuo em H , então as condições (A2) e (A3) são automaticamente satisfeitas, ou seja, a bifunção f é pseudomonótona em C e do tipo-Lipschitz em H , com constantes $c_1 = c_2 = \frac{L}{2}$. Observamos também que as condições (A1) e (A4) continuam válidas nesse contexto.

Assim, o resultado de convergência do Teorema 2.2.1 aplica-se diretamente, como sintetiza o corolário a seguir, sob hipóteses específicas impostas ao operador F .

Condição C

- (C1) Para algum $L > 0$, F é L -Lipschitz contínuo em H , isto é, $\|F(x) - F(y)\| \leq L\|x - y\|$, $\forall x, y \in H$;
- (C2) F é pseudomonótono em C , isto é, $\langle F(x), y - x \rangle \geq 0 \Rightarrow \langle F(y), x - y \rangle \leq 0$, $\forall x, y \in C$;
- (C3) $\limsup_{n \rightarrow \infty} \langle F(x_n), y - x_n \rangle \leq \langle F(x), y - x \rangle$, para todo $y \in C$ e toda sequência $\{x_n\} \subset C$ tal que $x_n \rightharpoonup x$;
- (C4) O conjunto solução $S_{PDV}(F, C)$ é não-vazio.

Corolário 2.2.1. *Seja $f : H \times H \rightarrow \mathbb{R}$ dada por $f(x, y) = \langle F(x), y - x \rangle$, onde $F : H \rightarrow H$ é um operador satisfazendo a Condição C. Escolha $u_n \in [0, 1)$ verificando a Condição B. Gere iterativamente as sequências $\{x_n\}$, $\{y_n\}$ e $\{z_n\}$ como segue:*

Sejam $x_0, x_1 \in H$, $u_1 \in [0, 1)$, $\lambda_1 > 0$, $\varphi = \frac{1 + \sqrt{5}}{2}$, $\theta = \varphi - 1$, $\mu \in \left(0, \frac{2}{\varphi}\right)$.

Dadas as iterações atuais x_{n-1} e x_n , construa o semiespaço

$$T_n = \{x \in H : \langle w_n - \lambda_n F(w_n) - y_n, x - y_n \rangle \leq 0\}.$$

Calcule

$$\begin{cases} w_n = x_n + u_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \lambda_n F(w_n)), \\ z_n = P_{T_n}(w_n - \lambda_n F(y_n)), \\ e_n = (1 - \theta)w_n + \theta x_n, \\ x_{n+1} = \alpha_n g(e_n) + (1 - \alpha_n)z_n. \end{cases}$$

Atualize o tamanho de passo λ_{n+1} da seguinte maneira:

$$\lambda_{n+1} = \begin{cases} \min \left\{ \lambda_n, \frac{\mu}{4\varphi} \frac{\|w_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2}{\langle F(w_n) - F(y_n), z_n - y_n \rangle} \right\}, & \text{se } \langle F(w_n) - F(y_n), z_n - y_n \rangle > 0 \\ \lambda_n, & \text{caso contrário,} \end{cases}$$

Então a sequência $\{x_n\}$ converge fortemente para $x^* = P_{SPDV(F,C)}(g(x^*))$.

2.3 Taxa de Convergência

Para obtermos a taxa de convergência do nosso algoritmo, vamos assumir que f é γ -fortemente pseudomonótona em C e que a sequência $\{u_n\} \subset [0, \alpha] \subset [0, 1)$ é não-decrescente. Os seguintes resultados podem ser obtidos com a ajuda do método apresentado em [48, 51].

Lema 2.3.1. *Suponha que as condições (A1), (A3), (A4) e (A6) sejam válidas e que f é γ -fortemente pseudomonótona em C . Seja $\{x_n\}$ gerada pelo Algoritmo 1. Então para todo $\sigma \in (0, 1)$, existem $N \in \mathbb{N}$, $\nu = \nu(\sigma)$ e $\epsilon = \epsilon(\sigma)$ tais que $\nu, \epsilon \in (0, 1)$ e*

$$\|x_{n+1} - x^*\|^2 \leq \nu \|w_n - x^*\|^2 - \epsilon \|x_{n+1} - w_n\|^2, \forall n \geq N, \quad (2.43)$$

onde x^* é a única solução de $SP_E(f, C)$.

Demonstração. De fato, como $x^* \in SP_E(f, C)$, $f(x^*, y_n) \geq 0 \forall n$. Pela pseudomonotonicidade forte de f , obtemos $f(y_n, x^*) \leq -\gamma \|y_n - x^*\|^2$. Isso implica de (2.7) que

$$\begin{aligned} -\lambda_n f(y_n, z_n) &\geq \langle w_n - z_n, x^* - z_n \rangle - \lambda_n f(y_n, x^*) \\ &\geq \langle w_n - z_n, x^* - z_n \rangle + \lambda_n \gamma \|y_n - x^*\|^2. \end{aligned} \quad (2.44)$$

A parte restante é semelhante às provas de (2.3) no Lema 2.2.1 e de (2.33) no Lema 2.2.4. Assim, podemos obter, respectivamente:

$$\begin{aligned} \|z_n - x^*\|^2 &\leq \|w_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 - 2\gamma\lambda_n \|y_n - x^*\|^2 \\ &\leq \|w_n - x^*\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 - 2\gamma\lambda_n \|y_n - x^*\|^2, \end{aligned} \quad (2.45)$$

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\leq (1 - \theta\alpha_n\rho^2) \|w_n - x^*\|^2 + \theta\alpha_n\rho^2 \|x_n - x^*\|^2 \\ &\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|w_n - y_n\|^2 \\ &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2 - 2\gamma\lambda_n (1 - \alpha_n) \|y_n - x^*\|^2. \end{aligned} \quad (2.46)$$

Agora, provaremos a desigualdade (2.43). Com efeito, sabemos que

$$(1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) > 0, \forall n \geq n_1.$$

Dado $\sigma \in (0, 1)$, graças a (2.46), obtemos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq (1 - \theta\alpha_n\rho^2)\|w_n - x^*\|^2 + \theta\alpha_n\rho^2\|x_n - x^*\|^2 \\
&\quad + 2\alpha_n\|g(e_n) - x^*\|\|g(x^*) - x^*\| - (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|^2 \\
&\quad - (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)(1 - \sigma)\|z_n - y_n\|^2 - 2\gamma\lambda_n(1 - \alpha_n)\|y_n - x^*\|^2 \\
&= (1 - \theta\alpha_n\rho^2)\|w_n - x^*\|^2 + \theta\alpha_n\rho^2\|x_n - x^*\|^2 \\
&\quad + 2\alpha_n\|g(e_n) - x^*\|\|g(x^*) - x^*\| - (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)\sigma\|w_n - y_n\|^2 \\
&\quad - (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)(1 - \sigma)\left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right] \\
&\quad - 2\gamma\lambda_n(1 - \alpha_n)\|y_n - x^*\|^2 \\
&\leq (1 - \theta\alpha_n\rho^2)\|w_n - x^*\|^2 + \theta\alpha_n\rho^2\|x_n - x^*\|^2 \\
&\quad + 2\alpha_n\|g(e_n) - x^*\|\|g(x^*) - x^*\| - (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)\sigma\|w_n - y_n\|^2 \\
&\quad - \frac{1}{2}(1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)(1 - \sigma)\|z_n - w_n\|^2 \\
&\quad - 2\gamma\lambda_n(1 - \alpha_n)\|y_n - x^*\|^2. \tag{2.47}
\end{aligned}$$

Usando a definição de x_{n+1} , podemos escrever

$$z_n - w_n = (x_{n+1} - w_n) - \alpha_n(g(e_n) - z_n).$$

Segue daí que

$$\begin{aligned}
\|z_n - w_n\|^2 &= \|(x_{n+1} - w_n) - \alpha_n(g(e_n) - z_n)\|^2 \\
&= \|x_{n+1} - w_n\|^2 + \alpha_n^2\|z_n - g(e_n)\|^2 - 2\alpha_n\langle x_{n+1} - w_n, g(e_n) - z_n \rangle \\
&\stackrel{DCS}{\geq} \|x_{n+1} - w_n\|^2 + \alpha_n^2\|z_n - g(e_n)\|^2 - 2\alpha_n\|x_{n+1} - w_n\|\|z_n - g(e_n)\| \tag{2.48}
\end{aligned}$$

Seja

$$\beta := \min\left\{\frac{2 - \mu\varphi}{4}\sigma, \lambda\gamma\right\}, \quad \text{onde } \lambda = \lim_{n \rightarrow \infty} \lambda_n (> 0).$$

Então

$$\begin{aligned}
\lim_{n \rightarrow \infty} (1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)\sigma &\geq 2\beta, \\
\lim_{n \rightarrow \infty} \frac{1}{2}(1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)(1 - \sigma) &\geq \frac{1}{2}\left(1 - \frac{\mu\varphi}{2}\right)(1 - \sigma)\sigma \geq \beta(1 - \sigma), \\
\lim_{n \rightarrow \infty} \gamma\lambda_n(1 - \alpha_n) &= \gamma\lambda \geq \beta.
\end{aligned}$$

Assim, existe $N \geq n_1$ tal que, para todo $n \geq N$,

$$\begin{aligned}
(1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)\sigma &\geq 2\beta, \\
\frac{1}{2}(1 - \alpha_n)\left(1 - \frac{\mu\varphi}{2}\frac{\lambda_n}{\lambda_{n+1}}\right)(1 - \sigma) &\geq \beta(1 - \sigma), \\
\gamma\lambda_n(1 - \alpha_n) &\geq \beta.
\end{aligned}$$

Usando essas estimativas, juntamente com (2.47) e (2.48), chegamos em

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\leq \|w_n - x^*\|^2 - 2\beta \|w_n - y_n\|^2 - \beta(1 - \sigma) \|x_{n+1} - w_n\|^2 - 2\beta \|y_n - x^*\|^2 \\ &= \|w_n - x^*\|^2 - 2\beta (\|w_n - y_n\|^2 + \|y_n - x^*\|^2) - \beta(1 - \sigma) \|x_{n+1} - w_n\|^2.\end{aligned}$$

Como $\|a\|^2 + \|b\|^2 \geq \frac{1}{2}\|a + b\|^2$ para quaisquer $a, b \in H$, tomando $a = w_n - y_n$ e $b = y_n - x^*$, obtemos

$$\|w_n - y_n\|^2 + \|y_n - x^*\|^2 \geq \frac{1}{2}\|w_n - x^*\|^2.$$

Logo,

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\leq \|w_n - x^*\|^2 - \beta \|w_n - x^*\|^2 - \beta(1 - \sigma) \|x_{n+1} - w_n\|^2 \\ &= (1 - \beta) \|w_n - x^*\|^2 - \beta(1 - \sigma) \|x_{n+1} - w_n\|^2, \quad \forall n \geq N.\end{aligned}$$

Por fim, defina $\nu := 1 - \beta \in (0, 1)$ e $\epsilon := \beta(1 - \sigma) \in (0, 1)$, pois $\sigma, \beta \in (0, 1)$. Isto prova (2.43). $\square$

A seguir, com o auxílio do lema anterior, enunciaremos e provaremos o teorema que garante a taxa de convergência R-linear do nosso algoritmo.

Teorema 2.3.1. *Suponha que a bifunção f é γ -fortemente pseudomonótona em C e satisfaça as condições (A3), (A4) e (A6). Seja $\sigma \in (0, 1)$ e $\alpha > 0$ satisfazendo*

$$\alpha \leq \frac{\nu\epsilon}{\nu\epsilon + 2\nu + \epsilon}, \quad (2.49)$$

onde $\nu = 1 - \beta$, $\epsilon = \beta(1 - \sigma)$, $\beta = \min \left\{ \frac{2 - \mu\varphi}{4} \sigma, \lambda\gamma \right\}$ e $\lambda = \lim_{n \rightarrow \infty} \lambda_n$. Então, a sequência gerada pelo Algoritmo 1 converge em norma para a única solução x^* do problema $PE(f, C)$, com taxa R-linear.

Demonstração. Pela definição de $\{w_n\}$, temos que

$$\begin{aligned}\|w_n - x^*\|^2 &= \|(1 + u_n)(x_n - x^*) - u_n(x_{n-1} - x^*)\|^2 \\ &= (1 + u_n)\|x_n - x^*\|^2 - u_n\|x_{n-1} - x^*\|^2 + u_n(1 + u_n)\|x_n - x_{n-1}\|^2\end{aligned}$$

e

$$\begin{aligned}\|x_{n+1} - w_n\|^2 &= \|x_{n+1} - x_n\|^2 + u_n^2\|x_n - x_{n-1}\|^2 - 2u_n \langle x_{n+1} - x_n, x_n - x_{n-1} \rangle \\ &\geq \|x_{n+1} - x_n\|^2 + u_n^2\|x_n - x_{n-1}\|^2 - 2u_n\|x_{n+1} - x_n\|\|x_n - x_{n-1}\| \\ &\geq \|x_{n+1} - x_n\|^2 + u_n^2\|x_n - x_{n-1}\|^2 - u_n\|x_{n+1} - x_n\|^2 - u_n\|x_n - x_{n-1}\|^2 \\ &= (1 - u_n)\|x_{n+1} - x_n\|^2 - u_n(1 - u_n)\|x_n - x_{n-1}\|^2.\end{aligned}$$

Combinando essas desigualdades com (2.43), obtemos, $\forall n \geq N$

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\leq \nu(1 + u_n)\|x_n - x^*\|^2 - \nu u_n\|x_{n-1} - x^*\|^2 + \nu u_n(1 + u_n)\|x_n - x_{n-1}\|^2 \\ &\quad - \epsilon(1 - u_n)\|x_{n+1} - x_n\|^2 + \epsilon u_n(1 - u_n)\|x_n - x_{n-1}\|^2.\end{aligned}$$

Como $\{u_n\}$ é não-decrescente, i.e., $u_n \leq u_{n+1}$, então

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \nu u_{n+1} \|x_n - x^*\|^2 + \epsilon(1 - u_{n+1}) \|x_{n+1} - x_n\|^2 \\ &\leq \sigma \left[\|x_n - x^*\|^2 - u_n \|x_{n-1} - x^*\|^2 + \epsilon(1 - u_n) \|x_n - x_{n-1}\|^2 \right] \\ &\quad - [\nu\epsilon(1 - u_n) - \nu u_n(1 + u_n) - \epsilon u_n(1 - u_n)] \|x_n - x_{n-1}\|^2. \end{aligned} \quad (2.50)$$

Definindo, $\forall n \geq N$, $\Gamma_n := \|x_n - x^*\|^2 - u_n \|x_{n-1} - x^*\|^2 + \epsilon(1 - u_n) \|x_n - x_{n-1}\|^2$, por (2.50) e $\nu \in (0, 1)$, obtemos

$$\begin{aligned} \Gamma_{n+1} &= \|x_{n+1} - x^*\|^2 - u_{n+1} \|x_n - x^*\|^2 + \epsilon(1 - u_{n+1}) \|x_{n+1} - x_n\|^2 \\ &\leq \|x_{n+1} - x^*\|^2 - \nu u_{n+1} \|x_n - x^*\|^2 + \epsilon(1 - u_{n+1}) \|x_{n+1} - x_n\|^2 \\ &\leq \nu \Gamma_n - [\nu\epsilon(1 - u_n) - \nu u_n(1 + u_n) - \epsilon u_n(1 - u_n)] \|x_n - x_{n-1}\|^2. \end{aligned} \quad (2.51)$$

Definindo $\Omega_n := \nu\epsilon(1 - u_n) - \nu u_n(1 + u_n) - \epsilon u_n(1 - u_n)$, mostraremos que $\Omega_n \geq 0, \forall n \geq N$. De fato, como $u_n \in [0, 1)$, então $u_n(1 + u_n) \leq 2u_n \leq 2\alpha$, $u_n(1 - u_n) \leq u_n \leq \alpha$ e $1 - u_n \geq 1 - \alpha$. Assim, de (2.49),

$$\Omega_n \geq \nu\epsilon(1 - \alpha) - 2\nu\alpha - \epsilon\alpha = \nu\epsilon - \alpha(\nu\epsilon + 2\nu + \epsilon) \geq 0, \forall n \geq N.$$

O que implica de (2.51) que $\Gamma_{n+1} \leq \nu\Gamma_n, \forall n \geq N$.

Agora, mostraremos que $\Gamma_n \geq 0, \forall n \geq N$.

De fato, como $u_n \leq \alpha \leq \frac{\nu\epsilon}{\nu\epsilon + 2\nu + \epsilon} \leq \frac{\epsilon}{2 + \epsilon}$, então $1 - u_n \geq \frac{2}{2 + \epsilon}$ e, consequentemente, $\frac{\epsilon(1 - u_n)}{2} \geq u_n$. Logo,

$$\begin{aligned} \Gamma_n &= [1 - \epsilon(1 - u_n)] \|x_n - x^*\|^2 + \epsilon(1 - u_n) \left(\|x_n - x^*\|^2 + \|x_n - x_{n-1}\|^2 \right) - u_n \|x_{n-1} - x^*\|^2 \\ &\geq [1 - \epsilon(1 - u_n)] \|x_n - x^*\|^2 + \frac{\epsilon(1 - u_n)}{2} \|x_{n-1} - x^*\|^2 - u_n \|x_{n-1} - x^*\|^2 \\ &\geq [1 - \epsilon(1 - u_n)] \|x_n - x^*\|^2 \geq 0, \forall n \geq N. \end{aligned}$$

Agora, como $\Gamma_{n+1} \leq \nu\Gamma_n, \forall n \geq N$, obtemos $\Gamma_{n+1} \leq \nu\Gamma_n \leq \dots \leq \nu^{n-N+1}\Gamma_N$. Portanto,

$$(1 - \epsilon) \|x_n - x^*\|^2 \leq [1 - \epsilon(1 - u_n)] \|x_n - x^*\|^2 \leq \frac{\Gamma_N}{\nu^N} \nu^n,$$

isto é,

$$\|x_n - x^*\|^2 \leq \frac{\Gamma_N}{(1 - \epsilon)\nu^N} \nu^n, \forall n \geq N,$$

o que implica que $\{x_n\}$ converge R -linearmente para x^* . □

Capítulo 3

Método Extragradiante Subgradiente Duplamente Acelerado para PE

Neste capítulo, propomos uma extensão do algoritmo apresentado no Capítulo 2, incorporando uma estratégia adicional de aceleração por meio de um segundo termo inercial. O objetivo é potencializar o desempenho computacional do método, preservando a convergência forte sob as mesmas hipóteses impostas à bifunção e mediante a escolha de uma nova sequência que satisfaz a Condição B. A nova versão, denominada algoritmo extragradiante subgradiente com duplo passo inercial, combina mecanismos de atualização autoadaptativa de passos, aproximação por viscosidade e aceleração inercial dupla. Essa combinação visa reduzir o número de iterações e o tempo de execução necessários para alcançar soluções de problemas de equilíbrio em espaços de Hilbert reais.

3.1 Descrição do Algoritmo 2

Nesta seção, formalizamos o algoritmo extragradiante subgradiente com duplo passo inercial, cuja estrutura é baseada no método apresentado anteriormente, enriquecido agora por dois termos inerciais distintos. O primeiro termo, w_n , segue as abordagens clássicas de inércia e é o mesmo utilizado no Algoritmo 1. Já o segundo termo, $\bar{w}_n$, é incorporado diretamente nas etapas de avaliação da bifunção e de construção do semiespaço auxiliar. Essa dupla contribuição inercial visa potencializar a aceleração do processo iterativo, preservando a estabilidade e a robustez da convergência.

Para isso, escolhemos uma nova sequência $\{s_n\} \subset [0, 1)$, a qual também satisfaz a Condição B, isto é, $\lim_{n \rightarrow +\infty} \frac{s_n}{\alpha_n} \|x_n - x_{n-1}\| = 0$. Essa sequência é utilizada na construção do ponto $\bar{w}_n$, que substitui w_n nas etapas de minimização do método (definições de y_n e z_n). Com isso, o algoritmo é descrito da seguinte forma:

Algoritmo 2 : Algoritmo extragradiante subgradiente com duplo passo inercial para (PE)

Passo 0 (Inicialização):

Escolha pontos iniciais $x_0, x_1 \in H$, $u_1, s_1 \in [0, 1)$, $\lambda_1 > 0$, $\varphi = \frac{1 + \sqrt{5}}{2}$, $\theta = \varphi - 1$, $\mu \in \left(0, \frac{2}{\varphi}\right)$ e defina $n = 1$.

Algoritmo 2 (continuação)

Passo 1:

Assuma que $x_{n-1}, x_n \in H$, $u_n, s_n \in [0, 1)$, $\lambda_n > 0$ ($n \geq 1$) são conhecidos.

Calcule

$$w_n = x_n + u_n(x_n - x_{n-1}), \quad \bar{w}_n = x_n + s_n(x_n - x_{n-1}),$$

e resolva o problema de minimização

$$y_n = \operatorname{argmin}_{y \in C} \left\{ \lambda_n f(\bar{w}_n, y) + \frac{1}{2} \|\bar{w}_n - y\|^2 \right\}.$$

Se $\bar{w}_n = y_n$, então **PARE**: y_n é solução do (PE). Caso contrário, continue.

Passo 2:

Escolha $v_n \in \partial_2 f(\bar{w}_n, y_n)$ e $q_n \in N_C(y_n)$ tais que $q_n = \bar{w}_n - \lambda_n v_n - y_n$ e construa o semiespaço

$$T_n = \{x \in H : \langle \bar{w}_n - \lambda_n v_n - y_n, x - y_n \rangle \leq 0\}.$$

Então, calcule

$$z_n = \operatorname{argmin}_{y \in T_n} \left\{ \lambda_n f(y_n, y) + \frac{1}{2} \|\bar{w}_n - y\|^2 \right\}.$$

Passo 3 (Atualização com viscosidade):

Calcule

$$e_n = (1 - \theta)w_n + \theta x_n, \quad x_{n+1} = \alpha_n g(e_n) + (1 - \alpha_n)z_n,$$

onde $g : H \rightarrow H$ é uma ρ -contração.

Atualize o tamanho de passo λ_{n+1} :

$$\lambda_{n+1} = \begin{cases} \min \left\{ \lambda_n, \frac{\mu(\|\bar{w}_n - y_n\|^2 + (1 + \varphi)\|z_n - y_n\|^2)}{4\varphi(f(\bar{w}_n, z_n) - f(\bar{w}_n, y_n) - f(y_n, z_n))} \right\}, & \text{se } d_n > 0, \\ \lambda_n, & \text{caso contrário,} \end{cases}$$

com $d_n = f(\bar{w}_n, z_n) - f(\bar{w}_n, y_n) - f(y_n, z_n)$.

Defina $n := n + 1$ e volte ao **Passo 1**.

Observação 3.1.1. A verificação da boa definição do Algoritmo 2 é análoga à do Algoritmo 1, uma vez que as alterações introduzidas não afetam a estrutura funcional dos subproblemas de minimização e demais etapas do algoritmo. Isso também vale para as propriedades da sequência $\{\lambda_n\}$.

3.2 Análise de Convergência

Para a análise de convergência do Algoritmo 2, procederemos de forma semelhante a que fizemos para o Algoritmo 1.

Os dois primeiros lemas auxiliares serão apenas enunciados, pois suas demonstrações são as mesmas dos lemas 2.2.1 e 2.2.2, trocando w_n por $\bar{w}_n$, visto que a expressão de w_n não é utilizada nas etapas das demonstrações.

Lema 3.2.1. *Suponha que a bifunção f satisfaz a Condição A e $\{z_n\}$ é uma sequência gerada pelo Algoritmo 2. Então, para todo $x^* \in S_{PE}(f, C)$, a seguinte desigualdade é válida:*

$$\|z_n - x^*\|^2 \leq \|\bar{w}_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}}\right) \|\bar{w}_n - y_n\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}}\right) \|z_n - y_n\|^2, \forall n \geq 1. \quad (3.1)$$

Lema 3.2.2. *Suponha que a Condição A seja válida e que $\{\bar{w}_n\}$ é uma sequência gerada pelo Algoritmo 2. Se existe uma subsequência $\{\bar{w}_{n_k}\}$ convergindo fracamente para $z \in H$ e $\lim_{k \rightarrow \infty} \|\bar{w}_{n_k} - y_{n_k}\| = 0$, $\lim_{k \rightarrow \infty} \|\bar{w}_{n_k} - z_{n_k}\| = 0$, $\lim_{k \rightarrow \infty} \|z_{n_k} - y_{n_k}\| = 0$, então $z \in S_{PE}(f, C)$.*

As demonstrações dos demais lemas envolvem diretamente a expressão de $\bar{w}_n$, o que as diferencia, em certos aspectos, das demonstrações correspondentes da Seção 2.2.

Lema 3.2.3. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 2 sob as condições A e B. Então, para qualquer $x^* \in S_{PE}(f, C)$, as sequências $\{x_n\}$, $\{w_n\}$, $\{\bar{w}_n\}$, $\{z_n\}$, $\{e_n\}$ e $\{g(e_n)\}$ são limitadas.*

Demonstração. Da definição de $\{w_n\}$ e desigualdade triangular (DT), temos:

$$\begin{aligned} \|w_n - x^*\| &= \|x_n + u_n(x_n - x_{n-1}) - x^*\| \\ &\leq \|x_n - x^*\| + u_n \|x_n - x_{n-1}\| \\ &= \|x_n - x^*\| + \alpha_n \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\|. \end{aligned} \quad (3.2)$$

Usando o mesmo argumento de (3.2) para $\bar{w}_n = x_n + s_n(x_n - x_{n-1})$, obtemos

$$\begin{aligned} \|\bar{w}_n - x^*\| &= \|x_n + s_n(x_n - x_{n-1}) - x^*\| \\ &\leq \|x_n - x^*\| + s_n \|x_n - x_{n-1}\| \\ &= \|x_n - x^*\| + \alpha_n \frac{s_n}{\alpha_n} \|x_n - x_{n-1}\|. \end{aligned} \quad (3.3)$$

Como

$$\lim_{n \rightarrow \infty} \frac{s_n}{\alpha_n} \|x_n - x_{n-1}\| = 0 \quad \text{e} \quad \lim_{n \rightarrow \infty} \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| = 0.$$

existem números positivos M_1, M_2 tais que

$$\frac{s_n}{\alpha_n} \|x_n - x_{n-1}\| \leq M_1 \quad \text{e} \quad \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \leq M_2, \quad \forall n \geq 1.$$

Logo, fazendo $M^* = \max\{M_1, M_2\}$, de (3.2) e (3.3) obtemos, respectivamente, que

$$\|w_n - x^*\| \leq \|x_n - x^*\| + \alpha_n M_2 \leq \|x_n - x^*\| + \alpha_n M^* \quad (3.4)$$

e

$$\|\bar{w}_n - x^*\| \leq \|x_n - x^*\| + \alpha_n M_1 \leq \|x_n - x^*\| + \alpha_n M^*. \quad (3.5)$$

Por outro lado, da definição de $\{e_n\}$ e (3.4),

$$\begin{aligned} \|e_n - x^*\| &= \|(1 - \theta)w_n + \theta x_n - x^*\| = \|(1 - \theta)(w_n - x^*) + \theta(x_n - x^*)\| \\ &\stackrel{DT}{\leq} (1 - \theta)\|w_n - x^*\| + \theta\|x_n - x^*\| \\ &\leq (1 - \theta)[\|x_n - x^*\| + \alpha_n M^*] + \theta\|x_n - x^*\| \\ &= \|x_n - x^*\| + (1 - \theta)\alpha_n M^*. \end{aligned} \quad (3.6)$$

Portanto, para todo $n \geq n_1$,

$$\begin{aligned}
\|x_{n+1} - x^*\| &= \|\alpha_n g(e_n) + (1 - \alpha_n)z_n - x^*\| \\
&= \|\alpha_n(g(e_n) - x^*) + (1 - \alpha_n)(z_n - x^*)\| \\
&\stackrel{DT}{\leq} \alpha_n \|g(e_n) - x^*\| + (1 - \alpha_n) \|z_n - x^*\| \\
&\stackrel{(3.1)}{\leq} \alpha_n \|g(e_n) - x^*\| + (1 - \alpha_n) \|\bar{w}_n - x^*\| \\
&= \alpha_n \|g(e_n) - g(x^*) + g(x^*) - x^*\| + (1 - \alpha_n) \|\bar{w}_n - x^*\| \\
&\stackrel{DT}{\leq} \alpha_n (\|g(e_n) - g(x^*)\| + \|g(x^*) - x^*\|) + (1 - \alpha_n) \|\bar{w}_n - x^*\| \\
&\stackrel{\rho\text{-contração}}{\leq} \rho \alpha_n \|e_n - x^*\| + \alpha_n \|g(x^*) - x^*\| + (1 - \alpha_n) \|\bar{w}_n - x^*\| \\
&\stackrel{(3.5), (3.6), 1-\theta \leq 1}{\leq} \rho \alpha_n (\|x_n - x^*\| + \alpha_n M^*) + \alpha_n \|g(x^*) - x^*\| \\
&\quad + (1 - \alpha_n) [\|x_n - x^*\| + \alpha_n M^*] \\
&= [\rho \alpha_n + (1 - \alpha_n)] \|x_n - x^*\| + \rho \alpha_n^2 M^* + (1 - \alpha_n) \alpha_n M^* \\
&\quad + \alpha_n \|g(x^*) - x^*\| \\
&= [\rho \alpha_n + (1 - \alpha_n)] \|x_n - x^*\| + (\rho - 1) \alpha_n^2 M^* + \alpha_n M^* \\
&\quad + \alpha_n \|g(x^*) - x^*\| \\
&\stackrel{\rho-1 < 0}{\leq} [1 - \alpha_n(1 - \rho)] \|x_n - x^*\| + \alpha_n M^* + \alpha_n \|g(x^*) - x^*\| \\
&= [1 - \alpha_n(1 - \rho)] \|x_n - x^*\| + \alpha_n(1 - \rho) \frac{M^* + \|g(x^*) - x^*\|}{1 - \rho} \\
&\leq \max \left\{ \|x_n - x^*\|, \frac{M^* + \|g(x^*) - x^*\|}{1 - \rho} \right\} \\
&\vdots \\
&\leq \max \left\{ \|x_{n_1} - x^*\|, \frac{M^* + \|g(x^*) - x^*\|}{1 - \rho} \right\},
\end{aligned}$$

isto é, $\{x_n\}$ é limitada. Além disso, por (3.4), $\{w_n\}$ é limitada; por (3.5), $\{\bar{w}_n\}$ é limitada; por (3.6), $\{e_n\}$ é limitada e, conseqüentemente, $\{g(e_n)\}$ é limitada. Por (3.5) e (3.1), $\{z_n\}$ é limitada. $\square$

Lema 3.2.4. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 2 sob as condições A e B. Então, para todo $x^* \in S_{PE}(f, C)$, existe $\bar{H}_3 > 0$ tal que:*

$$\begin{aligned}
(1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|\bar{w}_n - y_n\|^2 + (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \\
\leq \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\alpha_n \bar{H}_3.
\end{aligned} \tag{3.7}$$

Demonstração. Análogo à demonstração inicial do Lema 2.2.4, obtemos

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq \alpha_n \rho^2 [(1 - \theta) \|w_n - x^*\|^2 + \theta \|x_n - x^*\|^2] \\
&+ (1 - \alpha_n) \left[\|\bar{w}_n - x^*\|^2 - \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|\bar{w}_n - y_n\|^2 - \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \right] \\
&+ 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\|.
\end{aligned} \tag{3.8}$$

Pela definição de $\{w_n\}$ e Lema 1.1.4 (b), temos

$$\begin{aligned}
 \|w_n - x^*\|^2 &= \|x_n + u_n(x_n - x_{n-1}) - x^*\|^2 \\
 &\stackrel{\text{Lema 1.1.4(b)}}{\leq} \|x_n - x^*\|^2 + 2 \langle u_n(x_n - x_{n-1}), w_n - x^* \rangle \\
 &\stackrel{DCS}{\leq} \|x_n - x^*\|^2 + 2u_n \|x_n - x_{n-1}\| \|w_n - x^*\|.
 \end{aligned} \tag{3.9}$$

Agora, da Condição B e limitação de $\{w_n\}$, a sequência $\left\{ \frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \right\}$ é limitada.

Assim, existe $\bar{H}_2 > 0$ tal que $\frac{u_n}{\alpha_n} \|x_n - x_{n-1}\| \|w_n - x^*\| \leq \bar{H}_2, \forall n \geq 1$.

Logo,

$$\|w_n - x^*\|^2 \leq \|x_n - x^*\|^2 + 2\alpha_n \bar{H}_2, \forall n \geq 1. \tag{3.10}$$

Substituindo (3.5) e (3.10) em (3.8), segue que

$$\begin{aligned}
 \|x_{n+1} - x^*\|^2 &\leq \alpha_n \rho^2 (1 - \theta) [\|x_n - x^*\|^2 + 2\alpha_n \bar{H}_2] + \alpha_n \rho^2 \theta \|x_n - x^*\|^2 \\
 &\quad + (1 - \alpha_n) [\|x_n - x^*\| + \alpha_n M^*]^2 \\
 &\quad - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|\bar{w}_n - y_n\|^2 \\
 &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \\
 &\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\| \\
 &= \alpha_n \rho^2 \|x_n - x^*\|^2 + 2(1 - \theta) \alpha_n^2 \rho^2 \bar{H}_2 \\
 &\quad + (1 - \alpha_n) (\|x_n - x^*\|^2 + 2\alpha_n M^* \|x_n - x^*\| + \alpha_n^2 M^{*2}) \\
 &\quad - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|\bar{w}_n - y_n\|^2 \\
 &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 \\
 &\quad + 2\alpha_n \|g(e_n) - x^*\| \|g(x^*) - x^*\|.
 \end{aligned} \tag{3.11}$$

Note que, $\forall n \geq 1$, existe uma constante $\bar{H}_3 > 0$ tal que

$$(1 - \theta) \rho^2 \bar{H}_2 \alpha_n + (1 - \alpha_n) M^* \|x_n - x^*\| + \frac{M^{*2}}{2} \alpha_n (1 - \alpha_n) + \|g(x^*) - x^*\| \|g(e_n) - x^*\| \leq \bar{H}_3.$$

(De fato, $\{x_n\}, \{g(e_n)\}$ são limitadas pelo Lema 3.2.3 e $0 < \alpha_n < 1$.)

Usando essa estimativa e o fato que

$$\alpha_n \rho^2 + (1 - \alpha_n) \leq 1 \quad (\text{pois } \alpha_n \rho^2 + (1 - \alpha_n) = 1 - \alpha_n(1 - \rho^2)),$$

a relação (3.11) se escreve como:

$$\begin{aligned}
 \|x_{n+1} - x^*\|^2 &\leq \|x_n - x^*\|^2 - (1 - \alpha_n) \left(1 - \frac{\mu}{2\varphi} \frac{\lambda_n}{\lambda_{n+1}} \right) \|\bar{w}_n - y_n\|^2 \\
 &\quad - (1 - \alpha_n) \left(1 - \frac{\mu\varphi}{2} \frac{\lambda_n}{\lambda_{n+1}} \right) \|z_n - y_n\|^2 + 2\alpha_n \bar{H}_3,
 \end{aligned}$$

donde, obtemos (3.7). $\square$

Lema 3.2.5. *Seja $\{x_n\}$ uma sequência gerada pelo Algoritmo 2 sob as condições A e B. Então, para todo $x^* \in S_{PE}(f, C)$, vale:*

$$\|x_{n+1} - x^*\|^2 \leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 + \alpha_n(1 - \rho)\psi_n,$$

onde

$$\begin{aligned} \psi_n = & \frac{2s_n(1 - \alpha_n)\|x_n - x_{n-1}\|\|\bar{w}_n - x^*\|}{(1 - \rho)\alpha_n} + \frac{2\rho u_n\|x_n - x_{n-1}\|\|w_n - x^*\|}{1 - \rho} \\ & + \frac{2\langle g(x^*) - x^*, x_{n+1} - x^* \rangle}{1 - \rho}. \end{aligned}$$

Demonstração. De modo análogo ao início da demonstração do Lema 2.2.5, obtemos

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \leq & \alpha_n \rho^2 \left[(1 - \theta)\|w_n - x^*\|^2 + \theta\|x_n - x^*\|^2 \right] \\ & + (1 - \alpha_n)\|\bar{w}_n - x^*\|^2 + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle. \end{aligned} \quad (3.12)$$

Usando (3.9) em (3.12), obtemos

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \leq & \alpha_n \rho^2 \left[(1 - \theta)\|x_n - x^*\|^2 + 2(1 - \theta)u_n\|x_n - x_{n-1}\|\|w_n - x^*\| + \theta\|x_n - x^*\|^2 \right] \\ & + (1 - \alpha_n)\|\bar{w}_n - x^*\|^2 + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle. \end{aligned}$$

Utilizando $\rho^2 \leq \rho$ e $1 - \theta \leq 1$, segue que

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \leq & \alpha_n \rho \left[\|x_n - x^*\|^2 + 2u_n\|x_n - x_{n-1}\|\|w_n - x^*\| \right] \\ & + (1 - \alpha_n)\|\bar{w}_n - x^*\|^2 + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle. \end{aligned} \quad (3.13)$$

Por outro lado, pela definição de $\bar{w}_n$, pelo Lema 1.1.4 (b) e pela desigualdade de Cauchy-Schwarz, tem-se

$$\begin{aligned} \|\bar{w}_n - x^*\|^2 &= \|x_n + s_n(x_n - x_{n-1}) - x^*\|^2 \\ &= \|x_n - x^*\|^2 + 2\langle \bar{w}_n - x^*, s_n(x_n - x_{n-1}) \rangle \\ &\leq \|x_n - x^*\|^2 + 2s_n\|x_n - x_{n-1}\|\|\bar{w}_n - x^*\|. \end{aligned} \quad (3.14)$$

Substituindo (3.14) em (3.13), obtemos

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \leq & \alpha_n \rho \left[\|x_n - x^*\|^2 + 2u_n\|x_n - x_{n-1}\|\|w_n - x^*\| \right] \\ & + (1 - \alpha_n) \left(\|x_n - x^*\|^2 + 2s_n\|x_n - x_{n-1}\|\|\bar{w}_n - x^*\| \right) \\ & + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle \\ = & [\alpha_n \rho + (1 - \alpha_n)] \|x_n - x^*\|^2 + 2\rho u_n \alpha_n \|x_n - x_{n-1}\|\|w_n - x^*\| \\ & + 2s_n(1 - \alpha_n)\|x_n - x_{n-1}\|\|\bar{w}_n - x^*\| + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle. \end{aligned}$$

Como $[\alpha_n \rho + (1 - \alpha_n)] = 1 - \alpha_n(1 - \rho)$, segue que

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \leq & [1 - \alpha_n(1 - \rho)] \|x_n - x^*\|^2 + 2\rho u_n \alpha_n \|x_n - x_{n-1}\|\|w_n - x^*\| \\ & + 2s_n(1 - \alpha_n)\|x_n - x_{n-1}\|\|\bar{w}_n - x^*\| + 2\alpha_n \langle g(x^*) - x^*, x_{n+1} - x^* \rangle. \end{aligned}$$

Colocando em evidência $\alpha_n(1 - \rho)$ nos três últimos termos do lado direito, obtemos exatamente

$$\|x_{n+1} - x^*\|^2 \leq (1 - \alpha_n(1 - \rho)) \|x_n - x^*\|^2 + \alpha_n(1 - \rho)\psi_n,$$

onde ψ_n é dada na afirmação do lema. Isto conclui a prova. $\square$

A seguir, enunciamos e demonstramos o resultado que estabelece a convergência forte da sequência gerada pelo Algoritmo 2 para uma solução do problema de equilíbrio.

Teorema 3.2.1. *Seja C um subconjunto não vazio, convexo e fechado de um espaço de Hilbert real H , e seja $f : H \times H \rightarrow \mathbb{R}$ uma bifunção que satisfaz a Condição A. Escolha $u_n, s_n \in [0, 1]$ e $\alpha_n \in (0, 1)$, verificando a Condição B, e suponha que o conjunto solução $S_{PE}(f, C)$ seja não vazio. Então a sequência $\{x_n\}$ gerada pelo Algoritmo 2 converge fortemente para um elemento $x^* \in S_{PE}(f, C)$, onde $x^* = P_{S_{PE}(f, C)}(g(x^*))$.*

Demonstração. Pelo Lema 3.2.5, para todo $n \geq 1$ vale

$$\|x_{n+1} - x^*\|^2 \leq (1 - \alpha_n(1 - \rho))\|x_n - x^*\|^2 + \alpha_n(1 - \rho)\psi_n.$$

Defina

$$a_n := \|x_n - x^*\|^2, \quad \gamma_n := (1 - \rho)\alpha_n, \quad b_n := \psi_n.$$

Então

$$a_{n+1} \leq (1 - \gamma_n)a_n + \gamma_nb_n, \quad n \geq 1. \quad (3.15)$$

Pela Condição B, $0 < \gamma_n < 1$, $\gamma_n \rightarrow 0$ e $\sum_{n=1}^{\infty} \gamma_n = (1 - \rho) \sum_{n=1}^{\infty} \alpha_n = +\infty$.

Aplicaremos o Lema 1.1.3. Para isso, tome uma subsequência $\{a_{n_k}\}$ tal que

$$\liminf_{k \rightarrow \infty} (a_{n_k+1} - a_{n_k}) \geq 0. \quad (3.16)$$

Mostraremos que

$$\limsup_{k \rightarrow \infty} b_{n_k} \leq 0. \quad (3.17)$$

Primeiro, como $\bar{w}_n = x_n + s_n(x_n - x_{n-1})$, segue que

$$\|\bar{w}_n - x_n\| = s_n\|x_n - x_{n-1}\| = \alpha_n \frac{s_n}{\alpha_n} \|x_n - x_{n-1}\|.$$

Logo, pela Condição B,

$$\lim_{k \rightarrow \infty} \|\bar{w}_{n_k} - x_{n_k}\| = 0. \quad (3.18)$$

Pelo Lema 3.2.4 e (3.16), obtemos:

$$\|\bar{w}_{n_k} - y_{n_k}\| \rightarrow 0 \quad \text{e} \quad \|z_{n_k} - y_{n_k}\| \rightarrow 0, \quad (3.19)$$

logo

$$\|z_{n_k} - \bar{w}_{n_k}\| \rightarrow 0. \quad (3.20)$$

Da definição de x_{n+1} , desigualdade triangular e de $1 - \alpha_{n_k} \leq 1$ obtemos

$$\begin{aligned} \|x_{n_k+1} - x_{n_k}\| &= \|\alpha_{n_k}g(e_{n_k}) + (1 - \alpha_{n_k})z_{n_k} - x_{n_k}\| \\ &= \|\alpha_{n_k}(g(e_{n_k}) - x_{n_k}) + (1 - \alpha_{n_k})(z_{n_k} - x_{n_k})\| \\ &\leq \alpha_{n_k}\|g(e_{n_k}) - x_{n_k}\| + (1 - \alpha_{n_k})\|z_{n_k} - x_{n_k}\| \\ &\leq \alpha_{n_k}\|g(e_{n_k}) - x^*\| + \alpha_{n_k}\|x_{n_k} - x^*\| + \|z_{n_k} - \bar{w}_{n_k}\| + \|\bar{w}_{n_k} - x_{n_k}\|. \end{aligned}$$

Observe que os dois primeiros termos do lado direito convergem para zero, pois $\{x_n\}$ e $\{g(e_n)\}$ são limitadas e $\alpha_n \rightarrow 0$, enquanto os dois últimos convergem para zero por (3.20) e (3.18). Portanto,

$$\|x_{n_k+1} - x_{n_k}\| \rightarrow 0. \quad (3.21)$$

Como $\{x_{n_k}\}$ é limitada, existe uma subsequência $\{x_{n_{k_j}}\}$ tal que $x_{n_{k_j}} \rightharpoonup q \in H$. De (3.21), segue também $x_{n_{k_j}+1} \rightharpoonup q$. De (3.18), obtemos $\bar{w}_{n_{k_j}} \rightharpoonup q$. Com (3.19), (3.20) e o Lema 3.2.2, conclui-se que $q \in S_{PE}(f, C)$.

Agora, como $x^* = P_{S_{PE}(f, C)}(g(x^*))$, tem-se $\langle g(x^*) - x^*, x - x^* \rangle \leq 0, \forall x \in S_{PE}(f, C)$. Logo,

$$\begin{aligned} \limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k+1} - x^* \rangle &= \lim_{j \rightarrow \infty} \langle g(x^*) - x^*, x_{n_{k_j}+1} - x^* \rangle \\ &= \langle g(x^*) - x^*, q - x^* \rangle \leq 0. \end{aligned} \quad (3.22)$$

Portanto, da definição de ψ_n , da relação (3.22), Condição B e limitação das sequências $\{\|w_{n_k} - x^*\|\}$ e $\{\|\bar{w}_{n_k} - x^*\|\}$ (Lema 3.2.3), temos que

$$\begin{aligned} \limsup_{k \rightarrow \infty} b_{n_k} &= \limsup_{k \rightarrow \infty} \psi_{n_k} \\ &\leq \frac{2}{1-\rho} \limsup_{k \rightarrow \infty} (1 - \alpha_{n_k}) \frac{s_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_k-1}\| \|\bar{w}_{n_k} - x^*\| \\ &\quad + \frac{2\rho}{1-\rho} \limsup_{k \rightarrow \infty} \alpha_{n_k} \frac{u_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_k-1}\| \|w_{n_k} - x^*\| \\ &\quad + \frac{2}{1-\rho} \limsup_{k \rightarrow \infty} \langle g(x^*) - x^*, x_{n_k+1} - x^* \rangle \\ &\leq 0, \end{aligned}$$

isto é, vale (3.17).

E, conseqüentemente, do Lema 1.1.3, $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \|x_n - x^*\|^2 = 0$. □

Capítulo 4

Experimentos Numéricos

Neste capítulo, avaliamos o desempenho computacional do algoritmo proposto por meio de experimentos numéricos em diferentes configurações. Inicialmente, realizamos testes ilustrativos com o objetivo de analisar o impacto de parâmetros como o operador de contração, a sequência de relaxamento $\{\alpha_n\}$ e a dimensão do problema. Em seguida, conduzimos um estudo comparativo entre o Algoritmo 1 e sua extensão com duplo passo inercial (Algoritmo 2), explorando também a influência do parâmetro μ sobre a eficiência dos métodos.

4.1 Ilustrações Numéricas do Algoritmo 1

Nesta seção, apresentamos experimentos numéricos preliminares para avaliar o desempenho do algoritmo proposto. Todos os experimentos computacionais foram conduzidos usando MATLAB (R2017b) em um computador desktop equipado com uma CPU Intel(R) Core(TM) i5-7500 (3,40 GHz) e 15,9 GB de RAM.

É bem conhecido que um ponto $x \in H$ é uma solução para um problema de equilíbrio se, e somente se,

$$x = \arg \min_{y \in C} \left\{ f(x, y) + \frac{1}{2} \|y - x\|^2 \right\},$$

estabelecido em [44]. Com base nessa caracterização, avaliamos a precisão (acurácia) de uma solução candidata, para o problema de equilíbrio, calculando a função

$$A(x) = \left\| x - \arg \min_{y \in C} \left\{ f(x, y) + \frac{1}{2} \|y - x\|^2 \right\} \right\|^2,$$

que satisfaz $A(x) = 0$ se, e somente se, $x \in S_{PE}(f, C)$. Além disso, avaliamos a sequência

$$E_n = \frac{\|x_{n+1} - x_n\|}{\max\{1, \|x_n\|\}},$$

conforme foi usada em [51], que serve como um erro computacional relativo aproximado. Finalmente, também consideramos o critério de parada baseado na medida $\|y_n - w_n\|$.

Problema: Consideramos o problema de equilíbrio associado à bifunção $f : H \times H \rightarrow \mathbb{R}$ derivada do modelo de equilíbrio de Nash-Cournot (veja [43]), dada por

$$f(x, y) = \langle Px + Qy + q, y - x \rangle,$$

onde $q \in \mathbb{R}^m$ é escolhido randomicamente, com seus elementos distribuídos no intervalo $[-m, m]$. As matrizes P e Q são matrizes quadradas de ordem m , onde Q é simétrica semidefinida positiva e $Q - P$ é semidefinida negativa.

Nesta configuração, a bifunção f satisfaz as condições (A1)-(A6). As matrizes P e Q são geradas da seguinte forma: Primeiro, construímos duas matrizes diagonais randômicas A_1 e A_2 de ordem m , onde os autovalores de A_1 estão em $[0, 2]$, e os de A_2 estão em $[-2, 0]$. Em seguida, geramos duas matrizes ortogonais randômicas B_1 e B_2 usando decomposição QR. Elas são então usadas para construir uma matriz semidefinida positiva $M_1 = B_1 A_1 B_1^T$ e uma matriz semidefinida negativa $M_2 = B_2 A_2 B_2^T$. Finalmente, definimos $Q = M_1 + M_1^T$, $S = M_2 + M_2^T$ e $P = Q - S$.

Por outro lado, o conjunto viável $C \subset \mathbb{R}^m$ é definido como

$$C = \{x \in \mathbb{R}^m : Ax \leq b\},$$

onde A é uma matriz $l \times m$, cujas entradas são geradas randomicamente dentro do intervalo $[-2, 2]$, e $b \in \mathbb{R}^l$ é um vetor com elementos escolhidos randomicamente em $[1, 3]$. Sob essas condições, todos os problemas de otimização envolvidos podem ser resolvidos eficientemente usando a função `quadprog` do Toolbox de Otimização do MATLAB.

Em nossos experimentos, fixamos $\lambda_1 = 0.5$ e $\tau_u = 0.3$ (consequentemente, $u_1 \in [0, 1)$), onde $\{u_n\}$ é a sequência escolhida no exemplo da Observação 2.1.2. Além disso, definimos $g : H \rightarrow H$ por $g(x) = \rho \cdot x$, com $0 \leq \rho < 1$, de modo que g seja uma aplicação contração. Os demais parâmetros variam de acordo com cada experimento, conforme descrito a seguir.

Experimento 1. O objetivo deste experimento é investigar a influência do parâmetro de contração ρ no comportamento do algoritmo proposto. Inicializamos com $x_0 = (0, 0, \dots, 0, 0) \in \mathbb{R}^m$ e $x_1 = (1, 1, \dots, 1, 1) \in \mathbb{R}^m$. Consideramos a dimensão do problema $m \in \{5, 10, 20, 100\}$ e adotamos $l = m$. Fixamos $\mu = 1.2$ e $\alpha_n = \frac{1}{10n}$. O parâmetro de contração ρ assume valores do conjunto $\{0.5, 0.8, 0.99, 0.999, 1 - 10^{-6}\}$. Os resultados numéricos relativos ao erro computacional são apresentados nas Figuras 4.1 e 4.2, enquanto as Figuras 4.3 e 4.4 exibem o comportamento da acurácia (precisão da solução de equilíbrio).

Observa-se que o algoritmo apresenta melhor desempenho – em termos do erro computacional relativo E_n – à medida que ρ aumenta, considerando o mesmo número de iterações. Nota-se uma clara separação entre as curvas correspondentes aos diferentes valores de ρ , evidenciando a influência direta desse parâmetro na velocidade de redução do erro. Em particular, para valores de ρ muito próximos de 1, verifica-se uma redução bastante acentuada do erro nas primeiras iterações, seguida de estabilização em níveis significativamente menores.

Quanto à acurácia $A(x_n)$, observa-se comportamento qualitativamente semelhante. Valores maiores de ρ conduzem a reduções mais rápidas da medida de precisão da solução, permitindo atingir níveis significativamente menores de acurácia em menor número de iterações.

Além disso, o padrão de comportamento se mantém consistente para todas as dimensões consideradas, indicando robustez do método em relação ao aumento da dimensão do problema. De modo geral, os resultados confirmam que o parâmetro de contração ρ exerce papel relevante no desempenho computacional do algoritmo proposto.

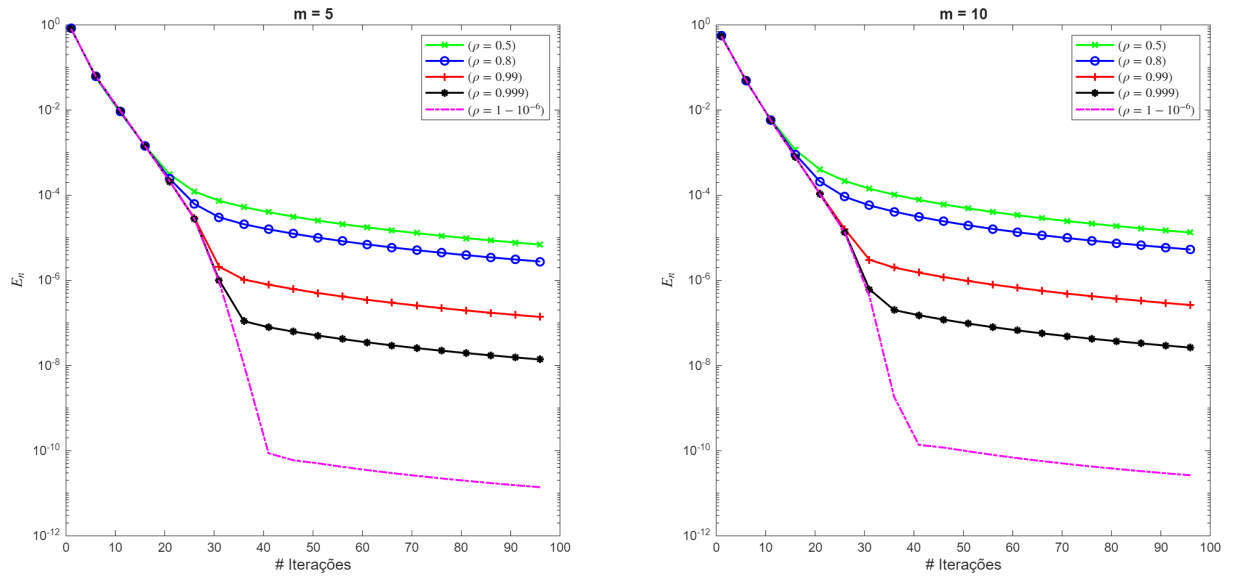


Figura 4.1: E_n em função do número de iterações, para diferentes valores de ρ .

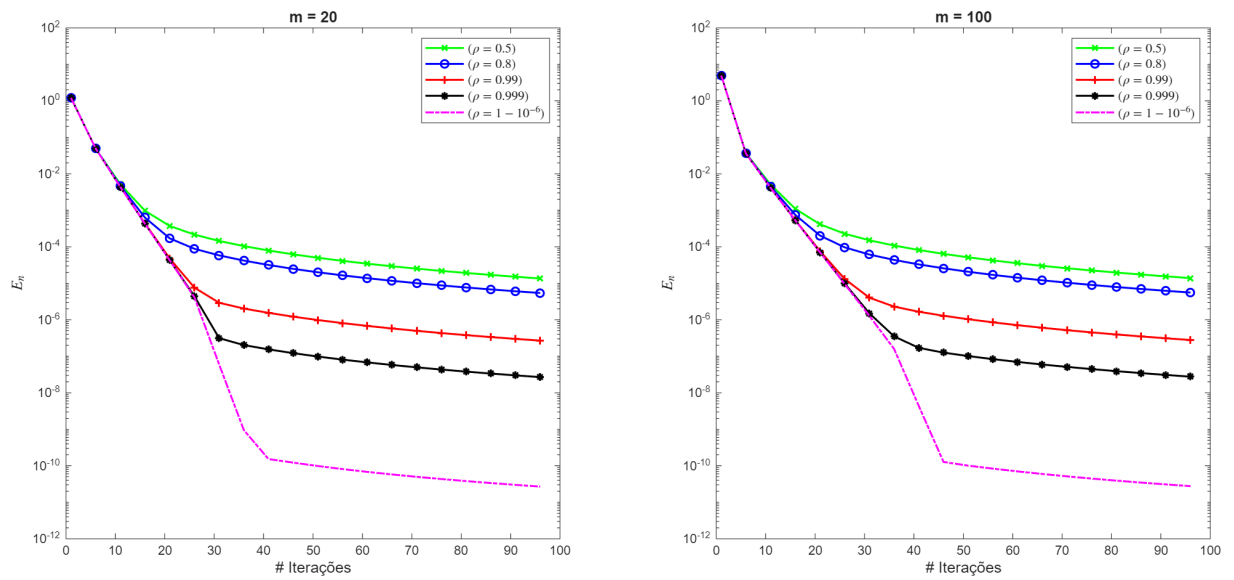


Figura 4.2: E_n em função do número de iterações, para diferentes valores de ρ .

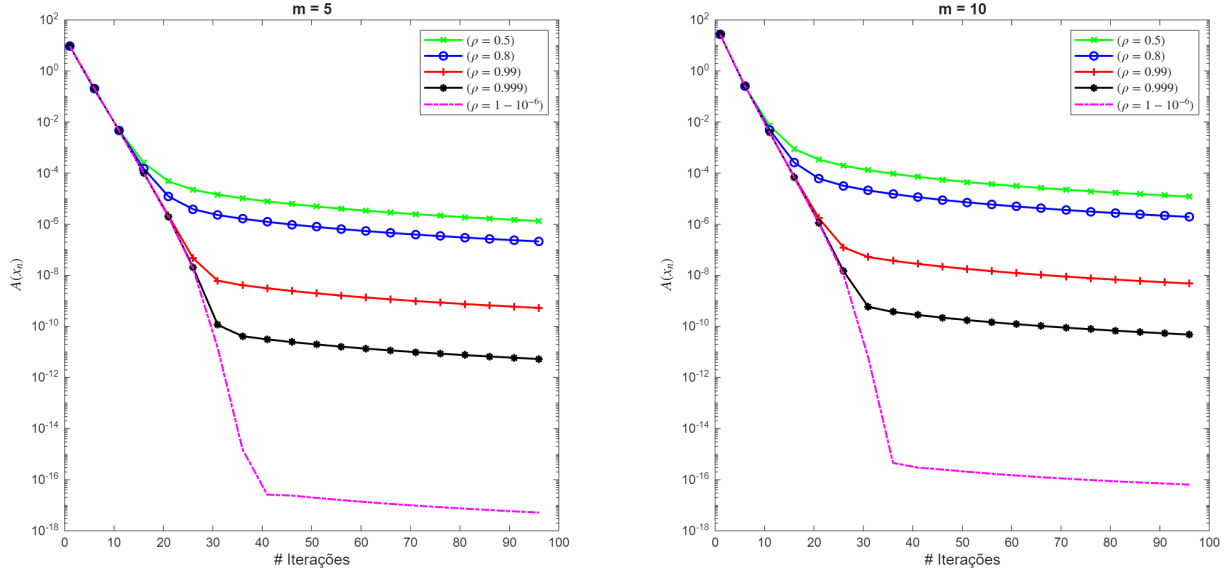


Figura 4.3: $A(x_n)$ em função do número de iterações, para diferentes valores de ρ .

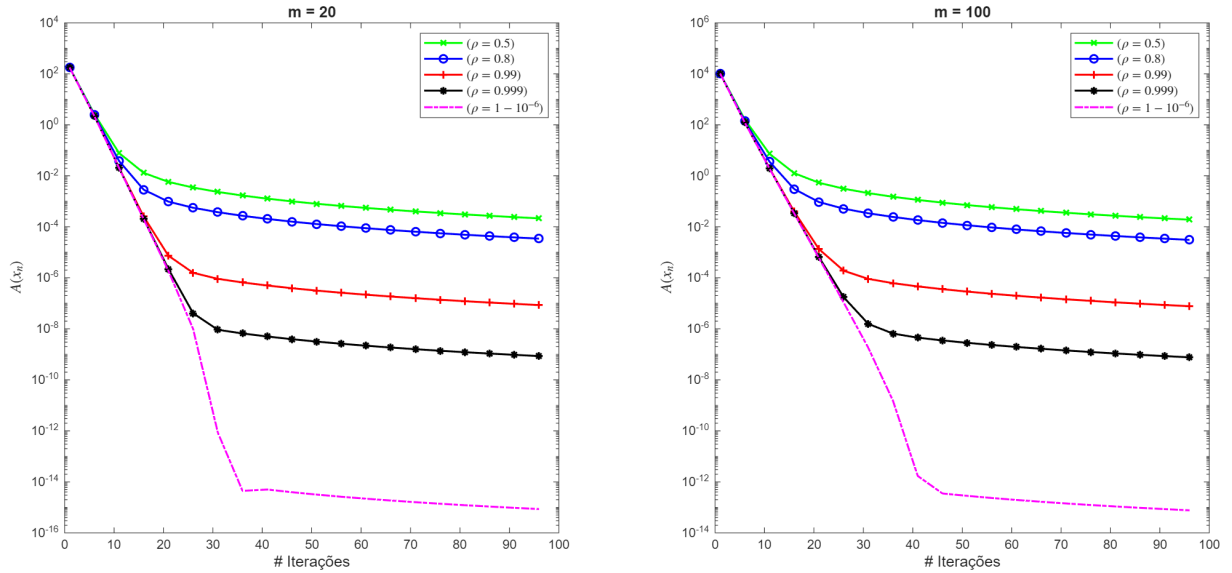


Figura 4.4: $A(x_n)$ em função do número de iterações, para diferentes valores de ρ .

Experimento 2. Este experimento tem como objetivo avaliar o impacto da escolha da sequência (α_n) no desempenho do método proposto. Os valores iniciais são definidos por $x_0 = (0, \dots, 0) \in \mathbb{R}^m$ e $x_1 = (1, \dots, 1) \in \mathbb{R}^m$, considerando $m \in \{5, 10, 20, 100\}$ e adotando $l = m$. Fixamos $\rho = 0.99$ e $\mu = 1.1$. A sequência (α_n) é escolhida no conjunto $\left\{ \frac{1}{10\sqrt{n}}, \frac{1}{10n}, \frac{1}{10^2n}, \frac{1}{10^3n}, \frac{1}{10^4n} \right\}$.

Os resultados numéricos são apresentados nas Figuras 4.5 e 4.6 para o erro computacional relativo E_n , e nas Figuras 4.7 e 4.8 para a acurácia $A(x_n)$.

Observa-se, pelas figuras correspondentes ao erro computacional, que o desempenho do algoritmo é sensivelmente influenciado pela escolha da sequência α_n . Sequências que decrescem mais rapidamente – isto é, com fatores de escala maiores no denominador – tendem a produzir reduções mais acentuadas do erro ao longo das iterações. Nota-se também uma separação clara entre as curvas associadas às diferentes escolhas de α_n , evidenciando o papel desse parâmetro no

comportamento dinâmico do método.

Quanto à acurácia $A(x_n)$, verifica-se tendência semelhante. A escolha de sequências α_n que assumem valores menores ao longo das iterações permite atingir níveis mais elevados de precisão em menor número de passos. Ainda que o ritmo de redução varie conforme a sequência adotada, o método apresenta comportamento estável e consistente em todas as dimensões consideradas.

De modo geral, os resultados indicam que a escolha da sequência α_n exerce influência significativa sobre o desempenho do algoritmo, tanto em termos de erro computacional quanto de precisão da solução, mantendo-se robusto em relação ao aumento da dimensão do problema.

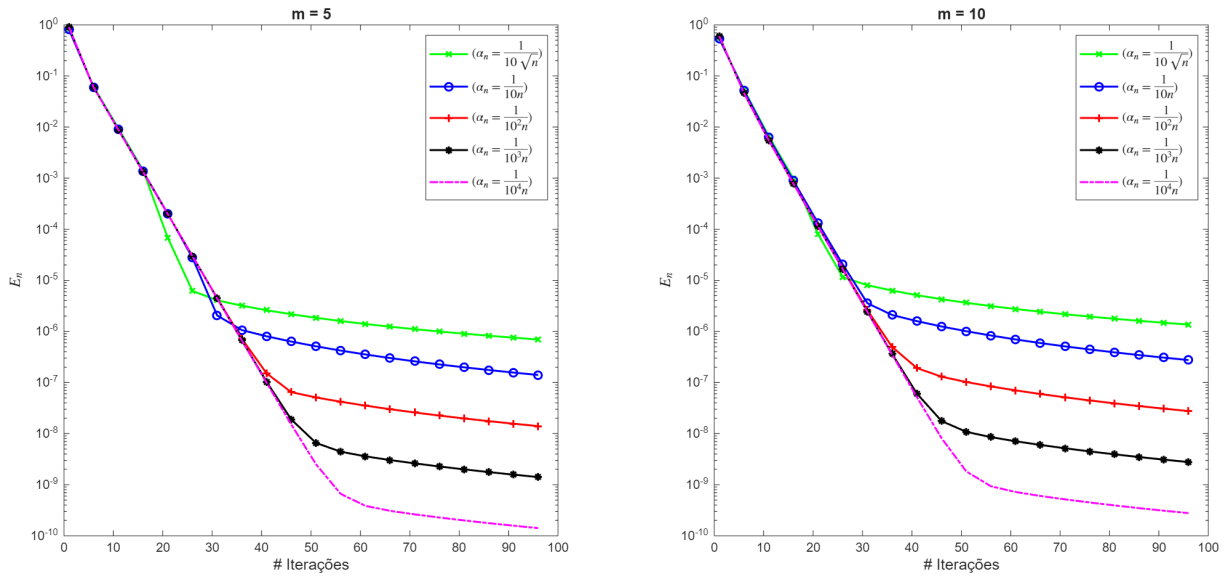


Figura 4.5: E_n em função do número de iterações, para diferentes sequências α_n .

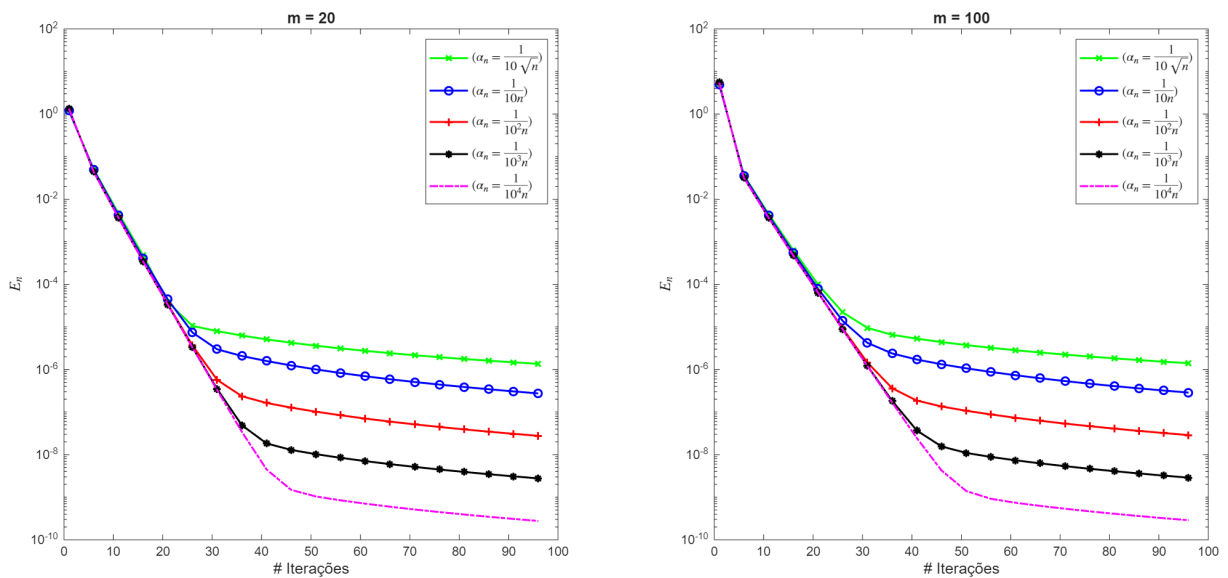


Figura 4.6: E_n em função do número de iterações, para diferentes sequências α_n .

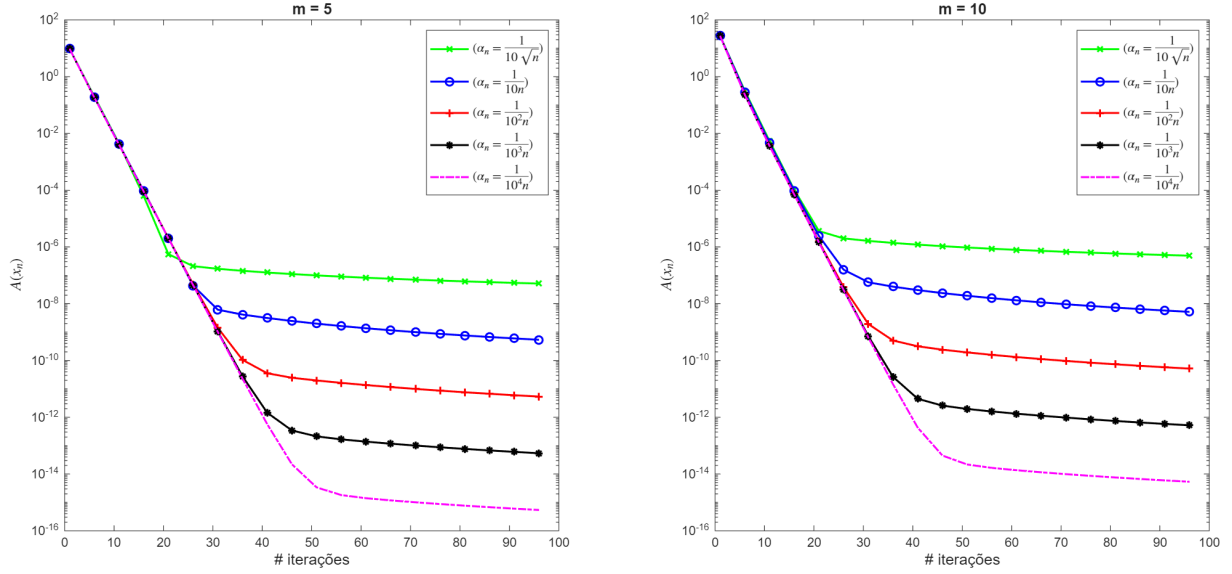


Figura 4.7: $A(x_n)$ em função do número de iterações, para diferentes sequências α_n .

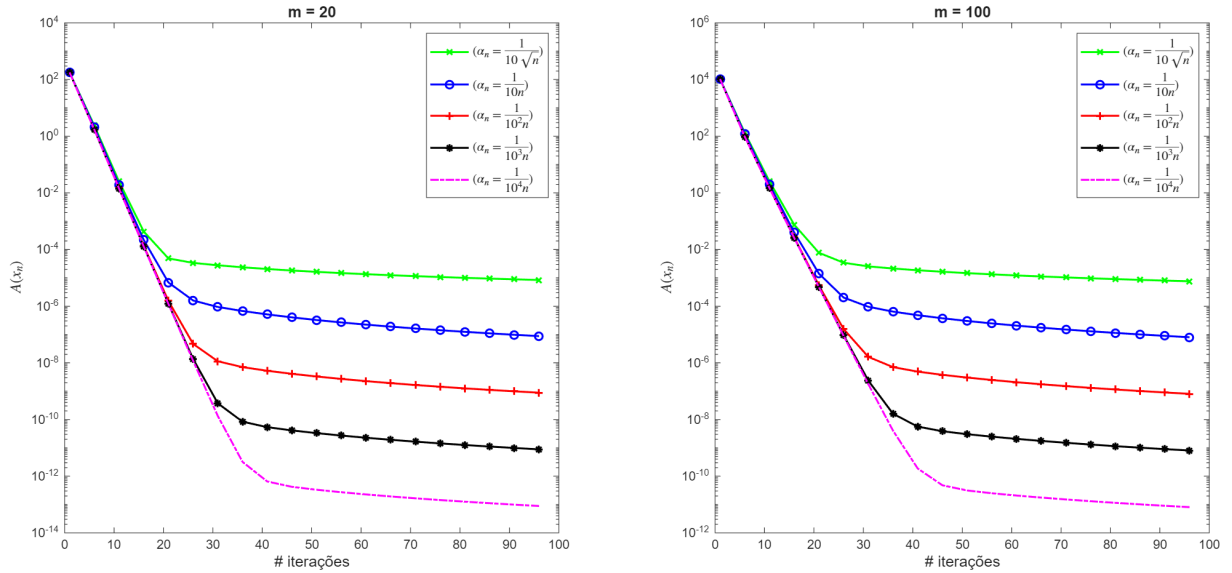


Figura 4.8: $A(x_n)$ em função do número de iterações, para diferentes sequências α_n .

Experimento 3. Este experimento examina o desempenho numérico do algoritmo proposto com relação ao número de iterações e ao tempo de CPU (em segundos) necessários para satisfazer o critério de parada $\|y_n - w_n\| \leq \varepsilon$.

Fixamos $\varepsilon = 10^{-4}$, $\rho = 0.99$, $\mu = 1.1$ e $\alpha_n = \frac{1}{10^3 n}$. O tempo de CPU é medido utilizando o comando `cputime` do MATLAB.

A Tabela 4.1 apresenta os resultados médios obtidos a partir de 100 execuções do algoritmo, com pontos iniciais gerados aleatoriamente na caixa $[0, 1]^m$. Para cada dimensão considerada, a tabela reporta o erro computacional relativo médio (E_n), o número médio de iterações e o tempo médio de CPU necessário para satisfazer o critério de parada.

Observa-se que o número médio de iterações permanece relativamente estável para $m = 5, 10$ e 20 , situando-se em torno de 24 iterações, com leve aumento para $m = 100$. O erro computacional relativo médio apresenta redução progressiva à medida que a dimensão aumenta,

indicando manutenção da precisão mesmo para problemas de maior porte.

Em relação ao tempo de CPU, verifica-se crescimento compatível com o aumento da dimensão do problema. Ainda assim, o algoritmo mantém desempenho computacional satisfatório, alcançando o critério de parada em tempo reduzido para dimensões moderadas.

De modo geral, os resultados confirmam a eficiência prática do método proposto, tanto em termos de número de iterações quanto de custo computacional, evidenciando sua aplicabilidade em diferentes dimensões.

Tabela 4.1: Resultados com 100 pontos iniciais aleatórios na caixa $[0, 1]^m$

m	Erro relativo médio	Iterações médias	CPU médio(s)
5	4.3893×10^{-5}	23.54	0.0273
10	1.9997×10^{-5}	23.97	0.0279
20	5.0495×10^{-6}	24.94	0.0457
100	3.9041×10^{-7}	34	2.0891

4.2 Experimentos Numéricos Comparativos

Nesta seção, conduzimos dois experimentos numéricos com o objetivo de avaliar o desempenho computacional dos Algoritmos 1 e 2.

No primeiro experimento, examinamos o impacto da introdução de um duplo passo inercial sobre a eficiência do processo iterativo, com base na comparação do número médio de iterações, tempo médio de CPU, acurácia da solução aproximada e residual médio no instante de parada.

No segundo experimento, investigamos a influência do parâmetro μ sobre a estabilidade e a eficiência dos algoritmos, observando como diferentes valores desse parâmetro afetam o tempo de execução e o comportamento da convergência, à luz da formulação teórica discutida no Capítulo 2.

4.2.1 Avaliação do impacto do duplo passo inercial

Experimento 4. Neste experimento, comparamos a performance numérica entre o Algoritmo 1 e o Algoritmo 2 com respeito ao número de iterações, tempo de CPU (em segundos), acurácia (ou precisão) da solução aproximada de equilíbrio e o residual no instante de parada. O critério de parada adotado foi o indicado em cada algoritmo: para o Algoritmo 1, $\|y_n - w_n\| \leq \varepsilon$, e para o Algoritmo 2, $\|y_n - \bar{w}_n\| \leq \varepsilon$, onde $\varepsilon = 10^{-4}$.

Os parâmetros utilizados foram: $\lambda_1 = 0.5$, $\tau_u = 0.3$, $\varepsilon = 10^{-4}$, $\rho = 0.99$ (com $g(x) = \rho \cdot x$), $\mu = 1.1$ e $\alpha_n = \frac{1}{10^3 n}$. O tempo de CPU foi medido utilizando o comando `cputime` do MATLAB.

Foram consideradas quatro escolhas distintas para a sequência (s_n) , dando origem às variantes do Algoritmo 2, denotadas, respectivamente, por Algoritmo 2-A, 2-B, 2-C e 2-D. Como o Algoritmo 2 é uma extensão do Algoritmo 1, as variantes diferem apenas na escolha de $(s_n) \subset [0, 1)$, utilizada na construção de $\bar{w}_n$, a qual é escolhida de modo a satisfazer a Condição B. No experimento, fixamos $\tau_s = 0.8$ e adotamos a regra

$$s_n = \min \left\{ \tau_s, \frac{\alpha_n^{1+\delta}}{\|x_n - x_{n-1}\|} \right\} \quad (\|x_n - x_{n-1}\| > 0), \quad s_n = \tau_s \quad (\|x_n - x_{n-1}\| = 0),$$

com $\delta \in \{0.01, 0.02, 0.03, 0.05\}$, correspondendo, respectivamente, às variantes 2-A, 2-B, 2-C e 2-D.

As tabelas a seguir apresentam os resultados médios obtidos a partir da execução dos algoritmos em 100 pontos iniciais gerados aleatoriamente na caixa $[0, 1]^m$, para $m = 10, 20$ e 50 . Os testes foram realizados de forma *pareada*: em cada repetição, o mesmo par inicial (x_0, x_1) foi utilizado em todas as execuções (Algoritmo 1 e Algoritmo 2-A/2-B/2-C/2-D). São exibidas as médias do número de iterações, do tempo de CPU (em segundos), da acurácia $A(x)$ e do residual no instante de parada.

Tabela 4.2: Resultados com 100 pontos aleatórios na caixa $[0, 1]^m$. ($m = 10$)

Algoritmo	Iterações médias	CPU médio (s)	Acurácia	Residual médio de parada
1	24.07	0.033906	1.9607×10^{-8}	8.1107×10^{-5}
2-A	22.67	0.031250	1.4410×10^{-8}	6.8222×10^{-5}
2-B	22.69	0.022500	1.6215×10^{-8}	7.2789×10^{-5}
2-C	22.88	0.022500	1.6711×10^{-8}	7.3521×10^{-5}
2-D	22.99	0.022969	1.6858×10^{-8}	7.4361×10^{-5}

Tabela 4.3: Resultados com 100 pontos aleatórios na caixa $[0, 1]^m$. ($m = 20$)

Algoritmo	Iterações médias	CPU médio (s)	Acurácia	Residual médio de parada
1	24.95	0.030625	1.2851×10^{-8}	7.5873×10^{-5}
2-A	23.62	0.028750	9.9321×10^{-9}	6.2918×10^{-5}
2-B	23.63	0.034844	1.0809×10^{-8}	6.6445×10^{-5}
2-C	23.88	0.031094	8.9039×10^{-9}	6.1228×10^{-5}
2-D	23.96	0.028594	1.0702×10^{-8}	6.8339×10^{-5}

Tabela 4.4: Resultados com 100 pontos aleatórios na caixa $[0, 1]^m$. ($m = 50$)

Algoritmo	Iterações médias	CPU médio (s)	Acurácia	Residual médio de parada
1	37.99	0.59844	5.3288×10^{-8}	8.2943×10^{-5}
2-A	36.81	0.57797	4.6151×10^{-8}	7.6217×10^{-5}
2-B	36.95	0.62469	4.3849×10^{-8}	7.4728×10^{-5}
2-C	36.99	0.59672	4.6572×10^{-8}	7.7206×10^{-5}
2-D	37.00	0.56766	5.4917×10^{-8}	8.3971×10^{-5}

Os resultados apresentados nas Tabelas 4.2–4.4 mostram que a introdução do duplo passo inercial no Algoritmo 2 resultou em melhorias nas métricas computacionais, especialmente em termos de número de iterações e tempo de CPU. Além disso, os algoritmos apresentaram valores comparáveis de acurácia, indicando que a melhoria de desempenho não comprometeu a qualidade da solução aproximada. Observa-se também que os valores do residual médio no instante de parada permaneceram coerentes com $\varepsilon = 10^{-4}$, como esperado pelo critério de parada adotado.

4.2.2 Influência do Parâmetro μ no Desempenho dos Algoritmos

Experimento 5. Neste experimento, avaliamos a influência do parâmetro μ no desempenho computacional dos Algoritmos 1 e 2. Como discutido no Capítulo 2, o parâmetro μ aparece na

regra de atualização do tamanho de passo λ_{n+1} e, com a técnica da proporção áurea, pode ser escolhido no intervalo ampliado $\mu \in (0, 2/\varphi)$, onde $\varphi = (1 + \sqrt{5})/2$. Essa ampliação constitui um diferencial em relação a métodos extragradiente subgradiente clássicos, nos quais usualmente adota-se $\mu \in (0, 1)$, como aqueles discutidos em [27, 51].

O principal objetivo deste experimento é investigar empiricamente se valores de μ próximos ou superiores a 1 mantêm a robustez do método e podem oferecer vantagens em termos de eficiência computacional (iterações e tempo de CPU). Fixamos os parâmetros $\lambda_1 = 0.5$, $\tau_u = 0.3$, $\varepsilon = 10^{-4}$, $\rho = 0.99$ (com $g(x) = \rho x$) e $\alpha_n = \frac{1}{10^3 n}$, e consideramos

$$\mu \in \{0.8, 0.9, 1.0, 1.1, 1.2\}.$$

Mantivemos as mesmas escolhas de sequências e os mesmos critérios de parada adotados no experimento anterior. Em particular, para o Algoritmo 2, fixamos $\tau_s = 0.8$ e $\delta = 0.02$ na construção de (s_n) , o que corresponde à variante 2-B do Experimento 4.

O tempo de CPU foi medido com o comando `cputime` do MATLAB. Os testes foram realizados de forma *pareada*: em cada repetição, ambos os algoritmos foram inicializados com o mesmo par (x_0, x_1) , gerado aleatoriamente na caixa $[0, 1]^m$, totalizando 100 execuções para cada dimensão $m \in \{10, 20, 50\}$.

A Tabela 4.5 apresenta, para cada valor de μ , o número médio de iterações e o tempo médio de CPU (em segundos) dos Algoritmos 1 e 2, destacando o menor tempo médio de CPU e o menor número médio de iterações em cada dimensão.

Tabela 4.5: Comparativo com diferentes valores de μ . Menor tempo médio de CPU em **negrito** e menor número médio de iterações em sublinhado.

		Algoritmo 1		Algoritmo 2	
		med. iter.	med. cpu(s)	med. iter.	med. cpu(s)
$m = 10$	$\mu = 0.8$	26.24	0.024688	24.99	0.023281
	$\mu = 0.9$	25.09	0.019375	23.84	0.017188
	$\mu = 1.0$	24.44	0.018594	23.12	0.017031
	$\mu = 1.1$	24.06	0.018281	22.78	0.017031
	$\mu = 1.2$	<u>23.78</u>	0.016875	<u>22.44</u>	0.018750
$m = 20$	$\mu = 0.8$	26.00	0.024688	25.00	0.024219
	$\mu = 0.9$	25.00	0.024219	24.00	0.020938
	$\mu = 1.0$	24.99	0.022812	23.87	0.022031
	$\mu = 1.1$	<u>24.96</u>	0.024531	<u>23.81</u>	0.018906
	$\mu = 1.2$	25.15	0.023594	23.98	0.020781
$m = 50$	$\mu = 0.8$	42.94	0.44500	41.93	0.43125
	$\mu = 0.9$	40.38	0.39828	39.26	0.39062
	$\mu = 1.0$	38.98	0.45047	37.94	0.44094
	$\mu = 1.1$	37.99	0.58000	36.96	0.55906
	$\mu = 1.2$	<u>37.21</u>	0.55453	<u>36.00</u>	0.52609

De modo geral, observa-se que o desempenho dos algoritmos permanece estável ao longo do conjunto testado, inclusive para valores $\mu \geq 1$. Além disso, em várias situações, valores de μ próximos de 1 ou ligeiramente maiores produziram reduções (ainda que moderadas) no número médio de iterações e/ou no tempo médio de CPU.

Em particular, para $m = 10$, o Algoritmo 1 apresentou o menor tempo médio de CPU e o menor número médio de iterações em $\mu = 1.2$, enquanto o Algoritmo 2 atingiu o menor número médio de iterações também em $\mu = 1.2$, e o menor tempo médio de CPU ocorreu por empate em $\mu \in \{1.0, 1.1\}$. Para $m = 20$, o Algoritmo 2 obteve o menor tempo médio de CPU e o menor número médio de iterações em $\mu = 1.1$, ao passo que, no Algoritmo 1, o menor número médio de iterações também ocorre em $\mu = 1.1$ (com diferenças pequenas em relação aos demais valores), e o menor tempo médio de CPU foi observado em $\mu = 1.0$. Por fim, para $m = 50$, ambos os algoritmos apresentaram os menores tempos médios de CPU em $\mu = 0.9$, enquanto os menores números médios de iterações ocorreram em $\mu = 1.2$.

Em síntese, os resultados sugerem que a escolha de μ no intervalo ampliado permitido pela técnica da proporção áurea não compromete a eficiência computacional e pode, em determinadas instâncias, trazer ganhos moderados em iterações e/ou tempo de CPU. Além disso, em todas as dimensões avaliadas, o Algoritmo 2 apresentou menor número médio de iterações do que o Algoritmo 1, indicando o efeito benéfico da introdução do duplo passo inercial também no cenário de variação de μ .

Capítulo 5

Considerações Finais

Nesta tese, propusemos e analisamos um método extragradiente subgradiente acelerado, de formulação explícita, para a resolução de problemas de equilíbrio pseudomonótonos em espaços de Hilbert reais. O algoritmo desenvolvido baseia-se na técnica da proporção áurea e incorpora estratégias inerciais e de aproximação por viscosidade, com o objetivo de assegurar a convergência forte da sequência gerada, mesmo em ambientes de dimensão infinita. Quando a bifunção de equilíbrio é adicionalmente fortemente pseudomonótona, foi possível estabelecer também uma taxa de convergência R-linear.

Dentre as principais vantagens do método proposto, destacamos a utilização de uma regra explícita para o tamanho de passo λ_n , construída de forma autoadaptativa com base na proporção áurea. Isso permite que o algoritmo seja implementado *sem requerer o conhecimento prévio* das constantes do tipo-Lipschitz associadas à bifunção de equilíbrio — uma exigência frequentemente presente em métodos da literatura, mas que, na prática, representa uma limitação significativa, dado o alto custo de estimar tais constantes.

Além disso, o uso do fator inercial confere ao método uma aceleração no processo iterativo, contribuindo para uma maior eficiência computacional. Essa característica foi estendida na construção do Algoritmo 2, que introduz um segundo passo inercial com o intuito de potencializar as propriedades de convergência do Algoritmo 1.

A robustez da formulação teórica proposta foi corroborada por experimentos numéricos abrangentes. Inicialmente, avaliamos o impacto do parâmetro de contração ρ e da sequência de relaxamento α_n sobre o comportamento do método, demonstrando a consistência e a estabilidade da abordagem sob diferentes regimes. Em seguida, conduzimos experimentos comparativos entre os algoritmos propostos, nos quais foi possível verificar que o Algoritmo 2 apresenta ganhos consistentes, em média, em termos de número de iterações e tempo de CPU, sem comprometer a precisão da solução aproximada. Por fim, exploramos a influência do parâmetro μ associado à atualização de passo, observando que a utilização de valores $\mu \geq 1$ (no intervalo admissível do método) não compromete a estabilidade computacional.

Em síntese, a adoção da técnica da proporção áurea revelou-se decisiva tanto do ponto de vista teórico quanto computacional. Ao possibilitar a ampliação do intervalo admissível para o parâmetro μ e eliminar a necessidade de buscas lineares, essa estratégia contribuiu para um algoritmo mais robusto, flexível e eficiente. Essa combinação de simplicidade estrutural e desempenho reforça o potencial da proporção áurea como um elemento fundamental na construção de métodos extragradiente subgradiente modernos.

Como perspectivas futuras, destacamos algumas direções promissoras:

- (i) Realizar comparações mais amplas com outros métodos extragradiente subgradiente

encontrados na literatura recente;

- (ii) Adaptar o algoritmo para resolver problemas de ponto comum entre o problema de equilíbrio e o problema de ponto fixo de operadores não expansivos;
- (iii) Investigar variantes baseadas na distância de Bregman, as quais podem melhorar a estabilidade do método e permitir uma adaptação mais eficiente à geometria do problema;
- (iv) Propor novos métodos extragradiente subgradiente com passo autoadaptativo baseados na proporção áurea, explorando diferentes formas de relaxamento e aceleração;
- (v) Investigar a extensão do algoritmo para variedades de Hadamard.

Os resultados obtidos nesta tese reforçam o potencial dos métodos extragradiente subgradiente com técnicas modernas de aceleração e autoadaptação, indicando um caminho promissor para a resolução eficiente de problemas de equilíbrio em ambientes generalizados.

Referências Bibliográficas

- [1] Alvarez, F.: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. *SIAM Journal on Optimization*, 14(3):773–782, 2004. 12
- [2] Anh, P.N., Hai, T., Tuan, P.: On ergodic algorithms for equilibrium problems. *Journal of Global Optimization*, 64:179–195, 2016. 8
- [3] Bauschke, H. H., Borwein, J. M.: On projection algorithms for solving convex feasibility problems. *SIAM review*, 38(3):367–426, 1996. 36, 37
- [4] Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2017. 19, 28
- [5] Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202, 2009. 12
- [6] Bigi, G., Castellani, M., Pappalardo, M., Passacantando, M.: *Nonlinear programming techniques for equilibria*. Springer, 2019. 8, 20
- [7] Blum, E., Oettli, W.: From optimization and variational inequalities to equilibrium problems. *Math. student* 63, 123–145, 1994. 8, 20
- [8] Boţ, R. I., Csetnek, E. R.: An inertial Tseng’s type proximal algorithm for nonsmooth and nonconvex optimization problems. *Journal of Optimization Theory and Applications*, 171(2):600–616, 2016. 12
- [9] Boţ, R. I., Csetnek, E. R., Hendrich, C.: Inertial Douglas–Rachford splitting for monotone inclusion problems. *Applied Mathematics and Computation*, 256:472–487, 2015. 12
- [10] Ceng, L. C., Ansari, Q. H., Yao, J. C.: An extragradient method for solving split feasibility and fixed point problems. *Computers & Mathematics with Applications*, 64(4):633–642, 2012. 35
- [11] Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalities in Hilbert space. *Journal of Optimization Theory and Applications* 148(2), 318–335, 2011. 9, 33, 34, 35, 38
- [12] Censor, Y., Gibali, A., Reich, S.: Extensions of Korpelevich’s extragradient method for the variational inequality problem in Euclidean space. *Optimization*, 61(9):1119–1132, 2012. 9, 35

- [13] Combettes, P.L., Hirstoaga, S.A.: Equilibrium programming in Hilbert spaces. *J. Nonlinear Convex Anal* **6**(1), 117-136, 2005. 8
- [14] Dadashi, V., Iyiola, O.S., Shehu, Y.: The subgradient extragradient method for pseudomonotone equilibrium problems. *Optimization* **69**, 901-923, 2020. 8
- [15] Dong, Q. L., He, S., Liu, L.: A general inertial projected gradient method for variational inequality problems. *Computational and Applied Mathematics*, **40**(5):168, 2021. 42
- [16] Facchinei, F., Pang, J.S.: Finite-dimensional variational inequalities and complementarity problems, vol. I and II. Springer, 2003. 8, 28, 29
- [17] Fan, K.: A minimax inequality and applications. *Inequalities* **3**, 103-113, 1972. 8
- [18] Flåm, S.D., Antipin, A.S.: Equilibrium programming using proximal-like algorithms. *Mathematical Programming* **78**(1), 29-41, 1996. 8
- [19] Halpern, B.: Fixed points of nonexpanding maps. 1967. 9
- [20] Hieu, D.V.: New extragradient method for a class of equilibrium problems in Hilbert spaces. *Applicable Analysis*, **97**(5):811–824, 2018. 35
- [21] Hieu, D.V., Cho, Y.J., Xiao, Y.B.: Modified extragradient algorithms for solving equilibrium problems. *Optimization* **67**(11), 2003-2029, 2018. 12, 35
- [22] Hieu, D.V., Strodiot, J.J., Muu, L.D.: Modified golden ratio algorithms for solving equilibrium problems. Preprint at <https://arxiv.org/abs/1907.04013>, 2019. 11, 12, 42
- [23] Hiriart-Urruty, J. B., Lemaréchal, C.: *Convex analysis and minimization algorithms I: Fundamentals*, vol. 305. Springer science & business media, 1993. 15
- [24] Iusem, A., Svaiter, B.: A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization*, **42**(4):309–321, 1997. 33
- [25] Iusem, A.N., Sosa, W.: New existence results for equilibrium problems. *Nonlinear Analysis: Theory, Methods & Applications*, **52**(2):621–635, 2003. 29
- [26] Izmailov, A., Solodov, M.: *Otimização, volume 1: condições de otimalidade, elementos de análise convexa e de dualidade*. Impa, 2020. 14
- [27] Izuchukwu, C., Ogwo, G.N., Zinsou, B.: A new inertial condition on the subgradient extragradient method for solving pseudomonotone equilibrium problem. *Communications in Nonlinear Science and Numerical Simulation*, **135**:108076, 2024. 79
- [28] Kassay, G., Rădulescu, V.: *Equilibrium problems and applications*. Academic Press, 2018. 20
- [29] Khobotov, E.N.: Modification of the extra-gradient method for solving variational inequalities and certain optimization problems. *USSR Computational Mathematics and Mathematical Physics*, **27**(5):120–127, 1987. 33

- [30] Konnov, I.: Combined relaxation methods for variational inequalities, vol. 495. Springer Science & Business Media, 2012. 8
- [31] Konnov, I.: Equilibrium models and variational inequalities. Elsevier, Amsterdam, 2007. 8
- [32] Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. *Matecon* 12, 747-756, 1976. 9, 33, 34
- [33] Kraikaew, R., Saejung, S.: Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces. *Journal of Optimization Theory and Applications*, 163:399–412, 2014. 10
- [34] Long, X.J., Yang, J., Cho, Y.J.: Modified subgradient extragradient algorithms with a new line-search rule for variational inequalities. *Bull. Malays. Math. Sci. Soc.* 46, 140, 2023. 11, 12, 23, 40, 41, 42
- [35] Malitsky, Y.: Golden ratio algorithms for variational inequalities. *Mathematical Programming* 184(1), 383-410, 2020. 10, 42
- [36] Mastroeni, G.: On auxiliary principle for equilibrium problems. In *Equilibrium problems and variational models*, pp. 289–298. Springer, 2003. 22
- [37] Mota, A. M.: Convergência de algoritmos para programação não linear. Tese de Mestrado, Universidade Federal do Paraná, Curitiba, PR, 2005. 20
- [38] Muu, L.D., Oettli, W.: Convergence of an adaptive penalty scheme for finding constrained equilibria. *Nonlinear Anal. TMA* 18(12), 1159-1166, 1992. 8
- [39] Muu, L. D., Quy, N.V.: On existence and solution methods for strongly pseudomonotone equilibrium problems. *Vietnam Journal of Mathematics*, 43(2):229–238, 2015. 8
- [40] Peypouquet, J.: *Convex optimization in normed spaces: theory, methods and examples*. Springer, 2015. 17, 18, 19
- [41] Quoc Tran, D., Le Dung, M., Nguyen, V.H.: Extragradient algorithms extended to equilibrium problems. *Optimization* 57(6), 749-776, 2008. 8, 9, 33, 41
- [42] Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. *Nonlinear Anal. TMA* 75(2), 742-750, 2012. 19
- [43] Santos, P.J.S., Santos, P.S.M., Scheimberg, S.: A Newton-type method for quasi-equilibrium problems and applications. *Optimization*, 71(1), 7-32, 2022. 71
- [44] Santos, P.J.S., Souza, J.C.O.: A proximal point method for quasi-equilibrium problems in Hilbert spaces. *Optimization* 71(1), 55-70, 2022. 71
- [45] Solodov, M.V., Svaiter, B. F.: A new projection method for variational inequality problems. *SIAM Journal on Control and Optimization*, 37(3):765–776, 1999. 33
- [46] Suantai, S., Pholasa, N., Cholamjiak, P.: The modified inertial relaxed CQ algorithm for solving the split feasibility problems. *J. Ind. Manag. Optim.* 14, 1595-1615, 2018. 41

- [47] Takahashi, S., Takahashi, W.: Viscosity approximation methods for equilibrium problems and fixed point problems in Hilbert spaces. *Journal of Mathematical Analysis and Applications* 331(1), 506-515, 2007. 8, 12
- [48] Tan, B., Cho, S.Y., Yao, J.C.: Accelerated inertial subgradient extragradient algorithms with non-monotonic step sizes for equilibrium problems and fixed point problems. *J. Nonlinear Var. Anal.* 6, 89-122, 2022. 59
- [49] Tan, B., Qin, X., Yao, J. C.: Extragradient algorithms for solving equilibrium problems on Hadamard manifolds. *Applied Numerical Mathematics*, 201:187–216, 2024. 10
- [50] Thong, D.V.: Extragradient method with a new adaptive step size for solving non-Lipschitzian pseudo-monotone variational inequalities. *Carpathian Journal of Mathematics*, 38(2):503–516, 2022. 33
- [51] Thong, D.V., Choleamjiak, P., Rassias, M.T., Cho, Y.J.: Strong convergence of inertial subgradient extragradient algorithm for solving pseudomonotone equilibrium problems. *Optimization Letters* 16(2), 545-573, 2022. 12, 19, 42, 59, 71, 79
- [52] Tiel, J. V.: *Convex Analysis: An Introductory Text*. John Wiley & Sons, 1984. 36
- [53] ur Rehman, H., Kumam, P., Cho, Y.J., Yordsorn, P.: Weak convergence of explicit extragradient algorithms for solving equilibrium problems. *J. Inequal. Appl.* 2019, 1-25, 2019. 12
- [54] ur Rehman, H., Kumam, P., Je Cho, Y., Suleiman, Y.I., Kumam, W.: Modified Popov's explicit iterative algorithms for solving pseudomonotone equilibrium problems. *Optimization Methods and Software* 36(1), 82-113, 2021. 20
- [55] Van Hieu, D.: Halpern subgradient extragradient method extended to equilibrium problems. *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* 111, 823-840, 2017. 8, 9
- [56] Van Hieu, D.: Convergence analysis of a new algorithm for strongly pseudomonotone equilibrium problems. *Numerical Algorithms*, 77:983–1001, 2018. 35
- [57] Van Hieu, D.: An inertial-like proximal algorithm for equilibrium problems. *Mathematical Methods of Operations Research*, 88(3):399–415, 2018. 12
- [58] Van Hieu, D., Quy, P.K., Van Vy, L.: Explicit iterative algorithms for solving equilibrium problems. *Calcolo* 56(2), 11, 2019. 12
- [59] Vinh, N.T.: Golden ratio algorithms for solving equilibrium problems in Hilbert spaces. Preprint at <https://arxiv.org/abs/1804.01829>, 2018. 10, 11, 42
- [60] Vinh, N.T., Muu, L.D.: Inertial extragradient algorithms for solving equilibrium problems. *Acta Mathematica Vietnamica* 44(3), 639-663, 2019. 9
- [61] Xu, H. K.: Viscosity approximation methods for nonexpansive mappings. *Journal of Mathematical Analysis and Applications*, 298(1):279–291, 2004. 12
- [62] Zaporozhets, D.N., Zykina, A.V., Melen'chuk, N.V.: Comparative analysis of the extragradient methods for solution of the variational inequalities of some problems. *Automation and Remote Control*, 73:626–636, 2012. 33