

O MÉTODO DO PROBLEMA AUXILIAR COM REGULARIZAÇÃO DE
BREGMAN PARA INEQUAÇÕES VARIACIONAIS

Simone Dutra Ramos

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE DOUTOR EM
CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



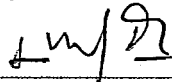
Prof.a. Susana Schemberg. de Makler, D.Sc.



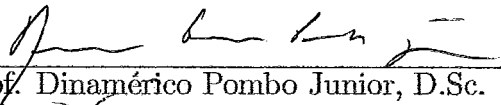
Prof.a. Maria Helena Cautiero H. Jardim, D.Sc.



Prof.a. Regina Sandra Burachick, PhD



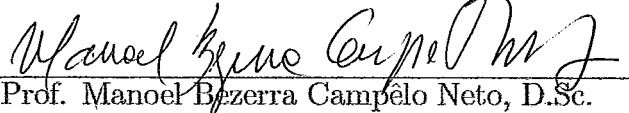
Prof. Luis Mauricio Graña Drumond, PhD



Prof. Dinamérico Pombo Junior, D.Sc.



Prof. Carlos Frederico F.B. e Vasconcellos, D.Sc.



Prof. Manoel Bezerra Campêlo Neto, D.Sc.

RIO DE JANEIRO, RJ - BRASIL
SETEMBRO DE 2001

RAMOS, SIMONE DUTRA

O Método do Problema Auxiliar com regularização de Bregman para inequações variacionais [Rio de Janeiro] 2001

IX, 109 p., 29,7 cm, (COPPE/UFRJ, D.Sc., ENGENHARIA DE SISTEMAS E COMPUTAÇÃO, 2001)

Tese – Universidade Federal do Rio de Janeiro, COPPE

1. Inequações Variacionais
2. Princípio do Problema Auxiliar
3. Regularização de Bregman
4. Método de Decomposição (“Splitting”)
I. COPPE/UFRJ II. Título(Série).

*Aos meus pais Manuel e Norma,
ao meu marido Edézio e
ao meu filho Rodrigo.*

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

THE AUXILIARY PROBLEM METHOD WITH BREGMAN
REGULARIZATION FOR THE VARIATIONAL INEQUALITY PROBLEM

Simone Dutra Ramos

September/2001

Advisor: Susana Scheimberg. de Makler

Department: Systems Engineering and Computer Science

In this work, we develop forward-backward splitting algorithms to solve variational inequality problems in Hilbert spaces. We present a procedure that generalizes the auxiliary principle problem given by Cohen for multivalued monotone operators. For single-valued operators, we introduce one class of algorithms that combines the auxiliary principle problem with the notion of Bregman regularization. We establish welldefinedness of the methods, as well as analyze their convergence. We show that the families of algorithms here introduced unify the existent methods from the theoretical point of view. Finally, we present some applications.

Resumo da Tese apresentada à COPPE como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

O MÉTODO DO PROBLEMA AUXILIAR COM REGULARIZAÇÃO DE BREGMAN PARA INEQUAÇÕES VARIACIONAIS

Simone Dutra Ramos

Setembro/2001

Orientadora: Susana Scheimberg de Makler

Programa: Engenharia de Sistemas e Computação

Neste trabalho, desenvolvemos algoritmos de decomposição do tipo “forward-backward” para resolver o problema de inequação variacional em espaços de Hilbert. Apresentamos um método que generaliza o princípio do problema auxiliar de Cohen para um operador ponto-conjunto. Para um operador ponto-ponto, introduzimos uma classe de algoritmos que combina o princípio do problema auxiliar com a noção de regularização de Bregman. Estabelecemos a boa definição dos métodos apresentados e analisamos as suas propriedades de convergência. Ao longo deste trabalho, mostramos que as famílias de algoritmos introduzidas unificam vários métodos existentes sob o mesmo ponto de vista teórico. Além disso, apresentamos aplicações desses métodos que se adequam as respectivas famílias.

Agradecimentos

Agradeço à Professora Susana Scheimberg de Makler pelo apoio, amizade e compreensão demonstradas ao longo da sua orientação. Não posso deixar de mencionar que, teria sido impossível a obtenção desta conquista sem a ajuda e dedicação da amiga Susana para superar os inúmeros problemas particulares surgidos no decorrer do meu curso de doutorado.

Agradeço ao meu marido, Edézio Pantoja Sacramento, pela compreensão, amor e dedicação com que leu e digitou o manuscrito.

Meus agradecimentos à CAPES pelo apoio financeiro e à Coordenação dos Programas de Pós-Graduação de Engenharia da Universidade Federal do Rio de Janeiro por ter me proporcionado boas condições de trabalho.

Agradeço à Deus, em especial, por ter tornado possível a realização desta conquista.

E também a todos aqueles que de alguma forma participaram deste trabalho, acompanhando, incentivando e apoiando nas horas em que foi necessário.

Conteúdo

1 Introdução	1
2 Preliminares	6
2.1 Operadores monótonos	6
3 O Problema de Inequação Variacional $PIV(\Psi, C, f)$	16
3.1 Introdução	16
3.2 Formulação do (PIV) em termos de uma equação generalizada	18
3.3 Formulações equivalentes	18
3.4 Existência de soluções	21
3.5 Algoritmos existentes	24
3.5.1 O Método do Problema Auxiliar	24
3.5.2 O Algoritmo do Ponto Proximal Generalizado	28
3.5.3 O Método de Perturbação	30
3.5.4 O Método de Decomposição (“Splitting”)	33
4 Generalização do Método do Problema Auxiliar	36
4.1 Introdução	36
4.2 O Problema Auxiliar Estendido	42
4.3 Boa definição da seqüência	42
4.4 Análise de convergência	43
4.5 Aplicações	53
5 A Condição de Dunn e Funções de Bregman	55
5.1 Introdução	55
5.2 A condição de Dunn	55
5.3 Funções de Bregman	57

6 O Método do Problema Auxiliar com regularização de Bregman	65
6.1 Introdução	65
6.2 O Problema Auxiliar com regularização de Bregman	66
6.3 O algoritmo	70
6.4 Boa definição da seqüência	71
6.5 Análise de convergência	72
6.6 Aplicações	83
Conclusões	87
Apêndice: Funções convexas com valores em $\overline{\mathbb{R}}$	88
Bibliografia	94

NOTAÇÕES

Notações gerais

\mathbf{R}_+^n	$\{u \in \mathbf{R}^n / u_j \geq 0 \forall j = 1, \dots, n\}$.
\mathbf{R}_{++}^n	$\{u \in \mathbf{R}^n / u_j > 0 \forall j = 1, \dots, n\}$.
$\overline{\mathbf{R}} = [-\infty, +\infty] = \mathbf{R} \cup \{-\infty, +\infty\}$	reais estendidos.
U	espaço de Hilbert real.
\langle , \rangle	produto interno em U .
$\ \cdot \ $	norma associada ao produto interno \langle , \rangle .
$B(u_0, \delta) = \{u; \ u - u_0\ < \delta\}$	bola aberta, centrada em u_0 , de raio δ .
A°	interior do conjunto A .
$r_i(A)$	interior relativo do conjunto A .
∂A	fronteira do conjunto A .
$clA = \overline{A}$	fecho do conjunto A .
$Co(A)$	envoltória convexa do conjunto A .
\overline{A}^w	fecho fraco do conjunto A .
u_i	convergência fraca.
s_i	convergência forte.
$epi(f)$	epígrafe da função f .
f^*	conjugada da função f .
I_C	indicatriz do conjunto C .
<i>s.c.i.</i>	semicontínua inferiormente.
G -diferencial	Gateaux-diferencial.
$\nabla f(u)$	G -diferencial de f em u .
$\nabla_1 f(u, v)$	G -diferencial de f em relação a u .
$\partial f(u)$	subdiferencial de f em u .
$\partial_1 f(u, v)$	subdiferencial de f em relação a u .
$P(U)$	conjunto das partes de U .
$Dom(\Psi)$	domínio do operador Ψ .
$Im(\Psi)$	imagem do operador Ψ .
$G(\Psi)$	gráfico do operador Ψ .
Ψ^{-1}	operador inverso de Ψ .
P_C	operador projeção sobre o conjunto C .
N_C	operador cone normal do conjunto C .
C^*	cone polar do conjunto C .
$dist(u, C)$	distância de u ao conjunto C .

1 Introdução

Sejam C um subconjunto de um espaço de Hilbert real U fechado, convexo e com o interior não-vazio, $\Psi : U \rightarrow P(U)$ um operador ponto-conjunto monótono e $f : U \rightarrow \overline{\mathbf{R}}$ uma função convexa, própria e semicontínua inferiormente. Nós consideramos o Problema de Inequação Variacional ($PIV(\Psi, C, f)$):

$$(PIV) \quad \text{Encontrar } u^* \in C \text{ e } r^* \in \Psi(u^*) : \langle r^*, u - u^* \rangle + f(u) - f(u^*) \geq 0 \quad \forall u \in C.$$

Esta desigualdade é conhecida como uma *inequação variacional generalizada* (veja [32]). Formulações deste tipo, nas quais Ψ é um operador ponto-conjunto ou ponto-ponto e a função f não é necessariamente identicamente nula, são encontradas em vários campos da matemática tais como programação convexa, equações diferenciais parciais e teoria dos jogos. Esta inequação variacional aparece frequentemente na modelagem de problemas provenientes de diversas áreas da engenharia, por exemplo, economia, transporte e mecânica. Veja, por exemplo, [3], [30], [33], [40] e [68] como referências para inúmeras aplicações do problema (PIV).

Uma grande variedade de problemas podem ser vistos como casos especiais do problema (PIV). Um deles é o caso em que a função f é identicamente nula. Neste caso, o problema (PIV) é reduzido ao *problema de inequação variacional* conhecido como $VIP(\Psi, C)$:

$$(VIP) \quad \text{Encontrar } u^* \in C \text{ e } r^* \in \Psi(u^*) : \langle r^*, u - u^* \rangle \geq 0 \quad \forall u \in C.$$

Tal problema tem sido extensivamente estudado na literatura. Veja, por exemplo, [18] e [33] para o caso ponto-ponto em \mathbf{R}^n . Para o caso ponto-conjunto, veja [14] e [36] em \mathbf{R}^n ; [15] e [16] em espaços de Banach e [13] em espaços de Hilbert.

No caso particular em que C é um cone convexo, não-vazio e fechado de U então o problema (VIP) é equivalente a:

$$(PCG) \quad \text{Encontrar } u^* \in C \text{ e } r^* \in \Psi(u^*) : r^* \in C^* \text{ e } \langle r^*, u^* \rangle = 0,$$

onde C^* denota o cone polar de C , definido por:

$$C^* = \{v \in U / \langle v, u \rangle \geq 0 \quad \forall u \in C\}.$$

Este problema é conhecido como o *problema de complementaridade generalizado*. Foi introduzido, em \mathbf{R}^n , por Karamardian [39] e é amplamente estudado na literatura.

Observemos que (PIV) é uma extensão natural do problema de Programação Convexa. De fato, sob as condições habituais, a saber,

$$(ia) \quad (Dom f)^o \cap C^o \neq \emptyset;$$

$$(ib) \quad Dom f \cap C \subset Dom(\Psi),$$

veremos, na seção 3.1, que se o operador Ψ é o subdiferencial de uma função convexa, própria e semicontínua inferiormente $J : U \rightarrow \bar{\mathbf{R}}$ então o (PIV) é equivalente ao problema de otimização convexa não-diferenciável:

$$(PO) \quad \underset{u \in C}{\text{minimizar}} \quad [J(u) + f(u)].$$

Neste caso, o problema (VIP) fica reduzido a

$$(PO)' \quad \underset{u \in C}{\text{minimizar}} \quad J(u).$$

Quando o operador Ψ não é um operador regular (ou seja, gradiente ou o subdiferencial de uma função convexa) como acabamos de ver, o problema de inequação variacional não se reduz a um problema de minimização. Assim sendo, um grande número de algoritmos foi criado para resolver o problema (PIV) . Enumeramos a seguir, algumas classes de métodos não excluentes conhecidos e amplamente estudados na literatura (seção 3.5).

- Métodos do Princípio Auxiliar

Para o caso ponto-ponto veja, por exemplo, [26] e [60] (em \mathbf{R}^n); [22] (em espaços de Hilbert); [45] (em espaços de Banach). Para o caso ponto-conjunto veja, por exemplo, [22] (em espaços de Hilbert).

- Algoritmos do Ponto Proximal Generalizado

Para o caso ponto-ponto veja, por exemplo, [18] (em \mathbf{R}^n). Para o caso ponto-conjunto veja, por exemplo, [14] e [62] (em \mathbf{R}^n); [13] e [29] (em espaços de Hilbert) e [15] (em espaços de Banach).

- Métodos de Decomposição (“Splitting”)

Para o caso ponto-ponto veja, por exemplo, [21] e [69] (em \mathbf{R}^n). Para o caso ponto-conjunto veja, por exemplo, [29] (em espaços de Hilbert).

- Métodos de Perturbação

Para o caso ponto-ponto veja, por exemplo, [60] (em \mathbf{R}^n) e [45] (em espaços de Banach). Para o caso ponto-conjunto veja, por exemplo, [61] (em Hilbert).

Nesta tese, introduzimos famílias de algoritmos para resolver o problema (*PIV*) que se enquadram nas estruturas dos procedimentos citados acima. Quando o operador é ponto-conjunto, desenvolvemos algoritmos pertencentes à classe de métodos que consideram o princípio do problema auxiliar. No caso ponto-ponto, as estruturas envolvidas são do princípio auxiliar e do ponto proximal generalizado.

O princípio do problema auxiliar foi originalmente introduzido por Cohen [23] e posteriormente utilizado por Cohen e Zhu [24] para resolver problemas de otimização do tipo (*PO*).

Em [22], este princípio é estendido para inequações variacionais. Cohen considera um funcional auxiliar $H : U \rightarrow \mathbf{R}$ fortemente convexo diferenciável e define o subproblema auxiliar (PA^k) a cada iteração k do seguinte modo: Dados $u^k \in C \cap \text{Dom } f \cap \text{Dom } (\Psi)$, $r^k \in \Psi(u^k)$ e $\lambda_k > 0$

(PA^k) Encontrar $u^{k+1} \in C$, solução de

$$\underset{u \in C}{\text{minimizar}} H(u) + \langle \lambda_k r^k - \nabla H^k(u^k), u \rangle + \lambda_k f(u).$$

O algoritmo do ponto proximal cujas propriedades fundamentais foram provadas por Rockafellar [58] foi estudado anteriormente por Martinet [47] para resolver problemas de otimização do tipo (*PO*)'. Burachik e Iusem [13] desenvolvem o método do ponto proximal generalizado. Eles consideram distâncias generalizadas, chamadas distâncias de Bregman, ao invés da distância associada ao produto interno e introduzem a cada iteração k do algoritmo um subproblema auxiliar que dentro do contexto considerado toma a seguinte forma: Dados $u^k \in C^o \cap \text{Dom } (\Psi)$ e $\lambda_k > 0$

($VIPB^k$) Encontrar $u^{k+1} \in C$ e $r^{k+1} \in \Psi(u^{k+1})$:

$$\langle \lambda_k r^{k+1} - \nabla g(u^k), u - u^{k+1} \rangle + g(u) - g(u^{k+1}) \geq 0 \quad \forall u \in C,$$

onde g denota a função de Bregman.

As famílias de algoritmos desenvolvidas neste trabalho são de decomposição do tipo “forward-backward” para resolver o problema (*PIV*) em espaços de Hilbert (veja,

por exemplo, [21]). Este tipo de método é aplicado quando o operador é decomposto na soma de operadores. Com efeito, o problema (PIV) pode ser reformulado como soma de operadores por meio da seguinte equação generalizada:

$$(EG) \quad \text{Encontrar } u^* \in C : 0 \in (\Psi + \partial f + N_C)(u^*).$$

Consideramos ao longo da tese a soma dos operadores Ψ e $(\partial f + N_C)$.

Para o operador ponto-conjunto, estendemos o algoritmo considerado por Cohen [22]. Variamos o funcional auxiliar de Cohen a cada iteração, generalizamos o subproblema auxiliar (PA^k) da seguinte forma: Dados $u^k \in C \cap \text{Dom } f \cap \text{Dom } (\Psi)$, $r^k \in \Psi(u^k)$ e $\lambda_k > 0$

$$(PA^k) \quad \text{Encontrar } u^{k+1} \in C, \text{ solução de}$$

$$\begin{aligned} & \text{minimizar } H^k(u) + \langle \lambda_k r^k - \nabla H^k(u^k), u \rangle + \lambda_k f(u) \\ & u \in C \end{aligned}$$

e analisamos as propriedades de convergência do método.

Para o operador ponto-ponto, introduzimos uma classe de algoritmos que combina o princípio do problema auxiliar com a noção de regularização de Bregman. Com efeito, elaboramos o seguinte subproblema auxiliar com regularização de Bregman: Dados $u^k \in C^o \cap \text{Dom } f \cap \text{Dom } (\Psi)$ e $\lambda_k > 0$

$$(PAB^k) \quad \text{Encontrar } u^{k+1} \in C, \text{ solução de}$$

$$\begin{aligned} & \text{minimizar } \alpha H(u) + \langle \lambda_k \Psi(u^k) - \alpha \nabla H(u^k), u \rangle + \lambda_k f(u) + \beta D_g(u, u^k), \\ & u \in C \end{aligned}$$

onde α e β são constantes positivas.

Para cada classe de algoritmos apresentada, provamos a boa definição, a sua convergência e apresentamos alguns algoritmos já existentes na literatura que se enquadram na família de métodos desenvolvidos nesta tese.

A seguir, descrevemos o conteúdo dos capítulos. No segundo capítulo e no apêndice, enunciamos definições e resultados clássicos da teoria de operadores monótonos e da análise convexa respectivamente, aos quais faremos referência ao longo deste trabalho.

No terceiro capítulo, apresentamos inicialmente o Problema de Inequação Variacional $PIV(\Psi, C; f)$ e provamos a sua equivalência com a equação generalizada (EG) . Algumas condições suficientes para a existência de soluções para (PIV) também são recordadas. Por fim, apresentamos uma síntese de algumas classes de algoritmos conhecidos e extensivamente estudados na literatura.

No quarto capítulo, estendemos o algoritmo do problema auxiliar para um operador ponto-conjunto proposto por Cohen [22]. Consideramos variações do funcional auxiliar introduzido por Cohen a cada iteração. Ampliamos também a família das funções f definidas em (PIV) , permitindo funções definidas em $\mathbf{R} \cup \{+\infty\}$ ao invés de apenas finitas como é suposto inicialmente. Convém ressaltar ainda que obtemos a boa definição do algoritmo sob condições mais fracas que as consideradas por Cohen. Na última seção, apresentamos aplicações que ilustram a nossa extensão.

Os resultados obtidos nesse capítulo foram apresentados e publicados nos anais do IX Congresso Latinoamericano de Investigación Operativa (CLAIO) realizado na Argentina em 1998 e do V Encontro Regional de Matemática Aplicada e Computacional (ERMAC) realizado na PUC - RIO em 1998. Além disso, se encontram também registrados como relatório técnico na COPPE - Sistemas, UFRJ.

O capítulo 5 destina-se à apresentação dos conceitos e propriedades da condição de Dunn e função (e distância) de Bregman que serão usados no sexto capítulo. Na seção 5.3, introduzimos resultados teóricos e apresentamos exemplos relacionados com a noção de função de Bregman.

No sexto capítulo, consideramos o caso de um operador ponto-ponto monótono maximal munido da condição de Dunn no conjunto C . Introduzimos o método do problema auxiliar combinado com a noção de regularização de Bregman. Estabelecemos a boa definição do método e estudamos a sua convergência. Um resultado de convergência fraca é obtido. Sob uma condição extra para função de Bregman, provamos a convergência fraca da seqüência a uma solução do problema. Provamos também a equivalência entre a limitação da seqüência gerada pelo algoritmo e a existência de soluções para (PIV) . Finalmente, apresentamos aplicações que ilustram o método.

2 Preliminares

Este capítulo contém conceitos e resultados clássicos da teoria de operadores monótonos que estão presentes praticamente em todo este trabalho. O material desenvolvido pode ser encontrado principalmente nos livros clássicos [53], [67] e [68] e no trabalho [57].

2.1 Operadores monótonos

Sejam U um espaço de Hilbert real com o produto escalar $\langle \cdot, \cdot \rangle$ e norma associada $\|\cdot\|$. Dado um operador ponto-conjunto $\Psi : U \rightarrow P(U)$, os conjuntos definidos por:

- $Dom(\Psi) = \{u \in U : \Psi(u) \neq \emptyset\}$;
- $Im(\Psi) = \bigcup_{u \in Dom(\Psi)} \Psi(u) = \{r \in \Psi(u) : u \in Dom(\Psi)\}$;
- $G(\Psi) = \{(u, r) \in U \times U : u \in Dom(\Psi), r \in \Psi(u)\}$;

são chamados o *domínio*, a *imagem* e o *gráfico* de Ψ respectivamente. O *operador inverso* de Ψ , $\Psi^{-1} : U \rightarrow P(U)$ é definido por

$$\Psi^{-1}(r) = \{u \in U : r \in \Psi(u)\}.$$

Assim, temos que

$$Dom(\Psi^{-1}) = Im(\Psi).$$

O valor de $r \in U$ aplicado em $u \in U$ é denotado por $\langle r, u \rangle$ e $\Psi(Q) := \bigcup_{u \in Q} \Psi(u)$, onde

Q denota um subconjunto de U .

Para operadores ponto-conjunto dados Ψ_1, Ψ_2 definidos em U e para escalares fixados $\alpha_1, \alpha_2 \in \mathbf{R}$, o operador soma $\alpha_1\Psi_1 + \alpha_2\Psi_2$ é definido por:

$$(\alpha_1\Psi_1 + \alpha_2\Psi_2)(u) = \begin{cases} \alpha_1\Psi_1(u) + \alpha_2\Psi_2(u) & \text{se } u \in Dom(\Psi_1) \cap Dom(\Psi_2), \\ \emptyset & \text{caso contrário.} \end{cases}$$

O conjunto dos operadores em U é ordenado pela inclusão gráfica, isto é,

$$\Psi_1 \subset \Psi_2 \Leftrightarrow \Psi_1(u) \subset \Psi_2(u), \forall u \in U.$$

Definição 2.1.1: Seja $\Psi : U \rightarrow P(U)$.

i) Ψ é monótono se $\forall u, u' \in \text{Dom}(\Psi), \forall r \in \Psi(u), r' \in \Psi(u')$ temos:

$$\langle r - r', u - u' \rangle \geq 0.$$

ii) Ψ é estritamente monótono se $\forall u, u' \in \text{Dom}(\Psi), u \neq u', \forall r \in \Psi(u), r' \in \Psi(u')$ temos:

$$\langle r - r', u - u' \rangle > 0.$$

iii) Ψ é fortemente monótono se existe $\alpha > 0$ tal que

$$\langle r - r', u - u' \rangle \geq \alpha \|u - u'\|^2 \quad \forall u, u' \in \text{Dom}(\Psi), \forall r \in \Psi(u), r' \in \Psi(u').$$

Neste caso, Ψ é dito fortemente monótono módulo α .

É fácil ver que (iii) \Rightarrow (ii) \Rightarrow (i). Além disso, se Ψ é estritamente monótono e $u \neq u'$ então $\Psi(u) \cap \Psi(u') = \emptyset$.

Definição 2.1.2: Seja $\Psi : U \rightarrow P(U)$ um operador monótono. Ψ é chamado *monótono maximal* se para qualquer operador monótono Ψ' tal que $G(\Psi) \subseteq G(\Psi')$ temos que $\Psi(u) = \Psi'(u) \quad \forall u \in \text{Dom}(\Psi)$.

Observação 2.1.1: É fácil ver que Ψ é um operador monótono maximal se, e somente se, Ψ é monótono e

$$\langle r - r', u - u' \rangle \geq 0 \quad \forall u' \in \text{Dom}(\Psi), \forall r' \in \Psi(u') \Rightarrow u \in \text{Dom}(\Psi) \text{ e } r \in \Psi(u).$$

Apresentamos a seguir um resultado que garante a monotonicidade e a monotonicidade maximal do subdiferencial de uma função.

Proposição 2.1.1: Sejam $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$ e $f : U \rightarrow \overline{\mathbf{R}}$ uma função convexa e própria. Então o operador subdiferencial ∂f é um operador monótono. Além disso, se f é semicontínua inferiormente então f é subdiferenciável no $(\text{Dom } f)^o$ e o subdiferencial é monótono maximal.

Demonstração: Veja [67], teorema 7.21 e [68], proposição 32.17.

Exemplo 2.1.1: Seja $f : \mathbf{R}^n \rightarrow (-\infty, +\infty]$ uma função definida por

$$f(u) = \begin{cases} \|u\| & \text{se } \|u\| < 1; \\ +\infty & \text{se } \|u\| \geq 1. \end{cases}$$

Observamos que f é uma função convexa e própria, mas não é semicontínua inferiormente. Além disso o subdiferencial da f é o operador definido por

$$\partial f(u) = \begin{cases} \left\{ \frac{u}{\|u\|} \right\} & \text{se } u \neq 0 \text{ e } \|u\| < 1; \\ \overline{B}(0, 1) & \text{se } u = 0; \\ \emptyset & \text{se } \|u\| \geq 1, \end{cases}$$

onde $\overline{B}(0, 1)$ denota o fecho da bola com centro na origem e raio 1.

É fácil ver que ∂f é um operador monótono e que o seu gráfico está estritamente contido no gráfico do operador monótono $\Psi : \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$ definido da seguinte forma:

$$\Psi(u) = \begin{cases} \left\{ \frac{u}{\|u\|} \right\} & \text{se } u \neq 0 \text{ e } \|u\| \leq 1; \\ \overline{B}(0, 1) & \text{se } u = 0; \\ \emptyset & \text{se } \|u\| > 1, \end{cases}$$

Assim, concluímos que ∂f não é monótono maximal.

Notamos ainda que Ψ também não é monótono maximal já que o seu gráfico está estritamente contido no gráfico do operador monótono $\overline{\Psi} : \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$ definido por

$$\overline{\Psi}(u) = \begin{cases} \Psi(u) & \text{se } \|u\| \neq 1; \\ \{u\} + N_{\overline{B}(0,1)}(u) & \text{se } \|u\| = 1. \end{cases}$$

Observação 2.1.2: A Proposição 2.1.1 não se aplica ao exemplo anterior, já que a hipótese de semicontinuidade inferiormente da função f não é satisfeita.

Exemplo 2.1.2: Seja $\overline{\Psi}$ o operador definido no Exemplo 2.1.1. Como $\overline{\Psi}$ é o subdiferencial da função $g : \mathbf{R}^n \rightarrow (-\infty, +\infty]$ definida por (veja [56], teorema 25.6)

$$g(u) = \begin{cases} \|u\| & \text{se } \|u\| \leq 1; \\ +\infty & \text{se } \|u\| > 1, \end{cases}$$

segue novamente da Proposição 2.1.1 que $\overline{\Psi}$ é um operador nonótono maximal.

Exemplo 2.1.3: O operador *cone normal* de um conjunto convexo, fechado e não-vazio $C \subseteq U$, N_C , é definido por:

$$N_C(u) = \begin{cases} \{w \in U : \langle w, v - u \rangle \leq 0, \forall v \in C\} & , u \in C; \\ \emptyset & , u \notin C. \end{cases}$$

É fácil ver que $N_C \equiv \partial I_C$, onde I_C denota a função indicatriz em C (veja apêndice em anexo, definição 5 e exemplo 1). Então segue da Proposição 2.1.1 que N_C é um operador monótono maximal.

O resultado a seguir mostra que a limitação do domínio de um operador monótono maximal Ψ é condição suficiente para a sua sobrejetividade.

Proposição 2.1.2: Seja $\Psi : U \rightarrow P(U)$ um operador monótono maximal. Se $Dom(\Psi)$ é limitado então Ψ é sobrejetivo, isto é, $Im(\Psi) = U$.

Demonstração: Veja [6], corolário 2.2.

Para dois operadores monótonos maximais Ψ_1 e Ψ_2 , temos que $\Psi_1 + \Psi_2$ é monótono, mas não necessariamente monótono maximal. De fato, podemos ter, por exemplo, que $Dom(\Psi_1) \cap Dom(\Psi_2) = \emptyset$. O resultado seguinte é conhecido e muito útil por estabelecer a maximalidade para soma de operadores monótonos maximais.

Proposição 2.1.3: Sejam $\Psi_1, \Psi_2 : U \rightarrow P(U)$ operadores monótonos maximais. Se $(Dom(\Psi_1))^o \cap Dom(\Psi_2) \neq \emptyset$ então $\Psi_1 + \Psi_2$ é um operador monótono maximal.

Demonstração: Veja [57], teorema 1.

Definição 2.1.3:

i) Uma seqüência $\{u^k\} \subset U$ é *fortemente convergente* para $u \in U$ ($u^k \xrightarrow{s} u$) se, e somente se, $\lim_{k \rightarrow \infty} \|u^k - u\| = 0$;

ii) Uma seqüência $\{u^k\} \subset U$ é *fracamente convergente* para $u \in U$ ($u^k \xrightarrow{w} u$) se, e somente se, $\lim_{k \rightarrow \infty} \langle r, u^k - u \rangle = 0, \forall r \in U$.

A seguinte definição pode ser encontrada em [3].

Definição 2.1.4: Sejam $\beta : U \rightarrow [0, +\infty]$ uma função convexa, própria, semicontínua inferiormente e $\Psi : U \rightarrow P(U)$. Ψ é chamado *β -monótono* se $\forall u, u' \in Dom(\Psi), \forall r \in \Psi(u), r' \in \Psi(u')$ temos:

$$\langle r - r', u - u' \rangle \geq \beta(u - u').$$

É fácil ver que se Ψ é β -monótono então é monótono. Por outro lado, se Ψ é monótono então é β -monótono com $\beta \equiv 0$. Já para Ψ fortemente monótono módulo α temos a β -monotonicidade com $\beta(u) := \alpha \|u\|^2, \forall u \in U$.

Definição 2.1.5: Um operador $\Psi : U \rightarrow P(U)$ é *localmente limitado* em $u \in U$ se existe uma vizinhança V de u tal que o conjunto $\Psi(V)$ é limitado. Além disso, Ψ é *localmente limitado* em um subconjunto $B \subset U$ se Ψ é localmente limitado em cada ponto de B .

Definição 2.1.6: Um operador $\Psi : U \rightarrow P(U)$ é *limitado em conjuntos limitados* se para qualquer subconjunto limitado $Q \subset U$ com o fecho \overline{Q} contido em $(Dom(\Psi))^o$ temos que $\Psi(Q)$ é limitado.

Observação 2.1.3: Se o espaço U é de dimensão finita então um argumento elementar de compacidade mostra que qualquer operador localmente limitado é limitado em conjuntos limitados. De fato, seja u um elemento arbitrário de \overline{Q} . As bolas $B(u, \delta_u)$ cobrem \overline{Q} . Portanto, se M_u é uma cota superior da norma de qualquer elemento de $\Psi(B(u, \delta_u))$ então o máximo dos M_{u_i} correspondentes à uma subcobertura finita é uma cota superior da $\|u\|$ para qualquer $u \in \Psi(Q)$.

A seguir, vemos que a limitação local é uma propriedade básica dos operadores monótonos. Além disso, a Proposição 2.1.5 estabelece novamente a maximalidade para a soma de operadores maximais a partir desta propriedade.

Proposição 2.1.4: Seja $\Psi : U \rightarrow P(U)$ um operador tal que $(Dom(\Psi))^o \neq \emptyset$. Se Ψ é monótono então Ψ é localmente limitado em $(Dom(\Psi))^o$. Se além da monotonicidade de Ψ , U é de dimensão finita então Ψ é limitado em conjuntos limitados.

Demonstração: Veja [53], teorema 2.2 capítulo III.

Proposição 2.1.5: Sejam $\Psi_1, \Psi_2 : U \rightarrow P(U)$ operadores monótonos maximais. Se existe um ponto $u \in cl(Dom(\Psi_1)) \cap cl(Dom(\Psi_2))$ tal que Ψ_1 é localmente limitado em u então $\Psi_1 + \Psi_2$ é monótono maximal.

Demonstração: Veja [57], teorema 1.

Definição 2.1.7: Seja $\Psi : U \rightarrow P(U)$. Ψ é dito um operador *semicontínuo superiormente* em $u \in Dom(\Psi)$ se para cada vizinhança aberta V de $\Psi(u)$, existe uma vizinhança correspondente W de u , tal que $\Psi(W) \subset V$. Ψ é *semicontínuo superiormente* se Ψ é semicontínuo superiormente em cada ponto $u \in Dom(\Psi)$.

Definição 2.1.8: Seja $\Psi : U \rightarrow P(U)$.

i) Ψ é chamado *demicontínuo em* $u \in \text{Dom}(\Psi)$ se Ψ é semicontínuo superiormente em u da topologia forte de U na topologia fraca de U . Ψ é *demicontínuo* se Ψ é demicontínuo em cada ponto $u \in \text{Dom}(\Psi)$.

ii) Ψ é dito *fortemente-fracamente* (resp. *fracamente-fortemente*) *fechado* se segue de $\{u^k\} \subset \text{Dom}(\Psi)$, $u^k \xrightarrow{s} u$ (resp. $u^k \xrightarrow{w} u$), $r^k \in \Psi(u^k)$ e $r^k \xrightarrow{w} r$ (resp. $r^k \xrightarrow{s} r$) que $r \in \Psi(u)$.

Definição 2.1.9: Um conjunto $M \subseteq U \times U$ é chamado *demifechado* se dadas as seqüências $\{u^k\}$, $\{r^k\} \subset U$ verificando uma das seguintes condições:

i) $u^k \xrightarrow{s} u$, $r^k \xrightarrow{w} r$;

ou

ii) $u^k \xrightarrow{w} u$, $r^k \xrightarrow{s} r$;

temos que $(u, r) \in M$.

Os conceitos de operador demicontínuo e conjunto demifechado que acabamos de apresentar são encontrados respectivamente em [68] e [53]. Mas estes mesmos conceitos aparecem em vários livros de análise convexa com possíveis mudanças em suas terminologias, por exemplo, em [3] termos como aplicação finitamente semicontínua superiormente e conjunto fortemente-fracamente (resp. fracamente-fortemente) fechado são empregados para estes conceitos.

Segue imediatamente das Definições 2.1.8 e 2.1.9 que o gráfico do operador Ψ é um conjunto demifechado se, e somente se, Ψ é fortemente-fracamente e fracamente-fortemente.

Para um operador ponto-ponto, usamos a seguinte terminologia:

Definição 2.1.10: Seja $\Psi : U \rightarrow U$.

i) Ψ é dito *contínuo* (resp. *fracamente contínuo*) se para qualquer seqüência $\{u^k\} \subset \text{Dom}(\Psi)$, $u^k \xrightarrow{s} u$ (resp. $u^k \xrightarrow{w} u$) temos que $\Psi(u^k) \xrightarrow{s} \Psi(u)$ (resp. $\Psi(u^k) \xrightarrow{w} \Psi(u)$).

ii) Ψ é dito *fortemente-fracamente* (resp. *fracamente-fortemente*) *contínuo* se $\{u^k\} \subset \text{Dom}(\Psi)$, $u^k \xrightarrow{s} u$ (resp. $u^k \xrightarrow{w} u$) implica que $\Psi(u^k) \xrightarrow{w} \Psi(u)$ (resp. $\Psi(u^k) \xrightarrow{s} \Psi(u)$).

Proposição 2.1.6: Seja $\Psi : U \rightarrow P(U)$ um operador monótono maximal. Então

- i) Ψ é demicontínuo;
- ii) O gráfico de Ψ , $G(\Psi)$, é um conjunto demifechado;
- iii) Os conjuntos $(Dom(\Psi))^o$ e $\overline{Dom(\Psi)}$ são convexos;
- iv) $\Psi(u)$ é um conjunto convexo e fechado (o que equivale à convexo e fracamente fechado) para cada $u \in Dom(\Psi)$.

Demonstração: Veja, por exemplo, [68], problema 32.6.

A definição abaixo pode ser encontrada em [53].

Definição 2.1.11: Sejam $\Psi : U \rightarrow P(U)$ e $f : U \rightarrow (-\infty, +\infty]$ uma função.

- i) Ψ é dito *coercivo em relação a um elemento fixo* $b \in U$ se existe um número $\delta > 0$ tal que

$$\langle r - b, u \rangle > 0 \quad \forall u \in Dom(\Psi), r \in \Psi(u) \text{ com } \|u\| > \delta.$$

- ii) Ψ é dito *coercivo* se

$$\lim_{\|u\| \rightarrow \infty} \frac{\langle r, u \rangle}{\|u\|} = +\infty,$$

onde $u \in Dom(\Psi)$ e $r \in \Psi(u)$.

É fácil ver que se Ψ é coercivo então é coercivo em relação a qualquer elemento $b \in U$ (veja [53], capítulo III, seção 2.8). Além disso, a condição de coercividade é satisfeita sempre que o operador esteja definido em um domínio limitado.

Em [68], podemos encontrar ainda a seguinte definição para o conceito de coercividade de um operador em relação a um elemento fixo:

- iii) Ψ é *coercivo em relação a um elemento fixo* $b \in U$ se existem $u_0 \in Dom(\Psi)$ e $\delta > 0$ tais que

$$\langle r - b, u - u_0 \rangle > 0 \quad \forall u \in Dom(\Psi), r \in \Psi(u) \text{ com } \|u\| > \delta.$$

Note que, quando $0 \in Dom(\Psi)$, a definição (i) implica (iii).

Ao lonto desta tese, usaremos também a noção de coercividade para funções que é análoga a noção correspondente para operadores.

iv) f é dita coerciva se

$$\lim_{\|u\| \rightarrow \infty} \frac{f(u)}{\|u\|} = +\infty,$$

onde $u \in \text{Dom } f$.

Definição 2.1.12: Um operador $\Psi : U \rightarrow P(U)$ é chamado *hemicontínuo em* $u \in \text{Dom}(\Psi)$ se $\text{Dom}(\Psi)$ é convexo e para cada $v \in \text{Dom}(\Psi)$ a aplicação definida por

$$t \mapsto \Psi[(1-t)u + tv],$$

é contínua de $[0, 1]$ na topologia fraca de U .

Definição 2.1.13: Seja G um subconjunto fechado e convexo de U . $\Psi : U \rightarrow P(U)$ é chamado *pseudomonótono de G para $P(U)$* se, e somente se, satisfaz a seguinte condição:

Considere uma seqüência $\{u^k\} \subset G$ convergindo fracamente para um elemento $u^o \in G$, e uma seqüência $\{w^k\} \subset U$, com $w^k \in \Psi(u^k)$, $\forall k$, tal que

$$\limsup_k \langle w^k, u^k - u^o \rangle \leq 0.$$

Então para cada $v \in G$ existe um elemento $w^o \in \Psi(u^o)$, tal que

$$\langle w^o, u^o - v \rangle \leq \liminf_k \langle w^k, u^k - v \rangle.$$

Proposição 2.1.7: Qualquer operador ponto-ponto contínuo em G é pseudomonótono de G para $P(U)$.

Demonstração: Veja [68], capítulo 32, problema 32.3a.

A proposição abaixo estabelece a pseudomonotonia para o operador subdiferencial da função indicatriz em \dot{C} e para a soma de operadores pseudomonótonos.

Proposição 2.1.8:

i) O operador N_C é pseudomonótono para qualquer conjunto $C \subset U$ fechado e convexo.

ii) A soma de operadores pseudomonótonos é um operador pseudomonótono.

Demonstração:

i) Veja [15], proposição 2.5.

ii) Veja [53], capítulo III, proposição 1.2.

A seguinte definição pode ser encontrada em [7] ou [15].

Definição 2.1.14: Um operador $\Psi : U \rightarrow P(U)$ é *regular* se

$$\forall u \in \text{Dom}(\Psi) \text{ e } \forall r \in \text{Im}(\Psi), \quad \sup_{(u', r') \in G(\Psi)} \langle r' - r, u - u' \rangle < \infty.$$

Note que o conceito de operador regular apresentado na definição acima é diferente do conceito mencionado na introdução (isto é, operador subdiferencial de uma função convexa) e usado, salvo especificações, ao longo desta tese.

Proposição 2.1.9: O subdiferencial de uma função convexa, própria e semicontínua inferiormente é um operador regular (no sentido da Definição 2.1.14).

Demonstração: Veja [7], página 167.

O resultado seguinte está provado em [7] para espaços de Hilbert e foi estendido para espaços de Banach em [15].

Proposição 2.1.10: Suponha U um espaço de Hilbert real. Sejam $\Psi_1, \Psi_2 : U \rightarrow P(U)$ operadores monótonos maximais tais que:

- a) Ψ é regular no sentido da Definição 2.1.14;
- b) $\text{Dom}(\Psi_1) \cap \text{Dom}(\Psi_2) \neq \emptyset$ e $\text{Im}(\Psi_1) = U$;
- c) $\Psi_1 + \Psi_2$ é monótono maximal.

Então $\text{Im}(\Psi_1 + \Psi_2) = U$.

Demonstração: Veja [15], lema 2.1.

A noção de paramonotonicidade foi introduzida em [18] e estudada posteriormente em [37]. A seguir, apresentamos a definição de paramonotonicidade em um subconjunto $C \subset U$ convexo e fechado.

Definição 2.1.15: Um operador $\Psi : U \rightarrow P(U)$ é *paramonótono em C* se é monótono em C e

$$\langle r - r', u - u' \rangle = 0 \text{ com } u, u' \in C, r \in \Psi(u), r' \in \Psi(u')$$

implica em $r \in \Psi(u'), r' \in \Psi(u)$.

Se Ψ é ponto-ponto, esta condição significa que

$$\langle \Psi(u) - \Psi(u'), u - u' \rangle = 0 \text{ com } u, u' \in C$$

implica em $\Psi(u) = \Psi(u')$.

De acordo com as definições, temos que monotonicidade estrita em C implica paramonotonicidade em C que, por sua vez, implica monotonicidade em C . É fácil ver que estas três classes são distintas até mesmo no caso das aplicações lineares. Por exemplo, se $\Psi(u, v) = (u - v, u)$, $\forall (u, v) \in \mathbf{R}^2$ temos que Ψ é monótono mas não é paramonótono. Já se $\Psi(u, v) = (u, 0)$, $\forall (u, v) \in \mathbf{R}^2$ temos que Ψ é paramonótono mas não é estritamente monótono.

A proposição abaixo provada em [37] apresenta as principais propriedades dos operadores paramonótonos.

Proposição 2.1.11:

- i)* O subdiferencial de uma função $f : U \rightarrow \overline{\mathbf{R}}$ convexa, própria e semicontínua inferiormente é um operador paramonótono;
- ii)* Se Ψ_1 e Ψ_2 são paramonótonos em C então $\Psi_1 + \Psi_2$ é paramonótono em C ;
- iii)* Suponha $\Psi : U \rightarrow P(U)$ paramonótono em C e pontos $u^* \in C$, $v^* \in \Psi(u^*)$ satisfazendo $\langle v^*, u - u^* \rangle \geq 0 \forall u \in C$. Se existem pontos $\bar{u} \in C$ e $\bar{v} \in \Psi(\bar{u})$ tais que $\langle \bar{v}, u^* - \bar{u} \rangle \geq 0$ então $v^* \in \Psi(\bar{u})$ e $\langle v^*, u - \bar{u} \rangle \geq 0 \forall u \in C$.

3 O Problema de Inequação Variacional $PIV(\Psi, C, f)$

3.1 Introdução

Neste capítulo, apresentamos inicialmente o Problema de Inequação Variacional $PIV(\Psi, C, f)$. Na seção 3.2, provamos a sua equivalência com uma equação generalizada. Na seção 3.3, mostramos que a sua formulação é equivalente, a menos da mudança de espaço, ao problema (VIP) . Algumas condições suficientes para a existência de soluções do (PIV) são dadas na seção 3.4. Por fim, na última seção, faremos uma síntese de algumas classes de algoritmos existentes para resolver (PIV) .

Sejam

- U um espaço de Hilbert real com o produto escalar $\langle \cdot, \cdot \rangle$ e norma associada $\|\cdot\|$;
- $\Psi : U \rightarrow P(U)$ um operador monótono;
- $f : U \rightarrow (-\infty, +\infty]$ uma função convexa, própria e semicontínua inferiormente;
- $C \subset U$ um subconjunto convexo e fechado tal que $C^\circ \neq \emptyset$.

Definimos o Problema de Inequação Variacional $PIV(\Psi, C, f)$ por:

$$(PIV) \text{ Encontrar } u^* \in C \text{ e } r^* \in \Psi(u^*) : \langle r^*, u - u^* \rangle + f(u) - f(u^*) \geq 0 \forall u \in C.$$

Problemas deste tipo, nos quais a função f não é identicamente nula, são encontrados em vários campos da matemática tais como programação convexa, equações diferenciais parciais, e teoria dos jogos. Além disso, aparecem também na modelagem de problemas provenientes de diversas áreas da engenharia, por exemplo, economia, transporte e mecânica. Veja, por exemplo, [3], [30], [33], [40] e [68], como referências para inúmeras aplicações do problema (PIV) .

A seguir, ilustramos um problema de otimização que pode ser reformulado como (PIV) .

Exemplo 3.1.1: Seja $J : U \rightarrow (-\infty, +\infty]$ uma função convexa, própria e semicontínua inferiormente. Suponha que:

i) $Dom J \cap (Dom f)^\circ \cap C^\circ \neq \emptyset$;

ou

ii) $(Dom J)^\circ \cap Dom f \cap C \neq \emptyset$.

O problema de otimização convexa

$$(PO) \quad \underset{u \in C}{\text{Minimizar}} \quad [J(u) + f(u)]$$

é equivalente ao problema $PIV(\partial J, C, f)$. De fato, existe u^* solução de (PO) se, e somente se, $u^* \in C$ e $0 \in \partial(J + f + I_C)(u^*)$ (veja [3], proposição 4), onde I_C é a função indicatriz do conjunto C . Mas pela hipótese (i) ou (ii) temos que (veja teoremas 2(i) e 5, apêndice):

$$\partial(J + f + I_C)(u^*) = \partial J(u^*) + \partial(f + I_C)(u^*).$$

Assim, u^* é solução de (PO)

$$\Leftrightarrow \exists u^* \in C \text{ e } \exists r^* \in \partial J(u^*) : -r^* \in \partial(f + I_C)(u^*)$$

$$\Leftrightarrow \exists u^* \in C \text{ e } \exists r^* \in \partial J(u^*) : \langle r^*, u - u^* \rangle + f(u) - f(u^*) \geq 0 \forall u \in C$$

$$\Leftrightarrow u^* \text{ é solução de } PIV(\partial J, C, f).$$

Obtemos assim uma extensão natural do problema de otimização convexa, (PO) , ao substituirmos o operador ∂J por um operador monótono qualquer. Quando o operador Ψ não é um operador regular (ou seja, gradiente ou um subdiferencial de uma função convexa) como acabamos de ver, o problema de inequação variacional não se reduz a um problema de minimização. Assim, um grande número de algoritmos iterativos foi criado para resolver o problema (PIV) .

As seguintes condições naturais para (PIV) são consideradas ao longo deste trabalho:

$$(i.a) \quad (Dom f)^o \cap C^o \neq \emptyset;$$

$$(i.b) \quad Dom f \cap C \subset Dom(\Psi).$$

Estas suposições são consideradas para a existência de soluções do (PIV) (veja, por exemplo, teoremas 3.4.2 e 3.4.3) e são usadas para a boa definição dos algoritmos apresentados ao longo deste trabalho. Veremos, nos capítulos seguintes, que (ia) e (ib) são usadas também para aproximações de Ψ por operadores regulares.

3.2 Formulação do (PIV) em termos de uma equação generalizada

Ao longo desta tese, usamos também a seguinte formulação generalizada para $PIV(\Psi, C, f)$:

(EG) Encontrar $u^* \in U : 0 \in (\Psi + \partial f + N_C)(u^*)$,

onde N_C denota o cone normal do conjunto C .

Com efeito, sejam u^* uma solução de $PIV(\Psi, C, f)$ e $r^* \in \Psi(u^*)$ verificando a inequação variacional. Então

$$\begin{aligned}
 u^* &\in C \text{ e } \langle r^*, u - u^* \rangle + f(u) - f(u^*) \geq 0 \quad \forall u \in C \\
 &\Leftrightarrow \langle r^*, u - u^* \rangle + (f + I_C)(u) - (f + I_C)(u^*) \geq 0 \quad \forall u \in U \\
 &\Leftrightarrow (f + I_C)(u) \geq (f + I_C)(u^*) + \langle -r^*, u - u^* \rangle \quad \forall u \in U \\
 &\Leftrightarrow -r^* \in \partial(f + I_C)(u^*) \\
 &\Leftrightarrow 0 \in (\Psi + \partial(f + I_C))(u^*) \\
 &\Leftrightarrow 0 \in (\Psi + \partial f + N_C)(u^*),
 \end{aligned}$$

onde a última equivalência segue do Teorema 5 (apêndice).

A equivalência acima mostra que o problema (PIV) , pode ser reformulado em termos de uma equação generalizada, isto é, encontrar um zero de um operador $T = \Psi + \partial f + N_C$. Formulações deste tipo são consideradas, por exemplo, em [21], [43] e [51].

3.3 Formulações equivalentes

Uma grande variedade de problemas podem ser vistos como exemplos especiais do problema (PIV) . Um deles é o caso em que a função f é a função identicamente

nula. Neste caso, o problema (*PIV*) é reduzido ao *problema de inequação variacional* $VIP(\Psi, C)$:

$$(VIP) \quad \text{Encontrar } u^* \in C \text{ e } r^* \in \Psi(u^*) : \langle r^*, u - u^* \rangle \geq 0 \quad \forall u \in C.$$

Este problema vem sendo extensivamente estudado na literatura. Veja, por exemplo, [18] e [33] para o caso ponto-ponto, e [13], [15] e [36] para o caso ponto-conjunto.

Mostramos a seguir que a formulação do problema (*PIV*) é equivalente, a menos da mudança de espaço, à formulação dada ao problema (*VIP*). Com efeito, considere as seguintes definições:

- \tilde{U} é o espaço produto $\tilde{U} := U \times \mathbf{R}$, isto é, um espaço de Hilbert com norma induzida por U e \mathbf{R} , ou seja,

$$\langle \tilde{u}, \tilde{v} \rangle_{\tilde{U}} = \langle u, v \rangle + \mu\sigma \quad \forall \tilde{u} = (u, \mu), \tilde{v} = (v, \sigma) \in U \times \mathbf{R}.$$

- \tilde{C} é o conjunto obtido pela interseção do *epi* (f) (veja apêndice) e o “cilindro” $C \times \mathbf{R}$ em \tilde{U} , isto é,

$$\tilde{C} := \left\{ (u, t) \in \tilde{U} : u \in C \text{ e } f(u) \leq t \right\} = \text{epi}(f) \cap (C \times \mathbf{R});$$

- $\tilde{\Psi} : \tilde{U} \rightarrow P(\tilde{U})$ é o operador dado por:

$$\tilde{\Psi}(u, t) = \Psi(u) \times \{1\}, \quad \forall (u, t) \in \tilde{U}.$$

Então o problema $PIV(\Psi, C, f)$ é equivalente ao $VIP(\tilde{\Psi}, \tilde{C})$. De maneira mais precisa, temos o seguinte resultado.

Teorema 3.3.1 Considere U, Ψ, f e C como na seção 3.1. Se o ponto u^* é uma solução da inequação variacional $PIV(\Psi, C, f)$ então o ponto $\tilde{u}^* = (u^*, f(u^*)) \in \tilde{C}$ é uma solução da inequação $VIP(\tilde{\Psi}, \tilde{C})$. Reciprocamente, se o ponto $\tilde{u}^* = (u^*, \gamma)$ é uma solução de $VIP(\tilde{\Psi}, \tilde{C})$ então u^* é uma solução de $PIV(\Psi, C, f)$ e $\gamma = f(u^*)$.

Demonstração: Suponha u^* uma solução de $PIV(\Psi, C, f)$ com $r^* \in \Psi(u^*)$ verificando a inequação variacional. Seja $\tilde{u} = (u, \lambda) \in \tilde{C}$, isto é, $u \in C \cap \text{Dom } f$ e $\lambda \in \mathbf{R}$

tais que $f(u) \leq \lambda$. Então temos que,

$$\begin{aligned} 0 &\leq \langle r^*, u - u^* \rangle + f(u) - f(u^*) \leq \langle r^*, u - u^* \rangle + 1[\lambda - f(u^*)] \\ &= \langle (r^*, 1), (u, \lambda) - (u^*, f(u^*)) \rangle_{\tilde{U}} = \langle \tilde{r}^*, \tilde{u} - \tilde{u}^* \rangle_{\tilde{U}}, \end{aligned}$$

onde $\tilde{u}^* = (u^*, f(u^*)) \in \tilde{C}$ e $\tilde{r}^* = (r^*, 1) \in \tilde{\Psi}(\tilde{u}^*)$. Daí, \tilde{u}^* é uma solução de $VIP(\tilde{\Psi}, \tilde{C})$.

Para provar a recíproca, considere $\tilde{u}^* = (u^*, \gamma) \in \tilde{C}$ uma solução de $VIP(\tilde{\Psi}, \tilde{C})$ e $\tilde{r}^* \in \tilde{\Psi}(\tilde{u}^*)$ verificando a respectiva inequação. Então $f(u^*) \leq \gamma$. Além disso, para cada $u \in C$ e $\lambda \in \mathbf{R}$ tais que $f(u) \leq \lambda$ temos que $\tilde{u} = (u, \lambda) \in \tilde{C}$. Portanto,

$$0 \leq \langle \tilde{r}^*, \tilde{u} - \tilde{u}^* \rangle = \langle r^*, u - u^* \rangle + \lambda - \gamma,$$

onde $r^* \in \Psi(u^*)$ é tal que $\tilde{r}^* = (r^*, 1)$.

Daí, considerando, em particular, $\tilde{u} = (u^*, f(u^*))$ obtemos $f(u^*) \geq \gamma$. Logo provamos que $\gamma = f(u^*)$. Por outro lado, se $u \in C \cap \text{Dom } f$ e $\lambda = f(u)$ então $(u, f(u)) \in \tilde{C}$ e a inequação acima implica, em particular, que

$$\langle r^*, u - u^* \rangle + f(u) - f(u^*) \geq 0.$$

Como esta inequação é evidente para $u \notin \text{Dom } f$, resulta que u^* é solução de $PIV(\Psi, C, f)$.

Note que $\tilde{\Psi}$ é monótono maximal se Ψ é maximal e \tilde{C} é convexo e fechado (veja apêndice, teorema 1 e proposição 1). Além disso, \tilde{C}^o é não-vazio pela condição (ia) suposta para $PIV(\Psi, C, f)$.

Note que, como o problema $VIP(\Psi, C)$ é, em particular, um problema (PIV), qualquer método iterativo aplicável à resolução do (PIV) também se aplica a de (VIP). A equivalência estabelecida no resultado acima mostra que, a menos da mudança de espaço, a recíproca é verdadeira. Entretanto, sob o ponto de vista numérico, a releitura de um método definido para (VIP) pode dar lugar a um método não interessante para o (PIV). Por exemplo, um procedimento desenvolvido para (VIP) que considere modificações no conjunto de restrições, quando aplicado a \tilde{C} acarreta modificações no $\text{Dom } f$ e no conjunto C e não apenas de f .

No caso ponto-ponto, o resultado é análogo ao que acabamos de apresentar e pode ser encontrado em [53].

Outra formulação de (PIV) em termos de (VIP) é obtida considerando-se uma modificação do operador da inequação variacional. Com efeito, temos a seguinte propriedade:

Teorema 3.3.2: O problema $PIV(\Psi, C, f)$ é equivalente ao problema $VIP(\Psi + \partial f, C)$ sob as condições (ia) e (ib) .

Demonstração: De fato, vimos que u^* é solução do $PIV(\Psi, C, f)$ se, e somente se,

$$0 \in (\Psi + f + N_C)(u^*) = (\Psi + f)(u^*) + N_C(u^*),$$

ou seja, u^* é solução do $VIP(\Psi + \partial f, C)$.

As formulações $PIV(\Psi, C, f)$ e $VIP(\Psi + \partial f, C)$ são “equivalentes”. Mas, enquanto na primeira os algoritmos aproximam o operador Ψ , a função f e o conjunto C em separado, na segunda os algoritmos aproximam o operador soma $\Psi + \partial f$. A opção de trabalharmos com a formulação (PIV) ao invés da (VIP) se deve ao fato de que em termos algorítmicos pode ser mais vantajoso trabalharmos com aproximações para o operador Ψ , para a função f e (ou) o conjunto C separadamente.

3.4 Existência de soluções

Como mostramos na seção 3.2, o problema de inequação variacional (PIV) é equivalente ao problema de encontrar um zero do operador T que é a soma de Ψ , ∂f e N_C . A seguir, enunciamos um resultado básico que garante a existência de um zero para um operador monótono maximal.

Teorema 3.4.1: Considere $T : U \rightarrow P(U)$ um operador monótono maximal. Suponha que existe $\alpha > 0$ tal que

$$\langle r, u \rangle \geq 0, \text{ quando } \|u\| > \alpha \text{ e } (u, r) \in G(T).$$

Então existe um ponto $u^* \in U$ tal que $0 \in T(u^*)$.

Demonstração: Veja [57], proposição 2.

É fácil ver que a condição principal deste teorema é satisfeita, por exemplo, se o domínio efetivo de T é um conjunto limitado ou se T é coercivo em U . Quando T é a soma de três operadores T_1 , T_2 e T_3 , T é coercivo, por exemplo, se $0 \in \text{Dom}(T_1) \cap \text{Dom}(T_2)$ e T_3 é coercivo ou se $\text{Dom}(T_1) \cap \text{Dom}(T_2) \cap \text{Dom}(T_3)$ é limitado.

Lembramos que no nosso caso, $T = \Psi + \partial f + N_C$ e o operador subdiferencial de uma função própria, convexa e s.c.i. é monótono maximal. Portanto, os operadores ∂f e N_C são maximais e a aplicação do Teorema 3.4.1 fornece o seguinte resultado:

Teorema 3.4.2: Suponha que as seguintes hipóteses são satisfeitas:

- (i) C é um subconjunto de U fechado, convexo e com o interior não-vazio;
- (ii) Ψ é monótono em U ;
- (iii) $f : U \rightarrow (-\infty, +\infty]$ uma função convexa, própria e s.c.i.;
- (iv) Uma das três seguintes condições se verifica:
 - Ψ é ponto-ponto e hemicontínuo;
 - Ψ é monótono maximal e $(\text{Dom}(\Psi))^o \cap \text{Dom}(\partial f) \cap C \neq \emptyset$;
 - Ψ é monótono maximal e $\text{Dom}(\Psi) \cap (\text{Dom}(\partial f) \cap C)^o \neq \emptyset$;
- (v) A soma $\Psi + \partial f + N_C$ é coercivo em relação a $b = 0$, isto é, que existem $u_0 \in \text{Dom}(\Psi) \cap \text{Dom}(\partial f) \cap C$ e $\delta > 0$ tais que

$$\langle r, u - u_0 \rangle > 0 \quad \forall (u, r) \in G(\Psi + \partial f + N_C) \text{ com } \|u\| > \delta.$$

Então o problema (PIV) admite uma solução.

Demonstração: Veja [68], proposição 32.36.

Proposição 3.4.1: Se o problema (PIV) tem solução e Ψ é estritamente monótono então a solução é única.

Demonstração: De fato, suponha $u_1^*, u_2^* \in C$ soluções do (PIV). Então

$$\exists r_1^* \in \Psi(u_1^*) : \langle r_1^*, u - u_1^* \rangle + f(u) - f(u_1^*) \geq 0 \quad \forall u \in C$$

e

$$\exists r_2^* \in \Psi(u_2^*) : \langle r_2^*, u - u_2^* \rangle + f(u) - f(u_2^*) \geq 0 \quad \forall u \in C.$$

Considerando $u = u_2^*$ na primeira desigualdade, $u = u_1^*$ na segunda desigualdade e somando-as temos:

$$\langle r_1^* - r_2^*, u_2^* - u_1^* \rangle \geq 0 \Leftrightarrow \langle r_2^* - r_1^*, u_2^* - u_1^* \rangle \leq 0 \Rightarrow u_2^* = u_1^*,$$

onde a última implicação na expressão segue da nonotonicidade estrita de Ψ .

Quando Ψ não é monótono maximal mas é pseudomonótono, obtemos o seguinte resultado de existência para o problema (PIV):

Teorema 3.4.3: Suponha que as seguintes hipóteses são satisfeitas:

- (i) C é um subconjunto de U fechado, convexo e com o interior não-vazio;
- (ii) Ψ é pseudomonótono e limitado em conjuntos limitados;
- (iii) $f : U \rightarrow (-\infty, +\infty]$ é convexa, própria e s.c.i. tal que $(\text{Dom } f \cap C) \subset \text{Dom } (\Psi)$;
- (iv) Para cada $u \in (\text{Dom } f \cap C)$, o conjunto $\Psi(u)$ é um subconjunto convexo, fechado e não-vazio de U ;
- (v) Para cada subconjunto $S \subset (\text{Dom } f \cap C)$, a aplicação $\Psi : \text{Co}(S) \rightarrow P(U)$ é demicontínua;
- (vi) Ψ é $(\partial f + N_C)$ - coercivo em relação a $b = 0$, isto é, existem $\delta > 0$ e $u_0 \in (\text{Dom } f \cap C)$ tais que

$$\langle r, u - u_0 \rangle > 0, \forall (u, r) \in G(\Psi) \text{ com } \|u\| > \delta.$$

Então o problema (PIV) tem solução.

Demonstração: Veja [68], problema 32.4, teorema 32.A.

Outros resultados de existência para o problema (PIV) (ou (VIP)) seguem como aplicações de vários teoremas encontrados na literatura. Por exemplo, o resultado enunciado abaixo é a aplicação do Teorema 12 (página 347) da referência [3].

Teorema 3.4.4: Suponha que as seguintes hipóteses são satisfeitas:

- (i) C é um subconjunto de U fechado, convexo e com o interior não-vazio;
- (ii) Ψ é β -monótono e demicontínuo com valores convexos, fechados e limitados;
- (iii) $f : U \rightarrow (-\infty, +\infty]$ é convexa, própria e s.c.i.;
- (iv) $(\text{Dom } f \cap C) \subset \text{Dom } (\Psi)$;
- (v) $0 \in (\text{Dom } (f + I_C)^* + \Psi(\text{Dom } f \cap C) + \text{Dom } \beta^*)^\circ$,

onde $*$ denota a conjugada da respectiva função.

Então o problema (PIV) tem solução.

Note que a condição (ii) do teorema acima é verificada, por exemplo, quando Ψ é monótono maximal com valores limitados (veja proposição 2.1.6) e fortemente monótono. De forma mais precisa, obtemos os seguintes corolários:

Corolário 3.4.1: Se no Teorema 3.4.4 a hipótese (ii) for substituída por (ii)' Ψ é monótono,

então a condição (v) deve ser substituída por

$$0 \in (Dom (f + I_c)^* + \Psi (Dom f \cap C))^o$$

De fato, Ψ monótono implica Ψ β -monótono com $\beta(u) := 0$ (e assim $\beta^* = I_{\{0\}}$, $Dom \beta^* = \{0\}$).

Corolário 3.4.2: Se no Teorema 3.4.4 a hipótese (ii) for substituída por (ii)'' Ψ é fortemente monótono módulo $\alpha > 0$

então a condição (v) é dispensada e a solução do (PIV) é única.

De fato, Ψ fortemente monótono implica Ψ β -monótono com $\beta(u) := \alpha \|u\|^2$ (e assim $\beta^*(\cdot) = \frac{1}{4\alpha} \|\cdot\|^2$, $Dom \beta^* = U$). Além disso, Ψ é estritamente monótono e conseqüentemente a unicidade da solução permanece.

3.5 Algoritmos existentes

Nesta seção, apresentamos uma síntese de algumas classes não excluentes de algoritmos conhecidos e extensivamente estudados na literatura para o problema (PIV).

3.5.1 O Método do Problema Auxiliar

O princípio do problema foi originalmente introduzido, em \mathbf{R}^n , por Cohen [23] e posteriormente utilizado por Cohen e Zhu [24] para resolver problemas de otimização do tipo (PO). A estrutura gerada por este princípio retrata e analisa algoritmos iterativos tais como métodos do gradiente ou subgradiente bem como métodos de decomposição e coordenação. Em [22], Cohen estende este princípio para resolver inequações variacionais generalizadas. Este conceito vem sendo estendido para operadores auxiliares não regulares por Renaud e Cohen [55]. A idéia, neste último trabalho, é considerar uma seqüência de operadores auxiliares ponto-ponto $\{\Omega^k\}$ supostos fortemente monótonos e Lipschitz, mas não necessariamente regulares, e uma seqüência de números positivos $\{\lambda_k\}$ de forma que o operador Ψ seja aproximado na iteração k por Ω^k . O subproblema auxiliar considerado na iteração k é definido por: Dados $u^k \in C \cap Dom(\Psi)$, $r^k \in \Psi(u^k)$ e $\lambda_k > 0$

($PIVA^k$) Encontrar $u^{k+1} \in C$:

$$\langle \lambda_k r^k + \Omega^k(u^{k+1}) - \Omega^k(u^k), u - u^{k+1} \rangle + \lambda_k (f(u) - f(u^{k+1})) \geq 0 \quad \forall u \in C.$$

O subproblema ($PIVA^k$) é equivalente a seguinte equação generalizada:

(EGA^k) Encontrar $u^{k+1} \in C$:

$$0 \in \lambda_k \Psi(u^k) + [\Omega^k(u^{k+1}) - \Omega^k(u^k)] + \lambda_k \partial f(u^{k+1}) + N_C(u^{k+1}).$$

Hipóteses para Ω^k e f são impostas para garantir a existência e unicidade da solução para cada subproblema.

Na primeira versão do método do problema auxiliar, cada operador auxiliar Ω^k foi escolhido como o gradiente de alguma função real direrenciável e fortemente convexa H (ou seja, de um funcional auxiliar). Neste caso, como mostra o Teorema 4.1.2, o subproblema ($PIVA^k$) é reduzido ao seguinte problema de minimização: Dados $u^k \in C \cap \text{Dom}(\Psi)$, $r^k \in \Psi(u^k)$ e $\lambda_k > 0$

(PA^k) Encontrar $u^{k+1} \in C$, solução de

$$\begin{aligned} & \text{minimizar } H(u) + \langle \lambda_k r^k - \nabla H(u^k), u \rangle + \lambda_k f(u). \\ & u \in C \end{aligned}$$

Note que problemas auxiliares deste tipo, permitem a obtenção de um algoritmo que a cada iteração aproxima o operador Ψ , que não é necessariamente regular, por um operador de regularização.

Além disso, o fato de que Ψ aparece apenas na parte linear $\langle \lambda_k r^k, u \rangle$ do subproblema de minimização (PA^k) tem motivado o uso deste método para a construção de algoritmos de decomposição. Decomposições paralelas podem ser feitas quando U é o produto de N espaços de Hilbert e f é aditiva em relação a esta decomposição. De fato, o problema (PA^k) se divide em N subproblemas de minimização independentes quando o funcional auxiliar H é aditivo em relação a esta estrutura (veja, por exemplo, [45]). A combinação de aproximações deste tipo com o princípio de relaxação gera algoritmos relaxados (veja, por exemplo, [23], página 283) para otimização diferenciável e não diferenciável.

Muitos algoritmos para resolver (PIV) são enquadrados dentro do método do problema auxiliar quando são feitas escolhas apropriadas para a seqüência $\{\Omega^k\}$ (ou H). A variação do operador auxiliar a cada iteração permite a obtenção de um

grau consideravelmente grande de flexibilidade e é fundamental para a aplicação de métodos como, por exemplo, o que descrevemos a seguir:

- Métodos de Aproximação Linear

Considere o problema (VIP) em $U = \mathbf{R}^n$ e a seqüência de operadores auxiliares $\{\Omega^k\}$ dada por

$$\Omega^k(u) = D(u^k)u, \quad \forall k \in \mathbf{N}, \quad \forall u \in \mathbf{R}^n,$$

onde $D(u^k)$ é uma matriz $(n \times n)$ definida positiva. Assim, com $\lambda_k = 1 \quad \forall k$, o subproblema (PIVA^k) se reduz a

(PIVA^k) Encontrar $u^{k+1} \in C$:

$$\langle r^k + D(u^k)(u^{k+1} - u^k), u - u^{k+1} \rangle \geq 0 \quad \forall u \in C.$$

Escolhas adequadas da matriz $D(u^k)$ sugerem diferentes métodos conhecidos tais como:

(i) Métodos da Projeção

Considerando $D(u^k) = B^k$, com $B_{n \times n}^k$ uma matriz simétrica e definida positiva, é fácil ver que o ponto u^{k+1} , solução de (PIVA^k), é exatamente a projeção do ponto $u^k - (B^k)^{-1}r^k$ sobre o conjunto C com respeito a norma $\|\cdot\|_{B^k}$, isto é,

$$u^{k+1} = P_C(u^k - (B^k)^{-1}r^k), \quad r^k \in \Psi(u^k),$$

onde para um dado vetor z , $\|z\|_{B^k} = \sqrt{\langle z, B^k z \rangle}$ e P_C^z é a única solução de

$$\begin{aligned} & \text{minimizar } \|z - u\|_{B^k}^2 \\ & u \in C \end{aligned} \tag{1}$$

Métodos da Projeção para (VIP) são estudados, por exemplo, em [46], [52] e [65]. Quando Ψ é o operador subdiferencial de uma função J e $B^k := I$ (matriz identidade), o subproblema (1) caracteriza o método do subgradiente projetado para resolver o problema de otimização restrito abaixo (ver, por exemplo, [1]):

$$\begin{aligned} (PO)' \quad & \text{minimizar } J(u) \\ & u \in C \end{aligned}$$

(ii) Métodos do tipo Newton

Para Ψ ponto-ponto e continuamente diferenciável, várias escolhas para $D(u^k)$ são possíveis. Por exemplo:

$$\begin{aligned} D(u^k) &= \nabla \Psi(u^k) \quad (\text{Newton}); \\ &\approx \nabla \Psi(u^k) \quad (\text{quase-Newton}); \\ &= \text{sym}(\nabla \Psi(u^k)) \quad (\text{Newton simetrizado}); \\ &= \text{diag}(\nabla \Psi(u^k)) \quad (\text{Jacobi linearizado}), \end{aligned}$$

onde as notações $\text{diag}(D)$ e $\text{sym}(D)$ denotam respectivamente as partes diagonal e simétrica de uma matriz $D = [d_{ij}]_{n \times n}$, isto é, $\text{diag}(D) = \sum_{i=1}^n d_{ii} I_n$ e $\text{sym}(D) = (D + D^T) / 2$.

Obviamente, quando Ψ é o gradiente de uma função J duas vezes diferenciável, nós reavemos os métodos do tipo Newton clássico para resolver o problema de otimização restrito $(PO)'$.

Veja [52] para mais detalhes e referência de outros métodos do tipo Newton.

A convergência do método do problema auxiliar vem sendo estudada de várias formas na literatura. O caso em que o operador Ψ é ponto-conjunto é tratado separadamente do caso em que Ψ é ponto-ponto. Quando Ψ é ponto-conjunto, os operadores auxiliares são geralmente considerados regulares e é necessário que a seqüência de parâmetros $\{\lambda_k\}$ convirja para zero. Em [22], Cohen prova a convergência forte do método em espaços de Hilbert (nos casos ponto-ponto e ponto-conjunto) quando Ψ é suposto fortemente monótono. No capítulo 4, consideramos variações do funcional auxiliar a cada iteração e estendemos este resultado no caso ponto-conjunto. Além disso, ampliamos a família das funções f definidas em (PIV) , permitindo funções definidas em $\mathbf{R} \cup \{+\infty\}$ ao invés de apenas finitas como é suposto por Cohen. No caso em que Ψ é ponto-ponto, a seqüência de parâmetros $\{\lambda_k\}$ é suposta limitada “longe de zero” (ou seja, $\exists \underline{\lambda} > 0 : \lambda_k > \underline{\lambda} \forall k \in \mathbf{N}$) e os resultados de convergência são de dois tipos. Por um lado, se os operadores auxiliares não são necessariamente regulares então o operador Ψ e a seqüência $\{\Omega^k\}$ devem estar relacionados por uma condição de contração (veja, por exemplo, [26] e [52]) ou por uma condição do tipo Dunn (veja, por exemplo, [65] e [71]). Por outro lado, se os operadores auxiliares são escolhidos regulares então Ψ é suposto fortemente monótono

(veja, por exemplo, [22], [45] e [59]) ou satisfazendo a condição (pseudo)Dunn (veja, por exemplo, [59] e [71]).

3.5.2 O Algoritmo do Ponto Proximal Generalizado

Vamos agora voltar a nossa atenção para o caso especial do problema (*PIV*) em que f é identicamente nula. Neste caso, a inequação variacional clássica $VIP(\Psi, C)$ é, por muitas vezes, tratada como a equação generalizada abaixo:

$$(EG)' \quad \text{Encontrar } u^* \in C : 0 \in (\Psi + N_C)(u^*).$$

Um algoritmo muito importante para encontrar zeros de operadores monótonos maximais, ou seja, para resolver $(EG)'$ é o algoritmo do ponto proximal. Este procedimento cujas propriedades fundamentais foram provadas por Rockafellar [58] foi estudado anteriormente por Martinet [47] para resolver problemas de otimização do tipo $(PO)'$. Para encontrar zero de operadores maximais considerou-se inicialmente $\Psi : \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$ um operador monótono maximal no problema abaixo:

$$(EG_0)' \quad \text{Encontrar } u^* \in \mathbf{R}^n : 0 \in \Psi(u^*),$$

O algoritmo do ponto proximal gera, para qualquer ponto inicial u^0 , uma seqüência $\{u^k\}$ da seguinte forma:

$$u^{k+1} \in \mathbf{R}^n : 0 \in \Psi(u^{k+1}) + \lambda_k(u^{k+1} - u^k), \quad (2)$$

onde $\{\lambda_k\}$ é uma seqüência de números positivos. A definição deste algoritmo está baseada no fato (veja em [49]) de que para todo $y \in \mathbf{R}^n$ e $\alpha > 0$, existe um único $x \in \mathbf{R}^n$ tal que

$$y \in (\Psi + \alpha I)(x),$$

onde I é a aplicação identidade.

Rockafellar [58] prova que a seqüência gerada por (2) converge para uma solução de $(EG_0)'$, isto é, para um zero de Ψ sempre que o conjunto de zeros é não-vazio e os parâmetros λ_k são limitados “longe de zero”. O algoritmo do ponto proximal pode ser visto como um método de regularização no qual o parâmetro de regularização não se aproxima de zero necessariamente, evitando assim um possível mal comportamento dos subproblemas regularizados.

O algoritmo do ponto proximal em um espaço de Hilbert U gera uma seqüência $\{u^k\} \subset U$ da seguinte forma:

$$u^0 \in U \tag{3}$$

$$u^{k+1} \in U : 0 \in \Psi(u^{k+1}) + \lambda_k (u^{k+1} - u^k), \tag{4}$$

onde

$$\lambda_k > \underline{\lambda}, \forall k \in \mathbf{N}, \underline{\lambda} > 0.$$

O algoritmo do ponto proximal é usado para resolver problemas irrestritos como o problema $(EG_0)'$. Burachik [12] considera um algoritmo similar que mantém todas as vantagens do algoritmo do ponto proximal, mas é apropriado para o caso “restrito”, isto é, para o problema $(EG)'$ ou $VIP(\Psi, C)$. A idéia central envolvida na estrutura deste algoritmo que Burachik chama de método do ponto proximal generalizado consiste no fato de que no caso “irrestrito”, isto é, do problema $(EG_0)'$, o algoritmo dado por (2) pode ser visto como:

$$u^{k+1} \in \mathbf{R}^n : 0 \in \Psi(u^{k+1}) + \frac{\lambda_k}{2} (\nabla \|\cdot - u^k\|^2)(u^{k+1}).$$

Esta distância é substituída por outra cujo comportamento no conjunto de restrições C seja análogo àquele da norma euclidiana ou da norma induzida pela distância em U . Assim, o algoritmo do ponto proximal generalizado pode ser descrito como:

$$u^{k+1} \in U : 0 \in \Psi(u^{k+1}) + \lambda_k \partial_1 D(u^{k+1}, u^k)$$

onde $\partial_1 D(u, v)$ denota o subdiferencial da distância generalizada em relação ao primeiro argumento.

Convém ressaltar que diferentes escolhas para a distância geram diferentes algoritmos generalizados. Além disso, condições convenientes são impostas a D a fim de assegurar a boa definição e a convergência do algoritmo. No quinto capítulo estudamos, em particular, a distância de Bregman D_g (seção 5.3)

3.5.3 O Método de Perturbação

Na otimização assim como em muitos problema de inequação variacional são usados métodos nos quais a seqüência gerada no algoritmo provêm de aproximações do conjunto admissível, da função objetiva e (ou) do operador. O método de perturbação permite aproximar o problema de inequação variacional por uma seqüência de subproblemas variacionais correspondentes com propriedades computacionais melhores. Tal procedimento é chamado de *convergência variacional*. Uma noção de convergência variacional bastante conhecida na literatura é a de Mosco [50] e se originou de dificuldades encontradas na aproximações de soluções de inequações variacionais. Primeiro, vamos recordar a Mosco-convergência para funções convexas.

Definição 3.5.3.1: Considere U um espaço de Hilbert real. Suponha $\{f^k\}$ uma seqüência de funções próprias, convexas e semicontínuas inferiormente de U em $(-\infty, +\infty]$ e $f : U \rightarrow (-\infty, +\infty]$ uma função própria, convexa e semicontínua inferiormente. A seqüência $\{f^k\}$ é *Mosco-convergente* para f (e denotamos $f^k \xrightarrow{M} f$) se as seguintes condições são satisfeitas:

$$(i) \forall u \in U, \forall u^k \xrightarrow{w} u, \liminf_{k \rightarrow \infty} f^k(u^k) \geq f(u);$$

$$(ii) \forall u \in U, \exists u^k \xrightarrow{s} u \text{ tal que } \limsup_{k \rightarrow \infty} f^k(u^k) \leq f(u).$$

Esta definição se encontra em [2], página 297.

Observação 3.5.3.1: Segue desta definição que se $f^k \xrightarrow{M} f$ então qualquer ponto $u \in U$ é o limite na topologia forte de uma seqüência $\{u^k\}$ tal que

$$\lim_{k \rightarrow \infty} f^k(u^k) = f(u).$$

Observe também que quando $f \leq f^k$ para todo k , a condição (i) é automaticamente satisfeita. Note que no caso em que a dimensão de U é finita, o conceito de Mosco-convergência de funções coincide com o de epi-convergência (veja, por exemplo, [2]).

A proposição seguinte apresenta o conceito de Mosco-convergência para subconjuntos fechados e convexos de U .

Proposição 3.5.3.1: Considere U um espaço de Hilbert real. Sejam $\{C_k\}$, C subconjuntos convexos, fechados e não-vazios de U . As seguintes propriedades são equivalentes:

(a) $I_{C_k} \xrightarrow{M} I_C$;

(b) $\left\{ \begin{array}{l} (b_1) \forall u \in C, \exists \{u^k\} \text{ tal que } u^k \in C_k \text{ para todo } k \in \mathbf{N} \text{ e } u^k \xrightarrow{s} u; \\ (b_2) \forall (k_j), \forall \{u^{k_j}\} \text{ com } u^{k_j} \in C_{k_j} \text{ para todo } j \in \mathbf{N}, \text{ tal que} \\ u^{k_j} \xrightarrow{w} u \text{ implica } u \in C. \end{array} \right.$

Demonstração: Veja [2], proposição 3.21.

Quando a propriedade (a) ou, conseqüentemente, (b) da proposição acima é satisfeita, a seqüência $\{C_k\}$ é *Mosco-convergente* para C (e denotamos $C_k \xrightarrow{M} C$).

A interpretação geométrica da Mosco-convergência de conjuntos convexos é dada pela seguinte proposição:

Proposição 3.5.3.2: Considere U um espaço de Hilbert real. Sejam $\{C_k\}$, C subconjuntos convexos, fechados e não-vazios de U . As seguintes propriedades são equivalentes:

(i) $C_k \xrightarrow{M} C$;

(ii) $\forall u \in U : P_{C_k} u \rightarrow P_C u$;

(iii) $\forall u \in U : dist(u, C_k) \rightarrow dist(u, C)$,

onde $P_{A}v$ é a projeção de v sobre o conjunto A e $dist(v, A)$ denota a distância de v a A .

Demonstração: Veja [2], teorema 3.33.

O uso da Mosco-convergência na área de otimização está baseado na seguinte propriedade: se $f^k \xrightarrow{M} f$ então a seqüência $\{u^k\}$ definida por $u^k \in argmin_{u \in U} f^k(u)$ para cada k , é tal que qualquer ponto de acumulação fraco é uma solução $\bar{u} \in argmin_{u \in U} f(u)$ (veja [2], teorema 1.10). Este resultado foi generalizado para o problema (PIV). Sob o ponto de vista da convergência variacional, o problema (PIV) pode ser perturbado substituindo-se a função original f por uma função aproximada f^k para obter o novo problema:

(PIVP^k) Encontrar $u^k \in C$ e $r^k \in \Psi(u^k)$:

$$\langle r^k, u - u^k \rangle + f^k(u) - f^k(u^k) \geq 0 \quad \forall u \in C$$

As condições que garantem a boa definição da seqüência $\{u^k\}$ e as suas propriedades de convergência são encontrados, por exemplo, em [50], teorema B e [59], teorema 3.1.

Vamos agora ilustrar o conceito de Mosco-convergência com as seguintes proposições:

Proposição 3.5.3.3: Considere U um espaço de Hilbert real. Seja $\{f^k\}$ uma seqüência de funções próprias, convexas e semicontínuas inferiormente de U em $(-\infty, +\infty]$.

(i) Se a seqüência $\{f^k\}$ é monótona crescente então

$$f^k \xrightarrow{M} \sup_{k \in \mathbf{N}} f^k;$$

(ii) Se a seqüência $\{f^k\}$ é monótona decrescente então

$$f^k \xrightarrow{M} cl\left(\inf_{k \in \mathbf{N}} f^k\right).$$

Demonstração: Veja [2], teorema 3.20.

Particularmente, quando f é a função indicatriz de um subconjunto não-vazio, fechado e convexo C de U , o resultado seguinte estabelece a aproximação de C por uma seqüência $\{C_k\}$ de subconjuntos convexos, fechados e não-vazios de U :

Proposição 3.5.3.4: Considere U um espaço de Hilbert real. Sejam $\{C_k\}$, C subconjuntos não-vazios, convexos e fechados de U .

(i) (aproximação interior de C)

Se a seqüência $\{C_k\}$ é crescente, isto é, $C_k \subset C_{k+1}$, $\forall k \in \mathbf{N}$ então

$$C_k \xrightarrow{M} C, \text{ onde } C = cl\left(\bigcup_{k \in \mathbf{N}} C_k\right);$$

(ii) (aproximação exterior de C)

Se a seqüência $\{C_k\}$ é decrescente, isto é, $C_{k+1} \subset C_k, \forall k \in \mathbf{N}$ então

$$C_k \xrightarrow{M} C, \text{ onde } C = \bigcap_{k \in \mathbf{N}} C_k.$$

Demonstração: Veja [2], teorema 3.22.

Uma família de métodos de perturbação para resolver o problema (*PIV*) pode ser obtida pela combinação da teoria da convergência variacional com o princípio do problema auxiliar. Mais precisamente, a cada iteração k , o subproblema auxiliar (PA^k) pode ser perturbado substituindo-se a função original f por uma função aproximada f^k . Obtém-se assim o seguinte subproblema auxiliar perturbado:

(PAP^k) Encontrar $u^{k+1} \in C$:

$$\langle \lambda_k \gamma^k + \Omega^k (u^{k+1}) - \Omega^k (u^k), u - u^{k+1} \rangle + \lambda_k (f^k (u) - f^k (u^{k+1})) \geq 0 \quad \forall u \in C.$$

O uso do método auxiliar perturbado ilustrado acima pode ser encontrado na literatura. Por exemplo, no caso ponto-ponto, Makler *et al.* [45] escolhem cada operador auxiliar Ω^k como o gradiente de alguma função continuamente diferenciável H^k e obtém resultados de convergência para a formulação equivalente $PIV(\Psi, U, f)$.

3.5.4 O Método de Decomposição (“Splitting”)

Métodos de decomposição do tipo “forward-backward” proporcionam uma série de aproximações para resolver problemas de otimização e inequações variacionais nas quais a estrutura de decomposição possa ser utilizada. Lions e Mercier [43] apresentam tais métodos da seguinte forma:

$$0 \in T(u^*) \text{ com } T(u) = T_1(u) + T_2(u),$$

onde $T : \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$. A representação $T = T_1 + T_2$, que pode ser estabelecida de diferentes modos, é chamada uma decomposição de T . Nos métodos de decomposição trabalha-se T_1 e T_2 separadamente. Por exemplo, dado um ponto inicial u^0 , um ponto u^k é gerado a cada iteração k , pela resolução do subproblema abaixo:

$$0 \in (T_{1k} + T_2)(u^{k+1}), \quad T_{1k}(u) = T_1(u_k) + \lambda_k^{-1} M_k [u - u_k] \quad (5)$$

onde $\lambda_k > 0$, para todo k e M_k é uma matriz $(n \times n)$ de implementação.

É fácil ver que as iterações (5) podem ser escritas na forma

$$u^{k+1} \in S_k(u^k) \text{ para } S_k := (M_k + \lambda_k T_2)^{-1} (M_k - \lambda_k T_1).$$

O termo “forward-backward” segue do fato de que quando M_k é não-singular, o operador S_k é equivalente a expressão

$$S_k = (I + \lambda_k M_k^{-1} T_2)^{-1} (I - \lambda_k M_k^{-1} T_1).$$

Na linguagem da análise numérica, $(I - \lambda_k M_k^{-1} T_1)$ realiza um passo à frente (forward) com medida λ_k e direção $d_k = -M_k^{-1} r^k$, $r^k \in T_1(u^k)$ (ou $r^k = T_1(u^k)$ se T_1 é ponto-ponto). Enquanto que o passo para trás (backward) é dado por $(I + \lambda_k M_k^{-1} T_2)^{-1}$. Implementações onde M_k é simétrica e definida positiva são fundamentais, entretanto condições mais fracas são de interesse em alguns casos. Mais detalhes para este algoritmo são encontrados, por exemplo, em [21], [43] e [66].

Particularmente, para o problema (PIV), o método de decomposição do tipo “forward-backward” apresentado por Chen e Rockafellar [21] pode ser aplicado a seguinte representação:

$$T = T_1 + T_2,$$

onde $T_1 = \Psi$ é um operador ponto-ponto e $T_2 = \partial f + N_C$ é ponto-conjunto. De fato, o (PIV) pode ser reformulado como a seguinte equação generalizada (seção 3.2):

$$(EG) \quad \text{Encontrar } u^* \in C : 0 \in \Psi(u^*) + (\partial f + N_C)(u^*).$$

O método do problema auxiliar pode ser visto como um método de decomposição quando estudamos a estrutura especial do problema (PIV) na qual os operadores Ψ e $\partial f + N_C$ são analisados separadamente. Com efeito, o subproblema $(PIVA^k)$ pode ser reescrito sob a forma “forward-backward” onde o passo “forward” dado por Ψ é seguido pelo passo “backward” dado por $\partial f + N_C$, isto é,

$$(PFB^k) \quad \text{Encontrar } u^{k+1} \in C : u^{k+1} \in S_k(u^k), \text{ onde}$$

$$S_k := (\Omega^k + \lambda_k (\partial f + N_C))^{-1} (\Omega^k - \lambda_k \Psi).$$

Outras classes de métodos podem ser obtidas escolhendo-se um operador auxiliar Ω^k que depende de Ψ . Por exemplo, se escolhermos $\Omega^k(u) = u + \lambda_k \Psi(u)$ (com Ψ ponto-ponto) então o algoritmo “forward-backward” é reduzido à um passo “backward” e portanto reavemos o algoritmo do ponto proximal para o operador $\Psi + \partial f + N_C$:

$$u^{k+1} = (I + \lambda_k (\Psi + \partial f + N_C))^{-1} u^k.$$

O algoritmo do ponto proximal pode ser visto como outros tipos de métodos de decomposição. Por exemplo, o método Douglas-Rachford (veja [29]).

Note que cada algoritmo apresentado nesta seção abrange uma grande variedade de métodos que dependem da formulação do problema e do contexto de aplicação.

4 Generalização do Método do Problema Auxiliar

4.1 Introdução

Neste capítulo, estendemos o método do problema auxiliar para um operador ponto-conjunto proposto por Cohen [22]. Na seção 4.2, consideramos variações do funcional auxiliar de Cohen a cada iteração. Ampliamos também a família das funções f definidas em (PIV) permitindo funções definidas em $\mathbf{R} \cup \{+\infty\}$ ao invés de apenas finitas como é suposto inicialmente. Na seção 4.3, estabelecemos a boa definição do algoritmo estendido sob condições mais fracas que as consideradas por Cohen. Na seção 4.4, provamos a convergência forte do algoritmo estendido e por fim, na última seção, apresentamos aplicações que ilustram a nossa extensão.

Em [22], Cohen considera um funcional auxiliar $H : U \rightarrow \mathbf{R}$ fortemente convexo e diferenciável e elabora o seguinte subproblema auxiliar (PA) para o operador ponto-conjunto monótono (Ψ): Dados $v \in C \cap \text{Dom}(\Psi)$, $r \in \Psi(v)$ e $\lambda > 0$

(PA) Encontrar $w = w(v, r) \in C$, solução de

$$\begin{aligned} & \text{minimizar } H(u) + \langle \lambda r - \nabla H(v), u \rangle + \lambda f(u). \\ & u \in C \end{aligned}$$

Problemas auxiliares deste tipo permitem a obtenção de um algoritmo que a cada iteração aproxima o operador ponto-conjunto Ψ que não é necessariamente regular (isto é, gradiente ou um subdiferencial de uma função convexa) por um operador de regularização. De forma mais precisa, temos o seguinte teorema:

Teorema 4.1.1: O problema auxiliar (PA) é equivalente ao problema de inequação variacional PIV (Ψ_H, C, f) onde $\Psi_H : U \rightarrow U$ é o operador regular definido por

$$\Psi_H(u) := r + \lambda^{-1}(\nabla H(u) - \nabla H(v)), \quad \forall u \in U.$$

Demonstração: De fato, considerando a função própria, convexa e s.c.i. $K : U \rightarrow (-\infty, +\infty]$ definida por

$$K(u) := H(u) + \langle \lambda r - \nabla H(v), u \rangle + \lambda f(u) \quad \forall u \in U,$$

notamos que resolver (PA) equivale a

$$\text{Encontrar } w = w(v, r) \in C : 0 \in \partial(K + I_C)(w). \quad (6)$$

Além disso, como $Dom H = U$ e $(Dom f)^o \cap C^o \neq \emptyset$, temos que,

$$Dom K \cap Dom I_C = Dom f \cap C \neq \emptyset.$$

Usando a continuidade de I_C em C^o , resulta que (veja teorema 5, apêndice),

$$\partial(K + I_C)(w) = \partial K(w) + \partial I_C(w) = \partial K(w) + N_C(w). \quad (7)$$

De forma análoga, usando agora a continuidade de H em C , temos ainda,

$$\partial K(w) = \partial(H + \langle \lambda r - \nabla H(v), \cdot \rangle)(w) + \lambda \partial f(w),$$

dado que $\lambda > 0$.

Como $Dom H \cap Dom(\langle \lambda r - \nabla H(v), \cdot \rangle) = U \neq \emptyset$ e as aplicações H e $\langle \lambda r - \nabla H(v), \cdot \rangle$ são diferenciáveis, ficamos com

$$\partial K(w) = \nabla H(w) + \lambda r - \nabla H(v) + \lambda \partial f(w). \quad (8)$$

Substituindo (7) e (8) na equação generalizada (6), obtemos:

$$\text{Encontrar } w = w(v, r) \in C : 0 \in \nabla H(w) + \lambda r - \nabla H(v) + \lambda \partial f(w) + N_C(w),$$

isto é,

$$\text{Encontrar } w = w(v, r) \in C : \nabla H(v) - \nabla H(w) - \lambda r \in \lambda \partial f(w) + N_C(w). \quad (9)$$

Mas,

$$\lambda \partial f(w) + N_C(w) = \partial(\lambda f + I_C)(w), \quad (10)$$

pois $(Dom f)^o \cap C^o \neq \emptyset$, I_C é contínua em C^o e $\lambda > 0$.

Assim a equação (6) pode ser reescrita como abaixo:

$$\text{Encontrar } w = w(v, r) \in C : \lambda f(u) \geq \lambda f(w) + \langle \nabla H(v) - \nabla H(w) - \lambda r, u - w \rangle \quad \forall u \in C,$$

isto é,

$$\text{Encontrar } w = w(v, r) \in C :$$

$$\langle \lambda r + \nabla H(w) - \nabla H(v), u - w \rangle + \lambda (f(u) - f(w)) \geq 0 \quad \forall u \in C,$$

o que equivale a

$$\text{Encontrar } w = w(v, r) \in C : \langle \Psi_H(w), u - w \rangle + f(u) - f(w) \geq 0 \quad \forall u \in C.$$

A equivalência acima conclui a prova.

Observação 4.1.1: O operador Ψ_H é monótono maximal, já que é o gradiente de uma função convexa, própria e s.c.i. (veja proposição 2.1.1).

Consideramos agora $H : U \rightarrow (-\infty, +\infty]$ uma função convexa, própria, G -diferenciável e contínua em $C \subset (\text{Dom } H)^o$ (isto é, um “funcional auxiliar”). Dados $v \in C \cap \text{Dom}(\Psi)$, $r \in \Psi(v)$ e $\lambda > 0$, ilustramos novamente o subproblema auxiliar:

(PA) Encontrar $w = w(v, r) \in C$, solução de

$$\begin{aligned} & \text{minimizar } H(u) + \langle \lambda r - \nabla H(v), u \rangle + \lambda f(u) . \\ & u \in C \end{aligned}$$

Apresentamos a seguir caracterizações para a solução do subproblema auxiliar acima.

Teorema 4.1.2: Dados os problemas PIV (Ψ, C, f) e (PA), as seguintes afirmações são equivalentes:

- (i) $w = w(v, r)$ é solução de (PA);
- (ii) $w \in C$ e $\langle \nabla H(w) + \lambda r - \nabla H(v), u - w \rangle + \lambda (f(u) - f(w)) \geq 0 \quad \forall u \in C$.

Demonstração: Suponha w uma solução de (PA). Isto significa que $w \in C$ e

$$\begin{aligned} H(w) + \langle \lambda r - \nabla H(v), w \rangle + \lambda f(w) + I_C(w) &\leq H(u) + \langle \lambda r - \nabla H(v), u \rangle \\ &+ \lambda f(u) + I_C(u) \quad \forall u \in U, \end{aligned}$$

isto é,

$$\langle \lambda r - \nabla H(v), u - w \rangle + (\lambda f + H + I_C)(u) - (\lambda f + H + I_C)(w) \geq 0 \quad \forall u \in U,$$

ou seja,

$$-\lambda r + \nabla H(v) \in \partial(H + \lambda f + I_C)(w).$$

Como $(\text{Dom } H) \cap (\text{Dom } f \cap C) = \text{Dom } f \cap C \neq \emptyset$, H é G -diferenciável e contínua em C , resulta que,

$$\partial(H + \lambda f + I_C)(w) = \nabla H(w) + \partial(\lambda f + I_C)(w).$$

Assim temos,

$$-\nabla H(w) - \lambda r + \nabla H(v) \in \partial(\lambda f + I_C)(w),$$

isto é,

$$\langle \nabla H(w) + \lambda r - \nabla H(v), u - w \rangle + \lambda(f(u) - f(w)) + I_C(u) - I_C(w) \geq 0 \quad \forall u \in U,$$

o que é equivalente a

$$\langle \nabla H(w) + \lambda r - \nabla H(v), u - w \rangle + \lambda(f(u) - f(w)) \geq 0 \quad \forall u \in C.$$

Teorema 4.1.3: Se o funcional auxiliar H considerado na definição do problema auxiliar (PA) é estritamente convexo em C então as afirmações (i) – (ii) apresentadas no Teorema 4.1.2 são equivalentes a

$$(iii) \quad w = (\nabla H + \lambda \partial f + N_C)^{-1}(\nabla H(v) - \lambda r).$$

Demonstração: Como (i) e (ii) são equivalentes, basta mostrarmos, por exemplo, que (ii) e (iii) são equivalentes. Seja $w \in C$ tal que

$$\langle \nabla H(w) + \lambda r - \nabla H(v), u - w \rangle + \lambda(f(u) - f(w)) \geq 0 \quad \forall u \in C, \quad (11)$$

o que equivale a

$$(\lambda f + I_C)(u) \geq (\lambda f + I_C)(w) + \langle \nabla H(v) - \lambda r - \nabla H(w), u - w \rangle \quad \forall u \in U,$$

isto é,

$$-\lambda r + \nabla H(v) \in \nabla H(w) + \partial(\lambda f + I_C)(w).$$

Como $\text{Dom } f \cap C^\circ \neq \emptyset$ e a função I_C é contínua em C° , temos que,

$$\partial(\lambda f + I_C)(w) = \partial \lambda f(w) + N_C(w) \stackrel{\lambda \geq 0}{=} \lambda \partial f(w) + N_C(w).$$

Portanto, (11) é equivalente a

$$\nabla H(v) - \lambda r \in \nabla H(w) + \lambda \partial f(w) + N_C(w).$$

Além disso, como H é estritamente convexa em C , para cada par (v, r) existe um único $w = w(v, r)$ satisfazendo (i) ou equivalentemente (ii). Logo o operador $(\nabla H + \lambda \partial f + N_C)^{-1}$ é ponto-ponto. Desta forma, podemos concluir que

$$w = (\nabla H + \lambda \partial f + N_C)^{-1}(\nabla H(v) - \lambda r).$$

Note que a conclusão (iii) no Teorema 4.1.3 mostra que o método do problema auxiliar para o problema (PIV) pode ser reestruturado como o algoritmo de decomposição do tipo “forward-backward” onde o passo “forward” realizado por Ψ é seguido pelo passo “backward” dado por $\partial f + N_C$ (veja subseção 3.5.4).

Apresentamos abaixo um resultado que garante a existência e unicidade para a solução do (PA).

Teorema 4.1.4: Considere os problemas $PIV(\Psi, C, f)$ e (PA). Suponha que as seguintes condições são satisfeitas:

(i) H é estritamente convexa em C ;

(ii) C é limitado ou H é coerciva em C (no sentido da definição 2.1.11(iv)).

Então $\forall v \in C \cap \text{Dom}(\Psi)$, $r \in \Psi(v)$ e $\lambda > 0$, existe uma única solução $w = w(v, r)$ do (PA).

Demonstração: Sejam $v \in C \cap \text{Dom}(\Psi)$, $r \in \Psi(v)$ e $\lambda > 0$. Como no Teorema 4.1.1, consideramos $K : U \rightarrow (-\infty, +\infty]$ definida por

$$K(u) := H(u) + \langle \lambda r - \nabla H(v), u \rangle + \lambda f(u) \quad \forall u \in U.$$

A função K é estritamente convexa, própria e semicontínua inferiormente em C . Quando o conjunto C é limitado, a prova segue imediatamente do resultado de minimização para funções convexas em espaços de Banach reflexivos (veja [30], capítulo 2, proposição 1.2). Portanto devemos mostrar que, quando C é ilimitado, a coercividade de H em C implica na coercividade de K em C . De fato, suponhamos C ilimitado. Como $\text{Dom}(\partial f) \neq \emptyset$ (veja proposição 2.1.1), podemos considerar $u_0 \in \text{Dom}(\partial f)$. Assim, para cada $u \in U$, temos

$$K(u) \geq H(u) + \langle \lambda r - \nabla H(v), u \rangle + \lambda f(u_0) + \lambda \langle g_0, u - u_0 \rangle,$$

onde $g_0 \in \partial f(u_0)$, pois $f(u) \geq f(u_0) + \langle g_0, u - u_0 \rangle \quad \forall u \in U$. Logo, usando a desigualdade de Cauchy-Schwartz, temos que,

$$\begin{aligned} \frac{K(u)}{\|u\|} &\geq \frac{H(u)}{\|u\|} + \frac{\langle \lambda r - \nabla H(v), u \rangle}{\|u\|} + \frac{\lambda f(u_0)}{\|u\|} + \frac{\langle \lambda g_0, u - u_0 \rangle}{\|u\|} \\ &\geq \frac{H(u)}{\|u\|} - \|\lambda \mu r - \nabla H(v)\| + \frac{\lambda f(u_0)}{\|u\|} - \|\lambda g_0\| - \frac{\langle \lambda g_0, u_0 \rangle}{\|u\|} \quad \forall u \in C, u \neq 0. \end{aligned}$$

Como $\frac{H(u)}{\|u\|} \rightarrow +\infty$ quando $\|u\| \rightarrow +\infty$ (por hipótese), temos que,

$$\lim_{\substack{\|u\| \rightarrow +\infty \\ u \in C}} \frac{K(u)}{\|u\|} = +\infty,$$

e conseqüentemente,

$$\lim_{\substack{\|u\| \rightarrow +\infty \\ u \in C}} K(u) = +\infty.$$

Novamente, a existência e unicidade para a solução do (PA) segue da Proposição 1.2 em [30].

Observação 4.1.2: Cohen em [22], garante a existência e unicidade para a solução do (PA) sob hipóteses mais fortes do que as consideradas no Teorema 4.1.4. De fato, ele considera f real, convexa e semicontínua inferiormente e o funcional auxiliar H real, diferenciável e fortemente convexo.

A seguir, apresentamos o resultado que deu origem ao critério de parada no algoritmo.

Teorema 4.1.5: Dados os problemas $PIV(\Psi, C, f)$ e (PA) , considere $w = w(v, r)$ uma solução do (PA) . Então $w = v$ se, e somente se, v é solução do $PIV(\Psi, C, f)$.

Demonstração: Seja w uma solução de (PA) . Pelo Teorema 4.1.2, $w = v$ se, e somente se, $v \in C$ e

$$\langle \nabla H(v) + \lambda r - \nabla H(v), u - v \rangle + \lambda (f(u) - f(v)) \geq 0 \quad \forall u \in C,$$

isto é,

$$\langle \lambda r, u - v \rangle + \lambda (f(u) - f(v)) \geq 0 \quad \forall u \in C,$$

ou seja,

$$\langle r, u - v \rangle + f(u) - f(v) \geq 0 \quad \forall u \in C,$$

dado que $\lambda > 0$. A inequação acima significa que v é solução de $PIV(\Psi, C, f)$.

4.2 O Problema Auxiliar Estendido

Considere $\{\lambda_k\}$ uma seqüência de números reais positivos e $\{H^k\}$ uma seqüência de funcionais auxiliares como na seção 4.1. Estendemos o problema auxiliar a cada iteração k do seguinte modo: Dados $u^k \in C \cap \text{Dom}(\Psi)$, $r^k \in \Psi(u^k)$ e $\lambda_k > 0$, resolvemos o seguinte subproblema auxiliar:

(PA^k) Encontrar $u^{k+1} \in C$, solução de

$$\underset{u \in C}{\text{minimizar}} \quad H^k(u) + \langle \lambda_k r^k - \nabla H^k(u^k), u \rangle + \lambda_k f(u) .$$

Observemos que $u^{k+1} \in C \cap \text{Dom} f$.

O algoritmo estendido

Passo 1: Na iteração $k = 0$, iniciamos com algum ponto $u^0 \in C \cap \text{Dom}(\Psi)$ e $r^0 \in \Psi(u^0)$.

Passo 2: Na iteração k , encontramos a solução u^{k+1} do problema auxiliar (PA^k) .

Passo 3: Se $u^{k+1} = u^k$ então paramos. Caso contrário, selecionamos $r^{k+1} \in \Psi(u^{k+1})$ e repetimos o passo 2 com $k \leftarrow k + 1$.

Este algoritmo estende o princípio do problema auxiliar para um operador ponto-conjunto proposto por Cohen [22]. Com efeito, são consideradas variações do funcional auxiliar a cada passo. Esta extensão amplia a família de métodos computacionais usados na implementação do passo 2 permitindo a obtenção de melhores taxas de convergência para o algoritmo de Cohen. Variações deste tipo foram consideradas em [45] para o caso ponto-ponto.

4.3 Boa definição da seqüência

O Teorema 4.1.4 garante que a cada iteração k do algoritmo acima existe uma única u^{k+1} solução de (PA^k) desde que cada funcional H^k seja estritamente convexo e coercivo em C quando C for ilimitado. Além disso, a hipótese adicional *(ib)* considerada para $PIV(\Psi, C, f)$ garante que $u^{k+1} \in \text{Dom}(\Psi)$. Logo, a boa definição da seqüência $\{u^k\}$ pode ser estabelecida pela condição *(ib)* e pelo Teorema 4.1.4.

Observação 4.3.1: Podemos substituir, sem perda de generalidade, a condição (ib) considerada para PIV (Ψ, C, f) pela condição abaixo:

(ib)' $C' := \text{Dom } f \cap C \cap \text{Dom } (\Psi)$ é um subconjunto com o interior não-vazio, convexo e fechado de U ,

desde que C seja substituído por C' no algoritmo.

4.4 Análise de convergência

Apresentamos a seguir alguns resultados que são usados na prova da convergência forte do algoritmo estendido.

Lema 4.4.1: Seja U um espaço normado. A função $g : U \rightarrow \mathbf{R}$ definida por

$$g(u) := \|u\|^2, \forall u \in U$$

é Lipschitz sobre conjuntos limitados.

Demonstração: Seja $A \subset U$ um conjunto limitado. Considere $M > 0$ tal que $\|u\| \leq M \forall u \in A$. Sejam $u_1, u_2 \in A$. Usando a desigualdade triangular temos:

$$\begin{aligned} \|u_1\| - \|u_2\| &= \|(u_1 - u_2) + u_2\| - \|u_2\| \\ &\leq \|u_1 - u_2\| + \|u_2\| - \|u_2\| \\ &= \|u_1 - u_2\|. \end{aligned}$$

Assim,

$$\begin{aligned} \|u_1\|^2 - \|u_2\|^2 &= (\|u_1\| - \|u_2\|)(\|u_1\| + \|u_2\|) \\ &\leq \|u_1 - u_2\|(\|u_1\| + \|u_2\|) \leq 2M \|u_1 - u_2\|. \end{aligned}$$

Permutando u_1 e u_2 nas desigualdades acima, temos

$$\|u_2\|^2 - \|u_1\|^2 \leq 2M \|u_1 - u_2\|.$$

Logo,

$$|g(u_1) - g(u_2)| \leq 2M \|u_1 - u_2\|.$$

Lema 4.4.2: Sejam $\{x^k\}$ e $\{\alpha^k\}$ duas seqüências em \mathbf{R}_+ tais que

$$\sum_{k \in \mathbf{N}} \alpha^k < +\infty$$

e $x^N \leq \sum_{k=1}^{N-1} \alpha^k X^{k+1} + \delta^N$ onde $\delta^N \leq \delta$, $\forall N \in \mathbf{N}$ e $X^k := \sup_{l \leq k} x^l \forall k \in \mathbf{N}$. Então a

seqüência $\{x^k\}$ é limitada.

Demonstração: Veja [24].

Lema 4.4.3: Seja g uma função Lipschitz em um espaço de Hilbert U e considere as seqüências $\{u^k\} \subset U$ e $\{\varepsilon^k\} \subset \mathbf{R}_+$ tais que

$$\sum_{k \in \mathbf{N}} \varepsilon^k = +\infty;$$

$$\exists \delta : \forall k \in \mathbf{N}, \|u^{k+1} - u^k\| \leq \delta \varepsilon^k;$$

$$\exists \mu : \sum_{k \in \mathbf{N}} \varepsilon^k |g(u^k) - \mu| < +\infty.$$

Então $\lim g(u^k) = \mu$ quando $k \rightarrow +\infty$.

Demonstração: Veja [24].

Lema 4.4.4: Seja U um espaço normado. Então

$$\|u + v\|^2 \leq 2(\|u\|^2 + \|v\|^2) \quad \forall u, v \in U.$$

Demonstração: Usando a desigualdade triangular, $\forall u, v \in U$, temos

$$\begin{aligned} \|u + v\|^2 &\leq (\|u\| + \|v\|)^2 \\ &= \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &\leq \|u\|^2 + \|u\|^2 + \|v\|^2 + \|v\|^2 \\ &\leq 2(\|u\|^2 + \|v\|^2). \end{aligned}$$

Enunciamos a seguir um resultado bastante simples, mas que se faz necessário às provas deste capítulo.

Lema 4.4.5: Sejam a, b e c números reais. Então

$$abc \leq \frac{1}{2} \left(\frac{a^2}{\gamma} b^2 + \gamma c^2 \right), \quad \forall \gamma > 0.$$

Demonstração: Esta desigualdade é consequência imediata de:

$$\left(\frac{ab}{\sqrt{\gamma}} - \sqrt{\gamma} c \right)^2 \geq 0$$

e

$$\left(\frac{ab}{\sqrt{\gamma}} - \sqrt{\gamma} c \right)^2 = \frac{a^2}{\gamma} b^2 + \gamma c^2 - 2abc.$$

Observação 4.4.1: O resultado do Lema 4.4.3 permanece válido no caso mais geral em que U é um espaço de Banach real.

Lema 4.4.6: Sejam U um espaço de Hilbert e $\Psi : U \rightarrow P(U)$ um operador fortemente monótono (módulo $\alpha > 0$). Se o problema $PIV(\Psi, C, f)$ tem solução então a solução é única.

Demonstração: Segue da Proposição 3.4.1.

Lema 4.4.7: Sejam os subproblemas $(PA^k) \forall k \in \mathbf{N}$. Considere $\{H^k\}$ uma seqüência de funções tal que cada H^k é um funcional de Cohen, ou seja, $H^k : U \rightarrow \mathbf{R}$ é G -diferenciável, contínua e fortemente convexa (módulo $\beta_k > 0$) sobre C . Então cada problema auxiliar estendido (PA^k) possui uma única solução.

Demonstração: Esta prova segue como corolário do Teorema 4.1.4.

Teorema 4.4.1: Suponha que, além das hipóteses dos Lemas 4.4.6 e 4.4.7, são satisfeitas as seguintes condições:

(i) $\forall k \in \mathbf{N}, \exists \gamma_k > 0 : \forall u, v \in U,$

$$H^{k+1}(u) - H^{k+1}(v) - \langle \nabla H^{k+1}(v), u - v \rangle \leq \gamma_k (H^k(u) - H^k(v) - \langle \nabla H^k(v), u - v \rangle);$$

(ii) $\exists \beta > 0 : \frac{\beta_k}{M_k} \geq \beta \forall k \in \mathbf{N}$ onde $M_k := \prod_{i=1}^{k-1} \gamma_i \quad \forall k \geq 2;$

(iii) $\exists c > 0, \exists d > 0 : \forall u \in C, \forall r \in \Psi(u), \|r\| \leq c\|u\| + d$ (isto é, a norma de Ψ cresce no máximo linearmente com a norma de u);

(iv) A seqüência $\left\{ \frac{\lambda_k}{M_k} \right\}$ é tal que $\sum_{k=0}^{+\infty} \frac{\lambda_k}{M_k} = +\infty$ e $\sum_{k=0}^{+\infty} \left(\frac{\lambda_k}{M_k} \right)^2 < +\infty$.

Então a solução da inequação variacional $PIV(\Psi, C, f)$ é um ponto limite forte da seqüência $\{u^k\}$ gerada pelo algoritmo estendido.

Demonstração: Considere u^* a solução de $PIV(\Psi, C, f)$ e a família de funções $\Gamma^k : U \rightarrow \mathbf{R}$ definidas por

$$\Gamma^k(u) := \frac{1}{M_k} (H^k(u^*) - H^k(u) - \langle \nabla H^k(u), u^* - u \rangle), \quad \forall u \in U.$$

Como $\forall k \in \mathbf{N}$, H^k é fortemente convexo (módulo $\beta_k > 0$), temos que

$$H^k(u^*) - H^k(u^k) - \langle \nabla H^k(u^k), u^* - u^k \rangle \geq \frac{\beta_k}{2} \|u^* - u^k\|^2 \quad \forall k \in \mathbf{N}.$$

Então

$$\Gamma^k(u^k) \geq \frac{\beta_k}{2M_k} \|u^* - u^k\|^2 \stackrel{\text{por (ii)}}{\geq} \frac{\beta}{2} \|u^* - u^k\|^2 \quad \forall k \in \mathbf{N},$$

ou seja,

$$\Gamma^k(u^k) \geq \frac{\beta}{2} \|u^* - u^k\|^2 \quad \forall k \in \mathbf{N}.$$

Pela definição de Γ^k , temos também que,

$$\Gamma^{k+1}(u) \stackrel{\text{por (i)}}{\leq} \frac{1}{M_{k+1}} \gamma_k M_k \Gamma^k(u),$$

mas,

$$\frac{M_k}{M_{k+1}} \cdot \gamma_k = \frac{\gamma_1 \gamma_2 \dots \gamma_{k-1} \gamma_k}{\gamma_1 \gamma_2 \dots \gamma_k} = 1 \quad (\text{por (ii)}).$$

Daí, $\Gamma^{k+1}(u) \leq \Gamma^k(u) \quad \forall u \in U, k \in \mathbf{N}$. Assim, podemos escrever

$$\begin{aligned} \Gamma^k(u^k) - \Gamma^{k+1}(u^{k+1}) &\geq \Gamma^k(u^k) - \Gamma^k(u^{k+1}) \\ &= \frac{1}{M_k} [H^k(u^{k+1}) - H^k(u^k) - \langle \nabla H^k(u^k), u^* - u^k \rangle \\ &\quad + \langle \nabla H^k(u^{k+1}), u^* - u^{k+1} \rangle] \\ &= \frac{1}{M_k} (H^k(u^{k+1}) - H^k(u^k) - \langle \nabla H^k(u^k), u^{k+1} - u^k \rangle) \\ &\quad + \frac{1}{M_k} \langle \nabla H^k(u^{k+1}) - \nabla H^k(u^k), u^* - u^{k+1} \rangle \\ &\geq \frac{\beta}{2} \|u^{k+1} - u^k\|^2 + \frac{1}{M_k} \langle \nabla H^k(u^{k+1}) - \nabla H^k(u^k), u^* - u^{k+1} \rangle. \end{aligned}$$

Por outro lado, como u^{k+1} é solução do subproblema (PA^k) , usamos o Teorema 4.1.2 (para $v = u^k$, $w = u^{k+1}$ e $u = u^*$) e obtemos

$$\langle \nabla H^k(u^{k+1}) - \nabla H^k(u^k) + \lambda_k r_k, u^* - u^{k+1} \rangle + \lambda_k (f(u^*) - f(u^{k+1})) \geq 0.$$

Donde, resulta que

$$\begin{aligned} \frac{1}{M_k} \langle \nabla H^k(u^{k+1}) - \nabla H^k(u^k), u^* - u^{k+1} \rangle &\geq \frac{\lambda_k}{M_k} (f(u^{k+1}) - f(u^*)) + \frac{\lambda_k}{M_k} \langle r_k, u^{k+1} - u^* \rangle \\ &\geq -\frac{\lambda_k}{M_k} \langle r^*, u^{k+1} - u^* \rangle + \frac{\lambda_k}{M_k} \langle r_k, u^{k+1} - u^* \rangle, \end{aligned}$$

onde a última desigualdade resulta da substituição de u por u^{k+1} em (PIV).

Portanto

$$\frac{1}{M_k} \langle \nabla H^k(u^{k+1}) - \nabla H^k(u^k), u^* - u^{k+1} \rangle \geq \frac{\lambda_k}{M_k} \langle r_k - r^*, u^{k+1} - u^* \rangle.$$

Então podemos concluir que:

$$\begin{aligned} \Gamma^k(u^k) - \Gamma^{k+1}(u^{k+1}) &\geq \frac{\beta}{2} \|u^{k+1} - u^k\|^2 + \frac{\lambda_k}{M_k} \langle r_k - r^*, u^{k+1} - u^* \rangle \\ &= \frac{\beta}{2} \|u^{k+1} - u^k\|^2 + \frac{\lambda_k}{M_k} \langle r_k - r^*, u^{k+1} - u^k \rangle + \frac{\lambda_k}{M_k} \langle r_k - r^*, u^k - u^* \rangle \\ &\geq \frac{\beta}{2} \|u^{k+1} - u^k\|^2 - \frac{\lambda_k}{M_k} \|r_k - r^*\| \|u^{k+1} - u^k\| + \frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 \\ &\geq \frac{\beta}{2} \|u^{k+1} - u^k\|^2 + \frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 - \frac{\lambda_k^2}{2\beta M_k^2} \|r_k - r^*\|^2 - \frac{\beta}{2} \|u^{k+1} - u^k\|^2 \\ &= \frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 - \frac{\lambda_k^2}{2\beta M_k^2} \|r_k - r^*\|^2, \end{aligned}$$

onde as duas últimas desigualdades acima seguem da nonotonicidade forte de Ψ e do Lema 4.4.5 ($a = \frac{\lambda_k}{M_k}$, $b = \|r_k - r^*\|$, $c = \|u^{k+1} - u^k\|$ e $\gamma = \beta$) respectivamente.

Logo, obtemos

$$\Gamma^k(u^k) - \Gamma^{k+1}(u^{k+1}) \geq \frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 - \frac{\lambda_k^2}{2\beta M_k^2} \|r_k - r^*\|^2.$$

Agora, somando a desigualdade acima desde $k = 0$ até $k = N - 1$, temos que $\forall N \in \mathbb{N}$

$$\begin{aligned} \Gamma^N(u^N) - \Gamma^0(u^0) &= -\sum_{k=0}^{N-1} (\Gamma^k(u^k) - \Gamma^{k+1}(u^{k+1})) \\ &\leq \sum_{k=0}^{N-1} \left(-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + \frac{\lambda_k^2}{2\beta M_k^2} \|r_k - r^*\|^2 \right). \end{aligned}$$

Assim, usando o Lema 4.4.4 ($u = r^k$ e $v = r^*$), temos que

$$\begin{aligned}\Gamma^N(u^N) &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + \frac{\lambda_k^2}{2\beta M_k^2} \|r_k - r^*\|^2 \right) \\ &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + 2 \left(\frac{\lambda_k^2}{2\beta M_k^2} \right) (\|r_k\|^2 + \|r^*\|^2) \right).\end{aligned}$$

Mas,

$$\begin{aligned}\|r_k\| &\stackrel{\text{por (iii)}}{\leq} c \|u^k\| + d \Rightarrow \|r_k\|^2 \leq (c \|u^k\| + d)^2 \leq 2c^2 \|u^k\|^2 + 2d^2 \\ \text{e} \\ \|r^*\| &\leq c \|u^*\| + d \Rightarrow \|r^*\|^2 \leq (c \|u^*\| + d)^2 \leq 2c^2 \|u^*\|^2 + 2d^2\end{aligned}$$

implica que,

$$\|r_k\|^2 + \|r^*\|^2 \leq 2c^2 \|u^k\|^2 + 2c^2 \|u^*\|^2 + 4d^2.$$

Então, somando e subtraindo u^* no termo $\|u^k\|^2$ e usando o Lema 4.4.4, obtemos

$$\begin{aligned}\Gamma^N(u^N) &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + 4 \left(\frac{\lambda_k^2}{2\beta M_k^2} \right) (c^2 \|u^k\|^2 + c^2 \|u^*\|^2 + 2d^2) \right) \\ &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + 4 \left(\frac{\lambda_k^2}{2\beta M_k^2} \right) (2c^2 \|u^k - u^*\|^2 + 3c^2 \|u^*\|^2 + 2d^2) \right) \\ &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left\{ -\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + \left(\frac{\lambda_k}{M_k} \right)^2 \left[\frac{4}{\beta} c^2 \|u^k - u^*\|^2 + 6 \frac{c^2}{\beta} \|u^*\|^2 + 4 \frac{d^2}{\beta} \right] \right\} \\ &= \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left[-\frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 + \left(\frac{\lambda_k}{M_k} \right)^2 (\lambda \|u^k - u^*\|^2 + \tau) \right]\end{aligned}$$

onde λ e τ são constantes positivas. Logo $\forall N \in \mathbf{N}$, temos que

$$\begin{aligned}\frac{\beta}{2} \|u^N - u^*\|^2 &\leq \Gamma^N(u^N) \leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(\frac{\lambda_k}{M_k} \right)^2 (\lambda \|u^k - u^*\|^2 + \tau) \\ &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(\frac{\lambda_k}{M_k} \right)^2 (\lambda \|u^k - u^*\|^2 + \tau) \\ &\leq \Gamma^0(u^0) + \sum_{k=1}^{N-1} \left(\frac{\lambda_k}{M_k} \right)^2 \lambda \sup_{l \leq k+1} \|u^l - u^*\|^2 + \left(\frac{\lambda_0}{M_0} \right)^2 (\lambda \|u^0 - u^*\|^2 + \tau) \\ &\quad + \sum_{k=0}^{N-1} \left(\frac{\lambda_k}{M_k} \right)^2 \tau.\end{aligned}$$

Usando o Lema 4.4.2 ($x^k = \|u^k - u^*\|^2$, $\alpha^k = \left(\frac{\lambda_k}{M_k}\right)^2 \lambda_{\beta}^2$ e $\delta^k = \frac{2}{\beta}\Gamma^0(u^0) + \frac{2}{\beta} \sum_{l=1}^{N-1} \left(\frac{\lambda_l}{M_l}\right)^2 \tau + \frac{2}{\beta} \left(\frac{\lambda_0}{M_0}\right)^2 (\lambda \|u^0 - u^*\|^2 + \tau)$) e a condição (iv), podemos concluir que a seqüência $\{u^k\}$ é limitada.

Portanto,

$$\begin{aligned} \sum_{k=0}^{N-1} \frac{\lambda_k}{M_k} \alpha \|u^k - u^*\|^2 &\leq \Gamma^0(u^0) + \sum_{k=0}^{N-1} \left(\frac{\lambda_k}{M_k}\right)^2 (\lambda \|u^k - u^*\|^2 + \tau) \\ &\leq \Gamma^0(u^0) + L' \sum_{k=0}^{N-1} \left(\frac{\lambda_k}{M_k}\right)^2, \end{aligned}$$

onde $L' = L\lambda + \tau$ com $\|u^k - u^*\|^2 \leq L$, $\forall k \in \mathbf{N}$.

Então, usando (iv), obtemos

$$\sum_{k=0}^{+\infty} \frac{\lambda_k}{M_k} \|u^k - u^*\|^2 < +\infty. \quad (12)$$

Finalmente, supomos, por absurdo, que u^* não é ponto limite de $\{u^k\}$. Então temos que

$$d := \inf \{ \|u^k - u^*\| \mid k \in \mathbf{N} \} > 0.$$

Assim,

$$\begin{aligned} \|u^k - u^*\| \geq d > 0 \quad \forall k \in \mathbf{N} &\Rightarrow 0 < d^2 \leq \|u^k - u^*\|^2 \quad \forall k \in \mathbf{N}. \\ &\Rightarrow \frac{\lambda_k}{M_k} d^2 \leq \frac{\lambda_k}{M_k} \|u^k - u^*\|^2. \end{aligned}$$

Logo, usando (12) e o teste da comparação, garantimos a convergência da série

$$\sum_{k=0}^{+\infty} \frac{\lambda_k}{M_k} d^2$$

e conseqüentemente, temos que

$$\sum_{k=0}^{+\infty} \frac{\lambda_k}{M_k} < +\infty. \quad (13)$$

A condição (iv) e (13) geram uma contradição.

Salmon [61] enuncia um resultado de convergência fraca elaborado por Zhu (em 1999) para o problema particular (VIP) ao qual teve acesso por meio de “Private Communications”.

Supondo Ψ paramonótono, Zhu considera f identicamente nula no algoritmo estendido e prova que a seqüência gerada $\{u^k\}$ é limitada e pelo menos um de seus pontos limites fracos é solução do (VIP). Para isto, ele supõe as condições consideradas no teorema anterior, mas adiciona pelo menos uma das seguintes condições extra:

a) Para qualquer seqüência limitada $\{v^k\}$ de C , a seqüência $\{r^k\}$, $r^k \in \Psi(v^k)$ é limitada e existem subsequências $\{v^{k_j}\}$ e $\{r^{k_j}\}$ tais que

$$v^{k_j} \xrightarrow{w} \bar{v}, r^{k_j} \xrightarrow{w} \bar{r}, \bar{r} \in \Psi(\bar{v}) \quad \text{e} \quad \lim_{j \rightarrow +\infty} \langle r^{k_j}, v^{k_j} \rangle \geq \langle \bar{r}, \bar{v} \rangle.$$

b) Para $u^* \in C$ solução de (VIP), existem uma constante $\alpha > 0$ e uma função l Lipschitz em qualquer subconjunto compacto de C tais que:

- $\langle r, u - u^* \rangle \geq \alpha (l(u) - l(u^*)), \forall u \in C, r \in \Psi(u)$;
- $l(u) = l(u^*)$, se u é uma solução de (VIP);
- $l(u) > l(u^*)$, caso contrário.

Zhu observa que a condição (b) é verificada, por exemplo, para as seguintes classes de operadores:

- paramonótonos e Lipschitz;
- fortemente monótonos. Neste caso, seu resultado segue como corolário da nossa generalização;
- subdiferencias de funções próprias, convexas, semicontínuas inferiormente e Lipschitz em qualquer subconjunto compacto de C . Neste caso, o problema (VIP) é reduzido a um problema de otimização convexa.

Teorema 4.4.2: Suponha que, além das hipóteses do Teorema 4.4.1, são satisfeitas as seguintes condições para a função f definida em $PIV(\Psi, C, f)$:

(v) $\{u^k\} \subseteq (Dom(\partial f))^o$;

(vi) $\exists a > 0, \exists b > 0 : \forall u \in C \cap (Dom(\partial f))^o, \forall r \in \partial f(u), \|r\| \leq a \|u\| + b$ (isto é, a norma do operador ∂f cresce no máximo linearmente com a norma de u).

Então a seqüência $\{u^k\}$ converge forte para u^* .

Demonstração: Vamos mostrar primeiro que: $\exists \delta : \forall k \in \mathbf{N}, \|u^{k+1} - u^k\| \leq \delta \frac{\lambda_k}{M_k}$.

De fato, u^{k+1} é a solução do problema

$$(PA^k) \quad \underset{u \in C}{\text{minimizar}} \quad H^k(u) + \langle \lambda_k r^k - \nabla H^k(u^k), u \rangle + \lambda_k f(u).$$

Portanto,

$$\begin{aligned} H^k(u^{k+1}) + \langle \lambda_k r^k - \nabla H^k(u^k), u^{k+1} \rangle + \lambda_k f(u^{k+1}) \\ \leq H^k(u^k) + \langle \lambda_k r^k - \nabla H^k(u^k), u^k \rangle + \lambda_k f(u^k), \end{aligned}$$

ou seja,

$$\begin{aligned} \langle \lambda_k r^k, u^{k+1} - u^k \rangle \\ \leq \langle \nabla H^k(u^k), u^{k+1} - u^k \rangle + \lambda_k [f(u^k) - f(u^{k+1})] + H^k(u^k) - H^k(u^{k+1}) \\ \leq \lambda_k [f(u^k) - f(u^{k+1})] - \frac{\beta_k}{2} \|u^{k+1} - u^k\|^2 \end{aligned}$$

onde a desigualdade acima resulta da convexidade forte (módulo β_k) de cada funcional H^k .

Então

$$\|u^{k+1} - u^k\|^2 \leq 2 \frac{\lambda_k}{\beta_k} [f(u^k) - f(u^{k+1})] + 2 \frac{\lambda_k}{\beta_k} \langle r_k, u^k - u^{k+1} \rangle.$$

Por outro lado, a condição (v) implica que $\partial f(u^k) \neq \emptyset \forall k \in \mathbf{N}$. Considere $g^k \in \partial f(u^k)$. Então

$$f(u^{k+1}) \geq f(u^k) + \langle g^k, u^{k+1} - u^k \rangle,$$

ou seja,

$$f(u^k) - f(u^{k+1}) \leq \langle g^k, u^k - u^{k+1} \rangle.$$

Portanto,

$$\|u^{k+1} - u^k\|^2 \leq 2 \frac{\lambda_k}{\beta_k} \langle g^k + r_k, u^k - u^{k+1} \rangle \leq 2 \frac{\lambda_k}{\beta_k} \|g^k + r_k\| \|u^{k+1} - u^k\|.$$

Assim, usando a condição (ii) do Teorema 4.4.1, temos que,

$$\|u^{k+1} - u^k\| \leq 2 \frac{\lambda_k}{\beta_k} \|g^k + r_k\| = 2 \frac{M_k}{\beta_k} \frac{\lambda_k}{M_k} \|g^k + r_k\| \leq \frac{2}{\beta} \|g^k + r_k\| \frac{\lambda_k}{M_k}.$$

Logo

$$\|u^{k+1} - u^k\| \leq \delta \frac{\lambda_k}{M_k},$$

onde $\delta = \frac{2}{\beta} (L_r + L_g)$, $L_r, L_g > 0$ são tais que $\|r_k\| \leq L_r \forall k \in \mathbf{N}$ (isto segue da limitação de $\{u^k\}$ e condição (iii) do Teorema 4.4.1) e $\|g^k\| \leq L_g$ (isto segue da condição (vi)). Além disso, como a seqüência $\{u^k\}$ é limitada, temos que a função

$$\begin{aligned} g : Co(\{u^k\}) &\rightarrow \mathbf{R} \\ u &\mapsto \|u - u^*\|^2 \end{aligned}$$

é Lipschitz (veja lema 4.4.1). Finalmente, usando os resultados obtidos acima juntamente com a condição (iv) e as considerações feitas na demonstração do Teorema 4.4.1, aplicamos o Lema 4.4.3 (com $g(u) = \|u - u^*\|^2$, $\mu = 0$ e $\varepsilon^k = \frac{\lambda_k}{M_k}$) e concluimos que $\{u^k\}$ converge fortemente para u^* .

Observação 4.4.2: No caso em que $C \subset (Dom f)^o$, a suposição (v) do teorema anterior pode ser dispensada. De fato, como $f : U \rightarrow (-\infty, +\infty]$ é uma função convexa, própria e semicontínua inferiormente, a condição acima implica que $\partial f(u^k) \neq \emptyset \forall k \in \mathbf{N}$ (veja proposição 2.1.1).

Observação 4.4.3: No caso particular, em que o espaço U é de dimensão finita, a condição (vi) do Teorema 4.4.2 pode ser dispensada. Com efeito, como a função f é convexa, própria e semicontínua inferiormente, temos satisfeitas as hipóteses da Proposição 2.1.4 para o operador subdiferencial (∂f) que é portanto, limitado em conjuntos limitados.

Embora a condição (vi) não apareça explicitamente nos resultados apresentados em [22], observamos que esta condição é essencial à obtenção do resultado de convergência como confirma o próprio Cohen.

O resultado de convergência apresentado no Teorema 4.4.2 estende o método em [22] na medida em que considera variações do funcional auxiliar a cada iteração, e coincide com o teorema de Cohen no caso particular em que a função f é real, $H^k = H$, e $\gamma_k = 1, \forall k \in \mathbf{N}$. Além disso, ampliamos a família de métodos computacionais usados na implementação do passo 2 permitindo a obtenção de melhores taxas de convergência para o algoritmo.

4.5 Aplicações

É importante observar que o método apresentado neste capítulo inclui muitos algoritmos iterativos. Ilustramos alguns deles através dos exemplos abaixo.

Exemplo 4.5.1: Considere f identicamente nula em $PIV(\Psi, C, f)$ e $U = \mathbf{R}^n$. Defina a seguinte seqüência:

$$u^{k+1} := P_C(u^k - \lambda_k B_k^{-1} r_k), \quad r_k \in \Psi(u^k)$$

onde $B_{n \times n}$ é uma matriz simétrica e definida positiva, $\{\lambda_k\}$ é uma seqüência de números reais positivos e P_C é o operador projeção em C .

Notamos que $\{u^k\}$ está bem definida (veja [40], lema 2.1). Além disso, usando a caracterização da projeção obtemos que (veja [40], teorema 2.3):

$$\langle u^{k+1} + \lambda_k B_k^{-1} r_k - u^k, u - u^{k+1} \rangle \geq 0 \quad \forall u \in C.$$

A desigualdade acima significa que no algoritmo geral $\{H^k\}$ pode ser definida por:

$$H^k(u) := \frac{1}{2} \langle B_k u, u \rangle \quad \forall u \in U, \quad k \in \mathbf{N}.$$

Assim, o problema auxiliar a cada iteração é dado por

(PA^k) Encontrar $u^{k+1} \in C$, solução de

$$\underset{u \in C}{\text{minimizar}} \quad \frac{1}{2} \langle B_k u, u \rangle + \langle \lambda_k r^k - B_k u^k, u \rangle.$$

Exemplo 4.5.2: Seja U um espaço de Hilbert real. Defina a seguinte seqüência:

$$u^{k+1} := \arg \min_{u \in C} \left(f(u) + \frac{1}{2} \|u - (u^k - \lambda_k r^k)\|^2 \right).$$

Notamos que se $\Psi \equiv 0$, este método é reduzido ao método do ponto proximal para otimização convexa. Se, novamente, f é identicamente nula então a seqüência acima é definida como no exemplo anterior quando consideramos cada B_k , em particular, a matriz identidade. De modo análogo, temos

$$H^k(u) := \frac{1}{2} \langle u, u \rangle \quad \forall u \in U$$

ou

$$H^k(u) := \frac{1}{2} \|u - u^k\|^2 \quad \forall u \in U, k \in \mathbf{N}.$$

Assim, (PA^k) fica

(PA^k) Encontrar $u^{k+1} \in C$, solução de

$$\underset{u \in C}{\text{minimizar}} \quad \frac{1}{2} \|u - u^k\|^2 + \langle \lambda_k r_k, u \rangle.$$

Se, além disso, Ψ é um operador regular (isto é, um subdiferencial de uma função convexa) então nosso método é reduzido ao método clássico do subgradiente projetado (veja subseção 3.5.1).

Por fim, devemos ressaltar que a convergência do algoritmo para uma solução do (PIV) depende de hipóteses pré-estabelecidas para as seqüências dos parâmetros $\{\lambda_k\}$ e das matrizes $\{B_k\}$ (veja [11] para mais detalhes).

5 A Condição de Dunn e Funções de Bregman

5.1 Introdução

Este capítulo destina-se à apresentação dos conceitos e propriedades da condição de Dunn e função (e distância) de Bregman aos quais faremos referência no sexto capítulo. Na seção 5.3 introduzimos resultados teóricos e apresentamos exemplos relacionados com a noção de função de Bregman.

A partir de agora, vamos estudar o problema de inequação variacional $PIV(\Psi, C, f)$ quando o operador Ψ é ponto-ponto. Aplicações do caso ponto-ponto são estudadas, por exemplo, em [3], [30], [40], [53] e [68].

5.2 A condição de Dunn

Nesta seção, apresentamos a noção e as principais características da condição de Dunn. Mostramos que este conceito implica monotonicidade e paramonotonicidade mas não necessariamente monotonicidade forte. Entretanto, monotonicidade forte e a propriedade de Lipschitz implicam a condição de Dunn. Além disso, provamos que Dunn é uma condição mais forte que Lipschitz.

Nas provas de convergência que são apresentadas no capítulo seguinte, esta condição é adicionada à hipótese de maximalidade monótona para o operador Ψ . As definições e resultados apresentados abaixo podem ser encontrados em [54].

Definição 5.2.1: Seja U um espaço de Hilbert. O operador $\Psi : U \rightarrow U$ satisfaz a condição de Dunn com constante $\gamma > 0$ se

$$\langle \Psi(u) - \Psi(v), u - v \rangle \geq \gamma \|\Psi(u) - \Psi(v)\|^2 \quad \forall u, v \in U.$$

É fácil ver que um operador Ψ satisfaz a condição de Dunn com constante $\gamma > 0$ se, e somente se, seu operador inverso ponto-conjunto Ψ^{-1} é fortemente monótono módulo γ .

A condição de Dunn foi introduzida por Browder e Petryshyn [10] no contexto do cálculo de pontos fixos. Mais tarde esta condição foi usada em [11] para estabelecer a convergência do algoritmo da projeção e na elaboração da estrutura do princípio do problema auxiliar (veja, por exemplo, [55]). O nome Dunn foi atribuído a esta condição por Mataoui [48] a propósito do trabalho de Dunn [28]. Alguns nomes e

propriedades equivalentes aparecem na literatura. Por exemplo, o conceito de operador “firmemente não-expansivo” [29], “operador co-coercivo” ([46], [65], [70], e [71]) e “fortemente - f -monótono” [44].

Definição 5.2.2: Considere U_1 e U_2 espaços de Hilbert. Uma aplicação $\varphi : U_1 \rightarrow U_2$ é Lipschitz com constante $L > 0$ ou L -Lipschitz se

$$\|\varphi(u) - \varphi(v)\|_{U_2} \leq L \|u - v\|_{U_1} \quad \forall u, v \in U_1.$$

φ é não-expansiva se φ é Lipschitz com constante $L = 1$.

A proposição seguinte estabelece propriedades para a condição de Dunn.

Proposição 5.2.1: Considere U um espaço de Hilbert e $\Psi : U \rightarrow U$ um operador que satisfaz a condição de Dunn com constante γ . Então

- (i) Ψ verifica a condição de Lipschitz com constante γ^{-1} ;
- (ii) Ψ é paramonótono.

Convém ressaltar que paramonotonicidade não implica a condição de Dunn. Por exemplo, considere a aplicação definida por $\Psi(u) = u^3$, $\forall u \in \mathbf{R}$. É fácil ver que Ψ é paramonótono mas não satisfaz a propriedade de Dunn em \mathbf{R} (veja [59]). A condição (i) também é mais fraca que a condição de Dunn. De fato, em [48], encontra-se o seguinte exemplo: Considere Ψ o operador associado a rotação $\frac{\pi}{2}$ definido por

$$\Psi : \mathbf{R}^2 \rightarrow \mathbf{R}^2$$

$$(u, v) \mapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

O operador Ψ é Lipschitz mas não verifica a condição de Dunn.

No entanto, deve-se observar que quando Ψ é o gradiente de uma função convexa, real e G -diferenciável então a condição de Dunn é equivalente às suposições (i) e (ii) (veja [4], corolário 10). Contudo, isto não é verdade de um modo geral. Realmente, Salmon [59] considera $\Psi(u, v) = (\sqrt{u} + v, \sqrt{v} - u)$, $\forall (u, v) \in [1, +\infty) \times [1, +\infty)$. Neste caso o operador Ψ é paramonótono e Lipschitz em $[1, +\infty) \times [1, +\infty)$ mas não satisfaz a condição de Dunn neste conjunto.

O resultado seguinte mostra que a hipótese de convexidade forte é essencial para a estabelecer a equivalência entre as condições de Dunn e de Lipschitz. Sua prova pode ser encontrada em [54].

Proposição 5.2.2: Considere U um espaço de Hilbert. Se $\Psi : U \rightarrow U$ é um operador fortemente monótono módulo α e verifica a condição de Lipschitz com constante L então satisfaz a condição de Dunn com constante $\frac{\alpha}{L^2}$.

Finalmente, observamos que se Ψ é injetivo a condição de Dunn implica na monotonicidade estrita. Por outro lado, é fácil ver que operadores que verificam a condição de Dunn são monótonos mas não necessariamente fortemente monótonos (considere, por exemplo, a aplicação constante). Portanto, a condição de Dunn é um conceito intermediário entre monotonicidade e nonotonicidade forte.

5.3 Funções de Bregman

As noções de função e distância de Bregman foram introduzidas por Bregman [5] no estudo de algoritmos iterativos para o cálculo de pontos fixos de vários tipos de operadores. No entanto, o termo “função de Bregman” foi empregado primeiramente por Censor e Lent [19]. As funções de Bregman têm sido usadas extensivamente em algoritmos de otimização convexa (veja, por exemplo, [27], [35] e [41]) e em extensões do método do ponto proximal para inequações variacionais (veja, por exemplo, [12], [13], [15], [18] e [36]). Para o caso de dimensão infinita, a distância de Bregman tem sido usada em espaços de Hilbert (veja, por exemplo, [12] e [13]) e em espaços de Banach (veja, por exemplo, [15] e [16]).

Devemos notar que a definição original de função de Bregman [19] foi dada no espaço \mathbf{R}^n e generalizada para espaços de Banach (veja, por exemplo, [15] e [16]).

Seja U um espaço de Hilbert e seja $S \subset U$ um subconjunto aberto, convexo e não-vazio. Considere uma função $g : U \rightarrow (-\infty, +\infty]$ com as seguintes propriedades:

- (i) $\bar{S} \subset \text{Dom } g$;
 - (ii) g é G -diferenciável em S ;
 - (iii) g é s.c.i. e estritamente convexa em \bar{S} .
- (14)

Define-se $D_g : \bar{S} \times S \rightarrow \mathbf{R}_+$ por

$$D_g(u, v) = g(u) - g(v) - \langle \nabla g(v), u - v \rangle. \quad (15)$$

onde $\nabla g(\cdot)$ é o gradiente de g definida em S .

É fácil ver que para todo $u \in \bar{S}$, $v \in S$, $D_g(u, v) \geq 0$ e $D_g(u, v) = 0$ se, e somente se, $u = v$. Além disso, $D_g(\cdot, v)$ é estritamente convexa e s.c.i em \bar{S} para

cada $v \in S$. Supomos D_g estendida a $U \times U$ de maneira usual, ou seja, consideramos $D_g(u, v) = +\infty$ para todo par $(u, v) \notin \bar{S} \times S$. Notemos ainda que D_g não é a distância usual. Sua importância na otimização por substituir esta distância ou, mais precisamente, o quadrado da norma associada ao produto interno de U , foi primeiramente enfatizada por Bregman [5]. Em geral, D_g não é simétrica e não satisfaz a desigualdade triangular. Entretanto, D_g satisfaz a “igualdade dos três pontos” que enunciamos a seguir e é consequência imediata da definição (15).

Proposição 5.3.1:(Igualdade dos três pontos) Dados $u \in \bar{S}$, $v, w \in S$, verifica-se a seguinte igualdade:

$$\langle \nabla g(v) - \nabla g(w), w - u \rangle = D_g(u, v) - D_g(u, w) - D_g(w, v).$$

Definição 5.3.1: A função g verificando as condições (14) é uma *função de Bregman com zona S* e D_g é a *distância de Bregman relativa a g* se as seguintes condições são satisfeitas:

B_1 : O conjunto $S_{u,\delta} := \{v \in S / D_g(u, v) \leq \delta\}$ é limitado para cada $\delta \geq 0$ e para cada $u \in \bar{S}$;

B_2 : Se $\{u^k\}, \{v^k\} \subset S$, $u^k \xrightarrow{w} u$, $v^k \xrightarrow{w} u$ e $\lim_{k \rightarrow \infty} D_g(u^k, v^k) = 0$ então

$$\lim_{k \rightarrow \infty} (D_g(u, u^k) - D_g(u, v^k)) = 0.$$

A condição seguinte será utilizada para manter a seqüência gerada pelo método introduzido no capítulo 6 contida em S .

Definição 5.3.2: A função de Bregman g é *fronteira coerciva* se a seguinte condição é satisfeita:

B_3 : (fronteira coerciva)

Se $\{v^k\} \subset S$ é fracamente convergente para algum ponto na fronteira (∂S) de S então

$$\lim_{k \rightarrow \infty} \langle \nabla g(v^k), u - v^k \rangle = -\infty \quad \forall u \in S.$$

Consideramos ainda a seguinte condição alternativa a B_2 para a função g :

B_2^0 : Se $\{u^k\}, \{v^k\} \subset S, u^k \xrightarrow{w} u, v^k \xrightarrow{w} u$ e $\lim_{k \rightarrow \infty} \|u^k - v^k\| = 0$ então

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

Na proposição seguinte, apresentamos uma equivalência para a condição alternativa B_2^0 .

Proposição 5.3.2: Considere U um espaço de Hilbert e $g : U \rightarrow (-\infty, +\infty]$ uma função de Bregman com zona S . Então a condição B_2^0 é equivalente a seguinte condição:

B_2^* : Se $\{u^k\}, \{v^k\} \subset S$ são seqüências limitadas tais que $\lim_{k \rightarrow \infty} \|u^k - v^k\| = 0$ então

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

Demonstração: A prova que B_2^* implica B_2^0 é imediata. Suponha que g satisfaz B_2^0 e considere

$$\{u^k\}, \{v^k\} \subset S \text{ limitadas tais que } \lim_{k \rightarrow \infty} \|u^k - v^k\| = 0. \quad (16)$$

Devemos provar que

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

Observemos que é suficiente mostrar que toda subsequência de $\{\nabla g(u^k) - \nabla g(v^k)\}$ contém uma subsequência que converge para zero. A fim de facilitar a notação, suponha que $\{\nabla g(u^k) - \nabla g(v^k)\}$ é uma subsequência arbitrária. Da limitação de $\{u^k\}$, segue que existem uma subsequência $\{u^{k_j}\}$ de $\{u^k\}$ e um ponto $\bar{u} \in \bar{S}$ tais que

$$u^{k_j} \xrightarrow{w} \bar{u} \quad (17)$$

Afirmamos que

$$v^{k_j} \xrightarrow{w} \bar{u}, \text{ ou seja, } \lim_{j \rightarrow \infty} |\langle v^{k_j} - \bar{u}, z \rangle| = 0 \quad \forall z \in U.$$

De fato, como

$$|\langle v^{k_j} - \bar{u}, z \rangle| \leq |\langle v^{k_j} - u^{k_j}, z \rangle| + |\langle u^{k_j} - \bar{u}, z \rangle| \quad \forall z \in U$$

e cada termo do lado direito da desigualdade acima converge para zero (segue de (16), (17) e da desigualdade de Cauchy -Schwartz) temos que,

$$\lim_{j \rightarrow \infty} |\langle v^{k_j} - \bar{u}, z \rangle| = 0 \quad \forall z \in U.$$

Daí, temos $v^{k_j} \xrightarrow{w} \bar{u}$. Agora, usando B_2^0 , obtemos

$$\lim_{j \rightarrow \infty} (\nabla g(u^{k_j}) - \nabla g(v^{k_j})) = 0.$$

Portanto, a prova está concluída.

Devemos observar que o resultado da Proposição 5.3.2 permanece válida no caso mais geral em que U é um espaço de Banach, uma vez que sua demonstração se segue de forma análoga.

Burachik e Scheimberg [15], consideram a condição B_2^* juntamente com a hipótese de convexidade totalmente uniforme para provar a convergência fraca do método do ponto proximal generalizado para o problema $VIP(\Psi, C)$ em espaços de Banach reflexivos. Neste trabalho, usamos esta condição sob a forma de B_2^0

Finalmente, mencionamos a seguinte condição para a função de Bregman g , conhecida na literatura:

Definição 5.3.3: A função de Bregman g é *zona coerciva* se a seguinte condição é satisfeita:

B_4 : (zona coerciva)

$\forall v \in U, \exists u \in S$ tal que $\nabla g(u) = v$, ou seja, a função $\nabla g : S \rightarrow U$ é sobrejetora.

Muitas vezes, as condições B_3 ou B_4 são responsáveis por manterem a seqüência gerada pelo algoritmo do ponto proximal generalizado contida no interior do conjunto C (veja, por exemplo, [12], [13], [15] e [36]). Neste trabalho, usamos a condição mais fraca (a saber, B_3) com a mesma finalidade.

Observação 5.3.1:

(i) A condição B_3 implica a seguinte condição:

Se $\{v^k\}_k \subset S$ é fracamente convergente para algum ponto na fronteira (∂S) de S então

$$\lim_{k \rightarrow \infty} \|\nabla g(v^k)\| = +\infty.$$

De fato, isto segue da convergência fraca de $\{v^k\}$ para um ponto na fronteira, da sua consequente limitação e da desigualdade de Cauchy-Schwartz

$$\|\nabla g(v^k)\| \|v^k - v\| \geq \langle \nabla g(v^k), v^k - v \rangle,$$

onde v é um ponto fixo em S .

(ii) Se g é zona coerciva então é fronteira coerciva. De fato, esta prova segue de forma similar à da Proposição 4 em [18];

(iii) As condições $B_1 - B_4$ são extensões naturais das correspondentes condições em \mathbf{R}^n (veja, por exemplo, [14], [18], [34], [35], [36] e [62]);

(iv) Em [27], prova-se que quando $S = \mathbf{R}^n$, uma condição suficiente para que uma função convexa e diferenciável seja Bregman é

$$\lim_{\|x\| \rightarrow +\infty} \frac{g(x)}{\|x\|} = +\infty;$$

(v) Quando $U = \mathbf{R}^n$, é fácil ver que B_2^0 implica B_2 .

A seguir, mostramos a incompatibilidade das condições B_3 e B_2^0 . Este fato é de extrema relevância, na medida em que estabelece a inviabilidade de qualquer resultado que considere as duas condições simultaneamente. Portanto B_4 e B_2^0 também não podem ser consideradas simultaneamente.

Proposição 5.3.3: Se g é uma função de Bregman fronteira coerciva então a condição B_2^0 não se verifica.

Demonstração: Suponha g Bregman fronteira coerciva e considere $\{v^k\} \subset S$ uma seqüência fracamente convergente para um ponto \bar{v} na fronteira (∂S) de S . Como $\bar{v} \in \partial S$, existe uma seqüência em S convergindo forte para \bar{v} . Seja $\{u^k\}$ tal seqüência. Daí, por B_3 ,

$$\lim_{k \rightarrow +\infty} \langle \nabla g(u^k), u - u^k \rangle = -\infty \quad \forall u \in S.$$

Logo (pela observação 5.3.1 (i)),

$$\lim_{k \rightarrow +\infty} \|\nabla g(u^k)\| = +\infty.$$

Daí resulta a existência de uma subsequência $\{u^{k_j}\}$ de $\{u^k\}$ satisfazendo as seguintes condições:

$$(i) \quad \|\nabla g(u^{k_0})\| > 0;$$

$$(ii) \quad \|\nabla g(u^{k_j})\| > 2 \|\nabla g(u^{k_{j-1}})\|, \forall j \geq 1.$$

Consideramos agora as seqüências $\{w^j\}$ e $\{z^j\}$ definidas abaixo:

$$w^j := u^{k_j}, \forall j \geq 0;$$

$$z^j := w^{j-1}, \forall j \geq 1.$$

Assim, temos, $\{w^j\}, \{z^j\} \subset S$, $w^j \xrightarrow{s} \bar{v}$, $z^j \xrightarrow{s} \bar{v}$ e $\lim_{j \rightarrow +\infty} \|w^j - z^j\| = 0$.

Além disso, usando a condição (ii), obtemos

$$\begin{aligned} \|\nabla g(w^j) - \nabla g(z^j)\| &\geq \|\nabla g(w^j)\| - \|\nabla g(z^j)\| \\ &> 2 \|\nabla g(u^{k_{j-1}})\| - \|\nabla g(u^{k_{j-1}})\| \\ &= \|\nabla g(u^{k_{j-1}})\|. \end{aligned}$$

Como $\lim_{j \rightarrow +\infty} \|\nabla g(u^{k_{j-1}})\| = +\infty$, a desigualdade acima implica que

$$\lim_{j \rightarrow +\infty} \|\nabla g(w^j) - \nabla g(z^j)\| = +\infty.$$

Portanto, B_2^0 não se verifica e a prova está concluída.

Note que provamos um resultado mais forte que o da Proposição 5.3.3. De fato, mostramos que

$\forall v \in \partial S, \exists \{u^k\}, \{v^k\} \subset S, u^k \xrightarrow{s} v, v^k \xrightarrow{s} v$ tais que

$$\lim_{j \rightarrow +\infty} \|\nabla g(u^k) - \nabla g(v^k)\| = +\infty.$$

Lema 5.3.1: Se g é uma função de Bregman fronteira coerciva então $\forall v \in S$ fixo

$$\partial_1 D_g(u, v) = \begin{cases} \nabla g(u) - \nabla g(v) & \text{se } u \in S \\ \emptyset & \text{se } u \notin S, \end{cases}$$

onde $\partial_1 D_g(u, v)$ denota o subdiferencial de D_g como uma função do seu primeiro argumento.

Demonstração: Veja [13], lema 1.

Apresentamos a seguir alguns exemplos de funções de Bregman encontrados na literatura.

Exemplo 5.3.1: Seja U um espaço de Hilbert arbitrário. Considere $S = U$ e $g(u) = \|u\|^2$. Neste caso $D_g(u, v) = \|u - v\|^2$. A prova de que g satisfaz $B_1 - B_4$ pode ser encontrada em [12] e [13].

Exemplo 5.3.2: Seja $U = \mathbf{R}^n$. Considere $S = \mathbf{R}_{++}^n$, onde $\mathbf{R}_{++}^n = \{u \in \mathbf{R}^n / u_j > 0 \forall j = 1, \dots, n\}$ e $g(u) = \sum_{j=1}^n u_j \log u_j$ é estendida por continuidade para $\partial \mathbf{R}_{++}^n$ com a convenção que $0 \log 0 = 0$. Neste caso,

$$D_g(u, v) = \sum_{j=1}^n \left(u_j \log \frac{u_j}{v_j} + v_j - u_j \right),$$

e é a divergência de Kullback-Leibler (veja, por exemplo, [42]) usada frequentemente em estatística.

Exemplo 5.3.3: Seja $U = \mathbf{R}^n$. Considere S como no exemplo anterior e $g(u) = \sum_{j=1}^n (u_j^\alpha - u_j^\beta)$ com $\alpha \geq 1$, $0 < \beta < 1$. Para $\alpha = 2$ e $\beta = \frac{1}{2}$ temos

$$D_g(u, v) = \|u - v\|^2 + \sum_{j=1}^n \frac{1}{2\sqrt{v_j}} (\sqrt{u_j} - \sqrt{v_j})^2,$$

e para $\alpha = 1$, $\beta = \frac{1}{2}$ temos

$$D_g(u, v) = \sum_{j=1}^n \frac{1}{2\sqrt{v_j}} (\sqrt{u_j} - \sqrt{v_j})^2.$$

As funções de Bregman dos exemplos acima são fronteiras coerciva. Elas também são zonas coercivas, com a exceção do exemplo 5.3.3 quando $\alpha = 1$. A verificação das propriedades das funções de Bregman para os Exemplos 5.3.2 e 5.3.3 é uma questão de cálculos. O exemplo de função de Bregman onde $U = \mathbf{R}^n$ e \bar{S} é uma caixa será visto na seção 6.6. Além disso, apresentaremos um exemplo de função de Bregman em um espaço de Hilbert arbitrário. Ressaltamos ainda que, funções de Bregman para poliedros em \mathbf{R}^n foram sugeridas em [18] por B.F. Svaiter.

6 O Método do Problema Auxiliar com regularização de Bregman

6.1 Introdução

A resolução do problema (*PIV*), no caso ponto-ponto, é o objetivo de vários métodos iterativos (veja, por exemplo, [21], [22], [31], [43], [44], [45], [46], [70] e [71]). A condição de Dunn para o operador Ψ é considerada, por exemplo, em [29], [44], [46], [65], [70] e [71]. Contudo, é comum na literatura o uso de condições mais fortes que Dunn, a saber, a convexidade forte e a propriedade de Lipschitz (veja, por exemplo, [21], [22], [45] e [54]).

Neste capítulo, estudamos o caso de um operador ponto-ponto monótono maximal munido da condição de Dunn em C . Introduzimos o método do problema auxiliar combinado com a noção de regularização de Bregman. Desenvolvemos um algoritmo no qual o funcional auxiliar de Cohen é substituído pela combinação linear deste funcional com a função de Bregman, o que faz com que o método seja de pontos interiores. Por outro lado, a distância de Bregman é substituída pela distância relativa à combinação citada acima. Além disso, o método desenvolvido neste capítulo é de decomposição (“Splitting”) do tipo “forward-backward” estendendo portanto aquele apresentado por Chen e Rockafellar [21] em \mathbf{R}^n .

A partir de agora, consideramos as seguintes condições para (*PIV*):

- U é um espaço de Hilbert;
- $C \subset U$ é um subconjunto convexo e fechado tal que $C^\circ \neq \emptyset$;
- $f : U \rightarrow (-\infty, +\infty]$ é uma função convexa, própria e s.c.i.;
- $\Psi : U \rightarrow U$ é um operador monótono maximal satisfazendo a condição de Dunn em C ;
- (ia) $(\text{Dom } f)^\circ \cap C^\circ \neq \emptyset$;
- (ib) $(C \cap \text{Dom } f) \subset \text{Dom } (\Psi)$.

Este problema pertence à família de problemas considerada por Chen e Rockafellar [21] (veja subseção 3.5.4). Como já mencionamos, o problema (*PIV*) pode ser reformulado como um problema de encontrar um zero para a soma de dois operadores, onde o primeiro, Ψ , é ponto-ponto e o segundo, $\partial f + N_C$, é ponto-conjunto.

As hipóteses (ia) e (ib) descritas anteriormente são consideradas para a existência de soluções do (PIV), para a formulação (EG) e também são usadas para a boa definição dos métodos apresentados ao longo deste trabalho.

As três primeiras seções se destinam à introdução do Problema Auxiliar com regularização de Bregman e do respectivo algoritmo. Na seção 6.4, estabelecemos a boa definição do método. Na seção 6.5 é feita a análise de convergência na qual garantimos a limitação da seqüência gerada pelo algoritmo quando o conjunto de soluções para (PIV) é não-vazio. Além disso, obtemos um resultado de convergência fraca. Entretanto, com uma condição extra para a função de Bregman (norma compatível), provamos o resultado de convergência fraca, para a seqüência toda, a uma solução do (PIV). Estabelecemos ainda, nesta seção, a equivalência entre a limitação da seqüência e a existência de soluções para (PIV). Por fim, na última seção, apresentamos exemplos de funções de Bregman em um espaço de Hilbert arbitrário e em \mathbf{R}^n que se adequam ao método. Mostramos que em cada exemplo o método do problema auxiliar como regularização de Bregman aplicado ao problema (PIV) gera uma seqüência que converge para uma solução do problema.

6.2 O Problema Auxiliar com regularização de Bregman

Dadas as seguintes funções:

- $H : U \rightarrow (-\infty, +\infty]$ uma função convexa, própria, G-diferenciável e contínua em $C \subset (Dom H)^o$;

- $g : U \rightarrow (-\infty, +\infty]$ uma função de Bregman com zona C^o ,

consideramos $G : U \rightarrow (-\infty, +\infty]$ a função definida por

$$G(u) := \alpha H(u) + \beta g(u) \quad \forall u \in U,$$

onde $\alpha, \beta \in \mathbf{R}_+$ são dados.

Sejam $v \in C^o \cap Dom(\Psi)$ e $\lambda > 0$. Introduzimos o seguinte subproblema auxiliar com regularização de Bregman:

(PAB) Encontrar $w = w(v) \in C$, solução de

$$\begin{aligned} &\text{minimizar} \quad \alpha H(u) + \langle \lambda \Psi(v) - \alpha \nabla H(v), u \rangle + \lambda f(u) + \beta D_g(u, v) . \\ &u \in C \end{aligned}$$

É fácil ver que a função objetivo do subproblema (PAB) é própria, convexa e contínua no interior do seu domínio (veja [67], página 82, comentário 1). Além disso, observe que (PAB) é a regularização de Bregman do problema auxiliar (PA) elaborado por Cohen (veja capítulo 4). Em particular, para $\beta = 0$, obtemos o problema auxiliar clássico (PA) de Cohen. Por outro lado, quando $\Psi \equiv 0$ e $\alpha = 0$, o problema (PAB) se reduz a um problema de minimização com distância de Bregman e obtém-se o método do ponto proximal com distância de Bregman para o problema de minimização convexa (veja, por exemplo, [20], [35] e [63]).

Observação 6.2.1: É fácil ver que se a condição de Bregman B_1 é satisfeita para g então também se verifica para G . A recíproca não é verdadeira. De fato, considere, por exemplo, $H(u) = e^u$, $g(u) = e^{-u}$, $\forall u \in \mathbf{R}$. A função G definida por $G := \alpha H + \beta g$ com $\alpha = \beta = \frac{1}{2}$ é função co-seno hiperbólico e portanto, satisfaz a condição B_1 . Entretanto o mesmo não ocorre para g ou H .

A seguir, apresentamos as formulações equivalentes para o subproblema (PAB) .

Teorema 6.2.1: Seja $v \in C^o \cap \text{Dom}(\Psi)$ e $\lambda > 0$. Os seguintes problemas são equivalentes:

(i) (PAB) Encontrar $w = w(v) \in C$, solução de

$$\text{minimizar } \alpha H(u) + \langle \lambda \Psi(v) - \alpha \nabla H(v), u \rangle + \lambda f(u) + \beta D_g(u, v); \\ u \in C$$

(ii) (\overline{PAB}) Encontrar $w = w(v) \in C$, solução de

$$\text{minimizar } \lambda f(u) + \langle \lambda \Psi(v), u \rangle + D_G(u, v), \\ u \in C$$

onde D_G é a função distância definida em (15) relativa a G ;

(iii) (EGA) Encontrar $w = w(v) \in C$, solução de

$$0 \in \lambda \Psi(v) + [\lambda \partial f + N_C + \partial_1 D_G(\cdot, v)](w);$$

(iv) (PFB) Encontrar $w = w(v) \in C : w \in S_\lambda(v)$,

onde

$$S_\lambda := (\partial G + \lambda \partial f + N_C)^{-1} (\partial G - \lambda \Psi);$$

(v) (PIVPA) Encontrar $w = w(v) \in C$, solução de

$$\langle \lambda \Psi(v) + \partial_1 D_G(w, v), u - w \rangle + \lambda(f(u) - f(w)) \geq 0 \quad \forall u \in C.$$

Demonstração: Esta prova segue em quatro etapas:

- (a) (i) \Leftrightarrow (ii);
- (b) (ii) \Leftrightarrow (iii);
- (c) (iii) \Leftrightarrow (iv);
- (d) (iii) \Leftrightarrow (v).

(a) (i) \Leftrightarrow (ii): De fato, observamos que para cada $u \in C$, temos

$$\begin{aligned} & \alpha H(u) + \langle \lambda \Psi(v) - \alpha \nabla H(v), u \rangle + \lambda f(u) + \beta D_g(u, v) \\ &= \alpha [H(u) - \langle \nabla H(v), u \rangle] + \langle \lambda \Psi(v), u \rangle + \lambda f(u) + \beta D_g(u, v) \\ &= \alpha [H(u) - H(v) - \langle \nabla H(v), u - v \rangle] + \alpha [H(v) - \langle \nabla H(v), v \rangle] \\ & \quad + \langle \lambda \Psi(v), u \rangle + \lambda f(u) + \beta D_g(u, v) \\ &= D_{\alpha H}(u, v) + D_{\beta g}(u, v) + \lambda f(u) + \langle \lambda \Psi(v), u \rangle + \alpha [H(v) - \langle \nabla H(v), v \rangle] \\ &= \lambda f(u) + \langle \lambda \Psi(v), u \rangle + D_{\alpha H + \beta g}(u, v) + F(v) \end{aligned}$$

onde $F(v) = \alpha [H(v) - \langle \nabla H(v), v \rangle]$.

Portanto, o problema (PAB) é equivalente ao problema (\overline{PAB}) .

(b) (ii) \Leftrightarrow (iii): Segue de forma inteiramente análoga à feita na seção 3.2.

(c) (iii) \Leftrightarrow (iv): Esta equivalência resulta do fato de que $\partial_1 D_G(u, v) = \partial G(u) - \partial G(v) \quad \forall u \in C$. Realmente,

$$\begin{aligned} & w \in C \text{ e } 0 \in \lambda \Psi(v) + [\lambda \partial f + N_C + \partial_1 D_G(\cdot, v)](w) \\ & \Leftrightarrow w \in C \text{ e } 0 \in \lambda \Psi(v) + (\lambda \partial f + N_C + \partial G)(w) - \partial G(v) \\ & \Leftrightarrow w \in C \text{ e } (\lambda \partial f + N_C + \partial G)(w) \in (\partial G - \lambda \Psi)(v) \\ & \Leftrightarrow w \in C \text{ e } w \in (\partial G + \lambda \partial f + N_C)^{-1} (\partial G - \lambda \Psi)(v) \\ & \Leftrightarrow w \in C \text{ e } w \in S_\lambda(v) \end{aligned}$$

(d) (iii) \Leftrightarrow (v): Segue do fato de que $(\lambda\partial f + N_C)(u) = \partial(\lambda f + I_C)(u)$ (veja, teorema 5, apêndice). Com efeito,

$$w \in C \text{ e } 0 \in \lambda\Psi(v) + [\lambda\partial f + N_C + \partial_1 D_G(\cdot, v)](w)$$

$$w \in C \text{ e } 0 \in (-\lambda\Psi(v) - \partial_1 D_G(u, v)) \in (\lambda\partial f + N_C)(w)$$

$$w \in C \text{ e } 0 \in (-\lambda\Psi(v) - \partial_1 D_G(u, v)) \in \partial(\lambda f + I_C)(w)$$

$$w \in C \text{ e } \langle \lambda\Psi(v) + \partial_1 D_G(w, v), u - w \rangle + \lambda(f(u) - f(w)) \geq 0 \quad \forall u \in C.$$

Note que o subproblema auxiliar com regularização de Bregman (PAB) é equivalente a um problema de inequação variacional no qual o operador é uma aproximação regular de Ψ , isto é,

$$\Psi(u) \cong \lambda\Psi(v) + \partial_1 D_G(u, v) \quad \forall u \in C.$$

Por outro lado, a equivalência entre os subproblemas (PAB) e o (PFB) significa que o nosso algoritmo pertence à família de métodos de decomposição do tipo “forward-backward” onde o passo “forward” realizado por Ψ é seguido pelo passo “backward” dado por $\partial f + N_C$ (veja subseção 3.5.4). Em particular, se $\beta = 0$, $\alpha = 1$ e $H(u) = \langle Mu, u \rangle / 2 \quad \forall u \in \mathbf{R}^n$, onde M é a matriz $(n \times n)$ simétrica e definida positiva, então reavemos o algoritmo desenvolvido por Chen e Rockafellar [21].

Apresentamos a seguir um resultado que garante a existência e unicidade da solução para (PAB). O conceito de coercividade citado abaixo é no sentido da Definição 2.1.11(iv)

Proposição 6.2.1: Se o conjunto C é limitado ou a função G é coerciva em C então $\forall v \in C^\circ \cap \text{Dom}(\Psi)$ e $\lambda > 0$ existe uma única solução $w = w(v) \in C$ para o problema (PAB).

Demonstração: De fato, este resultado segue da equivalência dos problemas (PAB) e (\overline{PAB}) mostrado no teorema anterior e do resultado de minimização para funções convexas em espaços de Banach reflexivos (veja [30], proposição 1.2).

Observe que a função G considerada na Proposição acima compreende o funcional auxiliar fortemente convexo introduzido por Cohen [22] e considerado em vários algoritmos (veja, por exemplo, [45], [60], [61] e [71]).

A seguir, mostramos que a condição de coercividade na fronteira (B_3) mantém a solução do subproblema (PAB) no interior de C . Este resultado é fundamental para a boa definição da seqüência gerada pelo nosso algoritmo.

Teorema 6.2.2: Sejam $v \in C^o \cap Dom(\Psi)$ e $w = w(v) \in C$ uma solução do (PAB). Se g é fronteira coerciva e $\beta > 0$ então $w \in C^o$.

Demonstração: Este resultado é consequência imediata do Lema 5.3.1.

Proposição 6.2.2: Seja $w = w(v) \in C$ uma solução do subproblema (PAB). Se $w = v$ então v é solução de $PIV(\Psi, C, f)$.

Demonstração: Suponha que v é solução de (PAB). Daí, a condição (v) do Teorema 6.2.1 é satisfeita para $w = v$, isto é,

$$v \in C \text{ e } \langle \Psi(v), u - v \rangle + f(u) - f(v) \geq 0 \quad \forall u \in C.$$

Logo, v é solução de $PIV(\Psi, C, f)$.

6.3 O algoritmo

De forma análoga ao que foi feito anteriormente, introduzimos o subproblema auxiliar com regularização de Bregman a cada iteração k .

Dados $u^k \in C^o \cap Dom(\Psi)$ e $\lambda_k > 0$ no início de cada iteração k , resolvemos o seguinte subproblema auxiliar com regularização de Bregman:

(PAB^k) Encontrar $u^{k+1} \in C$, solução de

$$\begin{aligned} & \text{minimizar} \quad \alpha H(u) + \langle \lambda_k \Psi(u^k) - \alpha \nabla H(u^k), u \rangle + \lambda_k f(u) + \beta D_g(u, u^k). \\ & u \in C \end{aligned}$$

O que é equivalente, por exemplo, a

(EGA^k) Encontrar $u^{k+1} \in C$, solução de

$$0 \in (\lambda_k \Psi + \nabla G)(u^k) + [\lambda_k \partial f + N_C + \partial G](u^{k+1}).$$

O algoritmo segue o esquema abaixo:

Passo 1: Na iteração $k = 0$, iniciamos com algum ponto $u^0 \in C^o \cap Dom(\Psi)$;

Passo 2: Na iteração k , encontramos a solução u^{k+1} do subproblema auxiliar (PAB^k);

Passo 3: Se $u^{k+1} = u^k$ então paramos. Caso contrário, repetimos o passo 2 com $k \leftarrow k + 1$.

6.4 Boa definição da seqüência

As condições consideradas na Proposição 6.2.1 garantem que a cada iteração k do algoritmo, existe uma única u^{k+1} solução de (PAB^k) . Além disso, a condição de coercividade na fronteira para a função g mantém u^{k+1} no interior do conjunto C . Logo, a boa definição da seqüência $\{u^k\}$ pode ser estabelecida se, por exemplo, as seguintes condições são satisfeitas:

- (i) $\beta > 0$;
- (ii) G é coerciva em C ;
- (iii) g é Bregman fronteira coerciva.

Observação 6.4.1: Seja $u^* \in C$. Então u^* é uma solução do problema (PIV) se, e somente se, u^* é uma solução do seguinte problema de minimização:

$$(\overline{P}) \quad \underset{u \in C}{\text{minimizar}} \quad \langle \Psi(u^*), u \rangle + f(u).$$

Por outro lado, como os problemas (PAB^k) e $(\overline{PAB^k})$ são equivalentes (veja teorema 6.2.1), concluímos que o algoritmo considera a cada iteração k , o elemento $u^{k+1} \in U$ solução de

$$(\overline{PAB^k}) \quad \underset{u \in C}{\text{minimizar}} \quad \langle \lambda_k \Psi(u^k), u \rangle + \lambda_k f(u) + D_G(u, u^k),$$

onde $G := \alpha H + \beta g$.

Logo, podemos pensar que o algoritmo aproxima o problema de minimização (\overline{P}) (ou “ PIV ”) por problemas de otimização $(\overline{PAB^k})$

Observação 6.4.2: A cada iteração k , o operador Ψ é calculado em u^k , a saber, é solução do supproblema definido na iteração anterior. Isto não ocorre no método do ponto proximal generalizado, uma vez que nele Ψ é calculado em u^{k+1} que é solução do subproblema definido na própria iteração. Esta diferença favorece o nosso algoritmo na medida em que facilita a implementação do método quando trabalhamos com funções de Bregman implementáveis.

6.5 Análise de convergência

Se a seqüência $\{u^k\}$ gerada pelo algoritmo na seção 6.3 está bem definida e é finita então a Proposição 6.2.2 garante que o último termo é uma solução de $PIV(\Psi, C, f)$.

O resultado abaixo estabelece a limitação da seqüência $\{u^k\}$ quando o conjunto de soluções para $PIV(\Psi, C, f)$ é não-vazio. Além disso, o Teorema 6.5.4 apresenta hipóteses sob as quais a limitação de $\{u^k\}$ e a existência de soluções são equivalentes.

Teorema 6.5.1: Suponha que as seguintes condições são satisfeitas:

A_1 : O operador $\Psi : U \rightarrow U$ é monótono maximal e satisfaz a condição de Dunn em C com constante $\gamma > 0$;

A_2 : O funcional H é fortemente convexo (módulo $c > 0$);

A_3 : A função de Bregman g é fronteira coerciva;

A_4 : $\alpha, \beta > 0$ e $\lambda_k \in (0, \bar{\lambda}) \forall k \in \mathbf{N}$ com $\bar{\lambda} = 2\alpha c \gamma$.

Se o problema $PIV(\Psi, C, f)$ possui solução então

(i) A seqüência $\{u^k\}$ gerada pelo algoritmo na seção 6.3 está bem definida;

(ii) A seqüência $\{u^k\}$ é limitada;

(iii) $\lim_{k \rightarrow \infty} \|u^{k+1} - u^k\| = 0$;

(iv) $\sum_{k=0}^{\infty} D_{G_1}(u^{k+1}, u^k) < \infty$,

onde $G_1 : U \rightarrow (-\infty, +\infty]$ é a função definida por

$$G_1(u) := \alpha \left[H(u) - \frac{c}{2} \|u\|^2 \right] + \beta g(u) \quad \forall u \in U.$$

Demonstração: (i) Como H é fortemente convexo, a Proposição 6.2.1 garante a existência e unicidade da solução u^{k+1} do subproblema (PAB^k) para cada $k \in \mathbf{N}$. Além disso, a coercividade na fronteira para g mantém u^{k+1} em C^o (teorema 6.2.2).

(ii) Seja u^* uma solução de $PIV(\Psi, C, f)$. Como $u^{k+1} \in C^o, \forall k \in \mathbf{N}$, a equação generalizada (EGA^k) implica

$$\lambda_k (\Psi(u^k) + v^{k+1}) = \nabla G(u^k) - \nabla G(u^{k+1})$$

onde $v^{k+1} \in \partial f(u^{k+1}) = (\partial f + N_C)(u^{k+1})$, já que $u^{k+1} \in C^o$. Além disso, usando a Proposição 5.2.2 para a função $G = \alpha H + \beta g$, temos

$$D_G(u^*, u^{k+1}) = D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \langle \nabla G(u^k) - \nabla G(u^{k+1}), u^* - u^{k+1} \rangle.$$

Portanto, das duas igualdades acima, obtemos

$$\begin{aligned} D_G(u^*, u^{k+1}) &= D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \lambda_k \langle \Psi(u^k) + v^{k+1}, u^* - u^{k+1} \rangle \\ &= D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \lambda_k \langle \Psi(u^k) - \Psi(u^*), u^* - u^{k+1} \rangle \\ &\quad + \lambda_k \langle v^{k+1} + \Psi(u^*), u^* - u^{k+1} \rangle \quad \forall k \in \mathbf{N}. \end{aligned} \tag{18}$$

A última igualdade segue da soma e subtração de $\Psi(u^*)$ no produto escalar.

Por outro lado, como u^* é solução do (PIV), temos

$$\langle \bar{\Psi}(u^*), u - u^* \rangle + (f + I_C)(u) - (f + I_C)(u^*) \geq 0 \quad \forall u \in \bar{U},$$

isto é,

$$-\Psi(u^*) \in (\partial f + N_C)(u^*).$$

Assim, pela monotonicidade do operador $(\partial f + N_C)$ temos

$$\langle v^{k+1} + \Psi(u^*), u^* - u^{k+1} \rangle = \langle v^{k+1} - (-\Psi(u^*)), u^* - u^{k+1} \rangle \leq 0. \tag{19}$$

Logo, de (18) e (19) resulta

$$\begin{aligned} D_G(u^*, u^{k+1}) &\leq D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \lambda_k \langle \Psi(u^k) - \Psi(u^*), u^* - u^{k+1} \rangle \\ &= D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \lambda_k \langle \Psi(u^k) - \Psi(u^*), u^* - u^k \rangle \\ &\quad + \lambda_k \langle \Psi(u^k) - \Psi(u^*), u^k - u^{k+1} \rangle \quad \forall k \in \mathbf{N}. \end{aligned} \tag{20}$$

A igualdade acima resulta da soma e subtração de u^k no produto escalar.

Mas, por A_1 , temos

$$\langle \Psi(u^k) - \Psi(u^*), u^k - u^* \rangle \geq \gamma \|\Psi(u^k) - \Psi(u^*)\|^2$$

e usando a desigualdade acima e a desigualdade de Cauchy-Schartz no último termo em (20), obtemos

$$D_G(u^*, u^{k+1}) \leq D_G(u^*, u^k) - D_G(u^{k+1}, u^k) - \lambda_k \gamma \|\Psi(u^k) - \Psi(u^*)\|^2$$

$$+\lambda_k \|\Psi(u^k) - \Psi(u^*)\| \|u^k - u^{k+1}\| \quad \forall k \in \mathbf{N}. \quad (21)$$

Agora, considerando em (21) a desigualdade $-a^2 + 2ab \leq b^2$ ($a, b \in \mathbf{R}$) para

$$a = \sqrt{\gamma} \|\Psi(u^k) - \Psi(u^*)\|$$

e

$$b = \frac{1}{2\sqrt{\gamma}} \|u^k - u^{k+1}\|,$$

resulta

$$D_G(u^*, u^{k+1}) \leq D_G(u^*, u^k) - D_G(u^{k+1}, u^k) + \frac{\lambda_k \|u^k - u^{k+1}\|^2}{4\gamma}, \quad \forall k \in \mathbf{N}. \quad (22)$$

Além disso, como

$$\begin{aligned} G(u) &= \alpha H(u) + \beta g(u) \\ &= \alpha \left[H(u) - \frac{c}{2} \|u\|^2 \right] + \frac{\alpha c}{2} \|u\|^2 + \beta g(u) \\ &= G_1(u) + \frac{\alpha c}{2} \|u\|^2, \end{aligned}$$

onde $G_1(u) = \alpha \left[H(u) - \frac{c}{2} \|u\|^2 \right] + \beta g(u)$ é uma função convexa e s.c.i. em C , temos

$$D_G(u, v) = D_{G_1}(u, v) + \frac{\alpha c}{2} \|u - v\|^2.$$

Então

$$D_G(u^{k+1}, u^k) = D_{G_1}(u^{k+1}, u^k) + \frac{\alpha c}{2} \|u^{k+1} - u^k\|^2, \quad \forall k \in \mathbf{N}.$$

Daí, a desigualdade (22) fica

$$\begin{aligned} D_G(u^*, u^{k+1}) &\leq D_G(u^*, u^k) - D_{G_1}(u^{k+1}, u^k) - \frac{\alpha c}{2} \|u^{k+1} - u^k\|^2 + \frac{\lambda_k}{4\gamma} \|u^k - u^{k+1}\|^2 \\ &= D_G(u^*, u^k) - D_{G_1}(u^{k+1}, u^k) - \frac{1}{2} \left(\alpha c - \frac{\lambda_k}{2\gamma} \right) \|u^k - u^{k+1}\|^2 \\ &\leq D_G(u^*, u^k), \quad \forall k \in \mathbf{N} \end{aligned} \quad (23)$$

pois $\lambda_k < 2\alpha c\gamma$, $\forall k \in \mathbf{N}$ (por hipótese).

Portanto,

$$D_G(u^*, u^k) \leq D_G(u^*, u^0), \forall k \in \mathbf{N}.$$

Logo, $\{u^k\} \subset S_{u^*, \delta}$ para $\delta \geq D_G(u^*, u^0)$ e é limitada pela condição B_1 para G (veja observação 6.2.1).

(iii) Além disso, $\forall k \in \mathbf{N}$, temos

$$0 \leq D_G(u^*, u^{k+1}) \leq D_G(u^*, u^k) - D_{G_1}(u^{k+1}, u^k) \quad (24)$$

$$0 \leq \frac{1}{2} \left(\alpha c - \frac{\lambda_k}{2\gamma} \right) \|u^k - u^{k+1}\|^2 \leq D_G(u^*, u^k) - D_G(u^*, u^{k+1}) \quad (25)$$

As desigualdades em (24) implicam que $\{D_G(u^*, u^k)\}$ é uma seqüência decrescente e limitada inferiormente. Então $\{D_G(u^*, u^k)\}$ converge para um número positivo $\mu \geq 0$. Assim, por (25), temos

$$\lim_{k \rightarrow \infty} \|u^{k+1} - u^k\| = 0.$$

(iv) Finalmente, observe que

$$0 \leq D_{G_1}(u^{k+1}, u^k) \leq D_G(u^*, u^k) - D_G(u^*, u^{k+1}) \quad \forall k \in \mathbf{N} \quad (26)$$

e

$$\sum_{k=0}^{\infty} (D_G(u^*, u^k) - D_G(u^*, u^{k+1})) < \infty. \quad (27)$$

Portanto, segue de (26) e (27) que

$$\sum_{k=0}^{\infty} D_{G_1}(u^{k+1}, u^k) < \infty.$$

Observação 6.5.1: Note que o resultado do Teorema acima permanece válido se B_1 é satisfeita apenas para a função G .

O Teorema 6.5.1 estabelece a limitação da seqüência e portanto a existência de pelo menos um ponto de acumulação fraco. Enunciamos a seguir duas hipóteses sob as quais os pontos de acumulação fracos são soluções do problema original $PIV(\Psi, C, f)$. A primeira hipótese é a propriedade de Lipschitz para o gradiente do funcional H . A segunda é a pseudomonotonia para o operador $\Psi + \partial f$.

Teorema 6.5.2: Suponha que, além das hipóteses do Teorema 6.5.1, são satisfeitas as seguintes condições:

A_5 : O gradiente ∇H é Lipschitz em C^o com constante de Lipschitz $\Lambda > 0$;

A_6 : O operador $\Psi + \partial f$ é pseudomonótono com domínio fechado (e portanto convexo pela Proposição 2.1.6 (iii)).

Se existe um número $\underline{\lambda} > 0$ tal que $\lambda_k \in (\underline{\lambda}, \bar{\lambda}) \quad \forall k \in \mathbf{N}$ então qualquer ponto de acumulação fraco de $\{u^k\}$ é solução de PIV (Ψ, C, f) .

Demonstração: Seja \bar{u} um ponto de acumulação fraco de $\{u^k\}$. Como C é um subconjunto fechado e convexo de U e $\{u^k\} \subset C$ temos que

$$\bar{u} \in \overline{C}^w = \bar{C} = C,$$

sendo \overline{C}^w o fecho fraco do conjunto C .

De modo análogo, $\bar{u} \in \text{Dom}(\Psi + \partial f)$. Considere $\{u^{k_j}\} \subset \{u^k\}$ uma subsequência tal que

$$u^{k_j} \xrightarrow{w} \bar{u} \in C.$$

Como

$$\lim_{j \rightarrow \infty} \|u^{k_{j+1}} - u^{k_j}\| = 0,$$

resulta $u^{k_{j+1}} \xrightarrow{w} \bar{u}$. Além disso, temos

$$D_G(u^{k_{j+1}}, u^{k_j}) = D_{G_1}(u^{k_{j+1}}, u^{k_j}) + \frac{\alpha c}{2} \|u^{k_{j+1}} - u^{k_j}\|^2 \quad \forall j \in \mathbf{N}$$

e

$$\lim_{j \rightarrow \infty} D_{G_1}(u^{k_{j+1}}, u^{k_j}) = 0.$$

Então

$$\lim_{j \rightarrow \infty} D_G(u^{k_{j+1}}, u^{k_j}) = 0.$$

Mas

$$0 \leq \beta D_g(u^{k_{j+1}}, u^{k_j}) \leq D_G(u^{k_{j+1}}, u^{k_j}).$$

Então

$$\lim_{j \rightarrow \infty} D_g(u^{k_{j+1}}, u^{k_j}) = 0.$$

Assim, temos

- (i) $\{u^{k_j}\}, \{u^{k_{j+1}}\} \subset C^0$;
- (ii) $u^{k_j} \xrightarrow{w} \bar{u}, u^{k_{j+1}} \xrightarrow{w} \bar{u}$;
- (iii) $\lim_{j \rightarrow \infty} D_g(u^{k_{j+1}}, u^{k_j}) = 0$.

Então, por B_2 , obtemos

$$\lim_{j \rightarrow \infty} (D_g(\bar{u}, u^{k_{j+1}}) - D_g(\bar{u}, u^{k_j})) = 0. \quad (28)$$

Por outro lado, como $\bar{u} \in C$, temos que

$$\lambda_{k_j} \langle \Psi(u^{k_j}) + v^{k_{j+1}}, u^{k_{j+1}} - \bar{u} \rangle = D_G(\bar{u}, u^{k_j}) - D_G(\bar{u}, u^{k_{j+1}}) - D_G(u^{k_{j+1}}, u^{k_j}).$$

Daí,

$$\begin{aligned} \lambda_{k_j} \langle \Psi(u^{k_{j+1}}) + v^{k_{j+1}}, u^{k_{j+1}} - \bar{u} \rangle &= D_G(\bar{u}, u^{k_j}) - D_G(\bar{u}, u^{k_{j+1}}) - D_G(u^{k_{j+1}}, u^{k_j}) \\ &\quad + \lambda_{k_j} \langle \Psi(u^{k_{j+1}}) - \Psi(u^{k_j}), u^{k_{j+1}} - \bar{u} \rangle \\ &= \alpha [D_H(\bar{u}, u^{k_j}) - D_H(\bar{u}, u^{k_{j+1}})] + \beta [D_g(\bar{u}, u^{k_j}) - D_g(\bar{u}, u^{k_{j+1}})] \\ &\quad - D_G(u^{k_{j+1}}, u^{k_j}) + \lambda_{k_j} \langle \Psi(u^{k_{j+1}}) - \Psi(u^{k_j}), u^{k_{j+1}} - \bar{u} \rangle \\ &= \alpha D_H(u^{k_{j+1}}, u^{k_j}) + \alpha \langle \nabla H(u^{k_j}) - \nabla H(u^{k_{j+1}}), u^{k_{j+1}} - \bar{u} \rangle \\ &\quad + \beta [D_g(\bar{u}, u^{k_j}) - D_g(\bar{u}, u^{k_{j+1}})] - D_G(u^{k_{j+1}}, u^{k_j}) \\ &\quad + \lambda_{k_j} \langle \Psi(u^{k_{j+1}}) - \Psi(u^{k_j}), u^{k_{j+1}} - \bar{u} \rangle \\ &\leq \alpha D_H(u^{k_{j+1}}, u^{k_j}) - D_G(u^{k_{j+1}}, u^{k_j}) + \alpha \Lambda \|u^{k_{j+1}} - u^{k_j}\| \|u^{k_{j+1}} - \bar{u}\| \\ &\quad + \beta [D_g(\bar{u}, u^{k_j}) - D_g(\bar{u}, u^{k_{j+1}})] + \lambda_{k_j} L \|u^{k_{j+1}} - u^{k_j}\| \|u^{k_{j+1}} - \bar{u}\| \\ &\leq \beta [D_g(\bar{u}, u^{k_j}) - D_g(\bar{u}, u^{k_{j+1}})] + (\alpha \Lambda + \lambda_{k_j} L) \|u^{k_{j+1}} - u^{k_j}\| \|u^{k_{j+1}} - \bar{u}\|. \end{aligned}$$

A primeira desigualdade acima segue da propriedade de Lipschitz de Ψ e ∇H com respectivas constantes $L > 0$ e $\Lambda > 0$ juntamente com a desigualdade de Cauchy-Schwartz.

Portanto, usando (28) e a limitação das seqüências $\{\lambda_{k_j}\}$ e $\{u^{k_j+1}\}$ juntamente com a conclusão (iii) do Teorema 6.5.1, obtemos

$$\limsup_{j \rightarrow \infty} \lambda_{k_j} \langle \Psi(u^{k_j+1}) + v^{k_j+1}, u^{k_j+1} - \bar{u} \rangle \leq 0$$

e conseqüentemente, como $0 < \underline{\lambda} < \lambda_k, \forall k \in \mathbf{N}$, temos

$$\limsup_{j \rightarrow \infty} \langle \Psi(u^{k_j+1}) + v^{k_j+1}, u^{k_j+1} - \bar{u} \rangle \leq 0.$$

Além disso, $\{u^{k_j+1}\} \subset \text{Dom}(\Psi + \partial f)$, $\bar{u} \in \text{Dom}(\Psi + \partial f)$, $u^{k_j+1} \xrightarrow{w} \bar{u}$ e $\Psi(u^{k_j+1}) + v^{k_j+1} \in (\Psi + \partial f)(u^{k_j+1}) \forall j \in \mathbf{N}$. Logo, pela pseudomonotonia de $(\Psi + \partial f)$, temos que, em particular, para qualquer solução u^* de $PIV(\Psi, C, f)$, existe $v \in (\Psi + \partial f)(\bar{u})$, tal que

$$\langle v, \bar{u} - u^* \rangle \leq \liminf_{j \rightarrow \infty} \langle \Psi(u^{k_j+1}) + v^{k_j+1}, u^{k_j+1} - u^* \rangle.$$

Por outro lado, temos

$$\begin{aligned} \lambda_{k_j} \langle \Psi(u^{k_j+1}) + v^{k_j+1}, u^{k_j+1} - u^* \rangle &\leq D_G(u^*, u^{k_j}) - D_G(u^*, u^{k_j+1}) - D_G(u^{k_j+1}, u^{k_j}) \\ &\quad + \lambda_{k_j} L \|u^{k_j+1} - u^{k_j}\| \|u^{k_j+1} - u^*\|. \end{aligned}$$

Como $\{D_G(u^*, u^{k_j})\}$ converge para um número $\mu \geq 0$ e $\{D_G(u^{k_j+1}, u^{k_j})\}$ converge para 0, obtemos

$$\liminf_{j \rightarrow \infty} \lambda_{k_j} \langle \Psi(u^{k_j+1}) + v^{k_j+1}, u^{k_j+1} - u^* \rangle \leq 0$$

Portanto,

$$\langle v, \bar{u} - u^* \rangle \leq 0 \tag{29}$$

Finalmente, observando que o operador $\Psi + \partial f$ é paramonótono em C (proposição 5.2.1 (ii)) e usando a desigualdade acima o restante da prova segue imediatamente da Proposição 2.1.11 (iii).

Observação 6.5.2: Note que a condição A_6 é automaticamente satisfeita quando a função f é G -diferenciável. De fato, neste caso o operador ∇f é pseudomonótono (veja [12], corolário 2.1). Portanto, A_6 segue da pseudomonotonia de Ψ (proposição 2.1.7) e da Proposição 2.1.8(ii).

Observação 6.5.3: O resultado do Teorema acima permanece válido se a condição B_2 é satisfeita apenas para a função G . Além disso, é fácil ver que, sob a hipótese A_5 , a condição de coercividade na fronteira é válida para g se, e somente se, é válida para G .

O resultado de convergência fraca obtido no Teorema 6.5.2 é estabelecido para a seqüência toda se, por exemplo, o problema original $PIV(\Psi, C, f)$ possui uma única solução. Apresentamos a seguir outra condição sob a qual a seqüência toda converge fracamente para uma solução. Trata-se de uma condição adicional para a função G que enunciamos abaixo (veja [12], definição 5.3).

Definição 6.5.1: Uma função de Bregman g é *norma compatível* se satisfaz a seguinte condição:

Se $\{u^k\}, \{v^k\} \subset S$, $u^k \xrightarrow{w} u$, $v^k \xrightarrow{w} v$ e $u \neq v$ então

$$\liminf_{k \rightarrow +\infty} |\langle \nabla g(u^k) - \nabla g(v^k), u - v \rangle| > 0.$$

Teorema 6.5.3: Suponha que, além das hipóteses do Teorema 6.5.2, pelo menos uma das seguintes condições é satisfeita:

(i) O problema original $PIV(\Psi, C, f)$ possui uma única solução;

(ii) A função G é norma compatível.

Então a seqüência $\{u^k\}$ converge fracamente para uma solução de $PIV(\Psi, C, f)$, ou seja, existe um único ponto de acumulação fraco de $\{u^k\}$.

Demonstração: O resultado é evidente quando (i) se verifica. Suponha então a hipótese (ii). Devemos mostrar que existe apenas um ponto de acumulação fraco. Suponha que existam dois pontos \bar{u}_1 e \bar{u}_2 limites fracos de subsequências de $\{u^k\}$. Pelo Teorema 6.5.2, temos que \bar{u}_1 e \bar{u}_2 são soluções de $PIV(\Psi, C, f)$. A desigualdade (24) garante que as seqüências $\{D_G(\bar{u}_1, u^k)\}$ e $\{D_G(\bar{u}_2, u^k)\}$ são convergentes.

Considere μ_1 e μ_2 seus respectivos limites. Então

$$\begin{aligned} \lim_{k \rightarrow \infty} (D_G(\bar{u}_1, u^k) - D_G(\bar{u}_2, u^k)) &= \mu_1 - \mu_2 \\ &= G(\bar{u}_1) - G(\bar{u}_2) \\ &\quad + \lim_{k \rightarrow \infty} \langle \nabla G(u^k), \bar{u}_2 - \bar{u}_1 \rangle. \end{aligned} \tag{30}$$

Seja $\mu := \lim_{k \rightarrow \infty} \langle \nabla G(u^k), \bar{u}_2 - \bar{u}_1 \rangle$. Se $\{u^{k_j}\}$ e $\{u^{l_j}\}$ são subsequências de $\{u^k\}$ tais que $u^{k_j} \xrightarrow{w} \bar{u}_1$ e $u^{l_j} \xrightarrow{w} \bar{u}_2$ então considerando, em (30), $k = k_j$, $k = l_j$ e subtraindo as expressões resultantes, obtemos

$$\lim_{j \rightarrow \infty} \langle \nabla G(u^{k_j}) - \nabla G(u^{l_j}), \bar{u}_2 - \bar{u}_1 \rangle = 0.$$

Como $\bar{u}_1 \neq \bar{u}_2$, a igualdade acima contradiz a hipótese de norma compatível para a função G . Logo, a unicidade do ponto de acumulação fraco é estabelecida.

O resultado abaixo prova que a limitação da seqüência $\{u^k\}$ é uma condição suficiente para a existência de soluções do problema $PIV(\Psi, C, f)$.

Teorema 6.5.4: Suponha que as hipóteses do Teorema 6.5.2 são satisfeitas. Se o problema original $PIV(\Psi, C, f)$ não possui solução então a seqüência é ilimitada.

Demonstração: Suponha que (PIV) não possui solução e $\{u^k\}$ é limitada. Então existem um subconjunto convexo, fechado e limitado $S \subset U$ e um número positivo r tais que

$$\overline{\{u^k\}}^w \subset B(0, r) \subset S^o,$$

onde $\overline{\{u^k\}}^w$ denota o fecho fraco e $B(0, r) := \{y \in U : \|y\| \leq r\}$ é a bola fechada em U .

Considere a função \tilde{f} e os operadores T e \tilde{T} definidos por:

$$\tilde{f} := f + I_S, \quad T := \partial f + N_C \quad \text{e} \quad \tilde{T} := T + N_S,$$

onde I_S é a função indicatriz do conjunto S e N_S é o seu subdiferencial, ou seja, o operador cone normal associado ao conjunto S . O operador \tilde{T} é monótono maximal

pela Proposição 2.1.3 já que $Dom(T) \cap S^o \supset \{u^k\} \neq \emptyset$. Consideramos $\{\tilde{u}^k\}$ a seqüência gerada pelo algoritmo para o operador \tilde{T} e provamos o teorema em quatro etapas.

Passo 1: $\{\tilde{u}^k\}$ está bem definida, ou seja, existe um único $\tilde{u}^{k+1} \in C$ solução de

$$0 \in \lambda_k \Psi(\tilde{u}^k) + \left[\lambda_k \tilde{T} + \partial_1 D_G(\cdot, \tilde{u}^k) \right] (\tilde{u}^{k+1}) \quad (31)$$

De fato, o operador soma $(\lambda_k \tilde{T} + \nabla G)$ é sobrejetivo já que é monótono maximal e possui domínio limitado (proposição 2.1.2). Então existe $\tilde{u}^{k+1} \in C$ solução de

$$(\nabla G - \lambda_k \Psi)(\tilde{u}^k) \in (\lambda_k \tilde{T} + \nabla G)(\tilde{u}^{k+1}).$$

Portanto, \tilde{u}^{k+1} satisfaz a equação (31) e a unicidade de \tilde{u}^{k+1} segue da monotonicidade estrita do operador ∇G .

Passo 2: Se $\tilde{u}^0 = u^0$ no algoritmo aplicado para \tilde{T} então $\tilde{u}^k = u^k \forall k \in \mathbf{N}$.

Provamos esta afirmação por indução. De fato, para $k = 0$, a afirmação é óbvia. Suponha que $\tilde{u}^k = u^k$. O ponto u^{k+1} satisfaz

$$(\nabla G - \lambda_k \Psi)(u^k) \in (\lambda_k T + \nabla G)(u^{k+1}),$$

ou seja,

$$(\nabla G - \lambda_k \Psi)(\tilde{u}^k) \in (\lambda_k T + \nabla G)(u^{k+1}). \quad (32)$$

Como $u^{k+1} \in S^o$, temos que $N_S(u^{k+1}) = \{0\}$ e assim

$$\begin{aligned} (\lambda_k T + \nabla G)(u^{k+1}) &= (\lambda_k T + N_S + \nabla G)(u^{k+1}) \\ &= (\lambda_k \tilde{T} + \nabla G)(u^{k+1}). \end{aligned} \quad (33)$$

Agora, substituindo (33) em (32) resulta

$$(\nabla G - \lambda_k \Psi)(\tilde{u}^k) \in (\lambda_k \tilde{T} + \nabla G)(u^{k+1}).$$

Usando a unicidade de \tilde{u}^{k+1} na equação acima mostrada no passo 1 temos que $\tilde{u}^{k+1} = u^{k+1}$.

Passo 3: Se \bar{u} é um ponto de acumulação fraco de $\{\tilde{u}^k\}$ então \bar{u} é solução de $PIV(\Psi, C, \tilde{f})$.

De fato, vamos provar que as hipóteses do Teorema 6.5.2 são satisfeitas para a função \tilde{f} . Considere o operador $\tilde{\Psi} : U \rightarrow P(U)$ definido por

$$\tilde{\Psi} := \Psi + N_C + \partial f + N_S.$$

O operador $\tilde{\Psi}$ é monótono maximal e $Dom(\tilde{\Psi}) \subset S$ é limitado. Daí, $\tilde{\Psi}$ é sobrejetivo (proposição 2.1.2) e conseqüentemente possui zeros. Então $PIV(\Psi, C, \tilde{f})$ tem solução.

Por outro lado, as condições $A_1 - A_5$ são trivialmente satisfeitas, pois não dependem da função definida no problema (PIV). Para provar a validade de A_6 , basta observar que

$$\Psi + \partial(f + I_S) = \Psi + \partial f + \partial I_S$$

é pseudomonótono (proposição 2.1.8 (i) – (ii)) e seu domínio é fechado pois

$$Dom(\Psi + \partial f + \partial I_S) = Dom(\Psi + \partial f) \cap S.$$

Passo 4: Se \bar{u} é um ponto de acumulação fraco de $\{\tilde{u}^k\}$ então \bar{u} é solução de $PIV(\Psi, C, f)$.

Com efeito, suponha \bar{u} um ponto de acumulação fraco de $\{\tilde{u}^k\}$. Assim, usando o passo 2, temos que

$$\bar{u} \in \overline{\{\tilde{u}^k\}}^w = \overline{\{u^k\}}^w \subset S^\circ.$$

Além disso, pelo passo 3, \bar{u} é solução de $PIV(\Psi, C, \tilde{f})$. Portanto, satisfaz a seguinte inclusão (veja seção 3.2)

$$-\Psi(\bar{u}) \in (\partial \tilde{f} + N_C)(\bar{u}). \quad (34)$$

Mas, $\bar{u} \in S^\circ$. Então $N_S(\bar{u}) = 0$. Daí,

$$\begin{aligned} \partial \tilde{f}(\bar{u}) &= \partial f(\bar{u}) + N_S(\bar{u}) \\ &= \partial f(\bar{u}). \end{aligned}$$

Assim, a inclusão (34) fica

$$-\Psi(\bar{u}) \in (\partial f + N_C)(\bar{u}).$$

Isto significa que \bar{u} é solução de $PIV(\Psi, C, f)$.

Finalmente, como $\{\tilde{u}^k\}$ é limitada (pois $\tilde{u}^k = u^k, \forall k$), existe um ponto, \bar{u} , de acumulação fraco de $\{\tilde{u}^k\}$. Pelo passo 4, \bar{u} é solução de $PIV(\Psi, C, f)$. Isto contradiz a hipótese e portanto, segue que $\{u^k\}$ é ilimitada.

O seguinte corolário é consequência imediata dos Teoremas 6.5.1 e 6.5.4.

Corolário 6.5.1: Suponha que as hipóteses do Teorema 6.5.2 são satisfeitas. O problema (PIV) tem solução se, e somente se, a seqüência gerada pelo algoritmo na seção 6.3 é limitada.

6.6 Aplicações

Apresentamos abaixo dois exemplos para a função G . Mostramos que em cada um o Método do Problema Auxiliar com regularização de Bregman aplicado ao problema de inequação variacional (PIV) gera uma seqüência que converge para uma solução do problema.

Exemplo 6.6.1: Considere U um espaço de Hilbert e $\{a^1, \dots, a^m\} \subset U$. Seja

$$C := \{u \in U / \langle a^i, u \rangle \geq b_i, i = 1, \dots, m\}, \quad (35)$$

onde $b_1, \dots, b_m \in \mathbf{R}$. Suponha $C \cap \text{Dom } f \subset \text{Dom } (\Psi)$ e $C^\circ \cap (\text{Dom } f)^\circ \neq \emptyset$. Considere

$$G(u) := \beta \sum_{i=1}^m (\langle a^i, u \rangle - b_i) \log (\langle a^i, u \rangle - b_i) + \frac{\alpha \|u\|^2}{2}, \quad (36)$$

onde:

(i) $\alpha, \beta > 0$;

(ii) $g(u) := \sum_{i=1}^m (\langle a^i, u \rangle - b_i) \log (\langle a^i, u \rangle - b_i) \quad \forall u \in C^\circ$ estendida continuamente para a fronteira (∂C) de C com a convenção de que $0 \log 0 = 0$;

(iii) $H(u) := \frac{1}{2} \|u\|^2 \quad \forall u \in U$.

Observe que quando $\alpha = \beta = 1$ obtemos a função de Bregman considerada por Burachik [12]. Além disso, de forma inteiramente análoga ao feito em [12], prova-se que a função G é Bregman com zona C° fronteira coerciva e norma compatível. Logo, seguem os seguintes resultados:

Proposição 6.6.1: Considere G a função dada por (36). Então G é Gateaux-diferenciável em C^o e o seu gradiente é dado por

$$\nabla G(u) = \beta \sum_{i=1}^m (\log(\langle a^i, u \rangle - b_i) + 1) a^i + \alpha u.$$

Proposição 6.6.2: Se G é a função dada por (36) então satisfaz as condições $B_1 - B_3$ da definição 5.3.1 Além disso, a condição B_3 para a função g segue da observação 6.5.3.

Proposição 6.6.3: A função G dada por (36) é norma compatível.

Segue das Proposições 6.6.1 - 6.6.3, que as hipóteses do Teorema 6.5.3 são satisfeitas e assim o Método do Problema Auxiliar com regularização de Bregman para G dada por (36) aplicado ao problema de inequação variacional com C dado por (35) gera uma seqüência que converge fracamente para uma solução do problema.

Exemplo 6.6.2: Seja $C \subset U = \mathbf{R}^n$ uma caixa, isto é, um produto cartesiano de intervalos definidos por

$$C := [a_1, b_1] \times \dots \times [a_n, b_n], \quad (37)$$

onde $a_i < b_i$ ($1 \leq i \leq n$). Suponha, novamente satisfeitas as hipóteses habituais para o problema, ou seja, que $C \cap \text{Dom } f \subset \text{Dom } (\Psi)$ e $C^o \cap (\text{Dom } f)^o \neq \emptyset$. Considere, M uma matriz $n \times n$ simétrica definida positiva e

$$G(u) := \beta \sum_{i=1}^n [(u_i - a_i) \log(u_i - a_i) + (b_i - u_i) \log(b_i - u_i)] + \frac{\alpha}{2} \langle u, Mu \rangle, \quad (38)$$

onde:

(i) $\alpha, \beta > 0$;

(ii) $g(u) := \sum_{i=1}^n [(u_i - a_i) \log(u_i - a_i) + (b_i - u_i) \log(b_i - u_i)] \quad \forall u \in C^o$ estendida continuamente para a fronteira ∂C de C com a convenção de que $0 \log 0 = 0$;

(iii) $H(u) := \frac{1}{2} \|u\|_M^2 = \frac{1}{2} \langle u, Mu \rangle \quad \forall u \in U$.

Observação 6.6.1: A função g dada acima foi considerada por Iusem *et al* [18] e é um caso particular da função apresentada no Exemplo 6.6.1.

A seguir, mostramos que G e g são funções de Bregman com zona C^o fronteiras coercivas. Além disso, a função G é norma compatível.

Proposição 6.6.4: A função G dada por (38) é Gateaux-diferenciável em C^o e seu gradiente é dado por

$$\nabla G(u) = \beta \sum_{i=1}^n (\log(u_i - a_i) - \log(b_i - u_i)) e_i + \alpha M u,$$

onde $\{e_1, \dots, e_n\}$ é a base canônica do \mathbf{R}^n .

Demonstração: Esta prova segue de simples cálculos.

Note que a função G é continuamente diferenciável em C^o .

Proposição 6.6.5: As funções g e G dadas por (38) são de Bregman com zona C^o fronteiras coercivas.

Demonstração: A prova que g satisfaz as condições $B_1 - B_3$ da Definição 5.3.1 pode ser encontrada em [18], Exemplo 3. Além disso, as condições B_1 e B_3 para G seguem respectivamente das Observações 6.5.1 e 6.5.3.

Por outro lado, é fácil mostrar que a distância de Bregman em relação a G é dada por

$$D_G(u, v) = \beta D_g(u, v) + \frac{\alpha}{2} \langle u - v, M(u - v) \rangle = \beta D_g(u, v) + \frac{\alpha}{2} \|u - v\|_M^2. \quad (39)$$

Portanto, a condição B_2 para G decorre imediatamente de B_2 para g e de (39).

Proposição 6.6.6: A função G dada por (38) é norma compatível.

Demonstração: Suponha $\{u^k\}, \{v^k\} \subset C^o$ tais que $u^k \rightarrow u, v^k \rightarrow v$ e $u \neq v$. Considere a seguinte expressão:

$$\begin{aligned} \sigma &:= \liminf_{k \rightarrow \infty} |\langle \nabla G(u^k) - \nabla G(v^k), u - v \rangle| \\ &= \liminf_{k \rightarrow \infty} \left| \alpha \langle u^k - v^k, M(u - v) \rangle + \beta \sum_{i=1}^n \log \left[\frac{(u_i^k - a_i)(b_i - v_i^k)}{(v_i^k - a_i)(b_i - u_i^k)} \right] (u_i - v_i) \right|. \end{aligned} \quad (40)$$

Vamos mostrar agora que o termo à direita em (40) é positivo para k suficientemente grande. Para analisar cada termo da soma em (40), devemos considerar dois casos:

(i) $u_i = v_i$.

Neste caso o termo considerado é nulo.

(ii) $u_i > v_i$.

Então existe $k_0 \in \mathbb{N}$ tal que

$$\frac{(u_i^k - a_i)(b_i - v_i^k)}{(v_i^k - a_i)(b_i - u_i^k)} \geq 1, \forall k \geq k_0.$$

Note que para a desigualdade contrária, o resultado acima segue de forma análoga. Assim,

$$\log \left[\frac{(u_i^k - a_i)(b_i - v_i^k)}{(v_i^k - a_i)(b_i - u_i^k)} \right] (u_i - v_i) \geq 0, \forall k \geq k_0. \quad (41)$$

Daí,

$$\begin{aligned} \sigma &\geq \liminf_{k \rightarrow \infty} \left\{ \alpha \langle u^k - v^k, M(u - v) \rangle + \beta \sum_{i=1}^n \log \left[\frac{(u_i^k - a_i)(b_i - v_i^k)}{(v_i^k - a_i)(b_i - u_i^k)} \right] (u_i - v_i) \right\} \\ &\geq \alpha \liminf_{k \rightarrow \infty} \langle u^k - v^k, M(u - v) \rangle + \beta \liminf_{k \rightarrow \infty} \sum_{i=1}^n \log \left[\frac{(u_i^k - a_i)(b_i - v_i^k)}{(v_i^k - a_i)(b_i - u_i^k)} \right] (u_i - v_i) \\ &\geq \alpha \|u - v\|_M^2 > 0. \end{aligned}$$

onde a terceira desigualdade na expressão acima segue de (41).

Novamente, temos satisfeitas as hipóteses do Teorema 6.5.3. Portanto, o algoritmo do problema auxiliar com regularização de Bregman para G dado por (38) aplicado ao problema de inequação variacional com C dado por (37) gera uma seqüência que converge para uma solução do problema.

Conclusões

Neste trabalho, desenvolvemos algoritmos de decomposição do tipo “forward-backward” para resolver o (PIV) em espaços de Hilbert. Generalizamos o método do problema auxiliar proposto por Cohen para o operador ponto-conjunto e introduzimos uma classe de algoritmos que combina o princípio do problema auxiliar com a noção de regularização de Bregman para o operador ponto-ponto.

Se encontra em andamento a análise de convergência de um método para o (PIV) no caso em que a função f é identicamente nula e o operador é ponto-ponto. Este algoritmo aproxima o conjunto restrição usando a Mosco-convergência de uma seqüência de subconjuntos convexos, fechados e não-vazios.

Para trabalhos futuros, sugerimos a implementação dos métodos desenvolvidos na tese bem como a análise das respectivas taxas de convergências.

Apêndice - Funções convexas com valores em $\overline{\mathbf{R}}$

As noções apresentadas neste apêndice podem ser encontradas em [67].

Seja U um espaço de Hilbert real.

Considere $f : U \rightarrow \overline{\mathbf{R}}$ uma função. O conjunto

$$\text{epi}(f) := \{(u, \lambda) \in U \times \mathbf{R} / f(u) \leq \lambda\}$$

é chamado o *epígrafe* de f . O *domínio efetivo* de f é a projeção do $\text{epi}(f)$ em U , isto é,

$$\text{Dom}f := \{u \in U / \exists \lambda \in \mathbf{R} : (u, \lambda) \in \text{epi}(f)\}.$$

A função f é dita *própria* se $f(u) < +\infty$ para pelo menos um ponto $u \in U$ e $f(u) > -\infty$ para todo $u \in U$, isto é, se o seu epígrafe é não-vazio e não contém nenhuma reta.

Definição 1: Seja $f : U \rightarrow \mathbf{R} \cup \{+\infty\}$ uma função.

a) f é *convexa* em U se, para quaisquer $u, v \in U$ e $\lambda \in (0, 1)$,

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v);$$

b) f é *estritamente convexa* em U se, para quaisquer $u, v \in U$, $u \neq v$ e $\lambda \in (0, 1)$,

$$f(\lambda u + (1 - \lambda)v) < \lambda f(u) + (1 - \lambda)f(v);$$

c) f é *fortemente convexa* em U se, existe uma constante $c > 0$ tal que, para quaisquer $u, v \in U$ e $\lambda \in (0, 1)$,

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v) - c\lambda(1 - \lambda)\|u - v\|^2/2.$$

É fácil ver que $c) \Rightarrow b) \Rightarrow a)$.

O resultado abaixo estabelece a equivalência entre a convexidade de uma função e a convexidade de seu epígrafe.

Proposição 1: Seja $f : U \rightarrow \mathbf{R} \cup \{+\infty\}$ uma função. As seguintes condições são equivalentes:

- a) f é convexa;
 b) $\text{epi}(f)$ é convexo.

Demonstração: Veja [67], teorema 5.10.

Definição 2: Sejam $f : U \rightarrow \overline{\mathbf{R}}$ uma função e $u \in U$. f é dita *semicontínua inferiormente em u (s.c.i. em u)* se para cada $\lambda \in \mathbf{R}$ tal que $\lambda < f(u)$, existe uma vizinhança V de u tal que $f(V) > \lambda$. f é dita *semicontínua inferiormente (s.c.i.)* se f é semicontínua inferiormente em cada ponto de U .

A noção de semicontinuidade inferiormente também pode ser introduzida usando o conceito de \liminf , do seguinte modo:

Definição 3: Seja $u_0 \in U$.

$$\liminf_{u \rightarrow u_0} f(u) := \sup_V \{ \inf_{u \in V - \{u_0\}} f(u) \}$$

onde V percorre todo o sistema de vizinhança de u_0 . Se U é um espaço linear normado então temos que

$$\liminf_{u \rightarrow u_0} f(u) = \liminf_{\varepsilon \rightarrow 0^+} \{ f(u) / 0 < \|u - u_0\| < \varepsilon \}.$$

Proposição 2: Sejam $f : U \rightarrow \overline{\mathbf{R}}$ uma função e $u_0 \in U$. f é s.c.i. em u_0 se, e somente se,

$$\liminf_{u \rightarrow u_0} f(u) \geq f(u_0).$$

Demonstração: Veja [67], teorema 5.8.

Se consideramos a topologia fraca em U , a noção correspondente é a semicontinuidade inferiormente fraca. Obviamente, qualquer função semicontínua inferiormente fraca é semicontínua inferiormente. A recíproca, em geral, não é verdadeira. Contudo, temos o seguinte resultado:

Proposição 3: Uma função convexa $f : U \rightarrow \mathbf{R} \cup \{+\infty\}$ é semicontínua inferiormente (forte) se, e somente se, é semicontínua inferiormente fraca.

Demonstração: Veja [30], capítulo 1, corolário 2.2.

A seguir, mostramos que o conceito de semicontinuidade inferiormente para f também pode ser descrito em termos das propriedades de determinados conjuntos.

Teorema 1: Seja $f : U \rightarrow \overline{\mathbf{R}}$ uma função. As seguintes condições são equivalentes:

- a) f é s.c.i.;
- b) $\text{epi}(f)$ é fechado (como um subconjunto de $U \times \mathbf{R}$);
- c) $\{u \in U / f(u) \leq \lambda\}$ é fechado para cada $\lambda \in \mathbf{R}$;
- d) $\{u \in U / f(u) > \lambda\}$ é aberto para cada $\lambda \in \mathbf{R}$.

Demonstração: Veja [67], teorema 5.3.

Teorema 2: Seja $f : U \rightarrow \overline{\mathbf{R}}$ uma função convexa.

- (i) Se f é própria e s.c.i. então f é contínua no $(\text{Dom } f)^\circ$.
- (ii) Seja $u_0 \in U$ tal que $f(u_0) > -\infty$ e f é localmente limitado superiormente em u_0 . Então:
 - (a) f é própria e $(\text{Dom } f)^\circ \neq \emptyset$;
 - (b) f é localmente limitada superiormente no $(\text{Dom } f)^\circ$;
 - (c) f é contínua no $(\text{Dom } f)^\circ$.

Demonstração: (i) Veja [53], lema. 2.6. (ii) Veja [67], teorema 5.20.

Definição 4: Sejam $f : U \rightarrow \overline{\mathbf{R}}$ uma função e u_0 um ponto de U onde f é finita.

a) f é *Fréchet-diferenciável* (ou *F-diferenciável*) em u_0 se existe um elemento em U , que denotaremos $f'(u_0)$, tal que para cada $u \in U$

$$\lim_{u \rightarrow u_0} \frac{f(u) - f(u_0) - \langle f'(u_0), u - u_0 \rangle}{\|u - u_0\|} = 0.$$

$f'(u_0)$ é unicamente determinado pela igualdade acima e é chamado *Fréchet-diferencial* (ou *F-diferencial*) de f em u_0 ;

b) f é *Gateaux-diferenciável* (ou *G-diferenciável*) em u_0 se existe um elemento em U , que denotaremos $\nabla f(u_0)$, tal que para cada $u \in U$

$$\lim_{\varepsilon \rightarrow 0} \frac{f(u_0 + \varepsilon u) - f(u_0)}{\varepsilon} = \langle \nabla f(u_0), u \rangle.$$

$\nabla f(u_0)$ é unicamente determinado pela igualdade acima e é chamado *Gateaux-diferencial* (ou *G-diferencial*) de f em u_0 .

Se a Gateaux-diferencial $\nabla f(u_0)$ existe então $f'(u_0, u)$ (derivada direcional de f em u_0 na direção u) existe para qualquer $u \in U$ e

$$f'(u_0, u) = \langle \nabla f(u_0), u \rangle.$$

Mas a existência de todas as derivadas direcionais de f em u_0 não implica na Gateaux-diferenciabilidade de f em u_0 .

Observação 1: Fréchet-diferenciabilidade implica Gateaux-diferenciabilidade, mas a recíproca não é verdadeira em geral. Em particular, a recíproca é válida para funções convexas próprias em \mathbf{R}^n .

Teorema 3: Sejam $f : U \rightarrow \mathbf{R} \cup \{+\infty\}$ uma função e D um subconjunto aberto e convexo do $Dom f$. Se f é G -diferenciável em D então segue as seguintes equivalências:

a) f é convexa em D se, e somente se,

$$f(v) \geq f(u) + \langle \nabla f(u), v - u \rangle \quad \forall u, v \in D;$$

b) f é estritamente convexa em D se, e somente se,

$$f(v) > f(u) + \langle \nabla f(u), v - u \rangle \quad \forall u, v \in D, u \neq v;$$

c) f é fortemente convexa módulo $c > 0$ em D se, e somente se,

$$f(v) \geq f(u) + \langle \nabla f(u), v - u \rangle + c \|u - v\|^2 / 2 \quad \forall u, v \in D.$$

Demonstração: Veja [67], capítulo 5.

Definição 5: Seja $C \subseteq U$. A função *indicatriz* de C , $I_C : U \rightarrow \overline{\mathbf{R}}$, é definida por

$$I_C(u) := \begin{cases} 0 & \text{se } u \in C; \\ +\infty & \text{se } u \notin C. \end{cases}$$

Então segue imediatamente das respectivas definições que:

a) O subconjunto C é não-vazio se, e somente se, I_C é própria;

b) O subconjunto C é convexo se, e somente se, I_C é convexa;

c) O subconjunto C é fechado se, e somente se, I_C é s.c.i.

Definição 6: Sejam $f : U \rightarrow \overline{\mathbf{R}}$ uma função e $u_0 \in Dom f$.

a) $g \in U$ é dito *subgradiente* de f em u_0 se

$$f(u) \geq f(u_0) + \langle g, u - u_0 \rangle \quad \forall u \in U;$$

b) O conjunto de todos os subgradientes de f em u_0 é chamado o *subdiferencial* de f em u_0 e é denotado por $\partial f(u_0)$. O $\partial f(u_0)$ é um subconjunto convexo e fechado de U . f é dita *subdiferenciável* em u_0 se $\partial f(u_0) \neq \emptyset$. Se $u \notin \text{Dom } f$, definimos $\partial f(u) = \emptyset$;

c) O *subdiferencial* de f é o operador ponto-conjunto

$$\begin{aligned} \partial f : U &\rightarrow P(U) \\ u &\mapsto \partial f(u); \end{aligned}$$

d) O *domínio* do ∂f é o conjunto

$$\text{Dom}(\partial f) := \{u \in U : \partial f(u) \neq \emptyset\}.$$

Exemplo 1: Considere a função indicatriz, I_C , de um subconjunto convexo, fechado e não-vazio de U . Por definição, $w \in \partial I_C(u)$ se, e somente se, $u \in C$ e

$$I_C(v) \geq I_C(u) + \langle w, v - u \rangle \quad \forall v \in U.$$

Isto significa que $u \in C$ e $0 \geq \langle w, v - u \rangle \quad \forall v \in C$. Portanto, $\partial I_C(u)$ é o *cone normal* de C em u (denotado por $N_C(u)$):

$$\partial I_C(u) = \begin{cases} \{w \in U : \langle w, v - u \rangle \leq 0, \forall v \in C\} & , \quad u \in C \setminus C^\circ; \\ \{0\} & , \quad u \in C^\circ; \\ \emptyset & , \quad u \notin C. \end{cases}$$

Teorema 4: Sejam $f : U \rightarrow \overline{\mathbf{R}}$ uma função e u_0 um ponto de U onde f é finita. Se f é Gateaux-diferenciável em u_0 então

$$\partial f(u_0) = \{\nabla f(u_0)\}.$$

Reciprocamente, se f é contínua em u_0 e possui um único subgradiente em u_0 então f é Gateaux-diferenciável em u_0 e $\partial f(u_0) = \{\nabla f(u_0)\}$.

Demonstração: Veja [67], teorema 5.37.

Teorema 5: Sejam $f_1 : U \rightarrow \overline{\mathbf{R}}$ e $f_2 : U \rightarrow \overline{\mathbf{R}}$ duas funções convexas e próprias. Então

a) $\partial(f_1 + f_2)(u) \supseteq \partial f_1(u) + \partial f_2(u)$ para todo $u \in U$;

b) Se existe um ponto $u_0 \in \text{Dom } f_1 \cap \text{Dom } f_2$ com f_1 (ou f_2) contínua em u_0 então

$$\partial(f_1 + f_2)(u) = \partial f_1(u) + \partial f_2(u) \quad \forall u \in U.$$

Demonstração: Veja [67], teorema 5.38.

Teorema 6: Seja $f : U \rightarrow \overline{\mathbf{R}}$ uma função convexa e própria.

a) Se f é contínua em algum ponto $u_0 \in \text{Dom } f$ então f é subdiferenciável no $(\text{Dom } f)^\circ$. Portanto, a subdiferenciabilidade de f no $(\text{Dom } f)^\circ$ quando f é convexa, própria e s.c.i. segue imediatamente do teorema 2

b) Se $U = \mathbf{R}^n$ então f é subdiferenciável no $\text{ri}(\text{Dom } f)$ (interior relativo do $\text{Dom } f$).

Demonstração: Veja [67], teorema 5.35.

Bibliografia

- [1] ALBER, YA.I, IUSEN, A.N., SOLODOV, M.V., "On the Projected Subgradient Method for Nonsmooth Convex Optimization in a Hilbert Space" *Mathematical Programming*, n.81, pp.23-35, 1998.
- [2] ATTOUCH, H. *Variational Convergence for Functions and Operators*, British Library Cataloguing in Publication Data, 1984.
- [3] AUBIN, J.P., EKELAND, I., *Applied Nonlinear Analysis*, New York, Wiley, 1984.
- [4] BAILLON, J.B., HADDAD, G., "Quelques Propriétés des Opérateurs Angle-Bornés et n-Cycliquement Monotones", *Israel Journal of Mathematics*, n.26, pp.137-150, 1977.
- [5] BREGMAN, L.M. "The Relaxation Method of Finding the Common Points of Convex Sets and its Application to the Solution of Problems in Convex Programming". *USSR Computational Mathematics and Mathematical Physics* v. 7, n. 3, pp. 200 - 217, 1967.
- [6] BRÉZIS, H. *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, Amsterda, North-Holland, 1973.
- [7] BRÉZIS, H., HARAUX, A., "Images d'une Somme d'Opérateurs Monotones et Applications", *Israel Journal of Mathematics*, v. 23, n. 2, pp. 165-186, 1976.
- [8] BRÉZIS, H., *Analyse Fonctionnelle-Théorie et Applications*, Paris, Masson, 1987.
- [9] BROWDER, F.E., "Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces". In: *Proceedings of Symposia in Pure Mathematics*, American Mathematical Society, 18, 2, 1976.
- [10] BROWDER F.E., PETRYSHYN W.V. "Construction of Fixed Points of Nonlinear Mapping in Hilbert Space". *Journal of Mathematical Analysis and Applications*, v. 20, pp. 197-228, 1967.
- [11] BRUCK, Jr., R. E., "An Iterative Solution of a Variational Inequality for Certain Monotone Operators in Hilbert Space", *Bull. Amer. Math. Soc.*, v. 81, n.5, pp.890-892, 1975.
- [12] BURACHIK, R.S., *Generalized Proximal Point Methods for the Variational Inequality Problem*. PhD dissertation, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, Brasil, 1995.

- [13] BURACHIK, R.S., IUSEM, A.N. "A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Space", *SIAM Journal on Optimization*, v. 8, pp.197-216, 1998.
- [14] BURACHIK, R.S., IUSEM, A.N., SVAITER, B.F. "Enlargement of Maximal Monotone Operators with Application to Variational Inequalities", *Set-Valued Analysis*, v.5, pp. 159 - 180, 1997.
- [15] BURACHIK, R. S.,SCHEIMBERG, S., "A Proximal Point Method for the Variational Inequality Problem in Banach Spaces", *SIAM Journal on Optimization*, v. 39, n. 5, pp. 1633-1649.
- [16] BUTNARIU, D., IUSEM, A.N., "On a Proximal Point Method for Convex Optimization in Banach Spaces". *Numerical Functional Analysis and Optimization*.
- [17] BUTNARIU, D. IUSEM, A.N., "Local Moduli of Convexity and their Application to Finding Almost Common Fixed Points of Measurable Families of Operators". In: *Recent Developments in Optimization Theory and Nonlinear Analysis*, v. 204, Amer. Math. Soc., Contemporary Mathematics Series, Y. Censor and S. Reich, 1997.
- [18] CENSOR, Y., IUSEM, A.N., ZENIOS, S.A. "An Interior Point Method with Bregman Functions for the Variational Inequality Problem with Paramonotone Operators". *Mathematical Programming*. v. 81, pp. 373-400, 1998.
- [19] CENSOR, Y., LENT, A., "An Iterative Row Action Method for Interval Convex Programming", *Journal of Optimization theory and Applications*, v. 34, pp. 321-353
- [20] CENSOR, Y., ZENIOS, S.A., "The Proximal Minimization Algorithm with D-functions", *Journal of Optimization Theory and Applications*, v. 73, pp. 451 - 464, 1992.
- [21] CHEN, G.H.G., ROCKAFELLAR, R.T., "Convergence Rates in Forward - Backward Splitting Methods", *SIAM Journal on Optimization*, v. 7, pp. 421 - 444, 1997.
- [22] COHEN, G., "Auxiliary Problem Principle Extended to Variational Inequalities", *Journal of Optimization Theory and Applications*, v. 59, pp. 325-333, 1988.

- [23] COHEN, G., "Auxiliary Problem Principle and Decomposition of Optimization Problems", *Journal of Optimization Theory and Applications*, v. 32, pp. 277-305, 1980.
- [24] COHEN, G., ZHU, D.L., "Decomposition Coordination Methods in Large-Scale Optimization Problems: The Nondifferentiable Case and the Use of Augmented Lagrangians". In: Cruz, J.B., *Advances in Large-Scale Systems, Theory and Applications*, , Greenwich, v. 1, pp. 203-266, Connecticut, JAI Press, 1984.
- [25] CULIOLI, J.C., COHEN, G., "Decomposition / Coordination Algorithms in Stochastic Optimization", *SIAM Journal on Control and Optimization*, v. 28, pp.1372-1403, 1990.
- [26] DAFERMOS, S., "An Iterative Scheme for Variational Inequalities", *Mathematical Programming*, v. 26, pp. 40-47, 1983.
- [27] DE PIERRO , A.R., IUSEM, A.N. "A Relaxed Version of Bregman 's method for Convex Programming" , *Journal of Optimization Theory and Applications*, v. 51 pp. 421-440, 1986.
- [28] DUNN, J.C., "Convexity, Nonotonicity, and Gradient Process in Hilbert Spaces" , *Journal of Mathematical Analysis and Applications*, v. 53, pp. 145-158, 1976.
- [29] ECKSTEIN, J., BERTSEKAS, D. P., "On the Douglas-Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators" , *Mathematical Programming*, v. 55, pp. 293-318, 1992.
- [30] EKELAND, I., TEMAM, R., *Convex Analysis and Variational Inequalities Problems*, Amsterdam, Holland, North-Holland, 1976.
- [31] EL-FAROUQ, N., COHEN, G., "Progressive Regularization of Variational Inequalities and Decomposition Algorithms" , *Journal of Optimization Theory and Applications*, v. 97, pp.407-433, 1998.
- [32] FANG, S.C., PETERSON, E.,L., "Generalized Variational Inequalities" , *Journal of Optimization Theory and Applications*, v. 38, pp. 363-383, 1982.
- [33] HARKER, P.T.,PANG, J.S., "Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems, A Survey of Theory, Algorithms and Applications" , *Mathematical Programming*, v.48, pp.161-220, 1990.

- [34] IUSEM, A.N., "An Interior Point Method for the Nonlinear Complementarity", *Informes de Matemática, IMPA, Série B-081- Abril/94*.
- [35] IUSEM, A.N., "On Some Properties of Generalized Proximal Point Methods for Quadratic and Linear Programming", *Journal of Optimization Theory and Applications*, v. 85, pp. 593-612, 1995.
- [36] IUSEM, A.N., "On Some Properties of Generalized Proximal Point Methods for Variational Inequalities", *Journal of Optimization Theory and Applications*, v. 96, n. 2, pp. 337-362, 1998.
- [37] IUSEM, A.N., "On Some Properties of Paramonotone Operators", *Journal of Convex Analysis*, v. 5, pp. 269-278, 1998.
- [38] KAPLAN, A., TICHATSCHKEK, R., "Auxiliary Problem Principle and the Approximation of Variational Inequalities with Non-symmetric Multi-valued Operators". In: *Canadian Mathematical Society Conference Proceedings Series*, v. 185, M. Théra, pp. 185-209, 2000.
- [39] KARAMARDIAN, S., "Generalized Complementarity Problem", *Journal of Optimization Theory and Applications*, v. 8, pp. 161-168, 1971.
- [40] KINDERLEHER, D., STAMPACHIA, G., *An Introduction to Variational Inequalities and Their Applications*, New York, Academic Press, 1980.
- [41] KIWIEL, K., "Free - steering Relaxation Methods for Problems with Strictly Convex Costs and Linear Constraints". *Mathematics of Operations Research*, v. 22, pp. 326- 349, 1997.
- [42] LIESE, F., VAJDA, I. *Convex Statistical Distances*. Teubner, Leipzig, 1987.
- [43] LIONS, P.L., MERCIER, B., "Splitting Algorithms for the Sum of Two Nonlinear Operators", *SIAM Journal on Numerical Analysis*, v.16, pp. 964-979, 1979.
- [44] MAGNANTI, T.L., PERAKIS, G., "The Orthogonality Theorem and the Strong - f - Monotonicity Condition for Variational Inequality Algorithms", *SIAM Journal On Optimization*, v. 7, pp. 248 - 273, 1997.
- [45] MAKLER, S.C., S., NGUYEN, V. H., STRODIOT, J.J., "Family of Perturbation Methods for Variational Inequalities", *Journal of Optimization Theory and Applications*, v. 89, n. 2, pp. 423-452, 1996.

- [46] MARCOTTE, P., WU, J.H., "On the Convergence of Projection Methods: Application to the Decomposition of Affine Variational Inequalities", *Journal of Optimization Theory and Applications*, v. 85, pp. 347-362, 1995.
- [47] MARTINET, B., "Régularisation d'Inéquations Variationnelles par Approximations Successives", *Revue Française d'Informatique et de Recherche Opérationnelle*, v. 4, n.3, pp.154-159, 1970.
- [48] MATAOUI, M.A., *Contributions à la Décomposition et à L'Àgrégation des Problèmes Variationnels*, Ph.D. dissertation, École des Mines de Paris, Paris, France, 1990.
- [49] MINTY, G., "Monotone Nonlinear Operators in Hilbert Space", *Duke Mathematical Journal*, v. 29, pp.341-346, 1978.
- [50] MOSCO, U., "Convergence of Convex Sets and of Solutions of Variational Inequalities", *Advances in Mathematics*, v. 3, pp. 510-585, 1969.
- [51] MOUDAFI, A., THÉRA, M., "Finding a zero of the Sum of Two Maximal Monotone Operators", *Journal of Optimization Theory and Applications*, v. 94, pp. 425-448, 1997.
- [52] PANG, J.S., CHAN, D., "Iterative Methods for Variational and Complementarity Problems", *Mathematical Programming*, v. 24, pp.284-313, 1982.
- [53] PASCALI, D., SBURLAN, S., *Nonlinear Mappings of Monotone Type*. Bucuresti, România, Academiei, 1978.
- [54] RENAUD, A., COHEN, G., "Conditioning and Regularization of Nonsymmetric Operators", *J. Optim. Theory Appl.*, v. 92, pp. 127-148, 1997.
- [55] RENAUD, A., COHEN, G., "An Extension of the Auxiliary Problem Principle to Nonsymmetric Auxiliary Operators", *ESAIM: Control, Optimisation and Calculus of Variations*, v. 2, pp. 281-306, 1997.
- [56] ROCKAFELLAR, R.T, *Convex Analysis*. Princeton, New Jersey, Princeton University Press, 1970.
- [57] ROCKAFELLAR, R.T., "On the Maximality of Sums of Nonlinear monotone Operators", *Transactions of the American Mathematical Society*, v. 149, pp. 75-88, 1970.

- [58] ROCKAFELLAR, R.T., "Monotone Operators and the Proximal Point Algorithm", *SIAM Journal on Control and Optimization*, v. 14, pp. 877-898, 1976.
- [59] SALMON, G., *Perturbed Auxiliary Problem Methods to Solve Generalized Variational Inequalities*, Ph.D. dissertation, University of Namur, 2001.
- [60] SALMON, G., NGUYEN, V.H., STRODIOT, J.J., "Coupling the Auxiliary Problem Principle and the Epiconvergence Theory to Solve General Variational Inequalities", *Journal of Optimization Theory and Applications*, v. 104, pp. 629-657, 2000.
- [61] SALMON, G., NGUYEN, V.H. , STRODIOT, J.J., "A Perturbed Auxiliary Problem Method for Paramonotone Multivalued Mappings". In: *Advances in Convex Analysis and Global Optimization, Non Convex Optimization and Applications Series*, Hadjisavvas, N., and Pardalos, P., Kluwer Academic Publishers, 2001. to Appear.
- [62] SOLODOV, M.V., SVAITER, B.F., "An Inexact Hybrid Generalized Proximal Point Algorithm and Some New Results on the theory of Bregman Functions", *Mathematics of Operations Research*, v. 25, n. 2, pp.214-230, 2000.
- [63] TEBoulLE, M., "Entropic Proximal Mappings with Application to Nonlinear Programming", *Mathematics of Operations Research*, v. 17, pp.670 - 690, 1992.
- [64] TEBoulLE, M., "Convergence of Proximal - Like Algorithms", *SIAM Journal on Optimization*, v. 7, pp.1069-1083, 1997.
- [65] TSENG, P., "Further Applications of a Splitting Algorithm to Decomposition in Variational Inequalities and Convex Programming". *Mathematical Programming*, v. 48, pp. 249-263, 1990.
- [66] TSENG, P., "Applications of a Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities", *SIAM Journal on Control and Optimization*, v. 29, pp. 119-138, 1991.
- [67] VAN TIEL, J., *Convex Analysis, an Introductory Text*. John Wiley, New York, 1984.
- [68] ZEIDLER, E., *Nonlinear Functional Analysis and its Applications, II/B Nonlinear Monotone Operators*. New York, Springer-Verlag, 1990.

- [69] ZHU, C., “Asymptotic Convergence Analysis of the Forward-Backward Splitting Algorithm”, *Mathematics of Operations Research*, v. 20, pp. 449-464, 1995.
- [70] ZHU, D., MARCOTTE, P., “New Classes of Generalized Monotonicity”, *Journal of Optimization Theory and Applications*, v. 87, pp. 457-471, 1995.
- [71] ZHU, D., MARCOTTE, P., “Co-Coercivity and its Role in the Convergence of Iterative Schemes for Solving Variational Inequalities”, *SIAM Journal on Optimization*, v. 6, pp. 714-726, 1996.