

ESTUDO DO OPERADOR INVERSA PARCIAL APLICADO A UM CONE

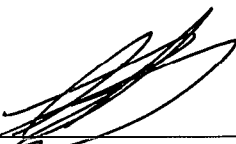
Felipe Antonio Garcia Moreno

DISSERTAÇÃO 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 A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

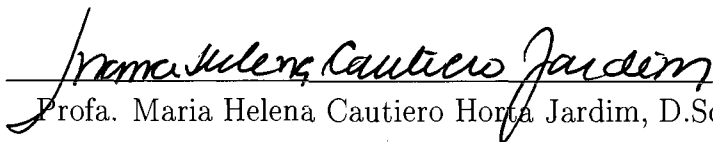
Aprovada por:



Prof. Susana Scheimberg de Makler, D.Sc.



Prof. Claudio Thomas Bornstein, Dr.Rer.Nat.



Prof. Maria Helena Cautiero Horta Jardim, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

FEVEREIRO DE 2006

MORENO, FELIPE ANTONIO GARCIA

Estudo do Operador Inversa Parcial Aplicado a um Cone [Rio de Janeiro] 2006

X, 91 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 2006)

Dissertação – Universidade Federal do Rio de Janeiro, COPPE

1 - Operadores Monótonos

2 - Algoritmo de Ponto Proximal

2 - Operador Inversa Parcial

I. COPPE/UFRJ II. Título (série)

*Aos meus pais Felipe e Isabel
e meu amor Kely.*

Agradecimentos

Agradeço a oportunidade brindada pelo Programa de Engenharia de Sistemas e Computação (PESC) da Coordenação dos Programas de Pós-Graduação em Engenharia (COPPE) da Universidade Federal do Rio de Janeiro (UFRJ) por ter me proporcionado boas condições de trabalho para realizar os estudos do mestrado na área de otimização.

Na parte acadêmica, agradeço a todos os Professores da área de otimização do PESC, em particular à Professora Susana Scheimberg pela paciência, orientação e conhecimentos transmitidos durante minha estadia no PESC.

Agradeço a minha mulher Kely, pelo amor incondicional.

A meus amigos Paulo Sergio, Flavia Morgana, Sissy e Nilomar pelas conversas e correções deste trabalho. Como também ao Gevane, Roberto e Disney.

Ao Professor Angel Coca pelo apoio e amistad.

À comunidade peruana a Mariela, Raúl, Mariella Janette, Erik, Marisa, Alvaro, Karl e Yalmar, pelas festinhas.

Meus agradecimentos à CNPq pelo apoio financiero.

Ao pessoal do PESC, em especial a Solange, Sônia, Lúcia, Claudia, Sueli, Mercedes, Itamar e Adilson pela atenção especial e também a todos aqueles que de alguma forma participaram deste trabalho, acompanhando, incentivando e apoiando nas horas em que foi necessario.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc)

ESTUDO DO OPERADOR INVERSA PARCIAL APLICADO A UM CONE

Felipe Antonio Garcia Moreno

Fevereiro/2006

Orientador: Susana Scheimberg de Makler

Programa: Engenharia de Sistemas e Computação

A inversa parcial de um operador com respeito a um subespaço, introduzido por Spingarn, permite estabelecer uma equivalência entre alguns problemas que envolvem operadores monótonos maximais com o problema de encontrar um zero da inversa parcial. Tal operador preserva a monotonicidade maximal do operador original, e portanto, viabiliza o uso do algoritmo de ponto proximal para encontrar um zero do problema equivalente. Este procedimento é conhecido como o método das inversas parciais.

Neste trabalho, estudamos uma generalização da inversa parcial de um operador com respeito a um cone. Discutimos algumas propriedades e mostramos por meio de contra-exemplos que este operador não preserva a monotonicidade nem a maximalidade do operador original como no caso dos subespaços impossibilitando a extensão dos resultados obtidos por Spingarn. Também são ilustrados através de contra-exemplos características de operadores que não são preservados pela sua inversa parcial com respeito a um subespaço.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

STUDY OF THE PARTIAL INVERSE OPERATOR APPLIED TO CONE

Felipe Antonio Garcia Moreno

February/2006

Advisor: Susana Scheimberg de Makler

Department: Computing and Systems Engineering

The partial inverse of an operator with respect to a subspace, introduced by Spingarn, allows to establish an equivalence between some problems that involve maximal monotone operators with the problem to find a zero of the partial inverse. Such operator preserves the maximal monotonicity of the original operator, and therefore, makes possible the use of the algorithm of proximal point to find a zero of the equivalent problem. This procedure is known as the method of partial inverses.

In this work, we study a generalization of the partial inverse of an operator with respect to a cone. We discuss some of its properties and show by means of contraexamples that this operator does not preserve monotonicity nor maximality of the original operator as in the case of the subspaces disabling the extension of the results gotten by Spingarn. Also they are illustrated through contraexamples characteristic of operators who are not preserved by its inverse partial with respect to one subespaço.

Sumário

Introdução	1
1 Preliminares	3
1.1 Análise convexa	3
1.2 Aplicações ponto-conjunto	8
1.3 Operadores monótonos	10
1.4 Operadores pseudomonótonos	26
1.5 Projeção sobre um conjunto convexo	28
1.5.1 Projeção sobre um subespaço	31
1.5.2 Decomposição de Moreau	32
2 O Problema	36
2.1 O problema primal-dual	36
2.2 Aplicações	37
2.2.1 Problema de otimização convexa	37
2.2.1.1 Minimização convexa sobre um subespaço	38
2.2.2 Soma de operadores monótonos	39
2.2.3 Viabilidade convexa	40
2.2.4 Desigualdades variacionais	41
3 Operador Inversa Parcial	43
3.1 Operador inversa parcial	43
3.2 Problemas duais	51
3.3 Operador Douglas-Rachford	54
3.4 Relação entre os operadores Inversa Parcial e Douglas-Rachford	55

4	Método Proximal	57
4.1	O conceito de regularização	57
4.2	Algoritmo de ponto proximal em otimização convexa	59
4.3	Algoritmo de ponto proximal para operadores monótonos maximais	62
5	Método da Inversa Parcial	65
5.1	Método da inversas parciais	65
5.2	Método de decomposição Douglas-Rachford	72
5.3	Relação entre os métodos de inversa parcial e Douglas-Rachford. . .	73
6	Uma Extensão do Operador Inversa Parcial	74
6.1	O operador inversa parcial generalizado	75
6.2	Propriedades de T não preservadas	82
	Conclusões	85
	Referências Bibliográficas	87

Lista de Figuras

1.1	10
1.2	Gráficos dos operadores T_1 , T_2 e \tilde{T} do Exemplo 1.3.2(iii).	14
1.3	Gráficos dos operadores ∂f e T do Exemplo 1.3.2(iv).	15
1.4	Gráficos dos conjuntos C_1 , C_2 , $N_{C_1(0)}$ e $N_{C_2(0)}$ do Exemplo 1.3.4.	20
1.5	Gráfico do ∂g definido no Exemplo 1.3.5(iii).	22
1.6	Gráficos dos operadores dos itens (i) e (iii) do Exemplo 1.4.1.	28
1.7	Projeção sobre um conjunto convexo.	29
1.8	Projeção sobre um subespaço	31
1.9	O item (i) do Exemplo 1.5.3.	34
1.10	Decomposição de Moreau para o cone K do Exemplo 1.5.3(ii)	35
3.1	Gráficos dos operadores T e T_A do Exemplo 3.1.2.	48
6.1	Gráficos dos operadores S e T do Exemplo 6.1.2.	77
6.2	Gráfico do operador S_K do Exemplo 6.1.2.	79

Notações

\mathbb{R}_+^n	O conjunto $\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0 \quad \forall i = 1, \dots, n\}$.
\mathbb{R}_{++}^n	O conjunto $\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i > 0 \quad \forall i = 1, \dots, n\}$.
\mathcal{H}	Espaço de Hilbert real.
$\mathcal{P}(\mathcal{H})$	Conjunto das partes de \mathcal{H} .
$\langle \cdot, \cdot \rangle$	Produto interno em \mathcal{H} .
$\ x\ = \sqrt{\langle x, x \rangle}$	Norma associada ao produto interno em \mathcal{H} .
$\overline{B}(a; r)$	O conjunto $\{x \in \mathcal{H} : \ x - a\ \leq r\}$.
$\text{int}(B)$	Interior topológico do conjunto $B \subseteq \mathcal{H}$.
$\text{ri}(A)$	Interior relativo do conjunto A .
$\text{dom}(f)$	Domínio efetivo da função f .
$\text{epi}(f)$	Epígrafe da função f .
δ_C	Função indicadora do conjunto $C \subseteq \mathcal{H}$.
$\partial f(u)$	Subdiferencial da função f em u .
$\nabla f(x)$	Gâteaux-diferencial da função f em x .
$\text{Im}(F)$	Imagem do operador F .
$\text{Dom}(F)$	Domínio do operador F .
$\text{Gr}(F)$	Gráfico do operador F .
$F^{-1}(x)$	O conjunto $\{x \in X : y \in F(x)\}$.
N_C	Operador cone normal do conjunto C .
A^\perp	Subespaço ortogonal de A .
$J_{\lambda T}$	Resolvente ($J_{\lambda T} = (I + \lambda T)^{-1}$).
x_C	Projeção do ponto x sobre o conjunto C ($x_C = \text{proj}_C(x)$).
\oplus	Soma direta.
K°	Cone polar de K . ($K^\circ = \{x \in \mathcal{H} : \langle x, y \rangle \leq 0 \quad \forall y \in K\}$).
S_{α, T_1, T_2}	Operador Douglas-Rachford.
T_A	Operador inversa parcial de T com respeito ao conjunto $A \subseteq \mathcal{H}$.
$\text{dist}(u, C)$	Distância de u ao conjunto C ($\text{dist}(u, C) = \inf_{s \in C} \ u - s\ $).
f_S^*	O valor $\inf_{x \in S} f(x)$.
X_S^*	O conjunto dos mínimos de f em S ($\{x \in \mathcal{H} : f(x) = f_S^*\}$).

Introdução

Nas últimas décadas, a teoria dos operadores monótonos tem sido muito explorada. Várias aplicações em diversas áreas da matemática têm sido estudadas, em particular na otimização. Um problema importante neste campo é encontrar zeros de operadores monótonos maximais. Um método de resolução para este problema bem conhecido na literatura é o algoritmo de ponto proximal estudado por Rockafellar [39].

Por outro lado, existem problemas que não são da forma de encontrar zeros de operadores, em particular:

$$P(T, A) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp \\ \text{tais que} \\ \bar{y} \in T(\bar{x}). \end{array} \right.$$

onde $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ é um operador monótono maximal e A um subespaço de um espaço de Hilbert \mathcal{H} . Devido a esta situação, surgem idéias para utilizar o método proximal de alguma forma na resolução deste tipo de problemas.

Neste contexto Spingarn [42], introduziu a inversa parcial de T com respeito a A , o qual permite estabelecer uma equivalência entre $P(T, A)$ com o problema de encontrar um zero da inversa parcial. Tal operador preserva a monotonicidade maximal do operador original, e portanto, viabiliza o uso do algoritmo de ponto proximal para encontrar um zero do problema equivalente. Este procedimento é conhecido como o método das inversas parciais.

Devido às boas propriedades do operador inversa parcial, nos motivamos a explorar uma possível generalização para este; utilizamos para isso um resultado que generaliza a decomposição para subespaços, dada por Moreau [29].

A motivação e justificação para generalizar o método das inversas parciais com respeito ao subespaço A , quando A é substituído por um cone, é devido à quantidade

de problemas que podem ser observados como instâncias particulares de $P(T, A)$ por substituir o subespaço A por um cone, como é estudado por Isac [17].

No decorrer do estudo, vemos a impossibilidade de generalizar os resultados obtidos por Spingarn [42] e [43], devido a que o resultado fundamental, com o qual ele consegue uma equivalência entre o problema $P(T, A)$ com o problema de encontrar um zero do operador inversa parcial, não é satisfeita quando substituímos o subespaço A por um cone convexo fechado K .

Com respeito às propriedades mostramos por meio de contra-exemplos que a inversa parcial generalizada, não mantém as propriedades de ser uma função, monotonicidade, maximalidade nem pseudomonotonicidade do operador original. A seguir descrevemos o conteúdo dos capítulos.

No capítulo 1, apresentamos a teoria básica necessária para o desenvolvimento desta dissertação, os resultados são apresentados sem prova, mas referenciados.

No capítulo 2, estudamos o problema $P(T, A)$ e estabelecemos equivalências com o problema da soma de operadores monótonos e seus casos particulares.

No capítulo 3, estudamos o operador inversa parcial e as importantes propriedades deste, e o relacionamos com o operador Douglas-Rachford.

No capítulo 4, apresentamos uma revisão do método de regularização e do algoritmo de ponto proximal.

No capítulo 5, discutimos o método das inversas parciais, e vemos a relação deste com o método de decomposição Douglas-Rachford.

No capítulo 6, generalizamos a definição do operador inversa parcial para um cone convexo e fechado e discutimos algumas propriedades deste onde são introduzidos contraexemplos.

Capítulo 1

Preliminares

Neste capítulo apresentamos os fundamentos teóricos para uma melhor compreensão desta dissertação. Lembramos algumas definições e resultados da análise convexa, estudamos brevemente as aplicações ponto-conjunto, os operadores monótonos e pseudomonótonos, e por fim a projeção sobre um conjunto convexo.

1.1 Análise convexa

Nesta seção listamos alguns resultados da análise convexa. Para maiores detalhes e referências, ver [12], [44] em dimensão infinita e [36] em dimensão finita. Denotamos por \mathcal{H} um espaço de Hilbert real com produto interno $\langle \cdot, \cdot \rangle$ e norma $\|x\| = \sqrt{\langle x, x \rangle}$.

Definição 1.1.1 [12] *Seja $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$.*

1. A função f é **convexa**, se para todo par de pontos $x, y \in \mathcal{H}$ temos que

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in [0, 1].$$

2. O **domínio efetivo** de f é $\text{dom}(f) := \{x \in \mathcal{H} : f(x) < +\infty\}$.

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

4. O **epígrafo** de f é $\text{epi}(f) := \{(x, y) \in \mathcal{H} \times \mathbb{R} : f(x) \leq y\}$.

Definição 1.1.2 [44] *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ e $\bar{x} \in \mathcal{H}$.*

1. A função f é **semicontínua inferior no ponto \bar{x}** se

$$f(\bar{x}) \leq \liminf_{x \rightarrow \bar{x}} f(x) := \liminf_{\epsilon \downarrow 0} \{f(x) : 0 < \|x - \bar{x}\| < \epsilon\}.$$

2. A função f é **semicontínua inferior** se f é semicontínua inferior em todo ponto $\bar{x} \in \mathcal{H}$.

Teorema 1.1.1 ([44], Teorema 5.3) *Seja $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função. As seguintes afirmações são equivalentes:*

1. f é semicontínua inferior.
2. $\{x \in \mathcal{H} : f(x) > \lambda\}$ é aberto para cada $\lambda \in \mathbb{R}$.
3. $\{x \in \mathcal{H} : f(x) \leq \lambda\}$ é fechado para cada $\lambda \in \mathbb{R}$.
4. $\text{epi}(f)$ é fechado (como um subconjunto de $\mathcal{H} \times \mathbb{R}$).

Definição 1.1.3 [44] *Seja C um subconjunto de \mathcal{H} . A função $\delta_C : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por:*

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

é chamada de **função indicadora** de C .

Definição 1.1.4 [44] *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e x_0 um ponto do $\text{dom}(f)$.*

1. O ponto $y \in \mathcal{H}$ é denominado um **subgradiente** da função f em x_0 se

$$f(x) \geq f(x_0) + \langle y, x - x_0 \rangle \quad \forall x \in \mathcal{H}. \quad (1.1)$$

2. O conjunto de todos os subgradientes da função f em x_0 é denominado o **subdiferencial** da função f em x_0 , e denotado por $\partial f(x_0)$. A função f é dita **subdiferenciável** em x_0 se $\partial f(x_0) \neq \emptyset$.

Observamos da definição acima que, se $x \notin \text{dom}(f)$, resulta $\partial f(x) = \emptyset$. Além disso, se $x \in \text{dom}(f)$ não necessariamente $\partial f(x) \neq \emptyset$, como vemos no exemplo abaixo.

Exemplo 1.1.1 ([30], pág. 283) *Seja $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função definida por*

$$f(x) = \begin{cases} -\sqrt{x} & \text{se } x \geq 0, \\ +\infty & \text{se } x < 0. \end{cases}$$

Afirmamos que $\partial f(0) = \emptyset$. Provamos por contradição, suponhamos que existe $y \in \partial f(0)$, então

$$-\sqrt{x} \geq \langle y, x \rangle \quad \forall x \geq 0,$$

assim, obtemos

$$\frac{1}{\sqrt{x}} \leq -y \quad \forall x > 0.$$

Fazendo $x \rightarrow 0^+$, resulta que $+\infty \leq -y$, o que é um absurdo. ■

Teorema 1.1.2 ([44], Teorema 5.37) *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e $x_0 \in \text{dom}(f)$. Se f é Gâteaux-diferenciável em x_0 então*

$$\partial f(x_0) = \{\nabla f(x_0)\},$$

onde $\nabla f(x_0)$ é a Gâteaux-diferencial da f em x_0 . Reciprocamente, se f é contínua em x_0 e possui um único subgradiente em x_0 então f é Gâteaux-diferenciável neste ponto e $\partial f(x_0) = \{\nabla f(x_0)\}$.

Exemplo 1.1.2

(i) [7, pág. 16] Seja $f : \mathcal{H} \rightarrow \mathbb{R}$ uma função definida por

$$f(x) = \frac{1}{2} \|x\|^2 \quad \forall x \in \mathcal{H}.$$

Observamos que f é convexa, contínua, com $\text{dom}(f) = \mathcal{H}$. Afirmamos que

$$\partial f(x) = \{x\} \quad \forall x \in \mathcal{H}.$$

De fato, seja $x \in \mathcal{H}$, então para qualquer $y \in \mathcal{H}$, temos

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

Portanto, $x \in \partial f(x)$. Por outro lado, se $w \in \partial f(x)$ então

$$\langle w, y - x \rangle \leq \frac{1}{2} (\|y\|^2 - \|x\|^2) \quad \forall y \in \mathcal{H}.$$

Em particular, considerando $y = x + \lambda z$, onde $\lambda \in \mathbb{R}$ e $z \in \mathcal{H}$, na desigualdade anterior resulta que

$$\lambda \langle w, z \rangle \leq \frac{1}{2} (\|x + \lambda z\|^2 - \|x\|^2) \leq \frac{1}{2} (\lambda^2 \|z\|^2 + 2|\lambda| \|x\| \|z\|). \quad (1.2)$$

Se $\lambda > 0$ em (1.2) e fazendo $\lambda \rightarrow 0^+$, obtemos

$$\langle w, z \rangle \leq \|x\| \|z\| \quad \forall z \in \mathcal{H}. \quad (1.3)$$

Por outro lado, se $\lambda < 0$ em (1.2) e fazendo $\lambda \rightarrow 0^-$, temos que

$$\langle w, z \rangle \geq -\|x\| \|z\| \quad \forall z \in \mathcal{H}. \quad (1.4)$$

Portanto, de (1.3) e (1.4)

$$|\langle w, z \rangle| \leq \|x\| \|z\| \quad \forall z \in \mathcal{H}.$$

Assim, para $z = x$ e $z = w$, segue que

$$|\langle w, x \rangle| \leq \|x\|^2 \text{ e } \|w\| \leq \|x\|. \quad (1.5)$$

Substituindo $z = x$ e $\lambda < 0$ na primeira desigualdade de (1.2), obtemos

$$\langle w, x \rangle \geq \frac{\lambda + 2}{2} \|x\|^2.$$

Fazendo $\lambda \rightarrow 0^-$, resulta

$$\langle w, x \rangle \geq \|x\|^2. \quad (1.6)$$

De (1.5) e (1.6), temos

$$\langle w, x \rangle = \|x\|^2,$$

logo

$$\|x - w\|^2 = \|x\|^2 - 2\langle x, w \rangle + \|w\|^2 = -\langle x, w \rangle + \|w\|^2 = 0.$$

Daqui, $w = x$.

(ii) [7, pág. 21] Seja $f : \mathcal{H} \rightarrow \mathbb{R}$ uma função definida por

$$f(x) = \|x\| \quad \forall x \in \mathcal{H}.$$

Notamos que f é convexa, própria e contínua. Afirmamos que

$$\partial f(x) = \begin{cases} \frac{x}{\|x\|} & \text{se } x \neq 0, \\ \bar{B}(0; 1) & \text{se } x = 0. \end{cases}$$

De fato, se $x \neq 0$, temos

$$\left\langle \frac{x}{\|x\|}, y - x \right\rangle = \left\langle \frac{x}{\|x\|}, y \right\rangle - \|x\| \leq \|y\| - \|x\| \quad \forall y \in \mathcal{H}.$$

Portanto, $\frac{x}{\|x\|} \in \partial f(x)$. Por outro lado, seja $w \in \partial f(x)$ então

$$\langle w, y - x \rangle \leq \|y\| - \|x\| \leq \|y - x\| \quad \forall y \in \mathcal{H}. \quad (1.7)$$

Fazendo $y = 0$, $y = 2x$ em (1.7), obtemos:

$$\langle w, x \rangle = \|x\| \text{ e } \|w\| = 1.$$

Assim,

$$\|w - \frac{x}{\|x\|}\|^2 = \|w\|^2 - 2 \left\langle w, \frac{x}{\|x\|} \right\rangle + 1 = 0.$$

Portanto $w = \frac{x}{\|x\|}$. Agora afirmamos que $\partial f(0) = \overline{B}(0; 1)$. Da Definição 1.1.4 temos que

$$y \in \partial f(0) \iff \|x\| \geq \langle y, x \rangle \quad \forall x \in \mathcal{H}.$$

Em particular, substituindo $x = y$ na desigualdade acima, obtemos

$$\|y\|(\|y\| - 1) \leq 0.$$

Portanto $\|y\| \leq 1$ o que equivale a $y \in \overline{B}(0; 1)$. Por outro lado, seja $y \in \overline{B}(0; 1)$, então

$$\langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\| \|y\| \leq \|x\| \quad \forall x \in \mathcal{H}.$$

Logo $y \in \partial f(0)$.

(iii) Seja $f : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função definida por

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

Pelo ítem (ii) obtemos que

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

Os próximos dois resultados são concernentes a operações com subdiferenciais.

Teorema 1.1.3 ([44], Teoremas 5.38 e 5.39) *Sejam $f_1, f_2 : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ funções convexas e próprias. Então*

1. $\partial f_1(x) + \partial f_2(x) \subseteq \partial(f_1 + f_2)(x) \quad \forall x \in \mathcal{H}.$

2. *Se uma das condições abaixo é verificada*

- (a) *Existe um ponto $\bar{u} \in \text{dom}(f_1) \cap \text{dom}(f_2)$ onde f_1 (ou f_2) é contínua.*

- (b) *O espaço \mathcal{H} tem dimensão finita e $\text{ri}(\text{dom}(f_1)) \cap \text{ri}(\text{dom}(f_2)) \neq \emptyset$.*

Então

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall x \in \mathcal{H}.$$

Proposição 1.1.1 ([12], pág. 27 e Proposição 5.7) *Sejam \mathcal{V} e \mathcal{Y} dois espaços de Hilbert reais, $\Lambda : \mathcal{V} \longrightarrow \mathcal{Y}$ uma aplicação linear contínua com adjunta $\Lambda^* : \mathcal{Y} \longrightarrow \mathcal{V}$ e $f : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e semicontínua inferior. Então*

1. *A função $f \circ \Lambda : \mathcal{V} \longrightarrow \mathbb{R}$ é convexa e semicontínua inferior.*

2. *Se existe um ponto $\bar{u} \in \mathcal{V}$, onde f é finita e contínua em $\Lambda(\bar{u})$, então para todo ponto u de \mathcal{V} , temos*

$$\partial(f \circ \Lambda)(u) = \Lambda^* \partial f(\Lambda(u)).$$

1.2 Aplicações ponto-conjunto

Definição 1.2.1 [1] *Consideremos dois conjuntos X e Y . Uma **aplicação ponto-conjunto** de X em Y , é uma relação que associa a cada elemento $x \in X$ um subconjunto $F(x)$ de Y dito imagem ou valor de F em x . Denotamos por $F : X \longrightarrow \mathcal{P}(Y)$ a aplicação ponto-conjunto F .*

As aplicações ponto-conjunto aparecem naturalmente em muitas áreas da matemática, como por exemplo na otimização. Um exemplo clássico é o subdiferencial de uma função convexa não diferenciável, isto é

$$\begin{aligned} \partial f : \mathcal{H} &\longrightarrow \mathcal{P}(\mathcal{H}) \\ x &\longmapsto \partial f(x). \end{aligned}$$

Para mais detalhes sobre aplicações ponto-conjunto, ver por exemplo [1], [2] em dimensão infinita e [40] em dimensão finita.

Definição 1.2.2 [1] *Seja $F : X \longrightarrow \mathcal{P}(Y)$.*

1. *O gráfico de F é o conjunto:*

$$\text{Gr}(F) := \{(x, y) \in X \times Y : y \in F(x)\}.$$

2. *O domínio de F é o conjunto:*

$$\text{Dom}(F) := \{x \in X : F(x) \neq \emptyset\}.$$

3. *A imagem de F é o conjunto:*

$$\text{Im}(F) := \bigcup_{x \in X} F(x).$$

4. *A inversa $F^{-1} : Y \longrightarrow \mathcal{P}(X)$ de F é definida por:*

$$F^{-1}(y) := \{x \in X : y \in F(x)\}.$$

5. *A aplicação ponto-conjunto F é dita **própria** se $\text{Gr}(F) \neq \emptyset$.*

O domínio de F é, portanto, a imagem de F^{-1} , e coincide com a projeção do gráfico de F sobre o espaço X . De forma simétrica, a imagem de F é igual ao domínio de F^{-1} , sendo a projeção do gráfico de F sobre o espaço Y .

Se, para cada $x \in \text{Dom}(F)$, o conjunto $F(x)$ reduz-se a um elemento, então F é uma aplicação no sentido clássico ou **ponto-ponto** o qual denotamos por $F(x) = y$ ou $F(x) = \{y\}$.

Na Figura 1.1 ilustramos o gráfico de uma aplicação ponto-conjunto $F : \mathbb{R}^n \longrightarrow \mathcal{P}(\mathbb{R}^m)$, o domínio e a imagem de F , a imagem $F(x)$ de um ponto x e a imagem da aplicação inversa F^{-1} de um ponto u .

Definição 1.2.3 [1] *Considere os conjuntos X, Y, Z , um espaço vetorial real V , os operadores A, B, C e D ; $A, B : X \longrightarrow \mathcal{P}(V)$, $C : X \longrightarrow \mathcal{P}(Y)$, $D : Y \longrightarrow \mathcal{P}(Z)$ e um escalar $\lambda \in \mathbb{R}$.*

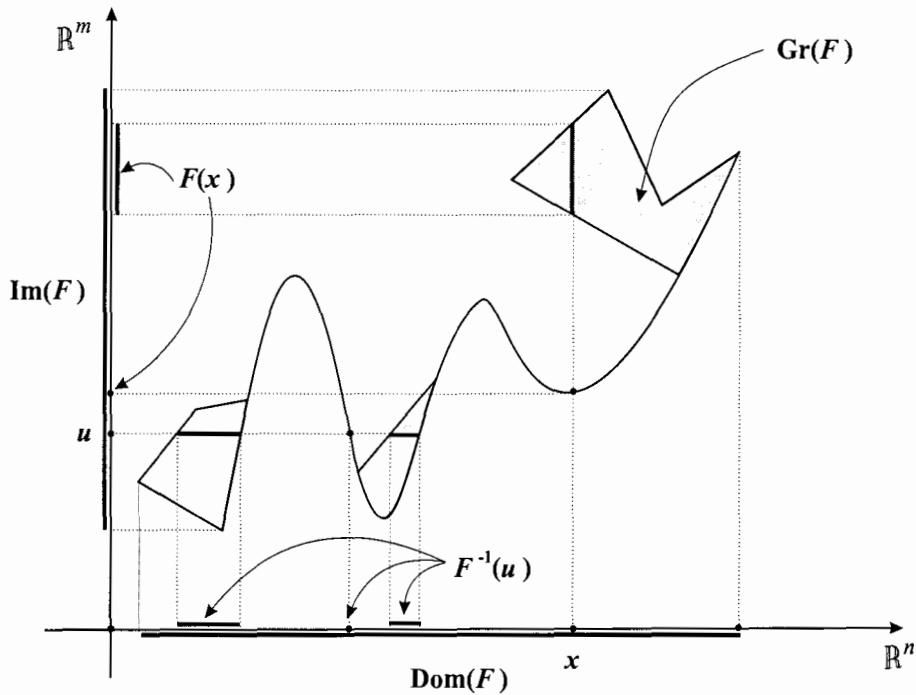


Figura 1.1:

1. A aplicação de **adição**, $A + B : X \rightarrow \mathcal{P}(V)$ é definida por:

$$(A + B)(x) := A(x) + B(x) = \{y + z \in V : y \in A(x), z \in B(x)\}.$$

2. A aplicação de **multiplicação por um escalar**, $\lambda A : X \rightarrow \mathcal{P}(V)$ é definida por:

$$(\lambda A)(x) := \lambda A(x) = \{\lambda y \in V : y \in A(x)\}.$$

3. A aplicação **composição** $D \circ C : X \rightarrow \mathcal{P}(Z)$ é definida por:

$$(D \circ C)(x) := \bigcup_{y \in C(x)} D(y).$$

Observamos da Definição 1.2.3(3) que

$$\text{Gr}(D \circ C) = \{(x, z) \in X \times Z : \exists y \in Y \text{ com } y \in C(x), z \in D(y)\}.$$

1.3 Operadores monótonos

Nesta seção, discutimos brevemente os operadores monótonos e algumas propriedades destes. A teoria de operadores monótonos tem aplicações em várias

áreas da matemática, tais como, desigualdades variacionais, equações diferenciais e otimização, ver por exemplo [32], [21], [4] e [12].

Definição 1.3.1 [33] *Seja $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ uma aplicação ponto-conjunto. T é dito um operador monótono se*

$$\langle y' - y, x' - x \rangle \geq 0 \quad \forall x, x' \in \mathcal{H}, \forall y' \in T(x'), \forall y \in T(x).$$

A seguinte propriedade é útil para provar a monotonicidade de um operador.

Propriedade 1.3.1 ([33], pág. 197) *Seja $T : \mathcal{H} \longrightarrow \mathcal{H}$ um operador linear. Então T é monótono se, e somente se, é um operador **positivo**, isto é,*

$$\langle T(x), x \rangle \geq 0 \quad \forall x \in \mathcal{H}.$$

Prova: Se T é monótono, considerando $x' = 0$ na Definição 1.3.1, e da linearidade de T , obtemos

$$0 \leq \langle T(0) - T(x), 0 - x \rangle = \langle T(x), x \rangle \quad \forall x \in \text{Dom}(T) = \mathcal{H}.$$

Por outro lado, se T é um operador positivo,

$$0 \leq \langle T(x' - x), x' - x \rangle = \langle T(x') - T(x), x' - x \rangle \quad \forall x', x \in \mathcal{H}. \quad \blacksquare$$

Exemplo 1.3.1

- (i) [33, pág. 194] *Seja D um subconjunto não vazio de \mathbb{R} . Uma função $\varphi : D \longrightarrow \mathbb{R}$ define um operador monótono se, e somente se, φ é monótona não decrescente no sentido usual, isto é*

$$[\varphi(t_2) - \varphi(t_1)](t_2 - t_1) \geq 0 \quad \forall t_1, t_2 \in D \iff \varphi(t_1) \leq \varphi(t_2) \text{ sempre que } t_1 \leq t_2.$$

- (ii) Os operadores $T_1, T_2 : \mathbb{R} \longrightarrow \mathcal{P}(\mathbb{R})$, definidos por

$$T_1(x) = \begin{cases} x - 1 & \text{se } x < 0, \\ \{-1, 1\} & \text{se } x = 0, \\ x + 1 & \text{se } x > 0, \end{cases} \quad \text{e} \quad T_2(x) = \begin{cases} x - 1 & \text{se } x < 0, \\ [0, 1] & \text{se } x = 0, \\ x + 1 & \text{se } x > 0, \end{cases}$$

são monótonos.

(iii) [32, pág. 102] Seja $f : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e subdiferenciável, então ∂f é um operador monótono, pois para qualquer $y \in \partial f(x)$ e $y' \in \partial f(x')$, temos

$$\begin{aligned} f(x') - f(x) &\geq \langle y, x' - x \rangle, \\ f(x) - f(x') &\geq \langle y', x - x' \rangle, \end{aligned}$$

somando ambas desigualdades obtemos

$$\langle y - y', x - x' \rangle \geq 0 \quad \forall x, x' \in \text{Dom}(\partial f).$$

Assim os subdiferenciais das funções definidas no Exemplo 1.1.2 são operadores monótonos.

(iv) Dado $\mathcal{H} = l_2 := \{\xi = (\xi_1, \xi_2, \dots) : \xi_i \in \mathbb{R} \forall i \in \mathbb{N} \text{ e } \sum_{i=1}^{\infty} |\xi_i|^2 < \infty\}$. Consideramos o seguinte produto interno em l_2

$$\langle \xi, \varphi \rangle = \sum_{i=1}^{\infty} \xi_i \varphi_i.$$

É fato bem conhecido que, com este produto interno, l_2 é um espaço de Hilbert real, ver por exemplo [26], pág. 49. Seja $T : l_2 \longrightarrow l_2$ definido por

$$T(\xi) = \left(\frac{\xi_1}{1}, \frac{\xi_2}{2}, \dots \right) \quad \forall \xi \in l_2.$$

Notamos que T é linear, além disso, para qualquer $\xi \in l_2$ temos

$$\langle T(\xi), \xi \rangle = \sum_{i=1}^{\infty} \frac{1}{i} \xi_i^2 \geq 0.$$

Portanto T é positivo e, pela Propriedade 1.3.1, concluímos que T é monótono. ■

Definição 1.3.2 [33] *Um operador monótono $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ é dito **monótono maximal** se não existe outro operador monótono cujo gráfico contenha estritamente o gráfico de T , de maneira equivalente, se $\tilde{T} : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ é outro operador monótono tal que $\tilde{T}(x) \supseteq T(x) \forall x \in \mathcal{H}$, implica que $\tilde{T} = T$.*

A próxima propriedade é útil para provar a maximalidade de um operador.

Propriedade 1.3.2 ([33], pág. 197) *Seja $T : \mathcal{H} \longrightarrow \mathcal{H}$ um operador linear e positivo. Então T é monótono maximal.*

Prova: Suponhamos que $(x, x^*) \in \text{Gr}(\tilde{T})$, onde \tilde{T} é outro operador monótono tal que $\text{Gr}(T) \subseteq \text{Gr}(\tilde{T})$, então para qualquer $z \in \mathcal{H}$ e $\lambda > 0$ temos

$$\begin{aligned} 0 \leq \langle T(x \pm \lambda z) - x^*, (x \pm \lambda z) - x \rangle &= \pm \lambda \langle T(x) \pm \lambda T(z) - x^*, z \rangle \\ &= \pm \lambda \langle T(x) - x^*, z \rangle + \lambda^2 \langle T(z), z \rangle. \end{aligned}$$

Dividindo por λ e fazendo $\lambda \longrightarrow 0$, obtemos que

$$\langle T(x) - x^*, z \rangle = 0 \quad \forall z \in \mathcal{H}.$$

Portanto, $x^* = T(x)$. ■

Exemplo 1.3.2

(i) [33, pág. 197] Dado φ uma função não decrescente em \mathbb{R} . Então φ é um operador monótono maximal se, e somente se, $\varphi(x) = [\varphi(x^-), \varphi(x^+)] \forall x \in \mathbb{R}$, onde $\varphi(x^-) := \lim_{t \rightarrow x^-} \varphi(t)$ e $\varphi(x^+) := \lim_{t \rightarrow x^+} \varphi(t)$.

(ii) O operador $\tilde{T} : \mathbb{R} \longrightarrow \mathcal{P}(\mathbb{R})$ definido por

$$\tilde{T}(x) = \begin{cases} x - 1 & \text{se } x < 0, \\ [-1, 1] & \text{se } x = 0, \\ x + 1 & \text{se } x > 0, \end{cases}$$

é monótono maximal.

(iii) Os operadores T_1 e T_2 definidos no item (ii) do Exemplo 1.3.1 não são maximais, posto que $\text{Gr}(T_i) \subsetneq \text{Gr}(\tilde{T})$ para $i = 1, 2$, onde \tilde{T} é o operador definido no item (ii).

(iv) O operador ∂f definido no item (iii) do Exemplo 1.1.2, não é maximal, pois $\text{Gr}(\partial f) \subsetneq \text{Gr}(T)$ onde $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ é o operador monótono definido por

$$T(x) = \begin{cases} \frac{x}{\|x\|} & \text{se } x \neq 0 \text{ e } \|x\| \leq 1, \\ \overline{B}(0; 1) & \text{se } x = 0, \\ \emptyset & \text{se } \|x\| > 1. \end{cases}$$

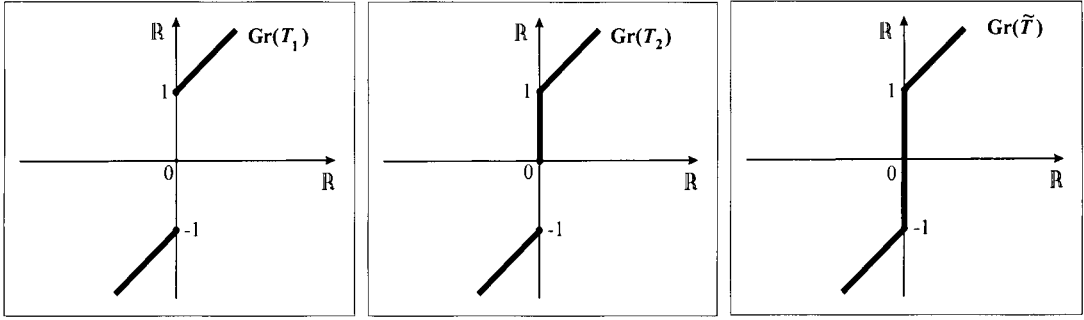


Figura 1.2: Gráficos dos operadores T_1 , T_2 e \tilde{T} do Exemplo 1.3.2(iii).

- (v) Pela Propriedade 1.3.2, os operadores ∂f e T definidos no ítem (i) do Exemplo 1.1.2 e no ítem (iv) do Exemplo 1.3.1 respectivamente, são maximais, pois são lineares e positivos, com $\text{Dom}(\partial f) = \mathcal{H}$ e $\text{Dom}(T) = l_2$.
- (vi) O operador $T : \mathbb{R} \longrightarrow \mathbb{R}$ definido por

$$T(x) = \begin{cases} -\frac{1}{x} & \text{se } x < 0, \\ \emptyset & \text{se } x \geq 0, \end{cases}$$

é monótono maximal com $\text{Dom}(T) = (-\infty, 0)$. De fato, suponhamos que $\text{Gr}(T) \subsetneq \text{Gr}(\tilde{T})$, onde \tilde{T} é outro operador monótono, portanto existe $(x, y) \in \text{Gr}(\tilde{T})$ tal que $(x, y) \notin \text{Gr}(T)$. Como T é contínua em $\text{Dom}(T)$, temos que $x \geq 0$. Em particular, consideramos $x = 0$ e

$$\left\langle -\frac{1}{x} - y, x - 0 \right\rangle \geq 0 \quad \forall x < 0.$$

Então

$$-\frac{1}{x} \leq y \quad \forall x < 0.$$

Fazemos $x \longrightarrow 0^-$ e obtemos $y = +\infty$, o que é um absurdo. Portanto $\tilde{T}(0) = \emptyset$. Assim, T é monótono maximal. ■

Em muitas situações é difícil verificar diretamente da definição, quando um operador monótono é maximal por tal motivo apresentamos o seguinte resultado.

Proposição 1.3.1 ([1], Proposição 6.7.2) *Seja $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$. Então T é monótono maximal se, e somente se, satisfaz*

$$\forall (x, y) \in \text{Gr}(T), \langle y' - y, x' - x \rangle \geq 0 \implies y' \in T(x'). \quad (1.8)$$

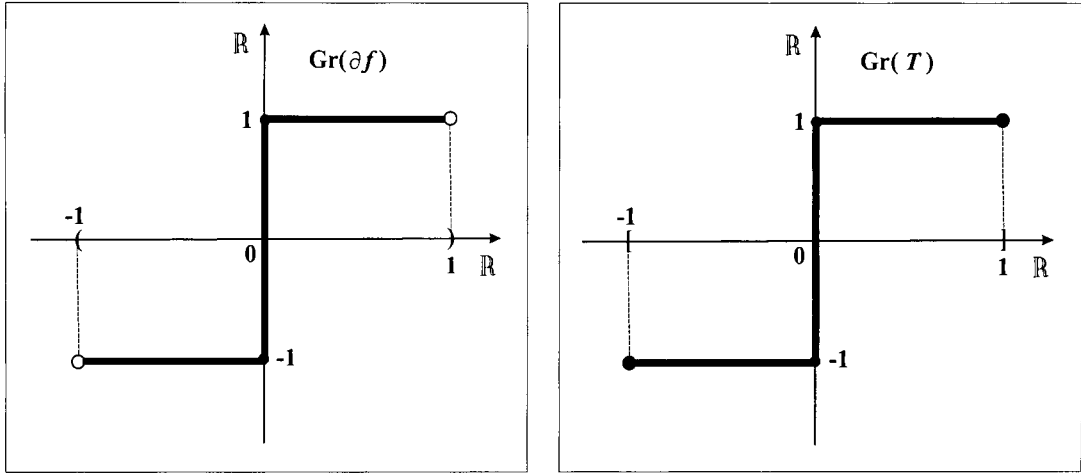


Figura 1.3: Gráficos dos operadores ∂f e T do Exemplo 1.3.2(iv).

Apresentamos uma propriedade importante dos operadores monótonos maximais a seguir.

Proposição 1.3.2 ([32], pág. 105) *Seja $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal. Então o conjunto $T(x)$ é convexo e fechado para todo $x \in \mathcal{H}$.*

Definimos a seguir um importante operador na teoria da otimização.

Definição 1.3.3 [37] *Seja C um subconjunto de \mathcal{H} . O operador $N_C : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ definido por*

$$N_C(x) = \begin{cases} \{w \in \mathcal{H} : \langle w, z - x \rangle \leq 0 \quad \forall z \in C\} & \text{se } x \in C, \\ \emptyset & \text{se } x \notin C, \end{cases}$$

é chamado de operador **cone normal** de C .

A seguir, apresentamos algumas propriedades básicas de N_C .

Proposição 1.3.3 ([10], Proposição 3.4) *Seja C um subconjunto convexo não vazio de \mathcal{H} . Então*

1. $N_C = \partial\delta_C$.
2. O operador N_C é monótono.
3. Se C é fechado então N_C é monótono maximal.

Corolário 1.3.1 ([32], pág. 134) *Seja A um subespaço de \mathcal{H} . Então N_A é um operador monótono maximal e*

$$N_A(x) = \begin{cases} A^\perp & \text{se } x \in A, \\ \emptyset & \text{se } x \notin A. \end{cases} \quad (1.9)$$

Abaixo, enunciamos algumas operações que preservam a monotonicidade.

Proposição 1.3.4 ([4], pág. 24) *Sejam $T_1, T_2 : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ operadores e um escalar $\lambda > 0$. Então*

1. *O operador T_1 é monótono (maximal) se, e somente se, λT_1 é monótono (maximal).*
2. *O operador T_1 é monótono (maximal) se, e somente se, T_1^{-1} é monótono (maximal).*
3. *Se T_1 e T_2 são monótonos, então $T_1 + T_2$ é monótono.*

A soma de dois operadores monótonos maximais não é, necessariamente, um operador monótono maximal. A seguir, apresentamos dois exemplos, um em dimensão infinita e outro em dimensão finita, onde é verificada tal afirmação.

Exemplo 1.3.3 ([34], Exemplo 7.4) Dado $\mathcal{H} = l_2$, consideramos os operadores $V, W : l_2 \longrightarrow l_2$ definidos por

$$\begin{aligned} V(x) &= (x_1, x_2 - x_1, x_3 - x_2, \dots), \\ W(x) &= (x_1 - x_2, x_2 - x_3, x_3 - x_4, \dots). \end{aligned}$$

Ambos os operadores estão bem definidos, pois

$$\begin{aligned} V(x) &= x - x' \\ W(x) &= x - x'', \end{aligned}$$

onde

$$\begin{aligned} x &= (x_1, x_2, x_3, \dots, x_n, \dots) \in l_2, \\ x' &= (0, x_1, x_2, \dots, x_{n-1}, \dots) \in l_2, \\ x'' &= (x_2, x_3, x_4, \dots, x_{n+1}, \dots) \in l_2. \end{aligned}$$

Observamos que V e W são lineares e para todo ponto $x \in l_2$ temos $\langle x, x \rangle = \langle x', x' \rangle$.

Assim

$$\begin{aligned}
 \langle V(x), x \rangle &= \langle x - x', x \rangle \\
 &= \langle x, x \rangle - \langle x', x \rangle \\
 &= \frac{1}{2} \langle x', x' \rangle - \langle x', x \rangle + \frac{1}{2} \langle x, x \rangle \\
 &= \frac{1}{2} \|x - x'\|^2 \\
 &= \frac{1}{2} \|V(x)\|^2.
 \end{aligned}$$

Além disso, como $\langle x'', x \rangle = \langle x', x \rangle$, obtemos

$$\langle W(x), x \rangle = \frac{1}{2} \|V(x)\|^2.$$

Logo, V e W são positivos em $\text{Dom}(V) = \text{Dom}(W) = l_2$, portanto, pela Propriedade 1.3.2, ambos são monótonos maximais. Como V e W são lineares, contínuos e injetivos, podemos definir os operadores lineares $V^{-1}, W^{-1} : l_2 \longrightarrow l_2$ por

$$\begin{aligned}
 V^{-1}(y_1, y_2, y_3, \dots, y_n, \dots) &= (y_1, y_1 + y_2, y_1 + y_2 + y_3, \dots, \sum_{i=1}^n y_i, \dots), \\
 W^{-1}(y_1, y_2, y_3, \dots, y_n, \dots) &= (\sum_{k \geq 1} y_k, \sum_{k \geq 2} y_k, \dots, \sum_{k \geq n} y_k, \dots),
 \end{aligned}$$

com $\text{Dom}(V^{-1}) = \text{Im}(V)$ e $\text{Dom}(W^{-1}) = \text{Im}(W)$. Assim, pela Proposição 1.3.4, resulta que os operadores V^{-1} e W^{-1} são monótonos maximais. Por outro lado, afirmamos que $\text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$ é denso em l_2 , como verificamos a seguir. Seja

$$D = \{x \in l_2 : x \neq 0 \text{ tem um número finito de componentes não nulas, } \sum x_k = 0\}.$$

Afirmamos que D é denso em l_2 . Para ver isto é suficiente mostrar que D é denso em E , onde

$$E = \{x \in l_2 : x \neq 0 \text{ tem um número finito de componentes não nulas}\}.$$

De fato, seja $y = (y_1, y_2, \dots, y_m, 0, \dots) \in E$ e definamos $s = \sum y_k$. Então, para qualquer $n \geq 1$, consideramos

$$y^n := (y_1, y_2, \dots, y_m, \underbrace{-s/n, -s/n, \dots, -s/n}_{n \text{ termos}}, 0, \dots) \in D.$$

Assim

$$\begin{aligned}
\lim_{n \rightarrow +\infty} \|y^n - y\|^2 &= \lim_{n \rightarrow +\infty} \langle y^n - y, y^n - y \rangle \\
&= \lim_{n \rightarrow +\infty} \left\langle \left(0, \dots, 0, -\frac{s}{n}, -\frac{s}{n}, \dots, -\frac{s}{n}, 0, 0, \dots\right), \right. \\
&\quad \left. \left(0, \dots, 0, -\frac{s}{n}, -\frac{s}{n}, \dots, -\frac{s}{n}, 0, 0, \dots\right) \right\rangle \\
&= \lim_{n \rightarrow +\infty} \left(\left(-\frac{s}{n}\right)^2 + \left(-\frac{s}{n}\right)^2 + \dots + \left(-\frac{s}{n}\right)^2 \right) \\
&= 0.
\end{aligned}$$

Portanto D é denso em E , que é denso em l_2 . Assim obtemos que D é denso em l_2 . Se $x \in D$ então $V^{-1}(x)$ e $W^{-1}(x)$ estão bem definidos. De fato, sejam os pontos $y = (y_1, y_2, \dots)$ e $z = (z_1, z_2, \dots)$ onde

$$y_n = \sum_{k \leq n} x_k \text{ e } z_n = \sum_{k \geq n} x_k \quad n = 1, 2, \dots$$

e que verificam $x = V(y)$ e $x = W(z)$. Assim $D \subset \text{Im}(V^{-1}) \cap \text{Im}(W^{-1})$, e portanto $\text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$ é denso em l_2 .

Por outro lado, se $x \in \text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$ então $\sum_{k=1}^{\infty} x_k = 0$. De fato, como $x \in \text{Dom}(V^{-1}) = \text{Im}(V)$, existe $y \in l_2$ tal que $x = V(y)$,

$$x = (x_1, x_2, x_3, \dots) = (y_1, y_2 - y_1, y_3 - y_2, \dots).$$

Portanto

$$\sum_{k=1}^{\infty} x_k = \lim_{n \rightarrow +\infty} \sum_{k=1}^n x_k = \lim_{n \rightarrow +\infty} y_n = 0.$$

Logo, $\text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1}) \subsetneq l_2$. Consideramos o operador $V^{-1} + W^{-1}$ definido em $\text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$. Dado $x \in \text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$, sejam $y = (y_1, y_2, \dots, y_n, \dots) = V^{-1}(x)$ e $z = (z_1, z_2, \dots, z_n, \dots) = W^{-1}(x)$ e temos que

$$\begin{aligned}
(V^{-1} + W^{-1})(x) &= V^{-1}(x) + W^{-1}(x) \\
&= (x_1, x_1 + x_2, \dots, \sum_{k \leq n} x_k, \dots) + \left(\sum_{k \geq 1} x_k, \sum_{k \geq 2} x_k, \dots, \sum_{k \geq n} x_k, \dots \right) \\
&= (x_1, x_2, \dots, x_n, \dots) \\
&= x.
\end{aligned}$$

Assim, $V^{-1} + W^{-1}$ é o operador identidade restrito ao $\text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1}) \subsetneq l_2$, pois para todo $x \in \text{Dom}(V^{-1}) \cap \text{Dom}(W^{-1})$ temos que $\sum_{i=1}^{\infty} x_i = 0$. Portanto $V^{-1} + W^{-1}$ não é maximal monótono. ■

Exemplo 1.3.4 Consideramos $\mathcal{H} = \mathbb{R}^2$ e os conjuntos

$$\begin{aligned} C_1 &= \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 + 1)^2 + x_2^2 \leq 1\}, \\ C_2 &= \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 \leq 1\}. \end{aligned}$$

Notamos que C_1 e C_2 são convexos e fechados. Assim pela Proposição 1.3.3, podemos definir os operadores monótonos maximais $T_1 = N_{C_1}$ e $T_2 = N_{C_2}$. Daqui,

$$\text{Dom}(T_1 + T_2) = \text{Dom}(N_{C_1}) \cap \text{Dom}(N_{C_2}) = C_1 \cap C_2 = \{0\}.$$

Por outro lado,

$$\begin{aligned} N_{C_1}(0) &= \{w \in \mathbb{R}^2 : \langle w, z \rangle \leq 0 \quad \forall z \in C_1\} \\ &= \{(w_1, 0) \in \mathbb{R}^2 : w_1 \geq 0\}. \end{aligned}$$

Analogamente,

$$N_{C_2}(0) = \{(w_1, 0) \in \mathbb{R}^2 : w_1 \leq 0\}.$$

Portanto

$$(N_{C_1} + N_{C_2})(0) = N_{C_1}(0) + N_{C_2}(0) = \mathbb{R} \times \{0\}.$$

Então

$$\text{Gr}(N_{C_1} + N_{C_2}) = \{0\} \times \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}.$$

Mas $N_{C_1} + N_{C_2}$ não é monótono maximal, pois pela Proposição 1.3.3, $N_{\{0\}}$ é monótono maximal com $N_{\{0\}}(0) = \mathbb{R}^2$ e

$$\text{Gr}(N_{C_1} + N_{C_2}) \subsetneq \text{Gr}(N_{\{0\}}) = \{0\} \times \mathbb{R}^2. \quad \blacksquare$$

Em vista dos exemplos anteriores, enunciamos um resultado que garante quando a soma de dois operadores monótonos maximais é monótono maximal.

Teorema 1.3.1 ([37], Teoremas 1 e 2) *Sejam $T_1, T_2 : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais. Se uma das condições abaixo é verificada*

1. $\text{Dom}(T_1) \cap \text{int}(\text{Dom}(T_2)) \neq \emptyset$,
2. *O espaço \mathcal{H} tem dimensão finita e $\text{ri}(\text{Dom}(T_1)) \cap \text{ri}(\text{Dom}(T_2)) \neq \emptyset$.*

Então $T_1 + T_2$ é monótono maximal.

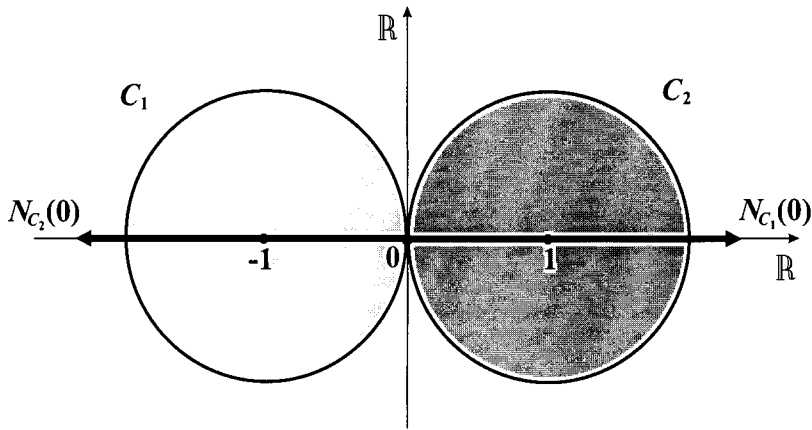


Figura 1.4: Gráficos dos conjuntos C_1 , C_2 , $N_{C_1(0)}$ e $N_{C_2(0)}$ do Exemplo 1.3.4.

Definição 1.3.4 [32] *O operador $T : \mathcal{H} \rightarrow \mathcal{H}$ é dito **hemicontínuo** no ponto $x \in \text{Dom}(T)$ se $\text{Dom}(T)$ é convexo e para qualquer $y \in \text{Dom}(T)$ a aplicação*

$$\begin{aligned} \tilde{T} : [0, 1] &\longrightarrow \mathcal{H} \\ t &\longmapsto \tilde{T}(t) := T((1-t)x + ty), \end{aligned}$$

é contínua na topologia fraca de \mathcal{H} .

O próximo resultado assegura quando a soma entre um operador hemicontínuo e o cone normal é maximal.

Teorema 1.3.2 ([37], Teorema 3) *Sejam K um subconjunto convexo, fechado e não vazio de \mathcal{H} , $T : \mathcal{H} \rightarrow \mathcal{H}$ um operador monótono e hemicontínuo em K tal que $K \subseteq \text{Dom}(T)$. Então o operador $T + N_K$ é monótono maximal.*

Definição 1.3.5 [1] *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador e um escalar $\lambda > 0$. O operador $J_{\lambda T} = (I + \lambda T)^{-1}$ é chamado o **resolvente** de T com respeito a λ .*

Um resultado importante na teoria dos operadores monótonos é o chamado Teorema de Minty, o qual apresentamos a seguir.

Teorema 1.3.3 ([1], pág. 380) *Um operador monótono $A : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ é maximal se, e somente se, o operador $I + A$ é sobrejetivo.*

Definição 1.3.6 *Sejam C um subconjunto convexo fechado de \mathcal{H} e $T : C \rightarrow \mathcal{H}$ um operador.*

1. [31] T é dito *não-expansivo* se

$$\|T(x) - T(y)\| \leq \|x - y\| \quad \forall x, y \in C.$$

2. [25] T é dito *firmemente não-expansivo* se

$$\langle T(x) - T(y), x - y \rangle \geq \|T(x) - T(y)\|^2 \quad \forall x, y \in C.$$

Teorema 1.3.4 ([1], pág. 383) *Seja $A : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal. Então, para todo $\lambda > 0$, o resolvente $(I + \lambda A)^{-1}$ é um operador ponto-ponto e não-expansivo de \mathcal{H} em \mathcal{H} .*

Definição 1.3.7 [11] *Seja $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador. Um zero de T é um ponto $x \in \mathcal{H}$ tal que $0 \in T(x)$. O conjunto dos zeros de T será denotado por $T^{-1}(0)$.*

O próximo resultado assegura que os zeros de um operador monótono maximal são os pontos fixos de seu resolvente.

Proposição 1.3.5 ([11], Lema 2) *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal, um ponto $x \in \mathcal{H}$ e um escalar $\lambda > 0$. Então $0 \in T(x)$ se, e somente se, $J_{\lambda T}(x) = x$.*

Em seguida apresentamos um resultado que assegura quando o subdiferencial de uma função convexa é maximal.

Teorema 1.3.5 ([38], Teorema A) *Seja $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semicontínua inferior. Então $\partial f : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ é um operador monótono maximal.*

Exemplo 1.3.5

- (i) Os subdiferenciais das funções $\frac{1}{2}\|\cdot\|^2$ e $\|\cdot\|$, definidas nos itens (i) e (ii) do Exemplo 1.1.2, satisfazem as condições do Teorema 1.3.5, e portanto são monótonos maximais, como notamos no Exemplo 1.3.2(v).
- (ii) A função f definida no item (iii) do Exemplo 1.1.2 não é semicontínua inferior nos pontos $x \in \mathcal{H}$ tais que $\|x\|=1$, e portanto, ∂f não é um operador monótono maximal como vimos no Exemplo 1.3.2(iv).

(iii) Seja $g : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função definida por

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

Observamos que g é convexa, própria e semicontínua inferior. Portanto, pelo Teorema 1.3.5, o subdiferencial ∂g é monótono maximal. Por outro lado, observamos que $g = \|\cdot\| + \delta_{B(0;1)}$, portanto, do Teorema 1.1.3 e a Proposição 1.3.3, obtemos

$$\partial g(x) = \partial \|x\| + N_{B(0;1)}(x) \quad \forall x \in \mathcal{H},$$

assim, pelo ítem (ii) do Exemplo 1.1.2, concluímos que

$$\partial g(x) = \begin{cases} \frac{x}{\|x\|} & \text{se } x \neq 0 \text{ e } \|x\| < 1, \\ \overline{B}(0;1) & \text{se } x = 0, \\ \{x\} + N_{B(0;1)}(x) & \text{se } \|x\| = 1, \\ \emptyset & \text{se } \|x\| > 1. \end{cases}$$

O que implica que $\text{Gr}(T) \subsetneq \text{Gr}(\partial g)$, onde T é o operador monótono definido no Exemplo 1.3.2(iv). Portanto T não é maximal. ■

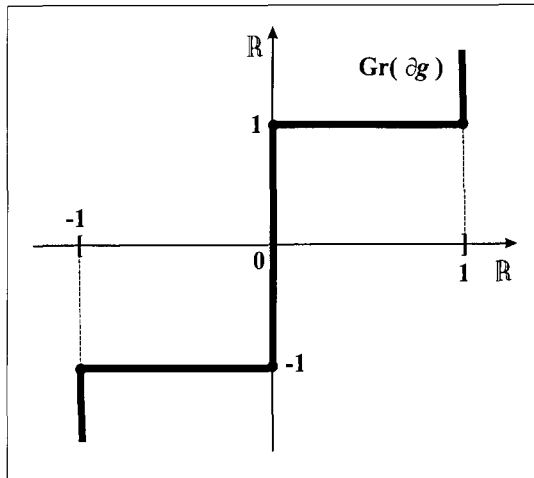


Figura 1.5: Gráfico do ∂g definido no Exemplo 1.3.5(iii).

A recíproca do Teorema 1.3.5 não é verdadeira; isto é, nem todo operador monótono maximal é subgradiente de uma função convexa.

Exemplo 1.3.6 Seja $\mathcal{H} = \mathbb{R}^2$ e consideramos o operador $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definido por

$$T(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x.$$

Observamos que T é um operador linear e positivo. Então, pela Propriedade 1.3.2, T é monótono maximal. Suponhamos que $T = \partial f$, para alguma função f convexa e própria. Como $\text{Dom}(T) = \mathbb{R}^2$, temos que para todo x_0 , $\partial f(x_0) = T(x_0) \neq \emptyset$. Assim de (1.1), obtemos

$$f(x) \geq f(x_0) + \langle T(x_0), x - x_0 \rangle \quad \forall x \in \mathbb{R}^2. \quad (1.10)$$

Em particular, consideramos $x_0 = (1, 1)^\top$ e $x = (-1, 0)^\top$; $x_0 = (-1, 0)^\top$ e $x = (0, -1)^\top$; $x_0 = (0, -1)^\top$ e $x = (1, 1)^\top$ em (1.10), respectivamente, e somando obtemos

$$0 \geq \langle (-1, 1)^\top, (-2, -1)^\top \rangle + \langle (0, -1)^\top, (1, -1)^\top \rangle + \langle (1, 0)^\top, (1, 2)^\top \rangle = 3,$$

o que é um absurdo, portanto T não é o subdiferencial de f . ■

Definição 1.3.8 [39] *Um operador $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ é fortemente monótono se existe um escalar $\alpha > 0$ tal que*

$$\langle y' - y, x' - x \rangle \geq \alpha \|x' - x\|^2 \quad \forall x, x' \in \mathcal{H}, \forall y \in T(x), \forall y' \in T(x'),$$

onde α é chamado o *módulo (ou coeficiente) de monotonicidade*.

Exemplo 1.3.7

(i) O operador identidade $I : \mathcal{H} \longrightarrow \mathcal{H}$ definido por

$$I(x) = x \quad \forall x \in \mathcal{H}$$

é fortemente monótono, sendo $\alpha \in (0, 1]$, posto que $\forall x, x' \in \mathcal{H}$

$$\|x' - x\|^2(1 - \alpha) \geq 0 \implies \|x' - x\|^2 = \langle x' - x, x' - x \rangle \geq \alpha \|x' - x\|^2.$$

(ii) [47, pág. 453] O operador $T : \mathbb{R} \longrightarrow \mathcal{P}(\mathbb{R})$ definido por

$$T(x) = \begin{cases} -e^{-x} & \text{se } x < 0, \\ [-1, 1] & \text{se } x = 0, \\ e^x & \text{se } x > 0, \end{cases}$$

é fortemente monótono com $\alpha = 1$. De fato, devido à convexidade e à diferenciabilidade da função exponencial, temos de (1.1) que

$$e^y - e^x \geq e^x(y - x) \quad \forall x, y \in \mathbb{R}. \quad (1.11)$$

Examinamos o seguintes casos:

- Se $x < y < 0$, então de (1.11) temos

$$(e^y - e^x)(y - x) \geq e^x(y - x)^2;$$

assim,

$$(e^{-x} - e^{-y})(y - x) = \frac{e^y - e^x}{e^{x+y}}(y - x) \geq \frac{e^x(y - x)^2}{e^{x+y}} = e^{-y}(y - x)^2 \geq (y - x)^2.$$

- Se $x < 0 < y$, então

$$\begin{aligned} e^{-x} \geq -x &\implies e^y + e^{-x} \geq e^y - x \geq y - x \\ &\implies (e^y + e^{-x})(y - x) \geq (y - x)^2. \end{aligned}$$

- Se $0 < x < y$, então de (1.11)

$$(e^y - e^x)(y - x) \geq e^x(y - x)^2 \geq (y - x)^2.$$

- Se $x < 0$ e $y = 0$, então, para $\alpha \in [-1, 1]$, temos que

$$-e^{-x} - 1 \leq -e^{-x} - \alpha \leq -e^{-x} + 1.$$

Substituímos $y = 0$ em (1.11) e obtemos

$$1 \geq e^x - e^x x \implies e^x - 1 \leq e^x x \implies \frac{e^x - 1}{e^x} \leq x \implies 1 - e^{-x} \leq x;$$

assim

$$-e^{-x} - \alpha \leq -e^{-x} + 1 \leq x,$$

portanto

$$(-e^{-x} - \alpha)x \geq (-e^{-x} + 1)x \geq x^2 \quad \forall \alpha \in [-1, 1].$$

- Se $y = 0$ e $x > 0$, então, para $\alpha \in [-1, 1]$, temos que

$$-1 - e^x \leq \alpha - e^x \leq 1 - e^x.$$

Substituímos $y = x$ e $x = 0$ em (1.11), portanto

$$e^x \geq 1 + x \implies -x \geq 1 - e^x,$$

assim

$$\alpha - e^x \leq 1 - e^x \leq -x,$$

o que implica

$$-x(\alpha - e^x) \geq -x(1 - e^x) \leq (-x)^2 = x^2 \quad \forall \alpha \in [-1, 1]. \quad \blacksquare$$

Observamos que um operador fortemente monótono é monótono, mas a recíproca não é verdadeira, para ver isto consideramos o operador T do Exemplo 1.3.6.

Proposição 1.3.6 ([10], Lema 3.1) *Seja $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador. Se T^{-1} é fortemente monótono então T é ponto-ponto.*

Prova: Consideramos $x \in \text{Dom}(T)$ e $y', y \in T(x)$, assim $x \in T^{-1}(y')$ e $x \in T^{-1}(y)$. Então

$$0 = \langle x - x, y' - y \rangle \geq \alpha \|y' - y\|^2.$$

Daqui $y' = y$. ■

Definição 1.3.9 [39] *Seja $h : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$. A função h é dita **fortemente convexa** (com módulo α) se $\alpha > 0$ e*

$$h((1-\lambda)x + \lambda y) \leq (1-\lambda)h(x) + \lambda h(y) - \frac{1}{2}\alpha\lambda(1-\lambda)\|x-y\|^2 \quad \forall x, y \in \mathcal{H}, \forall \lambda \in [0, 1].$$

Proposição 1.3.7 ([39], Proposição 6) *Seja $h : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria e subdiferenciável. Então h é fortemente convexa com módulo $\alpha > 0$ se, e somente se, o operador ∂h é fortemente monótono com módulo α .*

Exemplo 1.3.8

- (i) A função $f := \frac{1}{2}\|\cdot\|^2$ do Exemplo 1.1.2(i), é fortemente convexa, pois ∂f coincide com o operador identidade que é fortemente monótono, como provamos no Exemplo 1.3.7(i).

(ii) A função $h : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$h(x) = \begin{cases} e^{-x} & \text{se } x \leq 0, \\ e^x & \text{se } x > 0, \end{cases}$$

é fortemente convexa, visto que $\partial h = T$, onde T é o operador definido no Exemplo 1.3.7(ii). ■

1.4 Operadores pseudomonótonos

Karamardian [20], introduziu os operadores pseudomonótonos como uma caracterização dos gradientes de funções pseudoconvexas em dimensão finita, motivado pela caracterização das funções convexas via o gradiente. Tal conceito difere da definição encontrada, por exemplo em [32], pois não refere-se às propriedades topológicas, e sim às algébricas. Tais operadores têm sido objeto de intenso estudo na última década. As pesquisas incluem critérios para assegurar pseudomonotonicidade e problemas de desigualdades variacionais. Embora o caso monótono tenha sido muito explorado, o caso pseudomonótono ainda está em desenvolvimento, para mais detalhes ver [15], [14], [18] e as referências neles citadas. A seguir damos a definição de pseudomonotonicidade.

Definição 1.4.1 [14] *Seja $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$. O operador T é dito **pseudomonótono** se para todo $(x, y), (x', y') \in \text{Gr}(T)$, a condição abaixo é verificada*

$$\langle y, x' - x \rangle \geq 0 \implies \langle y', x' - x \rangle \geq 0.$$

Muitos exemplos de operadores pseudomonótonos são devido ao seguinte:

Proposição 1.4.1 *Se $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ é um operador monótono então T é pseudomonótono.*

Prova: Consideramos $(x, y), (x', y') \in \text{Gr}(T)$. Da monotonicidade de T ,

$$\langle y' - y, x' - x \rangle = \langle y', x' - x \rangle - \langle y, x' - x \rangle \geq 0,$$

assim

$$\langle y', x' - x \rangle \geq \langle y, x' - x \rangle,$$

obtemos a proposição. ■

Exemplo 1.4.1

(i) O operador $T : \mathbb{R} \longrightarrow \mathbb{R}$ definido por

$$T(x) = \begin{cases} \frac{1}{1+x} & \text{se } x > -1, \\ \emptyset & \text{se } x \leq -1, \end{cases}$$

é pseudomonótono. De fato, para $(x, \frac{1}{1+x}), (y, \frac{1}{1+y}) \in \text{Gr}(T)$ com $\langle \frac{1}{1+x}, y-x \rangle \geq 0$, implica que $y-x \geq 0$ visto que $\frac{1}{1+x} > 0$. Assim, obtemos que $\langle \frac{1}{1+y}, y-x \rangle \geq 0$.

(ii) [14, pág. 466] Sejam $a, b \in \mathbb{R}$ tais que $a \leq b$ e $f : \mathbb{R} \longrightarrow \mathbb{R}$ uma função continuamente diferenciável tal que

$$f'(x) = \begin{cases} < 0 & \text{se } x < a, \\ 0 & \text{se } x \in [a, b], \\ > 0 & \text{se } x > b. \end{cases}$$

Então $T(x) := \{f'(x)\}$ é um operador pseudomonótono.

(iii) [14, pág. 466] O operador $\tilde{T} : \mathbb{R} \longrightarrow \mathcal{P}(\mathbb{R})$ definido por

$$\tilde{T}(x) = \begin{cases} (-\infty, 0) & \text{se } x < a, \\ (-\infty, 0] & \text{se } x = a, \\ 0 & \text{se } x \in (a, b), \\ [0, +\infty) & \text{se } x = b, \\ (0, +\infty) & \text{se } x > b, \end{cases}$$

é um operador pseudomonótono, ver figura 1.6.

Observamos que o operador T definido no ítem (i) do Exemplo 1.4.1 não é monótono, e portanto a recíproca da Proposição 1.4.1 não é verdadeira. Além disso, $\text{Gr}(T)$ do ítem (ii) está contido em $\text{Gr}(\tilde{T})$ do ítem (iii) do Exemplo 1.4.1, isto sugere a maximalidade para este tipo de operadores, como no caso monótono, ver [14] para mais detalhes.

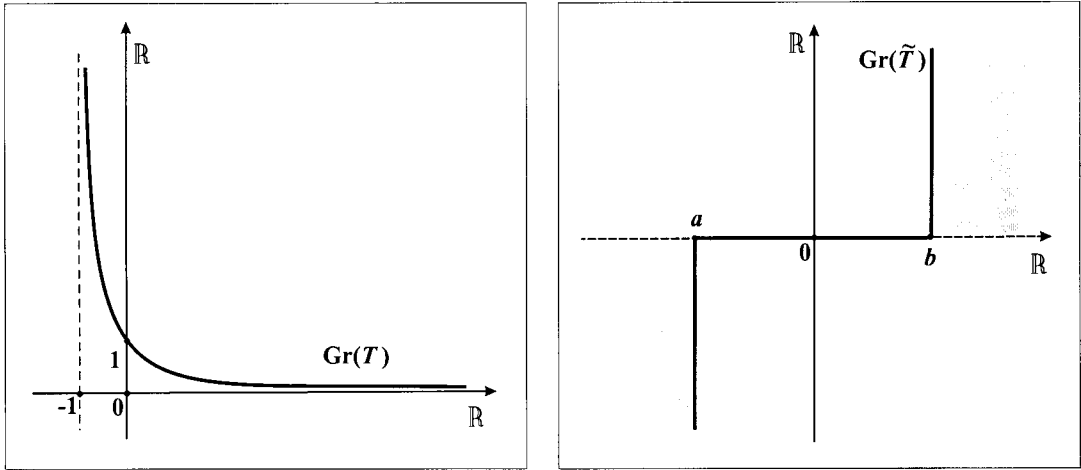


Figura 1.6: Gráficos dos operadores dos ítems (i) e (iii) do Exemplo 1.4.1.

1.5 Projção sobre um conjunto convexo

Nesta seção, apresentamos resultados com respeito à projeção sobre um conjunto convexo fechado e algumas propriedades importantes, além de discutir os casos particulares quando o conjunto é um subespaço ou um cone.

Teorema 1.5.1 ([45], Lema 2.1) *Sejam C um subconjunto convexo, fechado e não vazio de \mathcal{H} e um ponto x de \mathcal{H} . Então existe um único $\bar{x} \in C$ tal que*

$$\|x - \bar{x}\| = \min_{c \in C} \|x - c\|.$$

O resultado acima garante a boa definição do seguinte operador.

Definição 1.5.1 [45] *Seja C um subconjunto convexo, fechado não vazio de \mathcal{H} . O operador*

$$\begin{aligned} \text{proj}_C : \mathcal{H} &\longrightarrow C \\ x &\longmapsto \text{proj}_C(x) = \arg \min\{\|x - c\| : c \in C\} \end{aligned}$$

é denominado **projção sobre C** .

Diretamente do Teorema 1.5.1 observamos que $\text{proj}_C(x) = x$ para todo $x \in C$. A seguir, enunciamos algumas propriedades do operador proj_C .

Proposição 1.5.1 ([46], Lema 1.1 e pág. 241) *Sejam C um subconjunto convexo, fechado e não vazio de \mathcal{H} e um ponto $x \in \mathcal{H}$. Então*

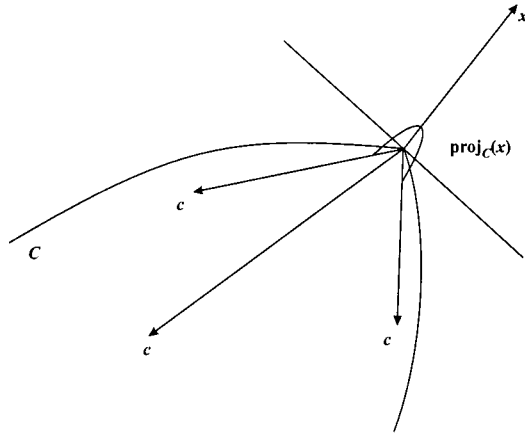


Figura 1.7: Projeção sobre um conjunto convexo.

1. $\bar{x} = \text{proj}_C(x)$ se, e somente se, $\bar{x} \in C$ e $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0 \quad \forall y \in C$.
2. $\langle \text{proj}_C(x) - \text{proj}_C(y), x - y \rangle \geq \|\text{proj}_C(x) - \text{proj}_C(y)\|^2 \quad \forall y \in C$.
3. $\|\text{proj}_C(x) - \text{proj}_C(y)\| \leq \|x - y\| \quad \forall x, y \in \mathcal{H}$.

Observamos que o ítem 1 da Proposição 1.5.1, fornece uma importante caracterização do operador projeção, o ítem 2 assegura que o operador proj_C é monótono, enquanto o ítem 3 garante que o operador proj_C é não-expansivo. Vemos agora alguns exemplos onde o operador de projeção tem uma representação explícita.

Exemplo 1.5.1

1. [3, pág. 41] Sejam $a \in \mathcal{H} \setminus \{0\}$ e $b \in \mathbb{R}$. Consideramos o *hiperplano* H

$$H = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

Como H é um conjunto convexo e fechado, o operador proj_H está bem definido. Seja $x \in \mathcal{H}$ e definamos

$$\bar{x} = x - \frac{\langle a, x \rangle - b}{\|a\|^2} a.$$

Assim,

$$\langle a, \bar{x} \rangle = \langle a, x \rangle - \frac{(\langle a, x \rangle - b)}{\|a\|^2} \langle a, a \rangle = b$$

e

$$\langle x - \bar{x}, y - \bar{x} \rangle = \frac{(\langle a, x \rangle - b)}{\|a\|^2} (\langle a, y \rangle - b) = 0 \quad \forall y \in H,$$

obtemos, da Proposição 1.5.1(1), que $\text{proj}_H(x) = \bar{x}$ para todo $x \in \mathcal{H}$.

2. [3, pág. 40] Seja um escalar $\alpha > 0$, definamos

$$K_\alpha = \{(x, r) \in \mathcal{H} \times \mathbb{R}_+ : \|x\| \leq \alpha r\}.$$

O operador proj_{K_α} é definido para todo $(x, r) \in \mathcal{H} \times \mathbb{R}$ por

$$\text{proj}_{K_\alpha}(x, r) = \begin{cases} (x, r) & \text{se } \|x\| \leq \alpha r, \\ (0, 0) & \text{se } \alpha \|x\| \leq -r, \\ \frac{\alpha \|x\| + r}{\alpha^2 + 1} \left(\alpha \frac{x}{\|x\|}, 1 \right) & \text{se } \|x\| > \alpha r \text{ ou } \alpha \|x\| > -r. \end{cases}$$

A primeira correspondência para a projeção é trivial. A segunda correspondência provém do fato que $(0, 0) \in K_\alpha$, e para qualquer $(y, s) \in K_\alpha$ temos

$$\langle (x, r), (y, s) \rangle = \langle x, y \rangle + rs \leq \|x\| \|y\| + rs \leq -\frac{r}{\alpha} \alpha s + rs = 0.$$

Portanto, da Proposição 1.5.1(1), obtemos a correspondência da projeção neste caso. A terceira correspondência verifica que

$$\left\| \frac{\alpha \|x\| + r}{\alpha^2 + 1} \alpha \frac{x}{\|x\|} \right\| = \alpha \left(\frac{\alpha \|x\| + r}{\alpha^2 + 1} \right),$$

assim, $\frac{\alpha \|x\| + r}{\alpha^2 + 1} \left(\alpha \frac{x}{\|x\|}, 1 \right) \in K_\alpha$. Se $\|x\| > \alpha r$ ou $\alpha \|x\| > -r$ então

$$\beta := \frac{\alpha \|x\| + r}{\alpha^2 + 1} > 0 \text{ e } \|x\| - \alpha\beta = \frac{\|x\| - \alpha r}{\alpha^2 + 1} > 0.$$

Logo

$$\begin{aligned} & \langle (x, r) - \beta \left(\alpha \frac{x}{\|x\|}, 1 \right), (y, s) - \beta \left(\alpha \frac{x}{\|x\|}, 1 \right) \rangle \\ &= \langle \left((\|x\| - \alpha\beta) \frac{x}{\|x\|}, r - \beta \right), \left(y - \alpha\beta \frac{x}{\|x\|}, s - \beta \right) \rangle \\ &= \langle (\|x\| - \alpha\beta) \frac{x}{\|x\|}, y - \alpha\beta \frac{x}{\|x\|} \rangle + (s - \beta)(r - \beta) \\ &= (\|x\| - \alpha\beta) \left\langle \frac{x}{\|x\|}, y \right\rangle + \alpha\beta(\alpha\beta - \|x\|) + (s - \beta)(r - \beta) \\ &\leq (\|x\| - \alpha\beta) \|y\| + \alpha\beta(\alpha\beta - \|x\|) + (s - \beta)(r - \beta) \\ &\leq (\|x\| - \alpha\beta) \alpha s + \alpha\beta(\alpha\beta - \|x\|) + (s - \beta)(r - \beta) \\ &= \alpha(\|x\| - \alpha\beta)(s - \beta) + (s - \beta)(r - \beta) \\ &= (s - \beta)[\alpha(\|x\| - \alpha\beta) + r - \beta] \\ &= (s - \beta)0 \\ &= 0. \end{aligned}$$

■

1.5.1 Projeção sobre um subespaço

Definição 1.5.2 [26] *Um espaço vetorial X é soma direta de dois subespaços M e N , $X = M \oplus N$, se qualquer $x \in X$ tem uma única representação da forma $x = m + n$, onde $m \in M$ e $n \in N$.*

Particularizando o Teorema 1.5.1 e a Proposição 1.5.1(1) quando C é um subespaço, recuperamos o Teorema clássico da projeção sobre um subespaço.

Teorema 1.5.2 ([26], Teorema 3.3.2) *Seja A um subespaço fechado de \mathcal{H} . Dado $x \in \mathcal{H}$, existe um único $\bar{x} \in A$ tal que $\|x - \bar{x}\| \leq \|x - a\|$ para todo $a \in A$.*

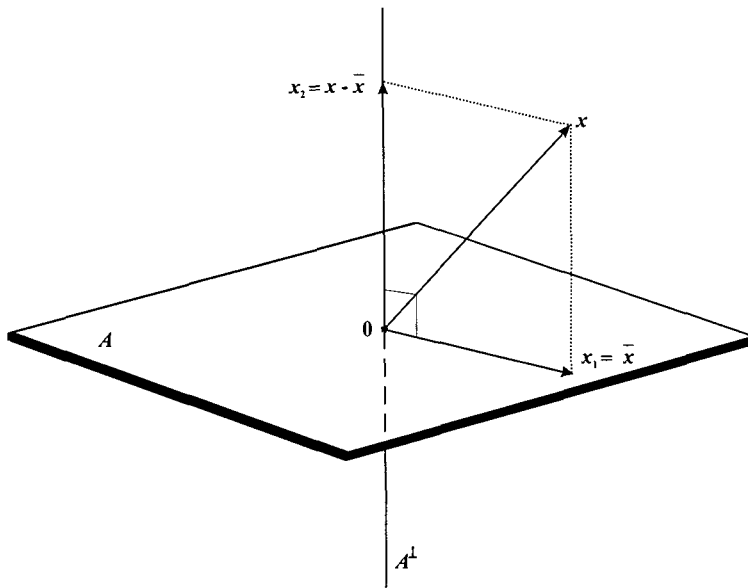


Figura 1.8: Projeção sobre um subespaço

Teorema 1.5.3 ([26], Teorema 3.4.1 e pág. 167) *Seja A um subespaço fechado de \mathcal{H} . Então*

1. $\mathcal{H} = A \oplus A^\perp$.
2. $(A^\perp)^\perp = A$.
3. O operador proj_A é linear.

1.5.2 Decomposição de Moreau

Entre o caso geral de considerar projeções sobre um convexo fechado C arbitrário, e o caso particular onde C é um subespaço, existe um caso intermediário importante para nosso estudo, que é quando consideramos C um cone, convexo e fechado.

Definição 1.5.3 Seja K um subconjunto não vazio de \mathcal{H} .

1. [26] O conjunto K é um *cone (com vértice na origem)* se $x \in K$ implica que $\alpha x \in K$ para todo $\alpha \geq 0$.

2. [29] O conjunto

$$K^\circ = \{x \in \mathcal{H} : \langle x, y \rangle \leq 0 \quad \forall y \in K\}$$

é chamado de *cone polar* de K .

Exemplo 1.5.2

(i) Sejam $\mathcal{H} = \mathbb{R}^n$ e K o cone definido por:

$$K = \mathbb{R}_+^n := \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0 \quad \forall i = 1, \dots, n\},$$

então

$$K^\circ = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \leq 0 \quad \forall i = 1, \dots, n\} = -K.$$

(ii) [3, pág. 40] Dado $\alpha > 0$. Seja o cone do Exemplo 1.5.1(ii), isto é

$$K_\alpha = \{(x, r) \in \mathcal{H} \times \mathbb{R}_+ : \|x\| \leq \alpha r\},$$

então

$$(K_\alpha)^\circ = -K_{1/\alpha}.$$

De fato, consideramos $(y, s) \in (K_\alpha)^\circ$, então $(\alpha y, \|y\|) \in K_\alpha$. Portanto

$$\langle (y, s), (\alpha y, \|y\|) \rangle = \alpha \|y\|^2 + s \|y\| \leq 0.$$

Se $y \neq 0$, significa que $\|y\| \leq -s/\alpha$. Logo, $(-y, -s) \in K_{1/\alpha}$. Se $y = 0$ e do fato que $(0, 1) \in K_\alpha$, concluímos que $\langle (0, s), (0, 1) \rangle = s \leq 0$. Assim, $\|y\| = 0 \leq -s/\alpha$ e portanto $(-0, -s) \in K_{1/\alpha}$. Por outro lado, consideramos $(y, s) \in -K_{1/\alpha}$, isto é, $\|y\| \leq -s/\alpha$ e um arbitrário $(x, r) \in K_\alpha$, então

$$\langle (x, r), (y, s) \rangle = \langle x, y \rangle + rs \leq \|x\| \|y\| + rs \leq (\alpha r)(-s/\alpha) + rs = 0.$$

Portanto, obtemos o que afirmamos. ■

A seguir destacamos alguns resultados sobre K° .

Proposição 1.5.2 ([16], Lema 2) *Seja K um subconjunto de \mathcal{H} . Então*

1. K° é um cone convexo e fechado de \mathcal{H} .
2. K é um cone convexo e fechado de \mathcal{H} se, e somente se, $K = (K^\circ)^\circ$.

Em geral, o operador de projeção não é linear a menos que K seja um subespaço. A linearidade pode ser substituída pelas seguintes propriedades.

Proposição 1.5.3 ([16], Lema 3) *Sejam K um cone convexo, fechado e não vazio de \mathcal{H} e $x \in \mathcal{H}$. Então*

1. O operador proj_K é **homogêneo não-negativo**, isto é

$$\text{proj}_K(\alpha x) = \alpha \text{proj}_K(x) \quad \forall \alpha \geq 0.$$

2. $\text{proj}_K(-x) = -\text{proj}_{-K}(x) \quad \forall x \in \mathcal{H}$.

3. O operador proj_K é **idempotente**, isto é, $\text{proj}_K \circ \text{proj}_K = \text{proj}_K$.

Assim, resultados muito usados de decomposição e projeção em um subespaço são preservados ao substituir o subespaço por um cone, ver [16] e [45]. Uma propriedade importante foi provada por Moreau [29] quando K é um cone convexo e fechado.

Proposição 1.5.4 ([29], Proposição 1) *Sejam K um cone convexo fechado de \mathcal{H} e $x, x_1, x_2 \in \mathcal{H}$. As seguintes propriedades são equivalentes:*

- (i) $x = x_1 + x_2$ onde $x_1 \in K, x_2 \in K^\circ$ e $\langle x_1, x_2 \rangle = 0$.
- (ii) $x_1 = \text{proj}_K(x)$ e $x_2 = \text{proj}_{K^\circ}(x)$.

Como consequência desta propriedade, temos que todo ponto x do espaço admite uma única decomposição $x = x_1 + x_2$, onde $x_1 \in K, x_2 \in K^\circ$, com $\langle x_1, x_2 \rangle = 0$. Se, em particular, K é um subespaço de \mathcal{H} , obtemos $\mathcal{H} = K \oplus K^\perp$. O próximo exemplo ilustra a decomposição de Moreau.

Exemplo 1.5.3

- (i) Sejam $\mathcal{H} = \mathbb{R}^n$ e $K = \mathbb{R}_+^n$, observamos que qualquer $x \in \mathbb{R}^n$ pode ser escrito como

$$x = \text{proj}_K(x) + \text{proj}_{K^\circ}(x) = (x_1^+, \dots, x_n^+) + (x_1^-, \dots, x_n^-),$$

onde

$$x_i^+ = \max\{x_i, 0\} \quad \text{e} \quad x_i^- = \min\{x_i, 0\} \quad \forall i = 1, \dots, n,$$

verificando a Proposição 1.5.4.

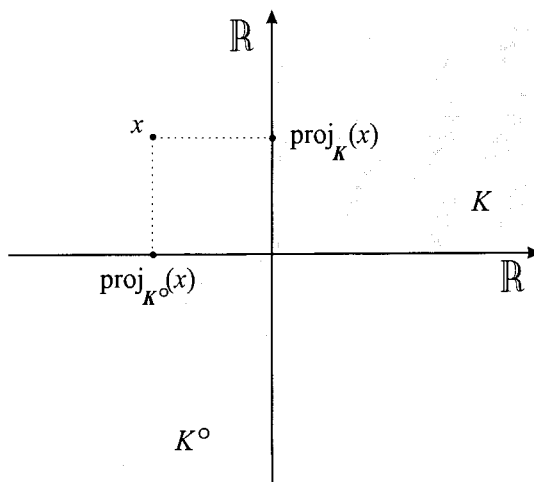


Figura 1.9: O ítem (i) do Exemplo 1.5.3.

- (ii) [3, pág. 63] Sejam $\mathcal{H} = \mathbb{R}^2$ e $\alpha = 1$ no ítem (ii) do Exemplo 1.5.1, isto é

$$K = K_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq x_3^2, x_3 \geq 0\},$$

então, pela caracterização da projeção, temos que para todo $x = (x_1, x_2, x_3) \in \mathbb{R}^3$

$$\text{proj}_K(x) = \begin{cases} (x_1, x_2, x_3) & \text{se } x_1^2 + x_2^2 \leq x_3^2 \text{ e } x_3 \geq 0, \\ (0, 0, 0) & \text{se } x_1^2 + x_2^2 \geq x_3^2 \text{ e } x_3 \leq 0, \\ \frac{\sqrt{x_1^2 + x_2^2} + x_3}{2\sqrt{x_1^2 + x_2^2}} (x_1, x_2, \sqrt{x_1^2 + x_2^2}) & \text{c.c.} \end{cases}$$

Pelo ítem (ii) do Exemplo 1.5.2 temos que $K^\circ = -K_1$. Assim, da Proposição 1.5.3(2), obtemos $\text{proj}_{K^\circ}(x) = \text{proj}_{-K}(x) = -\text{proj}_K(-x)$. Portanto a regra de correspondência de proj_K é útil para a projeção do polar. ■

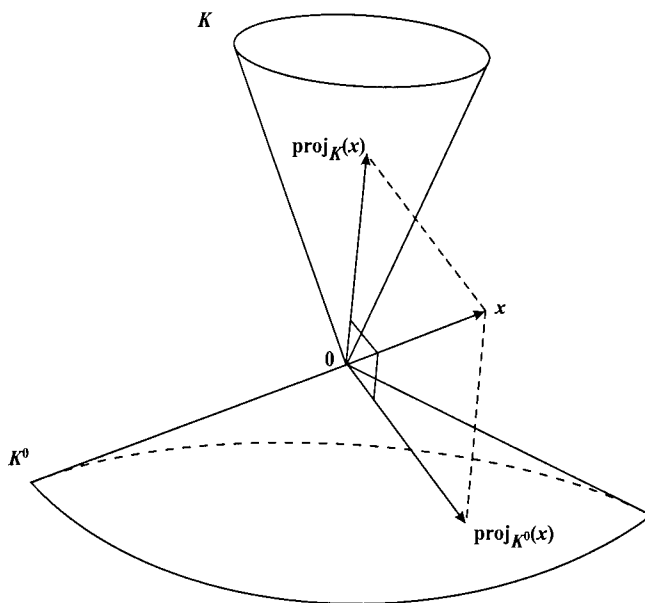


Figura 1.10: Decomposição de Moreau para o cone K do Exemplo 1.5.3(ii)

Capítulo 2

O Problema

Este capítulo serve de ponto central para a definição, referências e casos particulares do problema estudado nesta dissertação. Na seção 2.1, apresentamos o problema primal-dual $P(T, A)$. Na seção 2.2 mostramos que o problema $P(T, A)$ generaliza outros problemas da literatura, tais como o da soma de operadores, o da viabilidade convexa, o problema da desigualdade variacional e o multivaluado monótono da complementaridade. Os problemas citados neste capítulo podem ser encontrados em [42], [43], [5] e [17].

2.1 O problema primal-dual

Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal e A um subespaço de \mathcal{H} . Nosso interesse é estudar o seguinte problema *primal-dual*

$$P(T, A) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp \\ \text{tais que} \\ \bar{y} \in T(\bar{x}). \end{array} \right. \quad (2.1)$$

O nome primal-dual provém do fato de que tal problema unifica outros dos problemas de encontrar zeros de certos operadores, considerados cada um como a metade primal e dual de (2.1) respectivamente, isto será detalhado no próximo capítulo.

2.2 Aplicações

Muitos dos problemas encontrados na literatura podem ser reescritos sob a formulação (2.1) como mostramos a seguir.

2.2.1 Problema de otimização convexa

Seja $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria. O *problema de otimização convexa* é

$$\min_{x \in \mathcal{H}} f(x). \quad (2.2)$$

Segue-se que

$$f \text{ tem um mínimo (global) em } \bar{x} \text{ se, e somente se, } 0 \in \partial f(\bar{x}). \quad (2.3)$$

Muitos problemas de otimização com, ou sem restrições, são formulados como (2.2), por exemplo, quando $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função convexa e C é um conjunto convexo e não vazio de \mathcal{H} tal que $\text{dom}(f) \cap C \neq \emptyset$, o problema com restrições:

$$\begin{cases} \min f(x) \\ \text{s.a:} \\ x \in C \end{cases} \quad (2.4)$$

é um caso particular de (2.2), posto que definindo $\tilde{f} := f + \delta_C$. Claramente, \tilde{f} é uma função convexa e própria e resolver (2.4) é equivalente a resolver (2.2) para \tilde{f} . Deste modo, (2.3) implica que (2.4) tem um mínimo em $\bar{x} \in C$ se, e somente se, $0 \in \partial \tilde{f}(\bar{x})$. Pelo Teorema 1.1.3(2), a última condição pode ser escrita em termos de subdiferenciais como:

$$0 \in \partial f(\bar{x}) + N_C(\bar{x}), \quad (2.5)$$

em cada um dos seguintes casos:

1. existe um ponto em $C \cap \text{dom}(f)$ onde f é contínua.
2. $\text{int}(C) \cap \text{dom}(f) \neq \emptyset$.
3. \mathcal{H} é de dimensão finita e $\text{ri}(C) \cap \text{ri}(\text{dom}(f)) \neq \emptyset$.

2.2.1.1 Minimização convexa sobre um subespaço

Em particular, quando o conjunto C é um subespaço A de \mathcal{H} , o problema (2.4) é reescrito como

$$\begin{cases} \min f(x) \\ \text{s.a:} \\ x \in A. \end{cases} \quad (2.6)$$

A equivalência do problema (2.1) com (2.6) é obtida sobre a condição de ser f semicontínua inferior, como mostra o seguinte resultado.

Teorema 2.2.1 ([43], pág. 207) *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semicontínua inferior e A um subespaço de \mathcal{H} . Se uma das condições abaixo é verificada*

1. $A \cap \text{int}(\text{dom}(f)) \neq \emptyset$.
2. \mathcal{H} de dimensão finita e $A \cap \text{ri}(\text{dom}(f)) \neq \emptyset$.

Então $\bar{x} \in A$ é solução do problema (2.6) se, e somente se, existe $\bar{y} \in \mathcal{H}$ tal que (\bar{x}, \bar{y}) é solução do problema

$$\text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp : 0 \in \bar{y} + \partial f(\bar{x}). \quad (2.7)$$

Prova: Se a condição 1 é verificada, temos que existe $\bar{u} \in A \cap \text{int}(\text{dom}(f)) \subseteq \text{dom}(\delta_A) \cap \text{dom}(f)$ tal que f é contínua em \bar{u} . Por outro lado se a condição 2 é verdadeira, então temos que $\text{ri}(A) = A$, e portanto $\text{ri}(A) \cap \text{ri}(\text{dom}(f)) \neq \emptyset$. Assim, aplicamos o Teorema 1.1.3 para $f_1 = f$ e $f_2 = \delta_A$ e temos que:

$$\partial(f + \delta_A)(x) = \partial f(x) + \partial \delta_A(x) = \partial f(x) + N_A(x) = \partial f(x) + A^\perp \quad \forall x \in A.$$

Logo

$$\begin{aligned} \bar{x} \text{ é solução do problema (2.6)} &\iff 0 \in \partial(f + \delta_A)(\bar{x}) = \partial f(\bar{x}) + A^\perp \\ &\iff \exists \bar{y} \in A^\perp : 0 \in \partial f(\bar{x}) + \bar{y}. \end{aligned}$$

Por outro lado, da semicontinuidade inferior de f e pelo Teorema 1.3.5, temos que ∂f é um operador monótono maximal. Assim, concluímos que resolver o problema (2.6) é equivalente a resolver (2.7). ■

2.2.2 Soma de operadores monótonos

Sejam $T_1, \dots, T_n : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais e η_1, \dots, η_n números reais positivos não nulos. O *problema de soma de operadores monótonos* [42] é definido por:

$$\left\{ \begin{array}{l} \text{Encontrar } \bar{x}, \bar{y}_1, \dots, \bar{y}_n \in \mathcal{H} \\ \text{tais que} \\ \bar{y}_i \in T_i(\bar{x}) \quad \forall i = 1, \dots, n, \\ \sum_{i=1}^n \eta_i \bar{y}_i = 0. \end{array} \right. \quad (2.8)$$

Quando $n = 1$, o problema (2.8) reduz-se ao problema:

$$\text{Encontrar } \bar{x} \in \mathcal{H} \text{ tal que } 0 \in T(\bar{x}). \quad (2.9)$$

A seguir, apresentamos a equivalência entre os problemas (2.8) e $P(T, A)$, mostrada por Spingarn [42]. Denotamos por \mathcal{H}_i , o espaço de Hilbert H munido do produto interno $\langle x, y \rangle_i = \eta_i \langle x, y \rangle$, onde $\langle x, y \rangle$ é o produto interno em \mathcal{H} . Doravante, consideramos o produto interno em $\mathcal{H} \times \dots \times \mathcal{H}$ dado por:

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle_{\mathcal{H} \times \dots \times \mathcal{H}} := \sum_{i=1}^n \langle x_i, y_i \rangle_i.$$

Portanto $\mathcal{H} \times \dots \times \mathcal{H}$ é um espaço de Hilbert real. Definimos

$$\begin{aligned} T : \mathcal{H} \times \dots \times \mathcal{H} &\longrightarrow \mathcal{P}(\mathcal{H} \times \dots \times \mathcal{H}) \\ (x_1, \dots, x_n) &\longmapsto T(x_1, \dots, x_n) := T_1(x_1) \times \dots \times T_n(x_n). \end{aligned}$$

Devido à monotonicidade maximal dos operadores T_i , obtemos que o operador T é monótono maximal em $\mathcal{H} \times \dots \times \mathcal{H}$. Além disso, consideramos o subespaço

$$A = \{(x_1, \dots, x_n) \in \mathcal{H} \times \dots \times \mathcal{H} : x_1 = \dots = x_n\}.$$

Observamos que

$$A^\perp = \{(y_1, \dots, y_n) \in \mathcal{H} \times \dots \times \mathcal{H} : \sum_{i=1}^n \eta_i y_i = 0\}.$$

Portanto, o ponto $(\bar{x}, \bar{y}_1, \dots, \bar{y}_n)$ soluciona (2.8) se, e somente se, $(\bar{x}, \dots, \bar{x}) \in A$ e $(\bar{y}_1, \dots, \bar{y}_n) \in A^\perp$ solucionam $P(T, A)$, assim, temos a equivalência dos problemas $P(T, A)$ e (2.8).

A seguir, enunciamos problemas que são equivalentes ao problema de soma de operadores monótonos.

2.2.3 Viabilidade convexa

Sejam C_i subconjuntos convexos, fechados e não vazios de \mathcal{H} , $i = 1, \dots, n$. O *problema de viabilidade convexa* [42], é dado por:

$$\left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in \mathcal{H} \\ \text{tal que} \\ \bar{x} \in \bigcap_{i=1}^n C_i. \end{array} \right. \quad (2.10)$$

Para estabelecer a equivalência dos problemas (2.10) e (2.8), Spingarn [42] considerou os cones normais dos conjuntos C_i , isto é $T_i = N_{C_i}$ para todo $i = 1, \dots, n$. Pela Proposição 1.3.3, os operadores T_i são monótonos maximais para todo $i = 1, \dots, n$. Além disso,

$$x \in C_i \iff T_i(x) \neq \emptyset \iff 0 \in T_i(x). \quad (2.11)$$

Então

$$\bar{x} \text{ resolve (2.10)} \iff 0 \in T_1(\bar{x}) + \dots + T_n(\bar{x}).$$

De fato, se \bar{x} é solução de (2.10), então $\bar{x} \in \bigcap_{i=1}^n C_i$. Assim, $\bar{x} \in C_i$ para todo $i = 1, \dots, n$. Logo, de (2.11) obtemos que $0 \in T_i(\bar{x})$ para todo $i = 1, \dots, n$ com $\bar{x} \in \bigcap_{i=1}^n C_i = \bigcap_{i=1}^n \text{Dom}(N_{C_i}) = \text{Dom}(T_1 + \dots + T_n)$. O que resulta

$$0 \in T_1(\bar{x}) + \dots + T_n(\bar{x}).$$

Por outro lado, seja \bar{x} tal que $0 \in T_1(\bar{x}) + \dots + T_n(\bar{x})$, daqui $\bar{x} \in \text{Dom}(T_i)$ para todo $i = 1, \dots, n$, então $T_i(\bar{x}) \neq \emptyset$ de (2.11) temos que $\bar{x} \in C_i$ para todo $i = 1, \dots, n$, assim \bar{x} é solução de (2.10). Portanto, temos a equivalência dos problemas (2.10) e (2.8).

2.2.4 Desigualdades variacionais

Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal e C um subconjunto convexo, fechado e não vazio de \mathcal{H} . O *problema de desigualdade variacional* [5] é definido por:

$$\text{VIP}(T, C) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in C \text{ e } \bar{y} \in T(\bar{x}) \\ \text{tais que} \\ \langle \bar{y}, x - \bar{x} \rangle \geq 0 \quad \forall x \in C. \end{array} \right. \quad (2.12)$$

Observamos que \bar{x} resolve $\text{VIP}(T, C)$ se, e somente se,

$$\bar{y} \in T(\bar{x}) \quad \text{e} \quad -\bar{y} \in N_C(\bar{x}).$$

O que é equivalente a encontrar \bar{x} tal que

$$0 \in \tilde{T}(\bar{x}),$$

onde \tilde{T} está definido por:

$$\tilde{T}(x) := \left\{ \begin{array}{ll} T(x) + N_C(x) & \text{se } x \in C, \\ \emptyset & \text{se } x \notin C. \end{array} \right.$$

Portanto, o problema $\text{VIP}(T, C)$ é equivalente ao problema de soma (2.8) para T e N_C com $\eta_1 = \eta_2 = 1$. Além disso, se alguma das hipóteses do Teorema 1.3.1 é verificada para $T_1 = T$ e $T_2 = N_C$, temos que \tilde{T} é monótono maximal. Portanto, $\text{VIP}(T, V)$ reduz-se ao problema de encontrar um zero de \tilde{T} como em (2.9).

Observamos que o problema de otimização (2.5), no caso quando a função f é semicontínua inferior e C é fechado, é equivalente ao $\text{VIP}(\partial f, C)$, sobre uma das seguintes condições:

1. existe um ponto em $C \cap \text{dom}(f)$ onde f é contínua.
2. $\text{int}(C) \cap \text{dom}(f) \neq \emptyset$.
3. \mathcal{H} é de dimensão finita e $\text{ri}(C) \cap \text{ri}(\text{dom}(f)) \neq \emptyset$.

Quando o conjunto C é um cone, $\text{VIP}(T, C)$ reduz-se ao **problema monótono da complementaridade multivaluado** definido em [17] por:

$$\text{MCP}(T, C) \begin{cases} \text{Encontrar } \bar{x} \in C \text{ e } \bar{y} \in T(\bar{x}) \\ \text{tais que} \\ \bar{y} \in -C^\circ \text{ e } \langle \bar{x}, \bar{y} \rangle = 0. \end{cases}$$

De fato, se \bar{x} é solução de $\text{VIP}(T, C)$ com $\bar{y} \in T(\bar{x})$ então

$$\langle \bar{y}, x - \bar{x} \rangle \geq 0 \quad \forall x \in C.$$

Em particular, se consideramos $x = 0$ obtemos $\langle \bar{y}, \bar{x} \rangle \leq 0$, e se $x = 2\bar{x}$, resulta $\langle \bar{y}, \bar{x} \rangle \geq 0$. Assim, $\langle \bar{y}, \bar{x} \rangle = 0$. Logo, para todo $x \in C$ obtemos

$$\langle \bar{y}, x \rangle \geq \langle \bar{y}, \bar{x} \rangle = 0.$$

Portanto $-\bar{y} \in C^\circ$, ou seja, \bar{x} e \bar{y} resolvem $\text{MCP}(T, C)$. Reciprocamente, se $\bar{x} \in C$ e $\bar{y} \in T(\bar{x})$ são tais que $-\bar{y} \in C^\circ$ e $\langle \bar{y}, \bar{x} \rangle = 0$, então temos

$$0 = \langle \bar{y}, \bar{x} \rangle \leq \langle \bar{y}, x \rangle \quad \forall x \in C.$$

Isto é, $\langle \bar{y}, x - \bar{x} \rangle \geq 0$ para todo $x \in C$; quer dizer, \bar{x} resolve $\text{VIP}(T, C)$ com $\bar{y} \in T(\bar{x})$.

Capítulo 3

Operador Inversa Parcial

Neste capítulo estudamos as propriedades do operador inversa parcial, apresentado por Spingarn [42] e [43]. A principal motivação para o estudo deste operador, é a equivalência do problema primal-dual

$$P(T, A) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp \\ \text{tais que} \\ \bar{y} \in T(\bar{x}), \end{array} \right.$$

com o problema de encontrar um zero do operador inversa parcial. Na seção 3.1, definimos e estudamos as propriedades do operador inversa parcial, na seção seguinte formulamos um esquema de dualidade justificando o nome de problema primal-dual de (2.1) e nas seções 3.3 e 3.4 definimos e relacionamos o operador Douglas-Rachford com o operador inversa parcial. De agora em diante, usamos a notação, dado $z \in \mathcal{H}$, $z_B := \text{proj}_B(z)$ denota a projeção de z sobre B .

3.1 Operador inversa parcial

Seja A um subespaço de \mathcal{H} . Pelo Teorema 1.5.3 sabemos que todo ponto $z \in \mathcal{H}$ admite uma única decomposição do tipo $z = z_A + z_{A^\perp}$, onde $z_A \in A$ e $z_{A^\perp} \in A^\perp$. Com isso Spingarn [42] introduziu o operador inversa parcial:

Definição 3.1.1 [42] *Sejam um operador $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ e um subespaço A de \mathcal{H} . O operador inversa parcial de T com respeito a A , denotado por T_A , é definido pelo gráfico:*

$$\text{Gr}(T_A) = \{(x_A + y_B, y_A + x_B) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\},$$

onde $B \subseteq \mathcal{H}$ é o subespaço ortogonal a A , $B := A^\perp$.

Proposição 3.1.1 *Sejam um operador $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e um subespaço A de \mathcal{H} . Então $y \in T_A(x)$ se, e somente se, $y_A + x_B \in T(x_A + y_B)$, onde $B \subseteq \mathcal{H}$ é o subespaço ortogonal a A , $B := A^\perp$.*

Prova: O ponto $y \in T_A(x)$ se, e somente se, $(x, y) \in \text{Gr}(T_A)$ se, e somente se, existem u, v tais que $v \in T(u)$ e $x = u_A + v_B = x_A + x_B$, $y = v_A + u_B = y_A + y_B$. Como a decomposição é única, resulta

$$u = u_A + u_B = x_A + y_B,$$

$$v = v_A + v_B = y_A + x_B.$$

Logo, $y \in T_A(x)$ se, e somente se, $(y_A + x_B) \in T(x_A + y_B)$. ■

Observação 3.1.1 ([42], pág. 251) *Sejam um operador $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e um subespaço A de \mathcal{H} . No caso dos subespaços triviais de \mathcal{H} , isto é, $A = \{0\}$ e $A = \mathcal{H}$, obtemos*

$$T_{\{0\}} = T^{-1} \text{ e } T_{\mathcal{H}} = T.$$

De fato, se $A = \{0\}$, temos que para todo $x, y \in \mathcal{H}$, resulta $x_A = 0$ e $x_B = x$, $y_A = 0$, $y_B = y$ logo

$$\begin{aligned} \text{Gr}(T_A) &= \{(x_A + y_B, y_A + x_B) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \{(0 + y, 0 + x) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \{(y, x) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \text{Gr}(T^{-1}). \end{aligned}$$

Se $A = \mathcal{H}$, então para todo $x, y \in \mathcal{H}$ resulta $x_A = x$ e $x_B = 0$, $y_A = y$, $y_B = 0$, resultando $\text{Gr}(T_A) = \text{Gr}(T)$. Portanto, quando A é um subespaço não trivial, T_A “estará entre T e T^{-1} ”, isto motiva o nome de inversa parcial. ■

Lema 3.1.1 *Sejam um operador $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e um subespaço A de \mathcal{H} . Então verificam-se as seguintes igualdades*

1. $T_A = (T^{-1})_{A^\perp}$.

2. $(T^{-1})_{A^\perp} = (T_{A^\perp})^{-1}$.

Prova:

1. Vamos mostrar que $\text{Gr}(T_A) = \text{Gr}((T^{-1})_{A^\perp})$. De fato, considerando $B = A^\perp$

$$\begin{aligned}\text{Gr}(T_A) &= \{(x_A + y_{A^\perp}, y_A + x_{A^\perp}) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \{(y_B + x_{B^\perp}, x_B + y_{B^\perp}) \in \mathcal{H} \times \mathcal{H} : x \in T^{-1}(y)\} \\ &= \text{Gr}((T^{-1})_B) \\ &= \text{Gr}((T^{-1})_{A^\perp}).\end{aligned}$$

Portanto temos (1).

2. Vamos provar que $\text{Gr}((T^{-1})_{A^\perp}) = \text{Gr}((T_{A^\perp})^{-1})$. Diretamente,

$$\text{Gr}(T_{A^\perp}) = \{(x_{A^\perp} + y_A, y_{A^\perp} + x_A) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\},$$

daqui

$$\begin{aligned}\text{Gr}((T_{A^\perp})^{-1}) &= \{(y_{A^\perp} + x_A, x_{A^\perp} + y_A) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \text{Gr}((T^{-1})_{A^\perp}).\end{aligned}$$

Assim temos (2). ■

Uma pergunta que surgiu no decorrer do trabalho é se o operador inversa parcial de uma função preserva o fato de ser ponto-ponto, devido a que é menos complicado o trabalho numerico para estos operadores que para os ponto-conjunto, mas pelo seguinte exemplo provamos que a resposta é negativa.

Exemplo 3.1.1 Sejam A um subespaço de \mathcal{H} não trivial e $T : \mathcal{H} \longrightarrow A$ o operador definido por

$$T(x) = \text{proj}_A(x) = x_A.$$

Pelo Teorema 1.5.3, temos que proj_A é linear e, da Proposição 1.5.1(2), resulta que T é monótono. Da decomposição dos pontos x e $y := T(x)$ com respeito a A ,

$$\begin{aligned}x &= x_A + x_{A^\perp}, \\ y &= y_A + y_{A^\perp} = x_A + 0,\end{aligned}$$

obtemos que

$$(x_A + y_{A^\perp}, x_{A^\perp} + y_A) = (x_A + 0, x_{A^\perp} + x_A) = (x_A, x) \in \text{Gr}(T_A) \iff x \in T(x_A).$$

Seja $u \in A$ então $v \in T_A(u)$ se, e somente se, $v = u + v_B$ se, e somente se, $v_A = u$. Portanto, T_A não é um operador ponto-ponto, por exemplo seja $\mathcal{H} = \mathbb{R}^2$, $A = \mathbb{R} \times \{0\}$, $u = (5, 0) \in A$ então $T_A(u) = \{(5, t) : t \in \mathbb{R}\}$. ■

Uma característica não provada na literatura é que a inversa parcial seja gerada por um só operador, a seguir damos uma prova de tal fato.

Proposição 3.1.2 *Sejam $T, S : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ operadores e A um subespaço de \mathcal{H} . Então $T = S$ se, e somente se, $T_A = S_A$.*

Prova: Imediatamente temos que $T = S$ implica $T_A = S_A$. Provamos agora que $T = S$ se $T_A = S_A$. Seja $(x, y) \in \text{Gr}(T)$ ou equivalentemente $y \in T(x)$, então $(u, v) := (x_A + y_{A^\perp}, y_A + x_{A^\perp}) \in \text{Gr}(T_A) = \text{Gr}(S_A)$, assim existem $w, z \in \mathcal{H}$ tais que $u = w_A + z_{A^\perp}$, $v = z_A + w_{A^\perp}$ e $z \in S(w)$, portanto:

$$x_A + y_{A^\perp} = w_A + z_{A^\perp}, \quad y_A + x_{A^\perp} = z_A + w_{A^\perp}, \quad z \in S(w),$$

daqui

$$\begin{aligned} x_A - w_A &= z_{A^\perp} - y_{A^\perp} \in A \cap A^\perp = \{0\}, \\ y_A - z_A &= w_{A^\perp} - x_{A^\perp} \in A \cap A^\perp = \{0\}. \end{aligned}$$

A partir disto, obtemos $w = x$ e $z = y$, assim $y \in S(x)$, o que equivale a $(x, y) \in \text{Gr}(S)$ e vice-versa. Analogamente obtemos o outro contido. ■

Damos a seguir duas características importantes do operador T_A .

Proposição 3.1.3 ([42], pág. 251) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador e A um subespaço de \mathcal{H} . Então T_A é monótono se, e somente se, T é monótono.*

Prova: Consideramos $(x'_A + y'_{A^\perp}, y'_A + x'_{A^\perp}), (x_A + y_{A^\perp}, y_A + x_{A^\perp}) \in \text{Gr}(T_A)$ com $y' \in T(x')$ e $y \in T(x)$ então:

$$\begin{aligned} &\langle (y'_A + x'_{A^\perp}) - (y_A + x_{A^\perp}), (x'_A + y'_{A^\perp}) - (x_A + y_{A^\perp}) \rangle \\ &= \langle (y' - y)_A, (x' - x)_A \rangle + \langle (x' - x)_{A^\perp}, (y' - y)_{A^\perp} \rangle \\ &= \langle (y' - y)_A, (x' - x)_A \rangle + \langle (y' - y)_{A^\perp}, (x' - x)_{A^\perp} \rangle \\ &= \langle (y' - y)_A + (y' - y)_{A^\perp}, (x' - x)_A + (x' - x)_{A^\perp} \rangle \\ &= \langle y' - y, x' - x \rangle. \end{aligned}$$

Portanto T_A é monótono se, e somente se, T é monótono. ■

Proposição 3.1.4 ([42], Proposição 2.1) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador e A um subespaço de \mathcal{H} . Então T_A é monótono maximal se, e somente se, T é monótono maximal.*

Prova: Suponhamos que T_A é monótono maximal. Pelo Teorema 1.3.3 mostrar que T é monótono maximal é equivalente a provar que $\text{Im}(I + T) = \mathcal{H}$. Como $\text{Im}(I + T) \subseteq \mathcal{H}$, é suficiente provar que $\mathcal{H} \subset \text{Im}(I + T)$. Seja $y \in \mathcal{H}$, como T_A é monótono maximal; pelo Teorema 1.3.3 resulta $y \in \text{Im}(I + T_A)$. Portanto existe o ponto $x \in \mathcal{H}$ tal que $y \in \{x\} + T_A(x)$ o que equivale a dizer que $(y - x) \in T_A(x)$ ou $(x_A + x_{A^\perp}, (y - x)_A + (y - x)_{A^\perp}) \in \text{Gr}(T_A)$. Daí segue que $(y - x)_A + x_{A^\perp} \in T(x_A + (y - x)_{A^\perp})$. Portanto $y_A + y_{A^\perp} - x_A + x_{A^\perp} \in y_{A^\perp} + T(x_A + (y - x)_{A^\perp})$. Isto é, $y \in x_A + (y - x)_{A^\perp} + T(x_A + (y - x)_{A^\perp})$, assim $y \in (I + T)(x_A + (y - x)_{A^\perp}) \subset \text{Im}(I + T)$. A recíproca provamos de forma análoga. ■

As duas proposições anteriores garantem que o operador inversa parcial preserva as propriedades de monotonicidade e maximalidade do operador original. No entanto, notamos que o operador inversa parcial não mantêm necessariamente a monotonicidade forte do operador original, como ilustramos a seguir por meio de um exemplo.

Exemplo 3.1.2 Seja T o operador dado no Exemplo 1.3.7(ii):

$$T(x) = \begin{cases} -e^{-x} & \text{se } x < 0, \\ [-1, 1] & \text{se } x = 0, \\ e^x & \text{se } x > 0. \end{cases}$$

Consideramos o subespaço $A = \{0\}$. O operador T é fortemente monótono e $T_A = T^{-1}$, assim

$$T_A(x) = \begin{cases} \text{sinal}(x) \ln |x| & \text{se } |x| > 1, \\ 0 & \text{se } |x| \leq 1. \end{cases}$$

Mas T_A não é fortemente monótono. De fato, suponhamos que $\alpha > 0$ é o módulo de monotonicidade e consideramos $x' = -1$ e $x = \frac{1}{2}$ na Definição 1.3.8, isto é

$$0 = \left\langle 0 - 0, -1 - \frac{1}{2} \right\rangle \geq \alpha \left\| -1 - \frac{1}{2} \right\|^2 \geq 0,$$

daqui $\alpha = 0$, o qual é um absurdo. Portanto o operador inversa parcial não é fortemente monótono. ■

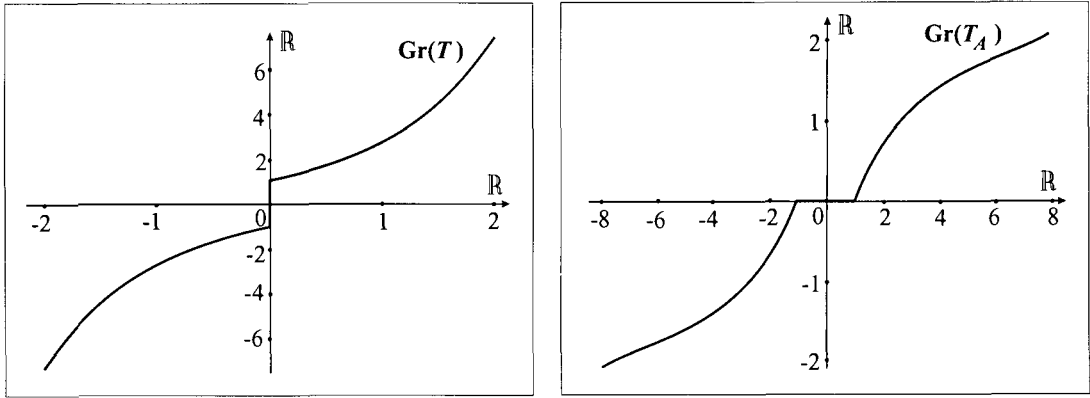


Figura 3.1: Gráficos dos operadores T e T_A do Exemplo 3.1.2.

A monotonicidade forte do operador inversa parcial está caracterizada pela seguinte propriedade.

Proposição 3.1.5 ([42], Proposição 2.2) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e A um subespaço de \mathcal{H} . Então T_A é fortemente monótono se, e somente se, existe um escalar $a > 0$ tal que*

$$\langle y' - y, x' - x \rangle \geq a(\|x'_A - x_A\|^2 + \|y'_{A^\perp} - y_{A^\perp}\|^2) \quad \forall y \in T(x), \forall y' \in T(x').$$

Prova: Consideramos $y' \in T(x')$ e $y \in T(x)$. Da Prova da Proposição 3.1.3 temos que

$$\langle y' - y, x' - x \rangle = \langle (x_A + y_{A^\perp}) - (x'_A + y'_{A^\perp}), (y_A + x_{A^\perp}) - (y'_A + x'_{A^\perp}) \rangle \quad (3.1)$$

Assim, se T_A é fortemente monótono, obtemos

$$\begin{aligned} \langle y' - y, x' - x \rangle &\geq a\|x_A + y_{A^\perp} - (x'_A + y'_{A^\perp})\|^2 \\ &= a\|(x_A - x'_A) + (y_{A^\perp} - y'_{A^\perp})\|^2 \\ &= a(\|x_A - x'_A\|^2 + \|y_{A^\perp} - y'_{A^\perp}\|^2). \end{aligned}$$

Reciprocamente

$$\begin{aligned} a(\|x'_A - x_A\|^2 + \|y'_{A^\perp} - y_{A^\perp}\|^2) &= a\|x_A + y_{A^\perp} - (x'_A + y'_{A^\perp})\|^2 \\ &\leq \langle y' - y, x' - x \rangle, \end{aligned}$$

e por (3.1), temos o resultado. ■

A seguinte característica do operador T_A mostra a utilidade do conceito de inversa parcial.

Teorema 3.1.1 ([42], Teorema 2.1) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador e A um subespaço de \mathcal{H} . Então $0 \in T_A(z)$ se, e somente se, $z_{A^\perp} \in T(z_A)$.*

Prova:

$$\begin{aligned}
0 \in T_A(z) &\iff (z, 0) \in \text{Gr}(T_A) \\
&\iff \exists y, x \in \mathcal{H} \text{ tal que } z = x_A + y_{A^\perp}, 0 = y_A + x_{A^\perp} \text{ e } y \in T(x) \\
&\iff z_A = x_A = x \in A, z_{A^\perp} = y_{A^\perp} = y \text{ e } y \in T(x) \\
&\iff z_{A^\perp} \in T(z_A). \quad \blacksquare
\end{aligned}$$

O teorema acima é a peça chave do trabalho de Spingarn [42], devido a que pela introdução do operador inversa parcial obtem uma equivalencia entre o problema primal-dual (2.1) e o problema de encontrar zeros para T_A . Também, é importante notar que o Teorema 3.1.1 é válido sem o requerimento de qualquer propriedade sobre o operador T .

O próximo resultado é de grande importância para nosso estudo.

Proposição 3.1.6 ([42], pág. 254) *Dados $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador, A um subespaço de \mathcal{H} e um escalar $c > 0$. Seja S_c o operador que tem como gráfico*

$$\text{Gr}(S_c) = \{(u, v) \in \mathcal{H} \times \mathcal{H} : \frac{1}{c}v_A + v_{A^\perp} \in T(u_A + \frac{1}{c}u_{A^\perp})\}. \quad (3.2)$$

Então

1. $S_c = (cT_A)_A$.
2. O operador T é monótono maximal se, e somente se, S_c é monótono maximal.

Prova:

1. Da definição de S_c :

$$v \in S_c(u) \iff \left(\frac{1}{c}v_A + v_{A^\perp}\right) \in T\left(u_A + \frac{1}{c}u_{A^\perp}\right),$$

pela Proposição 3.1.1 temos que

$$\left(\frac{1}{c}v_A + v_{A^\perp}\right) \in T\left(u_A + \frac{1}{c}u_{A^\perp}\right) \iff \left(\frac{1}{c}v_A + \frac{1}{c}u_{A^\perp}\right) \in T_A(u_A + v_{A^\perp}),$$

assim da definição de cT_A :

$$\left(\frac{1}{c}v_A + \frac{1}{c}u_{A^\perp}\right) \in T_A(u_A + v_{A^\perp}) \iff (v_A + u_{A^\perp}) \in cT_A(u_A + v_{A^\perp}),$$

novamente da Proposição 3.1.1:

$$(v_A + u_{A^\perp}) \in cT_A(u_A + v_{A^\perp}) \iff (v_A + v_{A^\perp}) \in (cT_A)_A(u_A + u_{A^\perp}).$$

Portanto, $S_c = (cT_A)_A$.

2. Pela Proposição 3.1.4, temos que T é monótono maximal se, e somente se, T_A é monótono maximal. Da Proposição 1.3.4, sabemos que T_A é monótono maximal se, e somente se, cT_A é monótono maximal. Aplicando novamente a Proposição 3.1.4, obtemos que cT_A é monótono maximal se, e somente se, $(cT_A)_A$ é monótono maximal. Logo, concluímos que T é monótono maximal se, e somente se, $(cT_A)_A$ é monótono maximal. ■

Observamos pela proposição acima, que no caso $c = 1$, temos $S_1 = T$ e, além disso, quando $A = \{0\} \subset \mathcal{H}$ então $S_c = \frac{1}{c}T$.

Na literatura, existem problemas que não envolvem operadores monótonos, em particular, os pseudomonótonos, os quais tem encontrado aplicações e estão sendo estudados na atualidade. Assim motivados pelas aplicações testamos se a inversa parcial preserva a pseudomonotonicidade do operador original. Não entanto notamos que isto é falso, como mostramos no exemplo.

Exemplo 3.1.3 Sejam T o operador pseudomonótono definido no Exemplo 1.4.1(i):

$$T(x) = \begin{cases} \frac{1}{1+x} & \text{se } x > -1, \\ \emptyset & \text{se } x \leq -1, \end{cases}$$

e o subespaço $A = \{0\}$. Portanto, da Observação 3.1.1(1), temos que $T_A = T^{-1}$, daqui

$$T_A(x) = \begin{cases} \frac{1-x}{x} & \text{se } x > 0, \\ \emptyset & \text{se } x \leq 0. \end{cases}$$

Mas T_A não é pseudomonótono. De fato, substituímos $x' = 2$ e $x = 1$ na Definição 1.4.1, pág. 25, e verificamos que

$$\langle T_A(x), x' - x \rangle = \left\langle \frac{1-1}{1}, 2-1 \right\rangle = 0,$$

porém,

$$\langle T_A(x'), x' - x \rangle = \left\langle \frac{1-2}{2}, 2-1 \right\rangle = -\frac{1}{2} < 0.$$

Portanto T_A não preserva a propriedade de pseudomonotonicidade de T . ■

3.2 Problemas duais

Nesta seção descrevemos um esquema de dualidade pela qual, dois problemas para encontrar zeros de certos operadores monótonos formam uma metade, respectivamente, do denominado problema primal-dual (2.1).

Definição 3.2.1 [42] *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador e A um subespaço de \mathcal{H} .*

1. *O operador seção- A de T , denotado por $T_0 : A \longrightarrow \mathcal{P}(A)$, é definido pelo gráfico:*

$$\text{Gr}(T_0) = \{(x, y_A) \in \mathcal{H} \times \mathcal{H} : y \in T(x), x \in A\}.$$

2. *O operador seção- B de T , denotado por $S_0 : B \longrightarrow \mathcal{P}(B)$, é definido pelo gráfico:*

$$\text{Gr}(S_0) = \{(y, x_B) \in \mathcal{H} \times \mathcal{H} : x \in T^{-1}(y), y \in B\}.$$

onde $B \subseteq \mathcal{H}$ é o subespaço ortogonal a A , $B := A^\perp$.

Observamos que $\text{Dom}(T_0) = A \cap \text{Dom}(T)$ e $\text{Dom}(S_0) = A^\perp \cap \text{Dom}(T^{-1})$. Estabelecemos o problema "primal" com respeito ao operador T_0 ([42])

$$\text{Encontrar } x \in A \text{ tal que } 0 \in T_0(x). \quad (3.3)$$

De forma semelhante, o problema "dual" a (3.3) mediante o operador S_0 ([42]) é definido a seguir:

$$\text{Encontrar } y \in A^\perp \text{ tal que } 0 \in S_0(y). \quad (3.4)$$

O seguinte reultado estabelece relações entre os problemas (3.3) e (3.4).

Proposição 3.2.1 ([42], Proposição 3.1) *Sejam um operador $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e um subespaço A de \mathcal{H} . Então*

1. O ponto $x \in A$ resolve o problema (3.3) se, e somente se, existe $y \in A^\perp$ tal que $y \in T(x)$.
2. O ponto $y \in A^\perp$ resolve o problema (3.4) se, e somente se, existe $x \in A$ tal que $x \in T^{-1}(y)$.

Prova: É consequência das definições de T_0 e S_0 . ■

Pela proposição anterior os problemas (3.3) e (3.4) são visualizados como as duas metades do problema primal-dual (2.1):

$$P(T, A) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp \\ \text{tais que} \\ \bar{y} \in T(\bar{x}). \end{array} \right.$$

Onde $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ é um operador e A um subespaço de \mathcal{H} .

Teorema 3.2.1 ([42], Teorema 3.1) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ e A um subespaço de \mathcal{H} . Então, para qualquer $x \in A$, $y \in A^\perp$, os seguinte itens são equivalentes:*

1. (x, y) é solução do problema (2.1).
2. $0 \in T_A(x + y)$.
3. $0 \in (T^{-1})_{A^\perp}(x + y)$.

Prova: Pelo Teorema 3.1.1 temos que $0 \in T_A(x + y)$ se, e somente se, $y \in T(x)$. Daqui, (1) e (2) são equivalentes. Por outro lado, pelo Lema 3.1.1 resulta que $T_A = (T^{-1})_{A^\perp}$, assim (2) e (3) são equivalentes. ■

Para os operadores T_0 e S_0 , Definição 3.2.1, temos a seguinte propriedade.

Proposição 3.2.2 ([42], Proposição 3.2) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal e A um subespaço de \mathcal{H} . Então*

1. Os operadores T_0 e S_0 são monótonos.
2. Se uma das condições abaixo é verificada
 - (a) $A \cap \text{int}(\text{Dom}(T)) \neq \emptyset$.
 - (b) O espaço \mathcal{H} tem dimensão finita e $A \cap \text{ri}(\text{Dom}(T)) \neq \emptyset$.

Então T_0 é maximal.

3. Se uma das condições abaixo é verificada

(a) $A^\perp \cap \text{int}(\text{Dom}(T)) \neq \emptyset$.

(b) O espaço \mathcal{H} tem dimensão finita e $A^\perp \cap \text{ri}(\text{Dom}(T)) \neq \emptyset$.

Então S_0 é maximal.

Prova:

1. Da Definição 3.2.1 temos que

$$\text{Gr}(T_0) = \{(x, y_A) \in \mathcal{H} \times \mathcal{H} : y \in T(x), x \in A\}.$$

Daqui consideramos $(x, y_A), (\bar{x}, \bar{y}_A) \in \text{Gr}(T_0)$ assim

$$\begin{aligned} \langle \bar{y}_A - y_A, \bar{x} - x \rangle &= \langle (\bar{y} - y)_A, \bar{x} - x \rangle + 0 \\ &= \langle (\bar{y} - y)_A, \bar{x} - x \rangle + \langle (\bar{y} - y)_{A^\perp}, \bar{x} - x \rangle \\ &= \langle \bar{y} - y, \bar{x} - x \rangle. \end{aligned}$$

Portanto T_0 é monótono. De forma análoga, pela Definição 3.2.1

$$\text{Gr}(S_0) = \{(y, x_{A^\perp}) \in \mathcal{H} \times \mathcal{H} : x \in T^{-1}(y), y \in A^\perp\}.$$

Daqui consideramos $(y, x_{A^\perp}), (\bar{y}, \bar{x}_{A^\perp}) \in \text{Gr}(S_0)$

$$\begin{aligned} \langle \bar{x}_{A^\perp} - x_{A^\perp}, \bar{y} - y \rangle &= \langle (\bar{x} - x)_{A^\perp}, \bar{y} - y \rangle + 0 \\ &= \langle (\bar{x} - x)_{A^\perp}, \bar{y} - y \rangle + \langle (\bar{x} - x)_A, \bar{y} - y \rangle \\ &= \langle \bar{x} - x, \bar{y} - y \rangle. \end{aligned}$$

Devido à monotonicidade de T , pela Proposição 1.3.4(1), resulta que T^{-1} é monótono, portanto S_0 é monótono.

2. Pela Proposição 1.3.3(3), resulta que o operador cone normal N_A é monótono maximal, portanto satisfaz-se as hipóteses do Teorema 1.3.1, o qual garante quando o operador $T'_0 := T + N_A$ é monótono maximal. Observamos que T'_0 é então um operador monótono maximal tal que $\text{Dom}(T'_0) \subseteq A$. Como T_0 é a projeção de T'_0 sobre A , isto é

$$\text{Gr}(T_0) = \{(x, y) \in \mathcal{H} \times \mathcal{H} : y \in T'_0(x), y \in A\},$$

obtemos que, T_0 é maximal.

3. Da mesma forma que em (2), temos que N_{A^\perp} é monótono maximal, portanto do Teorema 1.3.1, resulta que $S'_0 := T + N_{A^\perp}$ é monótono maximal. Assim, S'_0 é um operador monótono maximal tal que $\text{Dom}(S'_0) \subseteq A^\perp$. Como S_0 é a projeção de S'_0 sobre A^\perp , isto é

$$\text{Gr}(S_0) = \{ (x, y) \in \mathcal{H} \times \mathcal{H} : y \in S'_0(x), y \in A^\perp \},$$

então S_0 é maximal. ■

3.3 Operador Douglas-Rachford

A seguir fazemos uma breve revisão de um operador de muita importância no desenvolvimento de métodos para encontrar zeros de operadores, estudado em [10] e [11].

Definição 3.3.1 [10] Sejam $T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais e um escalar $\alpha > 0$. O *operador Douglas-Rachford* de T_1, T_2 e α , denotado por S_{α, T_1, T_2} é caracterizado por seu gráfico

$$\text{Gr}(S_{\alpha, T_1, T_2}) = \{ (v + \alpha b, u - v) \in \mathcal{H} \times \mathcal{H} : b \in T_2(u), a \in T_1(v), v + \alpha a = u - \alpha b \}.$$

Denotamos

$$G_{\alpha, T_1, T_2} = J_{\alpha T_1} \circ (2J_{\alpha T_2} - I) + (I - J_{\alpha T_2}).$$

No próximo resultado apresentamos algumas propriedades do operador G_{α, T_1, T_2} .

Proposição 3.3.1 ([10], Proposições 4.1-4.2 e Corolário 4.2.1) *Sejam $T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais e um escalar $\alpha > 0$. Então*

1. $\text{Gr}(G_{\alpha, T_1, T_2}) = \text{Gr}((I + S_{\alpha, T_1, T_2})^{-1})$
 $= \{ (u + \alpha b, v + \alpha b) \in \mathcal{H} \times \mathcal{H} :$
 $b \in T_2(u), a \in T_1(v), v + \alpha a = u - \alpha b \}.$
2. $(G_{\alpha, T_1, T_2})^{-1} - I = S_{\alpha, T_1, T_2}.$
3. O operador S_{α, T_1, T_2} é monótono maximal.
4. G_{α, T_1, T_2} é firmemente não-expansivo e $\text{Dom}(G_{\alpha, T_1, T_2}) = \mathcal{H}.$

3.4 Relação entre os operadores Inversa Parcial e Douglas-Rachford

Discutimos a seguir as relações entre o operador inversa parcial e o operador Douglas-Rachford, estabelecidas por Eckstein [10]. Devido ao fato que o problema $P(T, A)$, ver (2.1), é equivalente ao problema de soma de operadores monótonos (2.8) para T e N_C , Eckstein [10] estabeleceu o seguinte resultado:

Teorema 3.4.1 ([10], Proposição 4.11) *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal e A um subespaço de \mathcal{H} . Então $S_{1, N_A, T} = T_A$.*

Prova: Diretamente da definição do operador Douglas-Rachford (Definição 3.3.1), temos

$$\begin{aligned} \text{Gr}(S_{1, N_A, T}) &= \{(v + b, u - v) \in \mathcal{H} \times \mathcal{H} : b \in T(u), \\ &\quad a \in N_A(v), v + a = u - b\} \\ &= \{(v + b, u - v) \in \mathcal{H} \times \mathcal{H} : b \in T(u), \\ &\quad v \in A, a \in A^\perp, v + a = u - b\}. \end{aligned}$$

As condições $v \in A$, $a \in A^\perp$ e $v + a = u - b$ equivalem a

$$v = (u - b)_A \quad \text{e} \quad a = (u - b)_{A^\perp},$$

donde

$$\begin{aligned} \text{Gr}(S_{1, N_A, T}) &= \{((u - b)_A + b, u - (u - b)_A) \in \mathcal{H} \times \mathcal{H} : b \in T(u)\} \\ &= \{(u_A + b_{A^\perp}, b_A + u_{A^\perp}) \in \mathcal{H} \times \mathcal{H} : b \in T(u)\} \\ &= \text{Gr}(T_A). \end{aligned}$$

Portanto, $S_{1, N_A, T} = T_A$. ■

O resultado acima mostra a equivalência do operador Douglas-Rachford quando $\alpha = 1$, $T_1 = N_A$ e $T_2 = T$, ao operador inversa parcial. Uma pergunta natural é o que acontece quando escolhemos $T_1 = T$ e $T_2 = N_A$ no operador Douglas-Rachford, exploramos isto a seguir.

Teorema 3.4.2 ([10], pág. 153) *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal e A um subespaço de \mathcal{H} . Então*

$$S_{1, T, N_A} = \Lambda \circ T_A \circ \Lambda,$$

onde

$$\begin{aligned}\Lambda : \mathcal{H} &\longrightarrow \mathcal{H} \\ z &\longmapsto \Lambda(z) = z_A - z_{A^\perp}.\end{aligned}$$

Prova: Diretamente da Definição 3.3.1, temos

$$\text{Gr}(S_{1,T,N_A}) = \{(v+b, u-v) \in \mathcal{H} \times \mathcal{H} : b \in N_A(u), a \in T(v), v+a = u-b\}. \quad (3.5)$$

Da condição $v+a = u-b$, obtemos que

$$(v+a)_A = (u-b)_A = u \quad \text{e} \quad (v+a)_{A^\perp} = (u-b)_{A^\perp} = -b.$$

Substituímos em (3.5)

$$\begin{aligned}\text{Gr}(S_{1,T,N_A}) &= \{(v - (v+a)_{A^\perp}, (v+a)_A - v) \in \mathcal{H} \times \mathcal{H} : a \in T(v)\} \\ &= \{(v_A + v_{A^\perp} - v_{A^\perp} - a_{A^\perp}, v_A + a_A - v_A - v_{A^\perp}) \in \mathcal{H} \times \mathcal{H} : \\ &\hspace{15em} a \in T(v)\} \\ &= \{(v_A - a_{A^\perp}, a_A - v_{A^\perp}) \in \mathcal{H} \times \mathcal{H} : a \in T(v)\}.\end{aligned}$$

Assim, $\text{Gr}(S_{1,T,N_A})$ é quase idêntico a $\text{Gr}(T_A)$, exceto por uma diferença de sinal das A^\perp -componentes. Por outro lado, observamos que Λ é linear, contínua, autoadjunta e $\Lambda^{-1} = \Lambda$. Como conseqüência,

$$T_A = \Lambda \circ S_{1,T,N_A} \circ \Lambda.$$

Portanto, obtemos o resultado. ■

Do teorema anterior, temos que

$$\begin{aligned}(I + T_A)^{-1} &= (I + \Lambda \circ (S_{1,T,N_A}) \circ \Lambda)^{-1} \\ &= (\Lambda \circ \Lambda + \Lambda \circ (S_{1,T,N_A}) \circ \Lambda)^{-1} \\ &= (\Lambda \circ (I + S_{1,T,N_A}) \circ \Lambda)^{-1} \\ &= \Lambda \circ (I + S_{1,T,N_A})^{-1} \circ \Lambda \\ &= \Lambda \circ (G_{1,T,N_A}) \circ \Lambda,\end{aligned}$$

ou equivalentemente,

$$G_{1,T,N_A} = \Lambda \circ (I + T_A)^{-1} \circ \Lambda.$$

Assim, as duas aplicações $(I + T_A)^{-1}$ e G_{1,T,N_A} são equivalentes por meio da aplicação Λ .

Capítulo 4

Método Proximal

Neste capítulo apresentamos uma revisão do método de regularização e do algoritmo de ponto proximal, na forma exata, para a otimização convexa e para o cálculo dos zeros de operadores monótonos maximais. De agora em diante, usamos a seguinte notação, $f_S^* := \inf_{x \in S} f(x)$ e $X_S^* := \{x \in \mathcal{H} : f(x) = f_S^*\}$.

4.1 O conceito de regularização

A idéia de regularização surgiu em conexão com problemas mal-postos. Dado um problema da forma

$$L(f) = 0, \tag{4.1}$$

onde f é um elemento de um conjunto X (usualmente um espaço de funções) e $L : X \rightarrow X$ é um operador (usualmente diferencial ou integro-diferencial). (4.1) é dito *mal-posto* quando este não tem, ou tem mais de uma solução; ou no caso de ter uma única solução, esta solução não depende em um sentido contínuo de algum parâmetro do operador L . A idéia é substituir L por um operador regularizado $L + \lambda M$ (com $\lambda \in \mathbb{R}$, $M : X \rightarrow X$), onde M é tal que o problema

$$L(f) + \lambda M(f) = (L + \lambda M)(f) = 0 \tag{4.2}$$

seja *bem-posto*, isto é, que não seja mal-posto, para qualquer $\lambda > 0$. Em tal caso, (4.2) tem uma única solução f_λ , e espera-se que quando λ vai para 0, f_λ aproxima-se ate uma solução de (4.1).

Tal conceito aplica-se aos problemas de otimização se consideramos $X = \mathbb{R}^n$ e $L = \nabla f$, onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa; em tal caso, (4.1) torna-se

$$\nabla f(x) = 0, \tag{4.3}$$

ou equivalentemente,

$$\min_{x \in \mathbb{R}^n} f(x). \quad (4.4)$$

Assumimos que f é limitada inferiormente e consideramos $g : \mathbb{R}^n \rightarrow \mathbb{R}$ estritamente convexa e coerciva, isto é $\lim_{\|x\| \rightarrow \infty} g(x) = +\infty$. O problema (4.3) pode não ter solução, ou ter mais de uma solução, mas o problema regularizado

$$\min_{x \in \mathbb{R}^n} f(x) + \lambda g(x), \quad (4.5)$$

tem uma única solução para cada $\lambda > 0$, devido ao fato que a função $f + \lambda g$ é coerciva, (usando a hipótese que f é limitada inferiormente) o qual reduz o problema (4.5) a um conjunto compacto, garantindo assim a existência de soluções, e da convexidade estrita, obtemos a unicidade da solução. Portanto, (4.5) tem uma única solução x_λ e sobre algumas hipóteses razoáveis (incluindo a existência de soluções de (4.4)) pode ser provado que $\lim_{\lambda \rightarrow 0^+} x_\lambda$ existe e resolve o problema (4.3). Esta é a essência do método de regularização. O valor λ é dito de **parâmetro de regularização** e a função g é dita de **função regularizadora**.

Exemplo 4.1.1 *Sejam $f(x) := x$ e $g(x) = (x - 1)^2$. Observamos que neste caso não existe um ponto $x \in \mathbb{R}$ que solucione o problema (4.4), embora a função $f + \lambda g$, com $\lambda > 0$, tenha um único ponto de mínimo $\bar{x}_\lambda = 1 - \frac{1}{2\lambda}$. ■*

Consideremos o problema (2.4)

$$\begin{cases} \min f(x) \\ \text{s.a:} \\ x \in C, \end{cases}$$

onde $f : \mathcal{H} \rightarrow \mathbb{R}$ é uma função convexa, C é um subconjunto de \mathcal{H} . Uma técnica para resolver o problema (2.4) é usar a idéia de regularização, e assim gerar uma seqüência com os mínimos das funções regularizadas.

O próximo resultado estabelece o método de regularização no caso que o conjunto C é limitado:

Teorema 4.1.1 ([24], Teorema 5) *Sejam $f : \mathcal{H} \rightarrow \mathbb{R}$ uma função convexa semicontínua inferior, C um subconjunto convexo, fechado e limitado de \mathcal{H} . Se $g : \mathcal{H} \rightarrow \mathbb{R}$ é uma função não negativa, fortemente convexa, então, para cada*

$a_k > 0$, existe um ponto de mínimo, x^n , de $f + a_k g$ em C , onde para $a_k \rightarrow 0^+$ a seqüência $\{x^k\}_{k=0}^\infty \subset C$ é minimizadora, isto é

$$\lim_{k \rightarrow \infty} f(x^k) = f_C^*,$$

e converge a um ponto $x^* \in X_C^*$ (e $g(x^*)$ é o valor mínimo de g no conjunto de pontos de mínimo de f).

Em particular, o método de regularização para a função $g = \frac{1}{2} \|\cdot - z\|^2$ é

$$x^{k+1} = x_{a_k} = \arg \min_{x \in C} \left\{ f(x) + \frac{a_k}{2} \|x - z\|^2 \right\}, \quad a_k \rightarrow 0. \quad (4.6)$$

Embora na teoria a função $f_k := f + \frac{a_k}{2} \|\cdot - z\|^2$ tenha melhores propriedades, como a convexidade estrita e a existência de um único ponto de mínimo, que a própria função f , ver Exemplo 4.1.1, na prática tal função pode tornar-se numericamente mal comportada; claramente, para a_k muito pequenos, o ponto x_{a_k} estará mais próximo a uma solução, de modo que considerar a_k muito pequenos pareça ser mais apropriado. Entretanto, vemos que isto não pode ser feito devido aos erros ao calcular a função e o gradiente, como também aos erros de arredondamento ao resolver o problema auxiliar. Então, levanta-se a pergunta: se a solução obtida pelo método de regularização pode ser exata para a_k finitos. Ilustramos no próximo exemplo que $\|x_{a_k} - x^*\|$ pode ser grande para a_k pequenos.

Exemplo 4.1.2 ([35], pág. 175) Sejam $f(x) = \frac{x^p}{p}$, $x \in \mathbb{R}$, $p > 2$ par, $z = 1$. Então $\bar{x} = 0$, e temos que

$$|x^k - \bar{x}| = |x^k| \approx a_k^{p-1}.$$

Portanto se p é grande, então $|x^k - \bar{x}|$ é relativamente grande para a_k pequeno. Por exemplo se $p = 8$, $a_k = 10^{-7}$, temos $|x^k - \bar{x}| \approx 10^{-1}$. ■

4.2 Algoritmo de ponto proximal em otimização convexa

Consideremos o problema irrestrito (2.2),

$$\min_{x \in \mathcal{H}} f(x),$$

onde $f : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ é uma função convexa, própria e semicontínua inferior. Pelo Exemplo 4.1.2, o problema da aproximação por regularização é que, ao menos numericamente, o sistema $\nabla f(x) = 0$ é mal-posto quando λ aproxima-se ao 0, apesar do fato que este tem uma única solução para todo $\lambda > 0$.

Uma tentativa para evitar este possível mal condicionamento é que o parâmetro λ não vá a zero. Seguindo esta idéia, Martinet [27] propôr um método para resolver (4.2) baseado na noção da aplicação proximal dado por Moreau [30].

Definição 4.2.1 [30] *Seja $f : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semicontínua inferior. O único ponto x onde a função $u \longmapsto \frac{1}{2}\|u - z\|^2 + f(u)$ alcança seu mínimo é chamado de **ponto proximal** de z relativo à função f , denotado por $x = \text{prox}_f(z)$ isto é,*

$$\text{prox}_f(z) = \arg \min_{u \in \mathcal{H}} \left\{ f(u) + \frac{1}{2}\|u - z\|^2 \right\}.$$

Martinet [27] considerou o problema de minimização (2.4) quando f é uma função convexa e semicontínua inferior, e C é um subconjunto convexo, fechado de \mathcal{H} , e estabeleceu um algoritmo de regularização o qual apresentamos a seguir.

Teorema 4.2.1 ([27], Teorema 3) *Sejam f uma função convexa e semicontínua inferior, C um subconjunto convexo e fechado de \mathcal{H} . Se $\{x \in C : f(x) \leq a\}$ é limitado para todo $a \in \mathbb{R}$ então todo valor de aderência fraco da seqüência $\{x^k\}_{k=0}^\infty$ gerada por*

$$\begin{cases} x^0 \in C, \\ x^{k+1} = \arg \min_{x \in C} \{ f(x) + \|x - x^k\|^2 \}, \end{cases} \quad (4.7)$$

é solução de (2.4) e a seqüência $\{x^k\}_{k=0}^\infty$ é minimizante, isto é:

$$\lim_{k \rightarrow \infty} f(x^k) = \min_{x \in C} f(x).$$

Para o problema de minimização irrestrito (2.2), temos o seguinte:

Teorema 4.2.2 ([28], Teorema 2) *Sejam f uma função convexa e semicontínua inferior, C um subconjunto convexo e fechado de \mathcal{H} . Então a aplicação proximal prox_f é firmemente não-expansiva em $\text{dom}(f)$, além disso, se a função f é fortemente convexa então prox_f é contrativa. O conjunto dos pontos fixos de prox_f é o conjunto de soluções de (2.2), o qual é convexo, fechado, limitado e não vazio.*

Corolário 4.2.1 ([28], Corolário 1) *A seqüência $\{x^k\}_{k=0}^\infty$ gerada a partir de*

$$\begin{cases} x^0 \in \mathcal{H}, \\ x^{k+1} = \text{prox}_f(x^k), \end{cases} \quad (4.8)$$

converge fracamente a uma solução do problema (2.2). Além disso, tal seqüência é minimizante, isto é:

$$\lim_{k \rightarrow \infty} f(x^k) = \min_{x \in \mathcal{H}} f(x).$$

A conexão com a seção anterior, é evidente pois, ao invés de variar o parâmetro de regularização a_k , variamos o ponto z , substituindo-o por x^k em (4.6), observamos assim que (4.8) pode ser generalizado no método

$$\begin{cases} x^0 \in \mathcal{H}, \\ x^{k+1} := \arg \min_{x \in \mathcal{H}} \left\{ f(x) + \frac{1}{2\lambda_k} \|x - x^k\|^2 \right\}, \end{cases} \quad (4.9)$$

onde a seqüência de números reais positivos $\{\lambda_k\}_{k=0}^\infty \subset \mathbb{R}_{++}$ é dada. O procedimento (4.9) é denominado de **Algoritmo de Ponto Proximal** (veja, por exemplo [39], [13], [22], [23], [19]), devido a sua próxima relação com a aplicação proximal. Assim, o algoritmo de ponto proximal é um método de regularização no qual o parâmetro de regularização λ_k não vai para zero evitando problemas no cálculo dos pontos da seqüência. O próximo teorema contém propriedades de convergência do algoritmo de ponto proximal.

Teorema 4.2.3 ([13], Teorema 2.1) *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semicontínua inferior e $\{x^k\}_{k=0}^\infty$ uma seqüência gerada por (4.9). Para qualquer $u \in \mathcal{H}$, a seguinte estimativa de convergência global satisfaz-se:*

$$f(x^k) - f(u) \leq \frac{\|u - x^0\|^2}{2\delta_k},$$

onde $\delta_k := \sum_{n=1}^k \lambda_n$. Conseqüentemente, se $\delta_k \rightarrow \infty$, então $f(x^k) \downarrow f_{\mathcal{H}}^*$. Se $X_{\mathcal{H}}^* \neq \emptyset$, então $\{x^k\}_{k=0}^\infty$ converge fracamente a um minimizador de f . Além disso,

$$f(x^k) - f_{\mathcal{H}}^* \leq \frac{\text{dist}(x^0, X_{\mathcal{H}}^*)}{2\delta_k}.$$

Na literatura, as propriedades de convergência do algoritmo de ponto proximal, são somente estudadas no caso quando f tem um mínimo, e a velocidade de convergência do algoritmo de ponto proximal é dado somente no caso de f ser fortemente convexa. Para mais detalhes no assunto ver [39], [13], [22], [23], [19].

4.3 Algoritmo de ponto proximal para operadores monótonos maximais

Consideramos o problema (2.9):

$$\text{Encontrar } \bar{x} \in \mathcal{H} \text{ tal que } 0 \in T(\bar{x}),$$

onde $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ é um operador monótono maximal. Pela Proposição 1.3.5, o problema de encontrar zeros de T reduz-se ao problema de encontrar um ponto fixo do operador não-expansivo J_T . Daqui, o método o qual ocorre-nos naturalmente é o método iterativo

$$x^{k+1} = J_T(x^k), \quad (4.10)$$

para algum x^0 dado. O algoritmo de ponto proximal é dado pelo o esquema iterativo (4.10) para a resolvente $J_{\lambda T}$ associado com T . O operador $J_{\lambda T}$ é também denominado de *aplicação proximal* [39] associado com λT , segundo a terminologia de Moreau [30] para o caso de $T = \partial f$.

O *algoritmo de ponto proximal*, no contexto de operadores monótonos maximais, é definido pela seguinte regra:

$$\begin{cases} z^0 \in \mathcal{H}, \\ z^{k+1} := (I + \lambda_k T)^{-1}(z^k), \end{cases} \quad (4.11)$$

onde a seqüência $\{\lambda_k\}_{k=0}^{\infty}$ em \mathbb{R}_{++} é dada. A boa definição da seqüência deve-se ao Teorema de Minty, o qual afirma que para cada $z \in \mathcal{H}$ e $c > 0$ existe um único $u \in \mathcal{H}$ tal que $z - u \in cT(u)$, isto é

$$z \in (I + cT)(u).$$

Em vista do Teorema 1.3.4, temos que o operador $J_{cT} = (I + cT)^{-1}$ é ponto-ponto, não-expansivo, com $\text{Dom}(J_{cT}) = \mathcal{H}$, e devido à Proposição 1.3.5, $J_{cT}(z) = z$ se, e somente se, $0 \in T(z)$.

O algoritmo de ponto proximal é essencialmente um método de regularização, no sentido de ser um método de aproximação para resolver um problema, possivelmente mal-posto, por uma seqüência de problemas bem-postos para a qual a seqüência das soluções aproximadas é definida excepcionalmente. Se o conjunto das soluções do problema original for não vazio, com algumas hipóteses adicionais,

a seqüência converge a alguma solução. No caso de um problema de otimização, isto é, $T = \partial f$, o procedimento (4.11) reduz-se ao (4.9) e a seqüência $\{x^k\}_{k=0}^{\infty}$ converge ao ínfimo buscado.

Resultados de convergência do algoritmo de ponto proximal foram obtidos primeiro por Martinet [27] e [28] no caso $c_k = c$, ver Corólaro 4.2.1 no caso quando $T = \partial f$. Ele analisou em [27] um algoritmo para resolver o problema $\text{VIP}(T, C)$ quando o operador T é hemicontínuo em C (ver Definição 1.3.4). A seguir reproduzimos este resultado.

Teorema 4.3.1 ([27], Teorema 1) *Sejam C um subconjunto convexo, fechado e limitado de \mathcal{H} e $T : C \longrightarrow \mathcal{H}$ um operador monótono e hemicontínuo em C . Então a seqüência $\{x^k\}_{k=0}^{\infty}$ gerada por*

$$\begin{cases} \text{Dado } x^0 \in C \text{ encontrar o único } x^{k+1} \in C \text{ tal que} \\ \langle T(x^{k+1}) + x^{k+1} - x^k, x - x^{k+1} \rangle \geq 0 \quad \forall x \in C, \end{cases} \quad (4.12)$$

converge fracamente a um ponto \bar{x} solução do $\text{VIP}(T, C)$.

Observamos que (4.12) é um método de regularização no sentido que $T_k(u) := T(u) + \nabla(\frac{1}{2}\|u - x^k\|^2)$ é uma regularização para T , posto que a existência e unicidade de x^{k+1} é garantida pelas hipóteses, em particular, se $T = \nabla f$ temos $T_k(u) = \nabla(f(u) + \frac{1}{2}\|u - x^k\|^2)$. Rockafellar [39] observou que o método (4.12) é um caso particular de (4.11), pois (4.12) é equivalente a

$$x^{k+1} = (I + (T + N_C))^{-1}(x^k),$$

com $\lambda_k = 1$ para todo k , e como consequência do Teorema 1.3.2, o qual assegura que o operador $T + N_C$ é monótono maximal.

Os resultados de Martinet [27] e [28], são baseados em um teorema mais geral o qual envolve operadores firmemente não-expansivos:

Teorema 4.3.2 ([28], Teorema 1) *Seja $U : C \longrightarrow C$ um operador firmemente não-expansivo, onde C é um subconjunto não vazio, convexo, fechado e limitado de \mathcal{H} . Então a seqüência $\{x^k\}_{k=0}^{\infty}$ gerada pelo procedimento*

$$\begin{cases} x^0 \in C, \\ x^{k+1} := U(x^k), \end{cases} \quad (4.13)$$

converge fracamente a algum ponto fixo de U .

O teorema anterior é um corolário do teorema de Opial para $\lambda I + (1 - \lambda)U$, quando U é não-expansivo, e $0 < \lambda < 1$, para espaços de Hilbert.

Teorema 4.3.3 ([31], Teorema 3) *Sejam C um conjunto convexo, fechado de \mathcal{H} e $T : C \rightarrow C$ um operador não-expansivo com pelo menos um ponto fixo. Então para qualquer $\lambda \in (0, 1)$ a seqüência de aproximações sucessivas $\{x^k\}_{k=0}^\infty$ gerada pelo procedimento*

$$\begin{cases} x^0 \in C, \\ x^{k+1} := (\lambda I + (1 - \lambda)T)(x^k), \end{cases} \quad (4.14)$$

é fracamente convergente a algum ponto fixo de T .

Em resumo, o esquema (4.11) é uma forma elegante de uniformizar o cálculo dos zeros de operadores monótonos maximais por um esquema semelhante ao do ponto fixo (4.10). A seguir, estabelecemos o resultado sobre a convergência do algoritmo de ponto proximal.

Teorema 4.3.4 ([39], Teorema 1) *Seja $\{z^k\}_{k=0}^\infty$ uma seqüência gerada pelo algoritmo de ponto proximal (4.11) com $\{\lambda_k\}_{k=0}^\infty$ e $0 < \bar{\lambda} \leq \lambda_k$. Suponhamos que $\{z^k\}_{k=0}^\infty$ é limitada; isto é equivalente a dizer que $T^{-1}(0) \neq \emptyset$. Então $\{z^k\}_{k=0}^\infty$ converge na topologia fraca a um ponto z^∞ satisfazendo $0 \in T(z^\infty)$, e*

$$0 = \lim_{k \rightarrow \infty} \|(I + (\lambda_k T)^{-1})^{-1}(z^k)\| = \lim_{k \rightarrow \infty} \|z^{k+1} - z^k\|.$$

Rockafellar [39] também estabeleceu condições para garantir a convergência no caso em que o ponto x^{k+1} em (4.11) seja obtido inexatamente, devido ao fato que não é prático o cálculo exato das iterações, pois $(I + \lambda_k T)^{-1}(x^k)$ é um ponto a ser aproximado, e portanto não é disponível. Por tal motivo, critérios para o tratamento das soluções aproximadas de (4.11) são um importante objeto de estudo. Como exemplo disto, Solodov e Svaiter [41] apresentam uma família de algoritmos híbridos os quais combinam idéias de projeção com métodos proximais. A idéia consiste em, tendo x^k , obter um hiperplano que separa estritamente x^k do conjunto $T^{-1}(0)$. O novo iterado, x^{k+1} , será a projeção de x^k sobre este hiperplano, assim, x^{k+1} estará mais próximo do conjunto $T^{-1}(0)$ que x^k . Eles provaram que introduzindo o passo intermediário de projeção, é possível usar critérios de aceitação ainda mais permissíveis que os apresentados por Rockafellar [39], ver [41] para mais detalhes no assunto.

Capítulo 5

Método da Inversa Parcial

Devido à introdução do operador inversa parcial no Capítulo 3, o problema $P(T, A)$, ver (2.1), é equivalente a encontrar um zero de T_A , o que permite-nos a utilização do algoritmo de ponto proximal. Isto introduz um esquema para desenvolver um algoritmo de decomposição denominado o método das inversas parciais, apresentado por Spingarn [42] e [43], o qual é discutido na seção 5.1. Nas seções 5.2 e 5.3, estabelecemos e relacionamos o método de decomposição Douglas-Rachford com o método das inversas parciais.

5.1 Método da inversas parciais

Nesta seção é descrito um algoritmo para resolver o problema primal-dual (2.1). Pelo Teorema 3.1.1, temos que $0 \in T_A(z)$ se, e somente se, $x = z_A$ e $y = z_{A^\perp}$ resolvem o problema $P(T, A)$. Portanto, é natural resolver o problema $P(T, A)$ através da aplicação do algoritmo de ponto proximal, para encontrar um zero do operador inversa parcial T_A . Este procedimento, que detalhamos a seguir, recebe o nome de método das inversas parciais [42], [43].

A seguinte propriedade caracteriza o ponto proximal de T_A em um ponto em termos do ponto proximal do operador S_c , definido em (3.2), aplicado no mesmo ponto. Esta propriedade é usada no desenvolvimento do método estudado em [42]:

Proposição 5.1.1 ([42], pág. 254) *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal, A um subespaço de \mathcal{H} e $\lambda > 0$. Dado $z \in \mathcal{H}$, o ponto proximal $z^+ = (I + \lambda T_A)^{-1}(z)$ está definido por:*

$$z^+ = x'_A + y'_{A^\perp},$$

sendo

$$\begin{aligned}x' &= (I + S_\lambda)^{-1}(z) \\y' &= y + (x - x'), \quad z := x + y, \quad x \in A, \quad y \in A^\perp.\end{aligned}$$

Prova: Dado z , seja x' o ponto proximal definido por $x' = (I + S_\lambda)^{-1}(z)$ então, usando a decomposição de z , temos

$$(x + y) \in (I + S_\lambda)(x'). \quad (5.1)$$

Desta relação e da definição de y' , resulta

$$y' \in S_\lambda(x'),$$

assim, da definição de S_λ , (3.2), equivale a dizer que

$$\frac{1}{\lambda}y'_A + y'_{A^\perp} \in T\left(x'_A + \frac{1}{\lambda}x'_{A^\perp}\right) \quad (5.2)$$

sendo $x' = x'_A + x'_{A^\perp}$, $y' = y'_A + y'_{A^\perp}$. De (5.2), usando a caracterização de T_A (Proposição 3.1.1) deduz-se que

$$\frac{1}{\lambda}y'_A + \frac{1}{\lambda}x'_{A^\perp} \in T\left(x'_A + \frac{1}{\lambda}x'_{A^\perp}\right),$$

quer dizer

$$y'_A + x'_{A^\perp} \in \lambda T\left(x'_A + \frac{1}{\lambda}x'_{A^\perp}\right).$$

Como $z = x + y = x' + y' = (x'_A + y'_{A^\perp}) + (y'_A + x'_{A^\perp})$, obtemos

$$z \in (x'_A + y'_{A^\perp}) + \lambda T_A(x'_A + y'_{A^\perp}).$$

Logo, definindo $z^+ = x'_A + y'_{A^\perp}$, a relação anterior pode ser reescrita como

$$z \in (I + \lambda T_A)(z^+),$$

o qual, devido à monotonicidade maximal de T_A , equivale a

$$z^+ = (I + \lambda T_A)^{-1}(z),$$

concluindo a prova. ■

A proposição anterior mostra como z^{k+1} pode ser calculado a partir de z^k , ou equivalentemente, como $(z^{k+1})_A$ e $(z^{k+1})_{A^\perp}$ são calculados a partir de $(z^k)_A$ e $(z^k)_{A^\perp}$ por meio do operador S_{λ_k} , isto é

$$z^{k+1} = (I + \lambda_k T_A)^{-1}(z^k) = (\bar{x}^k)_A + (\bar{y}^k)_{A^\perp},$$

onde

$$\begin{aligned} z^k &:= x + y, \quad x \in A, \quad y \in A^\perp, \\ \bar{x}^k &= (I + S_{\lambda_k})^{-1}(z^k), \\ \bar{y}^k &= y + (x - \bar{x}^k). \end{aligned}$$

Assim podemos adaptar o algoritmo de ponto proximal para encontrar um zero de T_A , no seguinte procedimento.

Algoritmo 1 [42]

Início: Dados $z^0 \in \mathcal{H}$ e $\{\lambda_k\}_{k=0}^\infty \subset \mathbb{R}_{++}$.

Iteração k ($k = 0, 1, \dots$):

Passo proximal: Encontrar $\bar{x}^k, \bar{y}^k \in \mathcal{H}$ tais que

$$z^k = \bar{x}^k + \bar{y}^k \quad \text{e} \quad \frac{1}{\lambda_k}(\bar{y}^k)_A + (\bar{y}^k)_{A^\perp} \in T\left((\bar{x}^k)_A + \frac{1}{\lambda_k}(\bar{x}^k)_{A^\perp}\right).$$

Atualização:

$$z^{k+1} = (\bar{x}^k)_A + (\bar{y}^k)_{A^\perp}.$$

Denominamos o Algoritmo 1, o *método das inversas parciais*, devido a que o método foi derivado do algoritmo de ponto proximal aplicado ao operador inversa parcial; a convergência deste algoritmo é um caso particular do Teorema 4.3.4, para $T = T_A$. Assim, pela análise anterior, podemos particularizar o Teorema 4.3.4 no contexto do operador inversa parcial. Em resumo

Teorema 5.1.1 ([42], Teorema 4.1) *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador monótono maximal, A um subespaço de \mathcal{H} , $\{\lambda_k\}_{k=0}^\infty$ uma seqüência em \mathbb{R}_{++} tal*

que $\liminf_{k \rightarrow \infty} \lambda_k > 0$ e $\{z^k\}_{k=0}^{\infty}$ uma seqüência gerada pelo Algoritmo 1 aplicado ao problema $P(T, A)$. Então

1. O Algoritmo 1 é equivalente ao algoritmo de ponto proximal no seguinte sentido:

$$z^{k+1} = (I + \lambda_k T_A)^{-1}(z^k).$$

2. Se $\{z^k\}_{k=0}^{\infty}$ não é limitada então o problema $P(T, A)$ não tem solução.
3. Se $\{z^k\}_{k=0}^{\infty}$ é limitada então as seqüências $\{(z^k)_A\}_{k=0}^{\infty}$ e $\{(z^k)_{A^\perp}\}_{k=0}^{\infty}$ convergem fracamente aos pontos \bar{z}_A e \bar{z}_{A^\perp} respectivamente, onde $(\bar{z}_A, \bar{z}_{A^\perp})$ é uma solução do problema $P(T, A)$. ■

Como um caso particular do anterior, aplicaremos a seguir o método das inversas parciais ao problema de minimizar uma função sobre um subespaço, (2.6). Denominamos o procedimento resultante o denominamos de Algoritmo 2. Para isto, particularizamos o Algoritmo 1 para resolver o problema $P(\partial f, A)$.

Introduzimos um fator de escala $\eta > 0$. Pelo Teorema 2.2.1, sabemos que \bar{x} é solução do problema (2.6) se, e somente se, existe $\bar{y} \in A^\perp$ tal que

$$0 \in \bar{y} + \partial f(\bar{x}),$$

o qual é equivalente a encontrar $\bar{x} \in A$ e $\bar{y} \in A^\perp$ tais que

$$0 \in \eta(\bar{y} + \partial f(\bar{x})) = \eta\bar{y} + \eta\partial f(\bar{x}),$$

ou

$$-\eta\bar{y} \in \eta\partial f(\bar{x}).$$

Seja $\bar{w} := -\eta\bar{y}$, então, $\bar{w} \in A^\perp$ se, e somente se, $\bar{y} \in A^\perp$. Por outro lado, como f é uma função convexa, própria e semicontínua inferior, sabemos, pelo Teorema 1.3.5, que ∂f é monótono maximal e pela Proposição 1.3.4, o operador $\eta\partial f$ é também monótono maximal. Portanto o problema (2.7) é equivalente a

$$P(\eta\partial f, A) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{w} \in A^\perp \\ \text{tais que} \\ \bar{w} \in \eta\partial f(\bar{x}). \end{array} \right. \quad (5.3)$$

Esta reformulação do problema (2.7) deixa o seu conjunto solução inalterado, mas muda a forma do algoritmo. Reescrevemos o Algoritmo 1 para o problema (5.3):

Início: Dado $\{\lambda_k\}_{k=0}^{\infty} \subset \mathbb{R}_{++}$, considerar $z^0 = x^0 - \eta y^0$ com $x^0 \in A$ e $y^0 \in A^\perp$.

Iteração k ($k = 0, 1, \dots$):

Passo proximal: encontrar $\bar{x}^k, \bar{w}^k \in \mathcal{H}$ tais que

$$z^k = \bar{x}^k + \bar{w}^k \quad \text{e} \quad \frac{1}{\lambda_k}(\bar{w}^k)_A + (\bar{w}^k)_{A^\perp} \in \eta \partial f\left((\bar{x}^k)_A + \frac{1}{\lambda_k}(\bar{x}^k)_{A^\perp}\right),$$

o qual é equivalente a

$$z^k = \bar{x}^k - \eta \left(\frac{-\bar{w}^k}{\eta}\right) \quad \text{e} \quad 0 \in \frac{1}{\lambda_k} \left(\frac{-\bar{w}^k}{\eta}\right)_A + \left(\frac{-\bar{w}^k}{\eta}\right)_{A^\perp} + \partial f\left((\bar{x}^k)_A + \frac{1}{\lambda_k}(\bar{x}^k)_{A^\perp}\right).$$

Portanto, devemos encontrar \bar{x}^k e $\bar{y}^k = -\frac{\bar{w}^k}{\eta}$, tais que

$$z^k = \bar{x}^k - \eta \bar{y}^k \quad \text{e} \quad 0 \in \frac{1}{\lambda_k}(\bar{y}^k)_A + (\bar{y}^k)_{A^\perp} + \partial f\left((\bar{x}^k)_A + \frac{1}{\lambda_k}(\bar{x}^k)_{A^\perp}\right). \quad (5.4)$$

Atualização:

$$z^{k+1} = (\bar{x}^k)_A - \eta(\bar{y}^k)_{A^\perp}. \quad (5.5)$$

Em resumo, o Algoritmo 1 transforma-se em:

Algoritmo 2

Início: Dado $\{\lambda_k\}_{k=0}^{\infty} \subset \mathbb{R}_{++}$, escolher $z^0 := x^0 - \eta y^0$ tal que $x^0 \in A$ e $y^0 \in A^\perp$.

Iteração k ($k = 0, 1, \dots$):

Passo proximal: Encontrar $\bar{x}^k, \bar{y}^k \in \mathcal{H}$ tais que

$$z^k = \bar{x}^k - \eta \bar{y}^k \quad \text{e} \quad 0 \in \frac{1}{\lambda_k}(\bar{y}^k)_A + (\bar{y}^k)_{A^\perp} + \partial f\left((\bar{x}^k)_A + \frac{1}{\lambda_k}(\bar{x}^k)_{A^\perp}\right).$$

Atualização:

$$z^{k+1} = (\bar{x}^k)_A - \eta(\bar{y}^k)_{A^\perp}.$$

Visto que o algoritmo acima é um caso particular do Algoritmo 1, este também é equivalente ao algoritmo de ponto proximal. Isto é:

$$x^{k+1} - \eta y^{k+1} = \left(I + \lambda_k (\eta \partial f)_A \right)^{-1} (x^k - \eta y^k).$$

A seguir escrevemos o Algoritmo 2 em termos da função f . Para isto definamos

$$\begin{aligned} g_k : \mathcal{H} &\longrightarrow \mathbb{R} \\ s &\longmapsto g_k(s) := f\left(s_A + \frac{1}{\lambda_k} s_{A^\perp}\right). \end{aligned}$$

Observamos que $g_k = f \circ \Lambda_k$, onde

$$\begin{aligned} \Lambda_k : \mathcal{H} &\longrightarrow \mathcal{H} \\ x &\longmapsto \Lambda_k(x) =: x_A + \frac{1}{\lambda_k} x_{A^\perp}, \end{aligned}$$

é um operador linear, contínuo com adjunta $\Lambda_k^* = \Lambda_k$. Temos que f e Λ satisfazem as hipóteses da Proposição 1.1.1, a qual assegura que g_k é convexa, semicontínua inferior e

$$\partial g_k(\bar{x}) = \Lambda_k\left(\partial f(\Lambda_k(\bar{x}))\right).$$

Assim

$$0 \in \bar{y} + \lambda_k \partial g_k(\bar{x}) \iff 0 \in \frac{1}{\lambda_k} \bar{y}_A + \bar{y}_{A^\perp} + \partial f\left(\bar{x}_A + \frac{1}{\lambda_k} \bar{x}_{A^\perp}\right). \quad (5.6)$$

Portanto na iteração k , temos

$$0 \in \bar{y}^k + \lambda_k \partial g_k(\bar{x}^k) = \frac{1}{\eta} (\bar{x}^k - z^k) + \lambda_k \partial g_k(\bar{x}^k),$$

logo $\bar{w} = \bar{x}^k$ é o único minimizador da função fortemente convexa (ver Definição 1.3.9)

$$\phi_k(w) := g_k(w) + \frac{1}{2\eta\lambda_k} \|w - z^k\|^2,$$

portanto, o passo proximal pode ser substituído pela discussão acima. Além disso, da atualização temos que:

$$z^{k+1} = (\bar{x}^k)_A - \eta (\bar{x}^k - z^k)_{A^\perp}.$$

Assim, o Algoritmo 2 toma a seguinte caracterização:

Algoritmo 3 [43]

Início: Dado $\{\lambda_k\}_{k=0}^\infty \subset \mathbb{R}_{++}$ escolher $z^0 := x^0 - \eta y^0$ tal que $x^0 \in A$ e $y^0 \in A^\perp$.

Iteração k ($k = 0, 1, \dots$):

Passo proximal: Encontrar o único $\bar{x}^k \in \mathcal{H}$ tal que $w = \bar{x}^k$ minimize

$$\phi_k(w) = f\left(w_A + \frac{1}{\lambda_k} w_{A^\perp}\right) + \frac{1}{2\eta\lambda_k} \|w - z^k\|^2.$$

Atualização:

$$z^{k+1} = (\bar{x}^k)_A - \eta(\bar{x}^k - z^k)_{A^\perp}.$$

Assim a convergência do Algoritmo 3 é um caso particular do Teorema 5.1.1.

Teorema 5.1.2 ([43], Teorema 2) *Sejam $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semicontínua inferior, A um subespaço de \mathcal{H} , $\{\lambda_k\}_{k=0}^\infty$ uma seqüência em \mathbb{R}_{++} tal que $\liminf_{k \rightarrow \infty} \lambda_k > 0$ e $\{z^k\}_{k=0}^\infty$ uma seqüência gerada pelo Algoritmo 3 aplicado ao problema (2.6). Então:*

1. *Se $\{z^k\}_{k=0}^\infty$ não é limitada então o problema (2.7) não tem solução.*
2. *Se uma das condições abaixo é verificada*

(a) $A \cap \text{int}(\text{dom}(f)) \neq \emptyset,$

(b) \mathcal{H} de dimensão finita e $A \cap \text{ri}(\text{dom}(f)) \neq \emptyset.$

E $\{z^k\}_{k=0}^\infty$ não é limitada então o problema (2.6) não tem solução.

3. *Se $\{z^k\}_{k=0}^\infty$ é limitada então as seqüências $\{(z^k)_A\}_{k=0}^\infty$ e $\{(z^k)_{A^\perp}\}_{k=0}^\infty$ convergem fracamente a \bar{z}_A e \bar{z}_{A^\perp} , respectivamente, onde $(\bar{z}_A, \bar{z}_{A^\perp})$ é uma solução do problema (2.7) e \bar{z}_A é uma solução do problema (2.6). ■*

Devido ao fato que o método das inversas parciais requer resolver o problema (2.9), uma ferramenta para isto é a aplicação do algoritmo de ponto proximal

para encontrar um zero do operador T_A , podemos utilizar técnicas inexatas, como algoritmos híbridos no sentido de Solodov e Svaiter [41], para encontrar um zero do operador inversa parcial. Segundo esta metodologia, recentemente Burachik, Sagastizábal e Scheimberg [6] forneceram um algoritmo híbrido para o contexto do operador inversa parcial, para mais detalhes sobre o assunto, ver [6].

5.2 Método de decomposição Douglas-Rachford

A principal dificuldade na aplicação do algoritmo de ponto proximal para a resolução do problema de encontrar zeros (2.9), é o cálculo do resolvente $J_{\alpha T} := (I + \lambda T)^{-1}$, para muitos operadores monótonos maximais isto pode ser muito complicado. Como uma saída para este problema, supomos que podemos representar o operador $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ como uma soma de outros dois operadores $T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ monótonos maximais, isto é

$$T = T_1 + T_2,$$

tais que o cálculo das resolventes $J_{\alpha T_1}$ e $J_{\alpha T_2}$ sejam mais fáceis de obter do que o resolvente $J_{\alpha T}$ do operador original.

Definição 5.2.1 [11] *Sejam $T, T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais tais que $T = T_1 + T_2$. Um Algoritmo de decomposição para o problema de encontrar zeros de T , é um método que emprega as resolventes $J_{\lambda T_1}$ e $J_{\lambda T_2}$ de T_1 e T_2 mas não usa a resolvente $J_{\lambda T}$ do operador original T .*

Um método que explora a definição anterior é o **método de decomposição Douglas-Rachford** definido, segundo [25], como:

$$\begin{cases} \text{Seja } z^0 := u^0 + \alpha b^0 \text{ onde } u^0 \in \text{Dom}(T_2), b^0 \in T_2(u^0), \\ z^{k+1} = J_{\alpha T_1}(2J_{\alpha T_2} - I)(z^k) + (I - J_{\alpha T_2})(z^k). \end{cases} \quad (5.7)$$

A seguir reproduzimos o resultado de convergência do método (5.7).

Teorema 5.2.1 ([25], Teorema 1) *Sejam $T, T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais e um escalar $\lambda > 0$. Se $T = T_1 + T_2$ e $T^{-1}(0) \neq \emptyset$ então a seqüência $\{z^k\}_{k=0}^{\infty}$ gerada por (5.7) converge fracamente a um ponto $\bar{z} \in \mathcal{H}$, tal que $\bar{u} = J_{\lambda T_2}(\bar{z})$ é uma solução de (2.9), isto é $0 \in T(\bar{u})$.*

Em [10] e [11] estuda-se, em particular, o método de decomposição Douglas-Rachford estabelecido pela regra

$$\begin{cases} z^0 \in \mathcal{H}, \\ z^{k+1} := G_{\alpha, T_1, T_2}(z^k), \end{cases} \quad (5.8)$$

onde G_{α, T_1, T_2} é o operador definido na Definição 3.3.1.

5.3 Relação entre os métodos de inversa parcial e Douglas-Rachford.

Devido à Proposição 3.3.1, temos que $G_{\alpha, T_1, T_2} = (I + S_{\alpha, T_1, T_2})^{-1}$, onde S_{α, T_1, T_2} é o operador Douglas-Rachford, o qual é monótono maximal. Portanto, estabelecemos primeiro a relação do método de decomposição Douglas-Rachford e do algoritmo de ponto proximal a seguir.

Teorema 5.3.1 ([11], Teorema 6) *Sejam $T, T_1, T_2 : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ operadores monótonos maximais e um escalar $\alpha > 0$. Se $T = T_1 + T_2$ e $T^{-1}(0) \neq \emptyset$ então o método de decomposição Douglas-Rachford (5.8) é equivalente a aplicar o algoritmo de ponto proximal ao operador S_{α, T_1, T_2} com $\lambda_k = 1$ para todo k .*

Por outro lado, sabemos pelo Teorema 3.4.1 que $S_{1, N_A, T} = T_A$, portanto, temos o seguinte:

Corolário 5.3.1 *Se $T^{-1}(0) \neq \emptyset$ então o método de decomposição Douglas-Rachford (5.8) é equivalente a aplicar o método das inversas parciais ao operador $S_{1, N_A, T}$ com $\lambda_k = 1$ para todo k .*

Capítulo 6

Uma Extensão do Operador Inversa Parcial

O objetivo deste capítulo é analisar a substituição da decomposição $\mathcal{H} = A \oplus A^\perp$ por uma outra do tipo $\mathcal{H} = K + K^\circ$, onde K é um cone convexo fechado e K° é o cone polar de K . Este estudo tem como finalidade examinar o problema seguinte:

$$P(T, K) \left\{ \begin{array}{l} \text{Encontrar } \bar{x} \in A \text{ e } \bar{y} \in A^\perp \\ \text{tais que} \\ \bar{y} \in T(\bar{x}) \text{ e } \langle \bar{x}, \bar{y} \rangle = 0. \end{array} \right. \quad (6.1)$$

além da possibilidade de estender o algoritmo das inversas parciais para cones.

A motivação do estudo provem dos problemas que podem ser considerados como instâncias particulares de (6.1):

$$\left\{ \begin{array}{l} \min f(x) \\ \text{s.a:} \\ x \in K, \end{array} \right. \quad (6.2)$$

isto estudado em Isac [17], e assim tentar uma abordagem como a feita por Spingarn [42] como no caso dos subespaços. Este questionamento surge do fato que cada ponto $x \in \mathcal{H}$ também pode ser decomposto de forma única em $x = x_K + x_{K^\circ}$.

Campos [8] estudou também este questionamento, apresentando um exemplo onde o operador inversa parcial generalizado não mantém a monotonicidade do operador original.

6.1 O operador inversa parcial generalizado

Dado K um cone convexo e fechado de \mathcal{H} , sabemos que todo ponto x do espaço admite uma única decomposição $x = x_1 + x_2$, onde $x_1 \in K$, $x_2 \in K^\circ$ e $\langle x_1, x_2 \rangle = 0$ (Proposição 1.5.4). Portanto, temos a boa definição do operador inversa parcial com respeito a K , que é definido a seguir.

Definição 6.1.1 [9] *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador, K um cone convexo fechado com vértice na origem de \mathcal{H} e K° o cone polar de K . O **operador inversa parcial generalizado** T_K de T com respeito a K é o operador obtido pelas trocas das K° -componentes de cada par no $\text{Gr}(T)$:*

$$\text{Gr}(T_K) = \{(x_K + y_{K^\circ}, y_K + x_{K^\circ}) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\}.$$

Observamos que as relações do Lema 3.1.1 ($T_A = (T^{-1})_{A^\perp} = (T_{A^\perp})^{-1}$), podem ser generalizadas quando o subespaço A é substituído por um cone convexo fechado de \mathcal{H} .

Propriedade 6.1.1 *Sejam $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$ um operador e K um cone convexo fechado de \mathcal{H} . Então*

$$T_K = (T^{-1})_{K^\circ} = (T_{K^\circ})^{-1}.$$

Prova: Da Definição de T^{-1} , temos

$$\text{Gr}(T^{-1}) = \{(y, x) \in \mathcal{H} \times \mathcal{H} : x \in T^{-1}(y)\}.$$

Daqui, da Definição 6.1.1 resulta

$$\begin{aligned} \text{Gr}((T^{-1})_{K^\circ}) &= \{(y_{K^\circ} + x_K, x_{K^\circ} + y_K) \in \mathcal{H} \times \mathcal{H} : x \in T^{-1}(y)\} \\ &= \{(x_K + y_{K^\circ}, y_K + x_{K^\circ}) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \text{Gr}(T_K). \end{aligned}$$

Por outro lado,

$$\text{Gr}(T_{K^\circ}) = \{(x_{K^\circ} + y_K, y_{K^\circ} + x_K) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\},$$

assim,

$$\begin{aligned} \text{Gr}((T_{K^\circ})^{-1}) &= \{(y_{K^\circ} + x_K, x_{K^\circ} + y_K) \in \mathcal{H} \times \mathcal{H} : y \in T(x)\} \\ &= \text{Gr}((T^{-1})_{K^\circ}). \end{aligned} \quad \blacksquare$$

Como no caso dos subespaços, o operador inversa parcial generalizado também não preserva a característica do operador original de ser ponto-ponto, como mostramos a seguir.

Exemplo 6.1.1 Considere $\mathcal{H} = \mathbb{R}$, $K = \mathbb{R}_+$, $K^\circ = -\mathbb{R}_+$. Definimos o operador $T : \mathbb{R} \rightarrow \mathbb{R}$ por

$$T(x) := x = 1.$$

De fato, $1 \in T_K(0)$ pois existem $x = 0 \in K$ e $y = T(x) = 1 \in K$ como também $-1 \in T_K(0)$ pois existem $x = -1 \in K^\circ$ e $y = T(x) = 0 \in K$, daqui $\{-1, +1\} \in T_K(0)$, portanto T_K é ponto-conjunto. ■

Notamos que o operador inversa parcial generalizado, não necessariamente provém de um só operador, como no caso dos subespaços, provado na Propriedade 3.1.2. Este fato é ilustrado no seguinte exemplo.

Exemplo 6.1.2 Considere $\mathcal{H} = \mathbb{R}$, $K = \mathbb{R}_+$, $K^\circ = -\mathbb{R}_+$. Definimos os operadores $S, T : \mathbb{R} \rightarrow \mathcal{P}(\mathbb{R})$ por

$$S(x) := x - 2$$

e

$$T(x) := \begin{cases} x - 2 & \text{se } x < 0, \\ [-2, 0] & \text{se } x = 0, \\ 0 & \text{se } x \in (0, 2], \\ x - 2 & \text{se } x > 2. \end{cases}$$

Vamos mostrar que $S_K = T_K$ sendo $T \neq S$.

Calculamos $\text{Gr}(S_K)$: dado $x \in \mathbb{R}$ resulta

$$(x_K + (x - 2)_{K^\circ}, (x - 2)_K + x_{K^\circ}) \in \text{Gr}(S_K). \quad (6.3)$$

- Se $x \geq 2$, então $S(x) \geq 0$ assim de (6.3) temos que $(x, x - 2) \in \text{Gr}(S_K)$.
- Se $0 \leq x < 2$, então $S(x) < 0$ assim de (6.3) temos que $(2x - 2, 0) \in \text{Gr}(S_K)$.
- Se $x < 0$, então $S(x) < 0$ assim de (6.3) temos que $(x - 2, x) \in \text{Gr}(S_K)$.

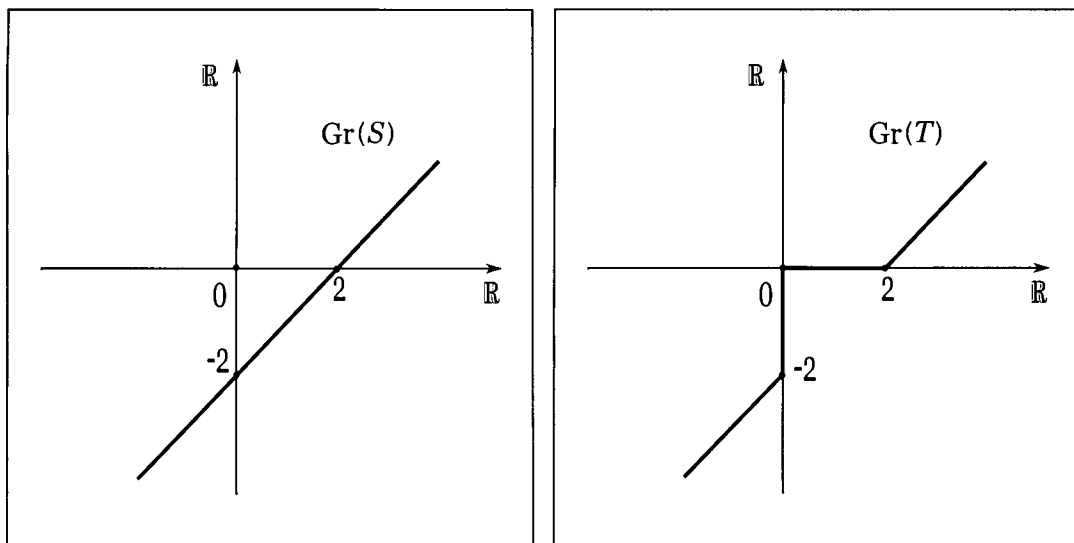


Figura 6.1: Gráficos dos operadores S e T do Exemplo 6.1.2.

1. Afirmamos que $S_K(x) = x - 2$ quando $x > 2$. Suponhamos que existe $w \in S_K(x)$ tal que $w \neq x - 2$, então existem $\hat{x}, \hat{y} \in \mathbb{R}$ tais que $x = \hat{x}_K + \hat{y}_{K^\circ}$, $w = \hat{y}_K + \hat{x}_{K^\circ}$ com $\hat{y} \in S(\hat{x})$. Como S é ponto-ponto, $\hat{y} = S(\hat{x}) = \hat{x} - 2$, daqui

$$(x, w) = (\hat{x}_K + (\hat{x} - 2)_{K^\circ}, (\hat{x} - 2)_K + \hat{x}_{K^\circ}). \quad (6.4)$$

- Se $\hat{x} > 2$ implica que $S(\hat{x}) > 0$, em (6.4) temos $(x, w) = (\hat{x}, \hat{x} - 2)$, absurdo com $w \neq x - 2$.
- Se $0 \leq \hat{x} \leq 2$ implica que $S(\hat{x}) \leq 0$, em (6.4) temos $(x, w) = (2\hat{x} - 2, 0)$, como $x > 2$ temos que $2\hat{x} - 2 > 2$, portanto $\hat{x} > 2$ o que é absurdo.
- Se $\hat{x} < 0$ implica que $S(\hat{x}) < 0$, em (6.4) temos $(x, w) = (\hat{x} - 2, \hat{x})$, como $x > 2$ temos que $\hat{x} - 2 > 2$, portanto $\hat{x} > 4$ o que é absurdo.

Portanto $S_K(x) = x - 2$ quando $x > 2$.

2. Afirmamos que $S_K(2x - 2) = 0$ quando $0 \leq x \leq 2$. Suponhamos que existe $w \in S_K(2x - 2)$ tal que $w \neq 0$, então existe $\hat{x} \in \mathbb{R}$ tais que

$$(2x - 2, w) = (\hat{x}_K + (\hat{x} - 2)_{K^\circ}, (\hat{x} - 2)_K + \hat{x}_{K^\circ}). \quad (6.5)$$

- Se $\hat{x} > 2$ implica que $S(\hat{x}) > 0$, em (6.5) temos $(2x - 2, w) = (\hat{x}, \hat{x} - 2)$, como $0 \leq x \leq 2$ obtemos que $\hat{x} = 2x - 2 \leq 2$ o que é um absurdo.

- Se $0 \leq \hat{x} \leq 2$ implica que $S(\hat{x}) \leq 0$, em (6.5) temos $(2x - 2, w) = (2\hat{x} - 2, 0)$, temos que $x = \hat{x}$ e $w = 0$, o que é um absurdo.
- Se $\hat{x} < 0$ implica que $S(\hat{x}) < 0$, em (6.5) temos $(2x - 2, w) = (\hat{x} - 2, \hat{x})$, temos que $x = \frac{\hat{x}}{2} < 0$, o que é absurdo.

Portanto $S_K(2x - 2) = 0$ quando $x \geq 2$. Fazemos uma mudança de variável e obtemos que $S_K(\bar{x}) = 0$ para $\bar{x} \in [-2, 2]$.

3. Afirmamos que $S_K(x - 2) = x$ quando $x < 0$. Suponhamos que existe $w \in S_K(x - 2)$ tal que $w \neq x$, então existe $\hat{x} \in \mathbb{R}$ tais que

$$(x - 2, w) = (\hat{x}_K + (\hat{x} - 2)_{K^\circ}, (\hat{x} - 2)_K + \hat{x}_{K^\circ}). \quad (6.6)$$

- Se $\hat{x} > 2$ implica que $S(\hat{x}) > 0$, em (6.6) temos $(x - 2, w) = (\hat{x}, \hat{x} - 2)$, então $0 > x = \hat{x} + 2 > 4$, absurdo.
- Se $0 \leq \hat{x} \leq 2$ implica que $S(\hat{x}) \leq 0$, em (6.6) temos $(x - 2, w) = (2\hat{x} - 2, 0)$, temos que $0 > x = 2\hat{x} \geq 0$, absurdo.
- Se $\hat{x} < 0$ implica que $S(\hat{x}) < 0$, em (6.6) temos $(x - 2, w) = (\hat{x} - 2, \hat{x})$, temos que $x = \hat{x}$ e $w = x$, absurdo.

Portanto $S_K(x - 2) = x$ quando $x < 0$. Fazemos uma mudança de variável e obtemos que $S_K(\bar{x}) = \bar{x} + 2$ para $\bar{x} < -2$.

Em resumo, obtemos que

$$S_K(\bar{x}) = \begin{cases} \bar{x} + 2 & \text{se } \bar{x} < -2, \\ 0 & \text{se } |\bar{x}| \leq 2, \\ \bar{x} - 2 & \text{se } \bar{x} > 2. \end{cases}$$

De forma análoga calculamos $\text{Gr}(T_K)$:

- Se $x > 2$, então $(x, x - 2) \in \text{Gr}(T_K)$.
- Se $0 \leq x < 2$, então $(x, 0) \in \text{Gr}(T_K)$.
- Se $x = 0$, então $(s, 0) \in \text{Gr}(T_K) \forall s \in [-2, 0]$.
- Se $x < 0$, então $(x - 2, x) \in \text{Gr}(T_K)$.

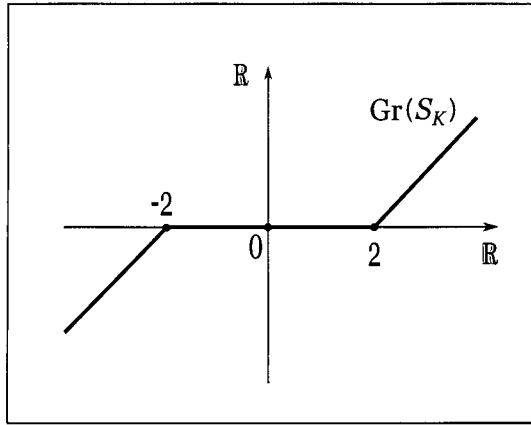


Figura 6.2: Gráfico do operador S_K do Exemplo 6.1.2.

1. Se $x > 0$, observamos que $\text{Gr}(T_K)$ coincide com $\text{Gr}(T)$.
2. Afirmamos $T_K(s) = 0$ para todo $s \in [-2, 0]$. Suponhamos que existe $w \in T_K(s)$ tal que $w \neq 0$, então existem $\hat{x}, \hat{y} \in \mathbb{R}$ tais que $s = \hat{x}_K + \hat{y}_{K^\circ}$, $w = \hat{y}_K + \hat{x}_{K^\circ}$, com $\hat{y} \in T(\hat{x})$.

- Se $\hat{x} > 2$ implica que $T(\hat{x}) > 0$, daqui

$$(s, w) = (\hat{x}_K + (\hat{x} - 2)_{K^\circ}, (\hat{x} - 2)_K + \hat{x}_{K^\circ}) = (\hat{x}, \hat{x} - 2).$$

Então $s > 2$, absurdo.

- Se $0 < \hat{x} \leq 2$ implica que $S(\hat{x}) = 0$, obtemos $(s, w) = (\hat{x}, 0)$, o que é absurdo.
- Se $\hat{x} = 0$ implica que $T(\hat{x}) \subset \mathbb{R}_-$, assim $(s, w) = (t, 0)$ com $t \in [-2, 0]$, implica que $w = 0$, absurdo.
- Se $\hat{x} < 0$ implica que $T(\hat{x}) < 0$, assim $(s, w) = (\hat{x} - 2, \hat{x})$, implica que $0 \leq \hat{x} \leq 2$, absurdo.

Portanto $T_K(s) = 0$ para todo $s \in [-2, 0]$.

3. Afirmamos que se $x < 0$ implica que $T_K(x - 2) = x$. Suponhamos que existe $w \in T_K(x - 2)$ tal que $w \neq x$, então existem $\hat{x}, \hat{y} \in \mathbb{R}$ tais que $x = \hat{x}_K + \hat{y}_{K^\circ}$, $y = \hat{y}_K + \hat{x}_{K^\circ}$, com $\hat{y} \in T(\hat{x})$.

- Se $\hat{x} > 2$ implica que $T(\hat{x}) > 0$, daqui

$$(x - 2, w) = (\hat{x}_K + (\hat{x} - 2)_{K^\circ}, (\hat{x} - 2)_K + \hat{x}_{K^\circ}) = (\hat{x}, \hat{x} - 2).$$

Portanto $x - 2 = \hat{x} > 0$ daqui $x > 2$ absurdo.

- Se $0 < \hat{x} \leq 2$ implica que $T(\hat{x}) = 0$, obtemos $(x - 2, w) = (\hat{x}, 0)$, daqui $0 < x - 2 \leq 2$ portanto $2 < x$, o que é absurdo.
- Se $\hat{x} = 0$ implica que $T(\hat{x}) \subset \mathbb{R}_-$, assim $(x - 2, w) = (t, 0)$ com $t \in [-2, 0]$, implica que $-2 \leq x - 2 \leq 0$, portanto $0 \leq x$, absurdo.
- Se $\hat{x} < 0$ implica que $T(\hat{x}) < 0$, assim $(x - 2, w) = (\hat{x} - 2, \hat{x})$, implica que $w = \hat{x}$, absurdo.

Portanto $T_K(x - 2) = x$ para todo $x < 0$. Fazemos uma mudança de variável e obtemos $T_K(\bar{x}) = \bar{x} + 2$ para todo $\bar{x} < -2$.

Em resumo,

$$T_K(\bar{x}) = \begin{cases} \bar{x} - 2 & \text{se } \bar{x} > 2, \\ 0 & \text{se } \bar{x} \in (0, 2], \\ 0 & \text{se } \bar{x} \in [-2, 0], \\ \bar{x} + 2 & \text{se } \bar{x} < -2. \end{cases}$$

Assim temos que $T_K = S_K$. Portanto temos a afirmação. ■

A importância do operador inversa parcial relativo a subespaços radica em que toda solução do problema (2.1):

$$\text{Encontrar } x \in A, y \in A^\perp \text{ tal que } y \in T(x)$$

pode ser obtido a partir de uma solução do problema de encontrar um zero do operador inversa parcial. Este fato foi estabelecido no Teorema 3.1.1.

Quando substituimos o subespaço por um cone, perdemos essa relação fundamental. Neste caso, verificamos a seguinte propriedade, que é mais fraca.

Teorema 6.1.1 *Sejam $T : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ um operador e K um cone, convexo e fechado em \mathcal{H} . Se $z_{K^\circ} \in T(z_K)$ então $0 \in T_K(z)$.*

Prova: Diretamente

$$\begin{aligned} (z_K, z_{K^\circ}) \in \text{Gr}(T) &\implies (z_K + z_{K^\circ}, 0 + 0) \in \text{Gr}(T_K) \\ &\implies 0 \in T_K(z). \end{aligned} \quad \blacksquare$$

A recíproca do teorema anterior não é verdadeira como mostramos no seguinte exemplo.

Exemplo 6.1.3 Seja $T : \mathbb{R} \longrightarrow \mathcal{P}(\mathbb{R})$ o operador definido por

$$T(x) = \begin{cases} 2(x - 2) & \text{se } x < 2, \\ [0, 1] & \text{se } x = 2, \\ 1 & \text{se } x > 2. \end{cases}$$

Consideramos $K = \mathbb{R}_+$ e seu respectivo cone polar $K^\circ = -\mathbb{R}_+$. temos que

$0 \in T_K(-1)$ pois existe $x = 1 \in K$, $y = T(x) = -2 \in K^\circ$ tal que

$$-1 = x_K + y_{K^\circ} = 1 - 2,$$

$$0 = y_K + x_{K^\circ}.$$

Seja $z = -1$. Então, $z_K = 0$, $z_{K^\circ} = -1$. Vimos que $0 \in T_K(z)$. Não entanto, $(z_K, z_{K^\circ}) = (0, -1) \notin \text{Gr}(T)$ pois $T(0) = -4$. ■

Pelo anterior, inferimos que não é possível encontrar uma solução do problema (6.1) através de zeros do operador inversa parcial generalizada e portanto não é possível a extensão do método das inversas parciais para cones.

Também temos o seguinte resultado relacionando zeros de um operador com a sua inversa parcial generalizada.

Proposição 6.1.1 *Seja $T : \mathcal{H} \longrightarrow \mathcal{P}(\mathcal{H})$. Se $\bar{x} \in T^{-1}(0)$ então $\bar{x}_{K^\circ} \in T_K(\bar{x}_K)$.*

Prova: É consequência imediata da definição de $\text{Gr}(T_K)$. ■

Observamos que a recíproca da Proposição 6.1.1 é falsa, a menos que K seja um subespaço.

Exemplo 6.1.4 Seja $f : \mathbb{R} \longrightarrow \mathbb{R}$ uma função definida por

$$f(x) = \begin{cases} (x - 2)^2 + 4 & \text{se } x \leq 2, \\ x + 2 & \text{se } x > 2. \end{cases}$$

Como f é convexa, própria e semicontínua inferior temos que ∂f é monótono maximal (Teorema 1.3.5). Além disso, ∂f é o operador

$$\partial f(x) = \begin{cases} 2(x-2) & \text{se } x < 2, \\ [0, 1] & \text{se } x = 2, \\ 1 & \text{se } x > 2. \end{cases}$$

Por outro lado, consideramos o cone $K = \mathbb{R}_+$, temos

$$(\partial f)_K = T_K.$$

Assim se $\bar{x} = 1 \in K$, verificamos que $0 = (\partial f)_K(1)$ pois $(\frac{5}{3}, -\frac{2}{3})$ verifica $-\frac{2}{3} \in \partial f(\frac{5}{3})$,

$$1 = x_K + y_{K^\circ} = -\frac{2}{3} + \frac{5}{3},$$

$$0 = y_K + x_{K^\circ} = 0 + 0,$$

mas isso não implica que $0 \in \partial f(1)$ pois $\partial f(1) = \{-2\}$. ■

6.2 Propriedades de T não preservadas

Em geral, quando K é um cone com vértice na origem, convexo fechado de \mathcal{H} , o operador T_K não preserva as propriedades, nem de monotonicidade nem de pseudomonotonicidade do operador original T , pois a projeção sobre um cone não é linear em geral. A seguir, apresentamos dois exemplos, um para monotonicidade e um para pseudomonotonicidade. Nesta seção consideramos $\mathcal{H} = \mathbb{R}^2$, $K = \mathbb{R}_+^2$ e portanto $K^\circ = -K$.

Exemplo 6.2.1 Consideramos um operador monótono tal que sua inversa parcial generalizada não verifica essa condição. Seja o operador linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por:

$$T(x) = (0.43x_1 - 0.80x_2, 0.80x_1 + 0.23x_2)^\top.$$

Este operador está definido pela matriz

$$T = \begin{pmatrix} 0.43 & -0.80 \\ 0.80 & 0.23 \end{pmatrix},$$

que é definida positiva. Com efeito,

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0.43 & -0.80 \\ 0.80 & 0.23 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.43x_1^2 + 0.23x_2^2.$$

Como conseqüência da Propriedade 1.3.2, resulta que o operador T é monótono maximal. Vemos que $\text{Gr}(T_K)$ não define o gráfico de um operador monótono. De fato, consideramos os pontos $x = (-0.75, 0.32)^\top$ e $y = (-0.27, 0.68)^\top$, então $T(x) = (-0.5785, -0.5264)^\top$ e $T(y) = (-0.6601, -0.0596)^\top$.

Calculamos a decomposição de Moreau dos pontos x , $T(x)$, y e $T(y)$:

$$x = x_K + x_{K^\circ} = (0, 0.32)^\top + (-0.75, 0)^\top,$$

$$T(x) = (T(x))_K + (T(x))_{K^\circ} = (0, 0)^\top + (-0.5785, -0.5264)^\top,$$

$$y = y_K + y_{K^\circ} = (0, 0.68)^\top + (-0.27, 0)^\top,$$

$$T(y) = (T(y))_K + (T(y))_{K^\circ} = (0, 0)^\top + (-0.6601, -0.0596)^\top.$$

Logo, temos que $\bar{x} = x_K + (T(x))_{K^\circ} = (-0.5785, -0.2064)^\top$, $\bar{v} = (T(x))_K + x_{K^\circ} = (-0.75, 0)^\top \in T_K(\bar{x})$ e $\bar{y} = y_K + (T(y))_{K^\circ} = (-0.6601, 0.6204)^\top$, $\bar{w} = (T(y))_K + y_{K^\circ} = (-0.27, 0)^\top \in T_K(\bar{y})$. Resulta,

$$\begin{aligned} \langle \bar{w} - \bar{v}, \bar{y} - \bar{x} \rangle &= \left\langle (0.48, 0)^\top, (-0.0816, 0.8268)^\top \right\rangle \\ &= -0.48(0.0816) < 0. \end{aligned}$$

Portanto, T_K não é monótono. ■

Embora existam exemplos onde o operador inversa parcial generalizado preserva a monotonicidade e a maximalidade do operador original, (o Exemplo 6.1.3 é monótono por ser não decrescente e maximal porque $\text{Im}(T) = \mathcal{H} = \mathbb{R}$), estes não necessariamente preservam a maximalidade deste. Em Campos e Scheimberg [9] é apresentado um exemplo de um operador monótono maximal tal que o operador inversa parcial generalizado preserva a monotonicidade, mas não é maximal.

Em [8] encontra-se um exemplo onde o operador inversa parcial generalizado não mantém a monotonicidade do operador original, para isto considerou-se um operador em \mathbb{R}^2 definido por uma matriz simétrica e definida positiva.

Como no caso dos subespaços, verificamos que o operador inversa parcial generalizado, não preserva necessariamente a pseudomonotonicidade do operador original T . A seguir ilustramos esta situação

Exemplo 6.2.2 Definamos o seguinte operador linear $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dado por:

$$T(x) = (4x_1 - 2x_2, -2x_1 + x_2)^\top$$

Este operador está definido pela matriz

$$T = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix},$$

que é semidefinida positiva. Em efeito,

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (2x_1 - x_2)^2.$$

Como conseqüência, resulta que o operador T é monótono e, portanto, T é pseudomonótono (Proposição 1.4.1). Afirmamos que o operador T_K não é pseudomonótono. De fato, seja $x_1 > 0$ fixo e consideramos os pontos $x = (x_1, 2x_1)^\top$ e $x' = (-1, 0)^\top$ então $T(x) = (0, 0)^\top$ e $T(x') = (-4, 2)^\top$.

Calculamos a decomposição de Moreau dos pontos x , $T(x)$, x' e $T(x')$:

$$\begin{aligned} x &= x_K + x_{K^\circ} = (x_1, 2x_1)^\top + (0, 0)^\top, \\ T(x) &= (T(x))_K + (T(x))_{K^\circ} = (0, 0)^\top + (0, 0)^\top, \\ x' &= x'_K + x'_{K^\circ} = (0, 0)^\top + (-1, 0)^\top, \\ T(x') &= (T(x'))_K + (T(x'))_{K^\circ} = (0, 2)^\top + (-4, 0)^\top. \end{aligned}$$

Logo, temos que $\bar{x} = x_K + (T(x))_{K^\circ} = (x_1, 2x_1)^\top$, $\bar{v} = (T(x))_K + x_{K^\circ} = (0, 0)^\top \in T_K(\bar{x})$ e $\bar{y} = y_K + (T(y))_{K^\circ} = (-4, 0)^\top$, $\bar{w} = (T(y))_K + y_{K^\circ} = (-1, 2)^\top \in T_K(\bar{y})$.

Resulta,

$$\begin{aligned} \langle \bar{v}, \bar{y} - \bar{x} \rangle &= \langle (0, 0)^\top, (-4 - x_1, -2x_1)^\top \rangle \\ &= 0 \end{aligned}$$

e

$$\begin{aligned} \langle \bar{w}, \bar{y} - \bar{x} \rangle &= \langle (-1, 2)^\top, (-4 - x_1, -2x_1)^\top \rangle \\ &= 4 - 3x_1, \end{aligned}$$

então, basta escolher $x_1 > 4/3$. Portanto, T_K não é pseudomonótono. ■

Observação 6.2.1 *O Exemplo 3.1.3 ilustra o caso em que o operador inversa parcial de T com respeito ao subespaço A , não preserva a pseudomonotonicidade do operador T .*

Conclusões

Como consequência da análise dos operadores inversa parcial e inversa parcial generalizado chegamos às seguintes conclusões.

I Operador inversa parcial com respeito a um subespaço:

- (a) Não preserva necessariamente todas as características do operador que o origina, como por exemplo, ser uma função, monotonicidade forte ou pseudomonotonicidade. (Exemplos 3.1.1, 3.1.2 e 3.1.3).
- (b) Se dois operadores tem o mesmo operador inversa parcial, então tais operadores são necessariamente iguais. (Proposição 3.1.2).

II Operador inversa parcial generalizado:

- (a) Estendemos o Lema 3.1.1 relativo à inversa parcial, isto é, $T_K = (T^{-1})_{K^\circ} = (T_{K^\circ})^{-1}$. (Propriedade 6.1.1).
- (b) Mostramos que dois operadores diferentes podem gerar a mesma inversa parcial generalizada. (Exemplo 6.1.2).
- (c) Verificamos que o Teorema 3.1.1 que é fundamental para caracterizar o problema primal-dual (2.1), não pode ser generalizado para cones. (Exemplo 6.1.3).
- (d) Não preserva necessariamente a propriedade de monotonicidade do operador original. (Exemplo 6.2.1).
- (e) Como no caso do operador inversa parcial, as propriedades de ser uma função, de monotonicidade forte e de pseudomonotonicidade não são preservadas. (Exemplos 6.1.1, 6.2.1 e 6.2.2).

Estas constatações impossibilitam resolver numericamente o problema $P(T, K)$, em particular o problema de otimização (6.2), através de métodos que encontram

zeros do operador inversa parcial generalizado. Isto é, a filosofia usada para resolver o problema $P(T, A)$ usando o conceito de inversa parcial não é aplicável quando o subespaço A e seu ortogonal A^\perp são substituídos por um cone convexo fechado e seu polar K° .

Um motivo pelo qual as boas propriedades do operador inversa parcial não são mantidas na sua generalização deve-se ao fato da não linearidade do operador projeção (usado por exemplo na monotonicidade). Um outro motivo que observamos é que embora a decomposição de Moreau seja única, $x = x_K + x_{K^\circ}$ onde $x_K = \text{proj}_K(x)$ e $x_{K^\circ} = \text{proj}_{K^\circ}(x)$ (e portanto $\langle x_K, x_{K^\circ} \rangle = 0$), tem-se outras decomposições do tipo, $x = x_1 + x_2$ com $x_1 \in K$ e $x_2 \in K^\circ$ mas $\langle x_1, x_2 \rangle \neq 0$, (usado no Exemplo 6.1.3).

Referências Bibliográficas

- [1] AUBIN, J.P., EKELAND, H., *Applied Nonlinear Analysis*. 1 ed. New York, John Wiley & Sons, 1984.
- [2] AUBIN, J.P., FRANKOWSKA, H., *Set-Valued Analysis*. 1 ed. Boston, Birkhäuser, 1990.
- [3] BAUSCHKE, H.H., *Projection Algorithms and Monotone Operators*. Ph.D. dissertation, Simon Fraser University, British Columbia, Canadá, 1996.
- [4] BRÉZIS, H., *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*. 1 ed. Amsterdam, North-Holland Publishing Company, 1973.
- [5] BURACHIK, R.S., IUSEM, A.N. "A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Space", *SIAM Optimization* v. 8, n. 1, pp. 197-216, 1998.
- [6] BURACHIK, R.S., SAGASTIZÁBAL, C., SCHEIMBERG, S. "An Inexact Method of Partial Inverses and a Parallel Bundle Method", *Optimization Methods and Software*, Published Online 2005.
- [7] CIORANESCU, I., *Geometry of Banach Spaces, Duality Mappings and Non-linear Problems*, v. 62, *Mathematics and Its Applications*, Kluwer Academic Publishers, 1990.
- [8] CAMPOS, C., *Sobre a Generalização do Método de Inversas Parciais para um Cone Convexo Fechado*. Dissertação de M.Sc., IM-NCE/UFRJ, Rio de Janeiro, RJ, 2004.

- [9] CAMPOS, M.H.J., MAKLER, S.S., *A Restriction on Generalizing the Orthogonal Decomposition on a Cone*. Trabalhos de Matemática n. 285, Universidade de Brasília, Brasília, 1994.
- [10] ECKSTEIN, J., *Splitting Methods for Monotone Operators with Applications to Parallel Optimization*. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 1989.
- [11] ECKSTEIN, J., BERTSEKAS, D. "On the Douglas-Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators", *Mathematical Programming* v. 55, n. 3, pp. 293-318, 1992.
- [12] EKELAND, I., TEMAM, R., *Convex Analysis and Variational Problems* . 1 ed. Amsterdam, North-Holland Publishing Company, 1976.
- [13] GÜLER, O., "On the Convergence of the Proximal Point Algorithm for Convex Minimization", *SIAM Control and Optimization* v.29, n. 2, pp. 403-419, 1991.
- [14] HADJISAVVAS, N., "Continuity and Maximality Properties of Pseudomonotone Operators", *Journal of Convex Analysis* v. 10, n. 2, pp. 465-475, 2003.
- [15] HADJISAVVAS, N., SCHAIBLE, S. "Quasimonotonicity and pseudomonotonicity in variational inequalities and equilibrium problems". In: *Generalized Convexity, Generalized Monotonicity Recent Results*, v. 27, *Nonconvex Optimization and Its Applications*, Kluwer Academic Publishers, 1998.
- [16] INGRAM, J.M., MARSH, M.M. "Projections onto convex cones in Hilbert space", *Journal of Approximation Theory* v. 64, n. 3, pp. 343-350, 1991.
- [17] ISAC, G., *Topological Methods in Complementarity Theory*, v. 41, *Nonconvex Optimization and Its Applications*, Kluwer Academic Publishers, 2000.
- [18] ISAC, G., COJOCARU, M.G. "Variational Inequalities, Complementarity Problems and Pseudo-monotonicity. Dynamical Aspects". In: *Proceedings of*

the International Conference on Nonlinear Operators, Differential Equations and Applications, v. 3, pp. 41-62, Babes-Bolyai University of Cluj-Napoca, 2002.

- [19] IUSEM, A.N., *Métodos de Ponto Proximal em Otimização*. 1 ed. Rio de Janeiro, IMPA, 1995.
- [20] KARAMARDIAN, S. "Complementarity Problems over Cones with Monotone and Pseudomonotone Maps", *Journal of Optimization Theory and Applications* v. 18, n. 4, pp. 445-454, 1976.
- [21] KINDERLEHRER, D., STAMPACCHIA, G., *An Introduction to Variational Inequalities and Their Applications*. 1 ed. New York, Academic Press, 1980.
- [22] LEMAIRE, B. "The Proximal Algorithm", *Boletim da SBMAC* v. 1, n. 2, pp. 1-16, 1989.
- [23] LEMAIRE, B. "About the Convergence of the Proximal Method". In *Lecture Notes in Economics and Mathematical Systems*, n. 382, pp. 39-51, Springer-Verlag, 1992.
- [24] LEVITIN, E.S., POLJAK, B.T. "Convergence of Minimizing Sequences in Conditional Extremum Problems", *Doklady* v. 168, n. 5, 1966.
- [25] LIONS, P., MERCIER, B. "Splitting Algorithms for the Sum of Two Nonlinear Operators", *SIAM Journal on Numerical Analysis* v. 16, n. 6, pp. 964-979, 1979.
- [26] LUEMBERGER, D.G., *Optimization by Vector Space Methods*. 1 ed. New York, John Wiley & Sons, 1969.
- [27] MARTINET, B., "Regularisation d'Inéquations Variationnelles par Approximations Successives", *Revue Française d'Informatique Recherche Opérationnelle* v. 4, S. R-3, pp. 154-158, 1970.
- [28] MARTINET, B., "Détermination Approchée d'un Point Fixe d'une Application Pseudo-contractante. Cas de l'application prox", *Comptes Rendus de l'Académie des Sciences de Paris* v. 274, pp. 163-165, 1972.

- [29] MOREAU, J.J. "Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires", *Comptes Rendus de l'Académie des Sciences de Paris* v. 255, pp. 238-240, 1962.
- [30] MOREAU, J.J. "Proximité et Dualité dans un espace Hilbertien", *Bulletin de la Société Mathématique de France* v. 93, pp. 273-299, 1965.
- [31] OPIAL, Z. "Weak Convergence of the Sequence of Successive Approximations for Nonexpansive Mappings", *Bulletin of the American Mathematical Society* v. 73, pp. 591-597, 1967.
- [32] PASCALI, D., SBURLAN, S., *Nonlinear Mappings of Monotone Type*. 1 ed. Bucuresti, Editura Academiei, 1978.
- [33] PHELPS, R.R. "Lectures on Maximal Monotone Operators", *Extracta Mathematicae* v. 12, n. 3, pp. 193-230, 1997.
- [34] PHELPS, R.R., SIMONS, S. "Unbounded Linear Monotone Operators on Nonreflexive Banach Spaces", *Journal of Convex Analysis* v. 5, n. 2, pp. 303-328, 1998.
- [35] POLYAK, B.T., *Introduction to Optimization*. 1 ed. New York, Optimization Software Inc., 1987.
- [36] ROCKAFELLAR, R.T., *Convex Analysis*. 1 ed. Princeton, Princeton University Press, 1970.
- [37] ROCKAFELLAR, R.T. "On the Maximality of Sum of nonlinear monotone operators", *Transactions of the American Mathematical Society* v. 149, n. 1, pp. 75-88, 1970.
- [38] ROCKAFELLAR, R.T. "On the Maximal Monotonicity of Subdifferential Mappings", *Pacific Journal of Mathematics* v. 33, n. 1, pp. 209-216, 1970.
- [39] ROCKAFELLAR, R.T., "Monotone Operators and the Proximal Point Algorithm", *SIAM Control and Optimization* v.14, n. 5, pp. 877-808, 1976.
- [40] ROCKAFELLAR, R.T., WETS, R., *Variational Analysis*. 1 ed. Berlin, Springer-Verlag, 1998.

- [41] SOLODOV, M.V., SVAITER, B.F. "A Unified Framework for Some Inexact Proximal Point Algorithms", *Numerical Functional Analysis and Optimization* v. 22, n. 7-8, pp. 1013-1035, 2001.
- [42] SPINGARN, J.E., "Partial Inverse of a Monotone Operator", *Applied Mathematics and Optimization* v. 10, n. 1, pp. 247-265, 1983.
- [43] SPINGARN, J.E., "Applications of the Method of Partial Inverses to Convex Programming: Decomposition", *Mathematical Programming* v. 32, n. 2, pp. 199-223, 1985.
- [44] VAN TIEL, J., *Convex Analysis: An introductory text*. 1 ed. New York, John Wiley & Sons, 1984.
- [45] WIERZBICKI, A., KURCYUSZ, S., "Projection on a Cone, Penalty Functionals and Duality Theory for Problems with Inequality Constraints in Hilbert Space", *SIAM Control and Optimization* v. 15, n. 1, pp. 25-56, 1977.
- [46] ZARANTONELLO, E.H., "Projections on Convex Sets in Hilbert Space and Spectral Theory", In Zarantonello, E.H. (ed), *Contributions to Nonlinear Functional Analysis*, 1 ed., New York, USA, Academic Press, 1971.
- [47] ZHU, C., "Asymptotic Convergence Analysis of the Forward-Backward Splitting Algorithm", *Mathematics of Operations Research* v. 20, n. 2, pp. 449-464, 1995.