Universidade Federal do Piauí

Centro de Ciências da Natureza

Pós-Graduação em Matemática

Mestrado em Matemática

Algoritmo do Ponto Proximal Generalizado em Espaços de Hilbert para o Problema de Desigualdade Variacional

Francisco Gilberto de Sousa Carvalho

Teresina - 2010

Francisco Gilberto de Sousa Carvalho

Dissertação de Mestrado:

Algoritmo do Ponto Proximal Generalizado em Espaços de Hilbert para o Problema de Desigualdade Variacional

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:
Prof. Dr. Jurandir de Oliveira Lopes

Teresina - 2010

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da UFPI
Biblioteca Comunitária Jornalista Carlos Castello Branco

C331a

Carvalho, Francisco Gilberto de Sousa.
Algoritmo do ponto proximal generalizado em espaços de Hilbert para o problema de desigualdade variacional / Francisco Gilberto de Sousa Carvalho. Teresina: 2010.
63 f.

Orientador: Prof. Dr. Jurandir de Oliveira Lopes

Dissertação (Mestrado em Matemática)–Universidade Federal do Piauí.

1. Algoritmo. 2. Ponto proximal. I. Título.

CDD: 005.1

Dedico este trabalho a meus pais, a minha esposa Helenita, aos meus filhos Vinícius e Viliane e a minha avó Ester (in memorian).

Agradecimentos

Agradeço primeiramente a Deus pelo dom da vida e por ter dado forças nos momentos mais difíceis.

Aos meus pais que sempre me incentivaram, me apoiaram e fizeram com que eu continuasse nesse caminho em busca de novas descobertas e de novos conhecimentos. Aos meus irmãos, tios, tias e primos que acreditaram em mim.

A minha esposa Helenita que também sempre me incentivou e me apoiou durante esta jornada, tendo paciência e sempre acreditando que juntos venceremos.

A minha filha Viliane que sentiu muito a minha ausência. Agradeço também ao meu filho Vinícius que sempre agiu com naturalidade mas que às vezes reclamava da minha falta.

A todos os meus amigos especialmente os alunos do mestrado, pelo companheirismo, amizade e solidariedade.

Aos professores do Departamento de Matemática da UFPI, especialmente aos professores: Marcondes que foi um dos primeiros a me incentivar a fazer o mestrado, professor João Xavier pelo apoio, pelas sugestões e correções, a professora Sissy pelas sugestões e o professor Jurandir que me orientou no desenvolvimento deste trabalho. Sempre paciente e compreensivo.

Ao professor Rolando Garciga Otero pelas correções e sugestões.

A todos os professores, coordenadores e funcionários da UESPI, campus Professor Possidônio Queiroz, da cidade de Oeiras- Pi, em especial a professora Francisca Inez, pelo apoio, pelos conselhos e pelo exemplo que é como pessoa e como profissional. Muito obrigado pois sem o seu apoio durante estes anos dificilmente teria conseguido mais uma conquista.

A FAPEPI pelo apoio financeiro.

A todos que contribuiram direta ou indiretamante para a realização deste trabalho.

"*O pensamento pode ter elevação sem ter elegância, e, na medida que não tiver elegância, perderá a ação sobre os outros. A força sem a destreza é uma simples massa* ".

Fernando Pessoa.

Resumo

Consideramos o Algoritmo o Ponto Proximal Generalizado APPG para resolver Problemas de Desigualdade Variacional PDV(T,C) relativo ao operador monótono maximal T no conjunto C. Ele difere do Algoritmo do Ponto Proximal APP no uso de distâncias de Bregman generalizadas onde a distância euclidiana é substituída por esta "distância". Esta "distância"faz com que a sequência gerada pelo APPG esteja bem definida, isto é, cada iterada existe e é única e permaneçe no interior do conjunto viável. Sob hipóteses adequadas aplicadas à distância de Bregman e nos operadores monótonos maximais, provaremos que a sequência converge fracamente se, e somente se, o PDV(T,C) tem soluções, neste caso o limite fraco é uma solução. Se o problema não tem soluções a sequência é não limitada.

Abstract

We consider the Generalized Proximal Point Algorithm (GPPA) for solving Variational Inequality Problems VIP(T,C) for a maximal monotone operator T an a set C. It differs from the Proximal Point Algorithm (PPA) in the use of generalized Bregman distances in dead of Euclidean distance. This "distance" allows us to prove that the sequence generated by the GPPA is well defined, i.e. each iterate exists, is unique and remains inside the feasible set. Under appropriate assumptions a the Bregman distance ad as the monotone operator T we prove that the sequence converges weakly if and only if, the VIP(T,C) has solutions. In this case, the weak aculation point is a solution. If the problem has no solution the sequence is unbounded.

Sumário

Introdução

Nosso principal objetivo neste trabalho é estudar Algoritmos de Pontos Proximais em espaços de Hilbert para resolver o Problema de Desigualdade Variacional, o qual consiste em determinar $\bar{x} \in C$ tal que existe $\bar{u} \in T(\bar{x})$ satisfazendo

$$\langle \bar{u}, x - \bar{x} \rangle \geqslant 0, \ \forall \ x \in C, \tag{1}$$

onde C è um subconjunto convexo e fechado de um espaço de Hilbert H e $T : H \rightrightarrows H$ é um operador monótono maximal. Denotamos por S^* o conjunto das soluções do Problema de Desigualdade Variacional relativo ao operador T no conjunto C e por simplicidade, a partir de agora, dizemos que S^* é o conjunto solução do PDV(T,C).

O Algoritmo do Ponto Proximal APP, é basicamente o método das aproximações sucessivas para determinar zeros de operadores monótonos maximais em espaços de Hilbert, o qual gera uma sequência $\{x^k\} \subset H$ da seguinte forma: começamos com qualquer elemento $x^0 \in H$ e, dado x^k, o elemento x^{k+1} é tomado de tal forma que $0 \in T_k(x^{k+1})$ onde $T_k(x) = T(x) + \lambda_k(x - x^k)$ e $\lambda_k \in (0, \bar{\lambda})$ para algum $\bar{\lambda} > 0$.

Mostra-se que a sequência $\{x^k\}$ converge fracamente para um zero de T, desde que o conjunto dos zeros seja não vazio.

Como podemos observar o APP é usado para resolver o problema de determinar zeros de operadores monótonos maximais, o qual chamamos de problema "irrestrito"associado a operadores monótonos maximais. Neste trabalho vamos mostrar um algoritmo semelhante ao APP com todas as suas vantagens, mas adequado ao problema restrito, ou seja, o PDV(T,C). No método do Ponto Proximal cada subproblema envolve uma distância quadrática. Mas no algoritmo que vamos considerar, esta distância quadrática é substituída por uma distância adaptada ao conjunto C, chamada distância de Bregman D_g. Esta distância será definida apenas na restrição ao conjunto C, de tal forma que a distância entre as iteradas consecutivas vai para "zero"e a sequência gerada pelo APPG converge fracamente para a solução do problema. Quando $S^* = \emptyset$ a sequência gerada pelo APPG é não limitada.

Mas antes de mostrarmos o APPG faremos uma breve revisão sobre alguns resultados básicos que serão de grande importância para alcançarmos nosso objetivo.

Como vamos trabalhar em espaços de Hilbert começamos, no primeiro capítulo, com algumas noções sobre convergência fraca e forte e enunciamos, sem demonstração, o Teorema de Bourbaki-Alaouglu, o qual será importante para mostrarmos que a sequência gerada pelo APPG tem pontos de acumulação. Mostraremos também alguns resultados sobre otimização convexa onde definimos conjuntos convexos, funções convexas e o subdiferencial de funções convexas, conceitos bastante utilizados no decorrer de nosso trabalho. Já no segundo capítulo introduzimos as noções de função e distância de Bregman, como já comentamos acima. Neste capítulo vamos mostrar que a distância de Bregman é estritamente convexa e que o seu subdiferencial é vazio para um elemento $x \notin \overline{S}$ e é igual a $\nabla g(x) - \nabla g(y)$ para $x \in S$, onde S é um subconjunto aberto convexo de H e é chamado de zona da função de Bregman g a qual está associada com a distância de Bregman.

No capítulo seguinte trataremos dos operadores monótonos, monótonos maximais, paramonótonos e pseudomonótonos. Demonstraremos alguns resultados e outros apenas enunciaremos sem demonstração para não perdermos o foco do nosso objetivo, que é de trabalhar o APPG. No quarto capítulo definiremos o APPG e apresentaremos os principais resultados de existência e de convergência do APPG. Deixamos como apêndice o último capítulo do nosso trabalho onde definimos a função indicadora e a função gap que relacionadas com alguns resultados de existência e de convergência tratado sno capítulo anterior.

Capítulo 1

Noções Preliminares

Neste capítulo, apresentaremos algumas ferramentas e notações que serão utilizadas nos demais capítulos desta dissertação.

Serão introduzidos conceitos matemáticos básicos, como noções de conjuntos convexos e de funções convexas, sendo muito importantes em otimização, pois estão relacionados com o conceito de mínimo global. Introduzimos também o conceito de convergência fraca, que será de fundamental importância nesse trabalho, principalmente para garantir a convergência da sequência gerada pelo Algoritmo do Ponto Proximal para espaços de Hilbert.

1.1 Elementos de Análise Funcional

sejam X, Y dois espaços lineares normados. Suponha que exista uma aplicação bijetiva de X em Y. Então dizemos que que X e Y são linearmente isomórficos e T é um isomorfismo linear.

Considere agora o conjugado de X^*, que será denotado por X^{**}. Para cada $x \in X$ corresponde um elemento $\hat{x} \in X^{**}$ tal que $\hat{x}(x^*) = x^*(x)\ \forall\ x^* \in X^*$. Escrevemos $\hat{x} = kx$ e chamamos k a imersão natural de X em X^{**}. Escrevemos também $\hat{X} = k(X)$.

Definição 1.1. *Um espaço de Bannach é reflexivo quando* $k(X) = X^{**}$.

Nosso trabalho vai ser realizado em espaços de Hilbert H, que são, em particular espaços de Bannach com a norma usual $\|\cdot\|$. Abaixo mostraremos alguns resultados em espaços de Hilbert. Começaremos com as noções de convergência forte e fraca.

Definição 1.2. *Uma sequência* $\{x^k\} \subset H$ *é fortemente convergente para* $x \in H$ *(*$x^k \to x$*) se, e somente se,*

$$0 = \lim_{k\to\infty} ||x^k - x||^2 = \lim_{k\to\infty} \langle x^k - x, x^k - x\rangle.$$

A topologia forte é também conhecida como topologia da norma.

Definição 1.3. *Uma sequência* $\{x^k\} \subset H$ *é fracamente convergente para* $x \in H$ *(*$x^k \rightharpoonup x$*) se, e somente se,*

$$0 = \lim_{k\to\infty} \langle x^k - x, y\rangle \ \forall \ y \in H.$$

A convergência forte $x^k \to x$ implica trivialmente a convergência fraca $x^k \rightharpoonup x$. De fato, dado $\varepsilon > 0$, seja $n_0 \in \mathbb{N}$ tal que $||x^k - x|| < \frac{\varepsilon}{||y||}$. Então, por (cauchy-Schwarz) temos que $|\langle x^k - x, y\rangle| \leqslant ||x^k - x||||y|| \leqslant \varepsilon \ \forall \ k \geqslant n_0$.
Mas a recíproca da afirmação acima não é verdadeira, pois seja

$$H = l_2 = \{x = \{x^k\};\ \sum_{k=1}^{\infty}(x^k)^2 < \infty\} \text{ com } \langle x, y\rangle = \sum_{k=1}^{\infty} x^k y^k$$

Tome $\{e^k\} \subset H$ com $e^k_n = \delta_{kn}$ (delta de Kronecker). Então $e^k \rightharpoonup 0$ pois $\lim_{k\to\infty}\langle e^k - 0, y\rangle = \lim_{k\to\infty}\langle e^k, y\rangle = \lim_{k\to\infty} y^k = 0$, onde usamos o fato de que $\sum_{k=1}^{\infty}(y^k)^2 < \infty$. Note que $||e^k|| = 1 \ \forall \ k$ e para todo $k \neq j$ temos que $\langle e^k, e^j\rangle = 0$ e $||e^k - e^j|| = \sqrt{2}$. Daí, temos que a sequência $\{e^k\}$ é fracamente convergente para zero mas não é fortemente convergente.
Com base nas definições acima, dizemos que um conjunto $C \subset H$ é chamado fracamente sequencialmente compacto se toda sequência $\{x^k\} \subset C$ contém uma susequ encia que converge fracamente para um ponto de C. E que o conjunto $C \subset H$ é fracamente fechado se o limite fraco de toda sequência $\{x^k\} \subset C$ fracamente convergente também está em C.

Corolário 1.1.1. *Todo espaço de Hilbert* H *é reflexivo. Em particular, um conjunto em* H *é fracamente compacto se, e somente se, ele é fracamente fechado.*

Demonstração. Ver [10]. □

Temos ainda que a bola unitária, em um espaço de Hilbert, é fracamente compacta. Com base nesse resultado temos o seguinte teorema:

Teorema 1.1. *(Bourbaki-Alaoglu). Se* $\{x^k\} \subset H$ *é limitada então* $\{x^k\}$ *tem uma subsequência fracamente convergente.*

Demonstração. Ver [10]. □

1.2 Existência de Soluções

Sejam dados um conjunto $C \subset H$, onde H é um espaço de Hilbert, e uma função $f : C \to \mathbb{R}$. O problema de achar um minimizador de f no conjunto C será escrito como

$$\min f(x) \text{ sujeito a } x \in Ck \tag{1.1}$$

O conjunto C será chamado de conjunto viável do problema e f será chamada função objetivo.

Definição 1.4. *Dizemos que um ponto $\bar{x} \in C$ é*
(i) minimizador global de (1.1), se

$$f(\bar{x}) \leqslant f(x) \ \forall \ x \in C \tag{1.2}$$

(ii) minimizador local de (1.1) se existe uma vizihança V de $\bar{x}$ tal que

$$f(\bar{x}) \leqslant f(x) \ \forall \ x \in C \cap V. \tag{1.3}$$

Se para todo $x \neq \bar{x}$ a desigualdade (1.2) ou (1.3) é estrita, dizemos que $\bar{x}$ será chamado minimizador estrito (global ou local respectivamente).

Antes de mostrarmos alguns resultados sobre a exitência de pontos de mínimo de uma função, apresentaremos algumas definições que serão importantes na demonstração dos resultados.

Definição 1.5. *Dizemos que uma função $f : C \to \mathbb{R}$ é coerciva no conjunto C quando para toda sequência $\{x^k\} \subset C$ e $\|x^k\| \to \infty$ ou $x^k \to x \in \bar{C}$, quando $k \to \infty$, tem-se* $\limsup_{k\to\infty} f(x^k) = \infty$.

Outra definição bastante utilizada em nosso trabalho é o de semicontinuidade inferior. Vejamos a definição:

Definição 1.6. *Dizemos que a função $f : C \to \mathbb{R}$ é fracamente semicontínua inferiormente no ponto $x \in C \subset H$, quando para qualquer sequência $\{x^k\} \subset H$ tal que $x^k \rightharpoonup x$ quando ($k \to \infty$), tem-se*

$$\liminf_{k\to\infty} f(x^k) \geqslant f(x).$$

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

$$\limsup_{k\to\infty} f(x^k) \leqslant f(x).$$

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

Teorema 1.2. . *Sejam* H *um espaço de Hilbert,* $C \subset H$ *um conjunto fracamente compacto e não vazio e* $f : C \to \mathbb{R}$ *uma função fracamente semicontínua inferiormente.*
Então, o problema de minimizar f *em* C *tem uma solução global, ou seja* f *é limitada inferiormente e existe* $x^0 \in C$ *tal que*

$$f(x^0) = \inf_{x\in C} f(x).$$

Demonstração. Podemos escrever $C = \cup_{n=1}^{\infty} f^{-1}(-n, \infty)$. Como cada $f^{-1}(-n, \infty)$ é aberto e C é fracamente compacto, então

$$C = \cup_{n=1}^{n_0} f^{-1}(-n, \infty)$$

para algum $n_0 \in \mathbb{N}$, logo $f(x) > -n_0$ para todo $x \in C$, de onde concluimos que f é limitada inferiormente.
Seja $c = \inf_{x\in C} f(x) > -\infty$, ou seja, $c \leqslant f(x)\ \forall\ x \in C$. Pela definição de ínfimo, existe uma sequência $\{x^k\} \subset C$ tal que $c = \lim_{k\to\infty} f(x^k)$. Como C é fracamente compacto, segue-se que a sequência $\{x^k\}$ é limitada. Logo ela possui uma subsequência x^{k_j} fracamente convergente para um ponto $x^0 \in C$ tal que $c = \lim_{k_j\to\infty} f(x^{k_j}) \geqslant f(x^0)$. Portanto $f(x^0) = c$. □

Teorema 1.3. *Sejam* H *um espaço de Hilbert e* $f : H \to \mathbb{R}$ *uma função fracamente semicontínua inferiormente e coerciva em* H*. Então, o problema de minimizar* f *em* H *possui uma solução global.*

Demonstração. Pela coercividade de f, escolhemos $R > 0$ tal que $f(x) \geqslant f(0)$ para todo $x \in H$ com $||x|| \geqslant R$. Uma vez que a bola fechada $\overline{B_R(0)}$ é fracamente compacta e, pela hipótese de f ser fracamente semicontínua inferiormente, a restrição $f : \overline{B_R(0)} \to \mathbb{R}$ também é semicontínua inferiormente, segue do Teorema 1.2 acima que existe um ponto $x^0 \in \overline{B_R(0)}$ tal que $f(x^0) = \inf_{x\in H} f(x)$, pela escolha de R. □

Uma consequência do Teorema 1.3 é mostrada abaixo.

Teorema 1.4. *Sejam* C *um subconjunto convexo e fracamente fechado de um espaço de Hilbert* H *e* $f : H \to \mathbb{R}$ *uma função fracamente semicontínua inferiormente e coerciva. Então existe* $\bar{x} \in C$ *tal que* $f(\bar{x}) = \inf_{x \in C} f(x)$.

Definição 1.7. *O conjunto de nível da função* $f : C \to \mathbb{R}$ *associado a* $c \in \mathbb{R}$, *é o conjunto dado por*

$$L_{f,C}(c) = \{x \in C \ ; \ f(x) \leqslant c\}.$$

Corolário 1.2.1. *Sejam* $C \subset H$ *e* $f : C \to \mathbb{R}$ *contínua no conjunto* C. *Suponhamos que existe* $c \in \mathbb{R}$ *tal que o conjunto de nível* $L_{f,C}(c)$ *seja não vazio e compacto.*
Então o problema (1.1) admite uma solução global.

1.3 Conjuntos Convexos

Um conjunto convexo se caracteriza por conter todos os segmentos cujos extremos são pontos do conjunto.

Definição 1.8. *Um conjunto* $C \subset H$ *é convexo se o segmento de reta que liga quaisquer dois pontos de* C *também pertencer a* C. *Equivalentemente, o conjunto* $C \subset H$ *é convexo se, e somente se, para todo* $x, y \in C$ *e* $\alpha \in \mathbb{R}$ *com* $0 \leqslant \alpha \leqslant 1$ *tem-se*

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

Um ponto da forma $\alpha_1 x_1 + ... + \alpha_k x_k$, $\alpha_i \geqslant 0$, $i = 1, ..., k$, *com* $\sum_{i=1}^{k} = 1$ *é chamado de combinação convexa dos pontos* $x_1, ..., x_k$.
Em particular, $\alpha x + (1 - \alpha) y$ *é uma combinação convexa de* C.

Alguns exemplos importantes de conjuntos convexos são os subespaços de H, os semi-espaços de H, isto é , os conjuntos da forma $\{x \in H; \langle a, x \rangle \leqslant c\}$ onde $a \in H$ e $c \in \mathbb{R}$ e os hiperplanos em H, ou seja os conjuntos $\{x \in H; \langle a, x \rangle = c\}$ com $a \in H$ $c \in \mathbb{R}$.

Teorema 1.5. *Um conjunto* $C \subset H$ *é convexo se, e somente se, para quaisquer* $p \in \mathbb{N}$, $x^i \in C$ *e* $\alpha \in [0, 1]$, $i = 1, ..., p$, *tais que* $\sum_{i=1}^{p} \alpha_i = 1$, *a combinação convexa* $\sum_{i=1}^{p} \alpha_i x^i \in C$.

Demonstração. ($\Rightarrow$) Suponha que $C \subset H$ seja convexo. A prova é feita por indução sobre $p \in \mathbb{N}$.

Dados $p \in \mathbb{N}$, $x^i \in C$ e $\alpha_i \in [0,1]$ tais que $\sum_{i=1}^{p} \alpha_i = 1$. Defina $x = \sum_{i=1}^{p} \alpha_i x^i$.

i) Se $p = 1$ tem-se que $\alpha_1 = 1$ e $x = 1x^1 \in C$.

ii) Suponha agora que vale para $p = n$. Mostraremos que vale para $p = n+1$.

De fato, temos $\sum_{i=1}^{n+1} \alpha_i = 1$, então $1 - \sum_{i=1}^{n} \alpha_i = \alpha_{n+1}$.

Se $\alpha_{n+1} = 1$, temos que $\alpha_i = 0, \forall\ i = 1, ..., n$.

Logo, $x = \sum_{i=1}^{p} 0.x^i + 1.x^{n+1} = x^{n+1} \in C$.

Se $\alpha_{n+1} \in [0,1)$. Temos, $(1 - \alpha_{n+1}) > 0$. Logo,

$$x = \sum_{i=1}^{n+1} \alpha_i x^i = \sum_{i=1}^{n} \alpha_i x^i + \alpha_{n+1} x^{n+1} = (1 - \alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{(1 - \alpha_{n+1})} x^i + \alpha_{n+1} x^{n+1}. \quad (1.4)$$

Tome $y = \sum_{i=1}^{n} \beta_i x^i$, onde $\beta_i = \frac{\alpha_i}{1-\ \alpha_{n+1}}$, $i = 1, ..., n$.

Como $\sum_{i=1}^{n+1} \alpha_i = 1$, tem-se $1 - \alpha_{n+1} = \sum_{i=1}^{n} \alpha_i$, então

$$\sum_{i=1}^{n} \beta_i = \frac{1}{1-\ \alpha_{n+1}} \sum_{i=1}^{n} \alpha_i = \frac{1}{1-\ \alpha_{n+1}} (1-\ \alpha_{n+1}) = 1$$

Portanto, por hipótese de indução $y \in C$.

Daí, substituindo y em (1.4), obtemos

$$x = (1-\ \alpha_{n+1})y + \alpha_{n+1} x^{n+1}.$$

Segue da convexidade de C que $x = \sum_{i=1}^{n+1} \alpha_i x^i \in C$ Logo, $x = \sum_{i=1}^{p} \alpha_i x^i \in C, \forall\ p \in \mathbb{N}$, com $x^i \in C$ e $\alpha_i \in [0,1]$ tais que $\sum_{i=1}^{p} \alpha_i = 1$.

($\Leftarrow$) Suponha que $\sum_{i=1}^{p} \alpha_i x^i \in C, \forall\ p \in \mathbb{N}$ com $x^i \in C$ e $\alpha_i \in [0,1]$ tais que $\sum_{i=1}^{p} \alpha_i = 1$.

Então, para $p = 2$, temos que

$$\sum_{i=1}^{2} \alpha_i x^i = \alpha_1 x^1 + \alpha_2 x^2 \in C,\ \text{com} x^1, x^2 \in C \text{ e } \alpha_1, \alpha_2 \in [0,1],\ \text{tal que } \alpha_1 + \alpha_2 = 1.$$

Assim, $\alpha_1 = 1 - \alpha_2$, logo $(1 - \alpha_2)x^1 + \alpha_2 x^2 = \alpha_1 x^1 + \alpha_2 x^2 \in C$.

Portanto, pela definição 1.8, C é um conjunto convexo.

□

Proposição 1.1. *Sejam* $C_j \subset H$, $j \in I$, *conjuntos convexos, onde* I *é um conjunto de índices. Então* $C = \cap_{j \in I} C_j$ *é um conjunto convexo.*

Demonstração. Sejam $x, y \in C$. Como C é a interseção dos conjuntos C_j e $x, y \in C$, temos que $x, y \in C_j$ para todo $j \in I$. Como C_j é convexo para todo $j \in I$, então para $\alpha \in [0, 1]$, tem-se $\alpha x + (1 - \alpha)y \in C_j$, para todo $j \in I$, portanto, $\alpha x + (1 - \alpha)y \in C$, logo C é convexo. □

Proposição 1.2. *Seja C um conjunto convexo. Então $\overline{C}$ e C^0 são conjuntos convexos.*

Demonstração. Sejam $x, y \in \overline{C}$ e $\alpha \in [0, 1]$. Então existem sequências $\{x^k\}_{k \in \mathbb{N}}$ e $\{y^k\}_{k \in \mathbb{N}}$ tais que $\lim_{k \to \infty} x^k = x$ e $\lim_{k \to \infty} y^k = y$. Pela convexidade de C temos que $\alpha x^k + (1 - \alpha)y^k \in C$, para todo $k \in \mathbb{R}$. Logo, $\alpha x + (1 - \alpha)y \in C$, pois $\alpha x + (1 - \alpha)y = \lim_{k \to \infty}[\alpha x^k + (1 - \alpha)y^k]$. Assim, $\alpha x + (1 - \alpha)y \in \overline{C}$. Portanto, $\overline{C}$ é convexo.
Sejam agora $x, y \in C^0$. Então existem $\varepsilon_1 > 0$ e $\varepsilon_2 > 0$ tais que $B(x, \varepsilon_1) \subset C$ e $B(y, \varepsilon_2) \subset C$. Considerando $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, tem-se que $B(x, \varepsilon) \subset C$ e $B(y, \varepsilon) \subset C$. Para monstrar que $\alpha x + (1 - \alpha y) \in C^0$, basta verificar que $B(\alpha x + (1 - \alpha)y, \varepsilon) \subset C$.
De fato,
seja $z \in B(\alpha x + (1 - \alpha)y, \varepsilon)$. Então $\|z - \alpha x - (1 - \alpha)y\| < \varepsilon$. Se $q = z - \alpha x - (1 - \alpha)y$, temos que $z = \alpha x + (1 - \alpha)y + q$ e $\|q\| < \varepsilon$. Logo, $z = \alpha x + \alpha q + (1 - \alpha)y + q - \alpha q = \alpha(x + q) + (1 - \alpha)(y + q)$. Como $\|x + q - x\| = \|q\| < \varepsilon$, temos que $x + q \in B(x, \varepsilon) \subset C$, da mesma forma, $\|y + q - y\| = \|q\| < \varepsilon$, então $y + q \in B(y, \varepsilon) \subset C$. Pela convexidade de C, $z = \alpha(x + q) + (1 - \alpha)(y + q) \in C$. Assim, $z \in C$ e tem-se $B(\alpha x + (1 - \alpha)y, \varepsilon) \subset C$. Logo, $\alpha x + (1 - \alpha)y \in C^0$. Portanto, C^0 é convexo. □

Definição 1.9. *Seja $C \subset H$ um conjunto qualquer. O fecho convexo de C, denotado por conv (C), é o menor conjunto convexo em H que contém C (ou, equivalentemente, a interseção de todos os conjuntos convexos de H que contém C).*

Como se trata de uma interseção de conjuntos convexos, temos, pela Proposição 1.1, que conv (C) é um conjunto convexo para qualquer $C \in H$. Além disso, se C é um conjunto convexo, então conv (C)=C.

Definição 1.10. *Um conjunto $K \subset H$ chama-se cone quando*

$$d \in K \Rightarrow td \in K, \forall\ t > 0.$$

Em adição, se K é convexo então K é um cone convexo.

Intuitivamente um cone é um conjunto de direções.

Definição 1.11. *Sejam* $C \subset H$ *um conjunto convexo e* $\bar{x} \in C$. *O cone normal (cone de direções normais) no ponto* $\bar{x}$ *em relação ao conjunto* C *é dado por*

$$N_C = \begin{cases} \{d \in H; \langle d, x - \bar{x} \rangle \leqslant 0 \ \forall \ x \in C\}, & \textit{se} \quad \bar{x} \in C \\ \emptyset, & \textit{se} \quad \bar{x} \notin C \end{cases}$$

Outra definição importante em Otimização é a de cone dual, que daremos a seguir.

Definição 1.12. *O cone dual de um cone* $K \subset H$ *é definido por*

$$K^* = \{y \in H; \ \langle y, d \rangle \leqslant 0 \ \forall \ d \ \in K\}$$

Definição 1.13. *(Direção Tangente) Dizemos que* $d \in H$ *é uma direção tangente em relação ao conjunto* C *no ponto* $\bar{x} \in C$ *quando*

$$\mathrm{dist}(\bar{x} + td, C) = o(t), \ t > 0.$$

Denotamos por $T_C(\bar{x})$ *o cone de todas as direções tangentes ao conjunto* C *no ponto* $\bar{x}$ *ou simplesmente cone tangente.*

Proposição 1.3. *Seja* $\emptyset \neq K \subset H$ *um cone qualquer. Tem-se que*

$$(K^*)^* = \overline{\mathrm{conv}\ (K)}.$$

Em particular, se K *é convexo e fechado, então* $K = (K^*)^*$.
Portanto, se $\emptyset \neq C \subset H$ *é um conjunto convexo, tem-se que*

$$(N_C(\bar{x}))^* = ((T_C(\bar{x}))^*)^* = T_C(\bar{x}).$$

1.4 Funções Convexas

Definição 1.14. *Se* $C \subset H$ *é um conjunto convexo não vazio, diz-se que a função* $f : C \to \mathbb{R} \cup \{\infty\}$ *é convexa em* C *quando para quaisquer* $x, y \in C$ *e* $\alpha \in [0, 1]$ *tem-se*

$$f(\alpha x + (1 - \alpha)y) \leqslant \alpha f(x) + (1 - \alpha) f(y)$$

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

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

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

É óbvio que uma função fortemente convexa é estritamente convexa, e uma função estritamente convexa é convexa. A função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = x^2$, é um exemplo de uma função fortemente convexa. A função $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = e^x$ é estritamente (mas não fortemente) convexa. Já a função $f : \mathbb{R} \to \mathbb{R}$, tal que $f(x) = x$ é convexa (mas não é estritamente convexa).

Definição 1.15. *O conjunto* $D(f) = \{x \in H; f(x) < \infty\}$ *é chamado de domínio efetivo de* f *. Já o conjunto*

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

é chamado de epígrado da função f.

A função f *é chamada de própria quando* $D(f) \neq \emptyset$.

Definição 1.16. *Uma função* $f : H \to \mathbb{R} \cup \{\infty\}$ *é chamada fechada quando seu epígrafo é fechado em* $H \times \mathbb{R}$ *ou equivalentemente se seus conjuntos de níveis são fechados.*

Podemos relacionar a convexidade de conjuntos e funções através do teorema que se segue.

Teorema 1.6. *Seja* $C \subset H$ *um conjunto convexo. Uma função* $f : C \to \mathbb{R} \cup \{\infty\}$ *é convexa em* C *se, e somente se, o epígrafo de* f *é um subconjunto convexo de* $H \times \mathbb{R}$

Demonstração. Primeiramente suponhamos que o epígrafo de f (E_f) seja convexo. Agora, sejam $x, y \in C$ quaiquer. Claramente, $(x, f(x)) \in E_f$ e $(y, f(y)) \in E_f$. Pela convexidade do epígrafo de f (E_f), para todo $\alpha \in [0.1]$, temos que

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

Pela definição do epígrafo, isto é equivalente a dizer que

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

ou seja, f é convexa.

$(\Leftarrow)$ Suponhamos agora que f seja convexa. Sejam $(x, c_1) \in E_f$ e $(y, c_2) \in E_f$. Como $f(x) \leqslant c_1$ e $f(y) \leqslant c_2$, pela convexidade de f, para todo $\alpha \in [0, 1]$ tem-se

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &\leqslant \alpha f(x) + (1-\alpha)f(y) \\ &\leqslant \alpha c_1 + (1-\alpha)c_2, \end{aligned}$$

o que significa que

$$\alpha(x, c_1) + (1-\alpha)(y, c_2) = (\alpha x + (1-\alpha)y, \alpha c_1 + (1-\alpha)c_2) \in E_f,$$

ou seja, o epígrado de f (E_f) è convexo. □

Pelo teorema 1.6 podemos definir a classe de funções convexas como as funções cujos epígrafos são convexos.

Dizemos que

$$\min f(x) \text{ sujeito a } x \in C \tag{1.5}$$

é um problema de minimização convexo quando $C \in H$ é convexo e $f : C \to \mathbb{R} \cup \{\infty\}$ é uma função convexa em C. A convexidade é extremamente importante pois ela garante que o minimizador local do problema (1.5) é na verdade um minimizador global, como pode ser visto no teorema seguinte.

Teorema 1.7. *Sejam $C \subset H$ um conjunto convexo e $f : C \to \mathbb{R}$ uma função convexa em C.*

Então todo minimizador local do problema (1.5), caso exista, é global. Além disso o conjunto dos minimizadores é convexo.

Se f é estritamente convexa, o minimizador, caso exista, é único.

Demonstração. Suponhamos que $x^* \in C$ seja um minimizador local que não é global. Então existe $y \in C$ tal que $f(y) < f(x^*)$.

Definamos $x(\alpha) = \alpha y + (1-\alpha)x^*$. Pela convexidade de f, para todo $\alpha \in (0, 1]$, tem-se

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

Tomando $\alpha > 0$ suficientemente pequeno, podemos garantir que o ponto $x(\alpha)$ é arbitrariamente próximo ao ponto x^*, e ainda tem-se que $f(x(\alpha)) < f(x^*)$ e $x \in C$. Isto contradiz o fato de que x^* é minimizador local do problema (1.5). Portanto qualquer solução local deve ser global.

Sejam $S \subset C$ o conjunto dos minimizadores (globais) e $v^* \in \mathbb{R}$ o valor ótimo do problema ($f(x^*) = v^*$ $\forall$ $x^* \in S$). Para quaisquer $x, x^* \in S$ e $\alpha \in [0, 1]$, pela convexidade de f

obtemos

$$\begin{aligned} f(\alpha x + (1-\alpha)x^*) &\leqslant \alpha f(x) + (1-\alpha)f(x^*) \\ &= \alpha v^* + (1-\alpha)v^* \\ &= v^*, \end{aligned}$$

o que implica que de fato $f(\alpha x + (1-\alpha)x^*) = v^*$ e, portanto, $\alpha x + (1-\alpha)x^* \in S$. Acabamos de mostrar que que S é convexo.

Suponhamos agora que f seja estritamente convexa e que existam $x, x^* \in S$, $x \neq x^*$. Seja $\alpha \in (0,1)$. Como x e x^* são minimizadores globais e $\alpha x + (1-\alpha)x^* \in C$, pela convexidade de C segue-se

$$f(\alpha x + (1-\alpha)x^*) \geqslant f(x) = f(x^*) = v^*.$$

No entanto, pela convexidade estrita,

$$\begin{aligned} f(\alpha x + (1-\alpha)x^*) &< \alpha f(x) + (1-\alpha)f(x^*) \\ &= \alpha v^* + (1-\alpha)v^*, \\ &= v^* \end{aligned} \tag{1.6}$$

o que resulta em contradição. Concluimos que neste caso o minimizador é único. □

Definição 1.17. *Uma aplicação* f *é Gateaux diferenciável em* x *se existe um elemento em* H*, que chamaremos de* $\nabla f(x)$*, tal que para todo* $d \in H$

$$\lim_{\alpha \to 0} \frac{f(x + \alpha d) - f(x)}{\alpha} = \langle d, \nabla f(x) \rangle. \tag{1.7}$$

$\nabla f(x)$ *é unicamente determinado por (1.7). Ele é chamado de derivada de Gateaux de* f *em* x*. A expressão do lado esquerdo de (1.7) é chamado de derivada direcional de* f *em* x *na direção de* d *e é denotada por* $f'(x, d)$*.*

Se $\nabla f(x)$ existe, então, por (1.7) $f'(x, d)$ existe para cada $d \in H$, e

$$f'(x, d) = \langle d, \nabla f(x) \rangle.$$

Mas a existência da derivada direcional de f em x não implica na diferenciabilidade de Gateux de f em x. Com esta última igualdade podemos observar que $f'(x, \cdot) : H \to \mathbb{R}$, quando existe, é linear e contínua.

Teorema 1.8. *(Caracterização de Funções Convexas Diferenciáveis). Sejam* $C \subset H$ *um conjunto convexo e aberto e* $f : C \to \mathbb{R}$ *uma função diferenciável em* C*. Então, as propriedades seguintes são equivalentes:*

a) A função f *é convexa em* C*;*

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

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

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

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

Demonstração. Inicialmente mostraremos que $(a) \Rightarrow (b) \Rightarrow (c)$.

Seja f convexa. Para $x, y \in C$ e $\alpha \in (0, 1]$ quaisquer, definindo $d = y - x$, temos que

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

donde

$$\alpha(f(y) - f(x)) \geqslant f(x + \alpha d) - f(x).$$

Dividindo os dois membros da desigualdade acima por $\alpha > 0$, e passando ao limite quando $\alpha \to 0_+$, obtemos

$$f(y) - f(x) \geqslant \lim_{\alpha \to 0_+} \frac{f(x + \alpha d) - f(x)}{\alpha} = \langle \nabla f(x), d \rangle = \langle \nabla f(x), y - x \rangle$$

ou seja, $f(y) - f(x) \geqslant \langle \nabla f(x), y - x \rangle$, portanto $f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle$, e desta forma obtemos (b).

Trocando agora o papel de x e y no item (b), temos

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

Somando esta desigualdade com a de (b), imediatamente obtemos (c).

Mostraremos agora que $(c) \Rightarrow (b) \Rightarrow (a)$.

Sejam $x, y \in C$. Pelo Teorema do Valor Médio, existe $\alpha \in (0, 1)$ tal que

$$f(y) - f(x) = \langle \nabla f(x + \alpha(y - x)), y - x \rangle. \tag{1.8}$$

Usando (c) para os pontos $(x + \alpha(y - x))$ e x, obtemos

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

Combinando esta desigualdade com (1.8), obtemos (b). Definindo novamente $d = y - x$, temos

$$\begin{aligned}f(x) &\geqslant f(x + \alpha d) - \alpha\langle \nabla f(x + \alpha d), d\rangle, \\ f(y) &\geqslant f(x + \alpha d) + (1 - \alpha)\langle \nabla f(x + \alpha d), d\rangle,\end{aligned}$$

onde usamos (b) para os pontos x e $x + \alpha d$; y e $x + \alpha d$, respectivamente. Multiplicando a primeira desigualdade por $1 - \alpha \geqslant 0$ e a segunda por $\alpha \geqslant 0$, e somando, obtemos

$$\begin{aligned}(1 - \alpha)f(x) + \alpha f(y) &\geqslant (1 - \alpha)(f(x + \alpha d) - \alpha\langle \nabla f(x + \alpha d), d\rangle) \\ &\quad + \alpha(f(x + \alpha d) + (1 - \alpha)\langle f(x + \alpha d), d\rangle) \\ &= f(x + \alpha d) \\ &= f((1 - \alpha)x + \alpha y),\end{aligned}$$

o que mostra que f é convexa.

$\square$

Lema 1.4.1. *Seja* $f : H \to \mathbb{R}$ *uma função diferenciável no ponto* $\bar{x} \in H$. *Então:*
(a) Para todo $d \in D(f(\bar{x}))$, *tem-se* $\langle \nabla f(\bar{x}), d\rangle \leqslant 0$.
(b) Se $d \in H$ *satisfaz* $\langle \nabla f(\bar{x}), d\rangle < 0$, *tem-se que* $d \in D(f(\bar{x}))$.
Onde $D(f(x)) = \{x \in H; f(x) < \infty\}$ *denota o domínio efetivo da função* f *no ponto* $\bar{x}$.

Demonstração. Seja $d \in D(f(\bar{x}))$. Para todo $t > 0$ suficientemente pequeno,

$$0 > f(\bar{x} + td) - f(\bar{x}) = t(\langle \nabla f(\bar{x}), d\rangle + o(t)/t).$$

Dividindo os dois lados da desigualdade acima por $t > 0$ e passando ao limite quando $t \to 0_+$, obtemos $\langle \nabla f(\bar{x}), d\rangle \leqslant 0$, o que prova o item (a).
Suponhamos agora que $\langle \nabla f(\bar{x}), d\rangle < 0$. Temos que

$$f(\bar{x} + td) - f(\bar{x}) = t(\langle \nabla f(\bar{x}), d\rangle + o(t)/t).$$

Em particular, para todo $t > 0$ suficientemente pequeno pequeno, temos

$$\langle \nabla f(\bar{x}), d\rangle + o(t)/t \leqslant \langle \nabla f(\bar{x}), d\rangle/2 < 0,$$

o que implica que $f(\bar{x} + td) - f(\bar{x}) < 0$, isto é $d \in D(f(\bar{x}))$. □

Proposição 1.4. *Uma função,* $f : H \to \mathbb{R}$*, convexa e semicontínua inferiormente é fracamente semicontínua inferiormente, ou seja,*

$$\liminf_{x^k \rightharpoonup x} f(x^k) \geqslant f(x)\ \forall\ x \in H. \tag{1.9}$$

Demonstração. Como f é convexa e semicontínua inferiormente, o E_f é convexo e fechado. Daí, E_f é fracamente fechado em $H \times \mathbb{R}$. A topologia em $H \times \mathbb{R}$ é a topologia do produto $H \times \mathbb{R}$ determinada pela topologia fraca em H e pela topologia usual em $\mathbb{R}$. Assim f é fracamente semicontínua inferiormente. □

A proposição 1.4 nos leva ao seguinte resultado:

Proposição 1.5. *Sejam* H *um espaço de Hilbert e* $f : H \to \mathbb{R} \cup \{\infty\}$ *uma função convexa e semicontínua inferiormente. Se* $C \subset H$ *é fechado, convexo e limitado com* $D(f) \cap C \neq \emptyset$*, então* f *atinge um ínfimo em* C*.*

Demonstração. Como todo espaço de Hilbert é reflexivo, então todo subconjunto limitado e fracamente fechado é sequencialmente fracamente relativamente compacto e pelo Teorema de Weierstrass, toda função fracamente semicontínua inferiormente atinge um ínfimo. E assim, provamos a proposição. □

Corolário 1.4.1. *Sejam* H *um espaço de Hilbert e* $f : H \to \mathbb{R} \cup \{\infty\}$ *uma função convexa semicontínua inferiormente tal que*

$$\lim_{||x|| \to \infty} f(x) = \infty. \tag{1.10}$$

Então f *atinge um mínimo em* H*.*

Demonstração. Pela condição (1.10) vemos que $\inf\{f(x); x \in H\} > -\infty$ e que para algum $R > 0$ temos

$$\inf\{f(x); x \in H\} = \inf\{f(x); ||x|| \leqslant R\}.$$

Então aplicamos a proposição 1.5 para concluirmos que f atinge seu ínfimo em $\{x; ||x|| \leqslant R\}$ e portanto em H. □

1.5 Subdiferencial de funções convexas

Nesta seção trbalharemos uma ferramenta fudamental em Problemas de Análise Convexa, que é o conceito de subdiferencial de uma função convexa.

Definição 1.18. *Seja* $f : H \to \mathbb{R}\cup\{\infty\}$ *uma função convexa e semicontínua inferiormente e* $x_0 \in D(f)$

a) $y \in H$ *é o subgradiente de* f *em* x_0 *se*

$$f(x) \geqslant f(x_0) + \langle y, x - x_0\rangle \ \forall \ x \in H.$$

b) O conjunto de todos os subgradientes de f *em* x_0 *é chamado de subdiferencial de* f *em* x_0*. Este conjunto é denotado por* $\partial f(x_0)$*. Se* $x \notin D(f)$*, definimos* $\partial f(x) = \emptyset$*.*

c) O domínio de ∂f *é o conjunto*

$$D(\partial f) = \{x \in H; \partial f(x) \neq \emptyset\}.$$

O subgradiente define uma aproximação linear de f cujo gráfico fica abaixo do gráfico de f e cujo valor coincide com f no ponto x.

Teorema 1.9. *(O subdiferencial de uma função convexa) Seja* $f : H \to \mathbb{R} \cup \{\infty\}$ *uma função convexa e contínua. Então para todo* $x \in H$*, o conjunto* $\partial f(x)$ *é convexo, compacto e não vazio. Além disso, para todo* $d \in H$*, tem-se*

$$f'(x, d) = \sup\{\langle y, d\rangle; y \in \partial f(x)\}. \tag{1.11}$$

Demonstração. Ver [13] □

Proposição 1.6. *Uma função* $f : H \to \mathbb{R}$ *convexa é Gateaux diferenciável no ponto* $x \in H$ *se, e somente se, o conjunto* $\partial f(x)$ *contém um único elemento. Neste caso,* $\partial f(x) = \{\nabla f(x)\}$*.*

Demonstração. Seja f diferenciável no ponto $x \in H$. Pelo Teorema 1.8,

$$f(z) \geqslant f(x) + \langle \nabla f(x), z - x\rangle \ \ \forall \ z \in H,$$

isto é, $\nabla f(x) \in \partial f(x)$. No Teorema 1.8 aparece a hipótese de que f é diferenciável em todos os pontos, mas para a desigualdade usada aqui é suficiente considerarmos a desigualdade somente no ponto x.

Seja agora $y \in \partial f(x)$. para todo $d \in H$,

$$f(x) + \langle y, d\rangle \leqslant f(x+d) = f(x) + \langle \nabla f(x), d\rangle + o(\|d\|).$$

Logo,

$$\langle y - \nabla f(x), d\rangle \leqslant o(\|d\|).$$

Daí,

$$\langle y - \nabla f(x), d\rangle \leqslant \frac{o(\|d\|)}{\|d\|}\|d\|.$$

Tomando limite quando $d \to 0$, temos

$$\langle y - \nabla f(x), d\rangle \leqslant 0 \ \forall \ d \in H$$

em particular para $d = y - \nabla f(x)$, logo

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

o que é possível somente quando $y - \nabla f(x) = 0$. Desta forma concluimos que $\partial f(x) = \{\nabla f(x)\}$.

Agora seja $\partial f(x) = \{y\}$. Pelo Teorema 1.8 temos que

$$\frac{\partial f(x)}{\partial d} = \langle y, d\rangle \ \forall \ d \in H.$$

Escolhendo d como sendo os elementos da base canônica de H, podemos observar que $y_i = \partial f(x)/\partial x_i$, $\forall \ i \in \mathbb{N}$. Temos que $\frac{\partial f(x)}{\partial d}$ é uma função linear em d, o que implica que f é diferenciável em x. $\square$

Teorema 1.10. *Sejam* f e g *funções convexas, semicontínuas inferiormente e próprias em* H *tais que* $(D(f))^0 \cap D(g) \neq \emptyset$. *Então*

$$\partial(f+g) = \partial f + \partial g.$$

Demonstração. A inclusão $\partial f + \partial g \subset \partial(f+g)$ é imediata, então basta provarmos que $\partial(f+g) \subset \partial f + \partial g$. Seja $x^0 \in D(\partial f) \cap D(\partial g)$ e $w \in \partial(f+g)(x^0)$ um ponto arbitrário, porém fixo. Vamos mostrar que $w = w_1 + w_2$ onde $w_1 \in \partial f(x^0)$ e $w_2 \in \partial g(x^0)$. Sem perda de generalidade podemos assumir que $x^0 = 0, w = 0$ e $f(0) = g(0) = 0$. Podemos conseguir isto substituindo f e g por $x \to f(x+x^0)-f(x^0)-\langle z_1, x\rangle$ e $x \to g(x+x^0)-g(x^0)-\langle z_2, x\rangle$, respectivamente. Para provarmos que $0 \in \partial f(0)+\partial g(0)$ consideremos os conjuntos

$$E_1 = \{(x,\lambda) \in H \times \mathbb{R}; f(x) \leqslant \lambda\}$$

$$E_2 = \{(x,\lambda) \in H \times \mathbb{R}; g(x) \leqslant -\lambda\}$$

Como $0 \in \partial(f+g)(0)$ segue que

$$0 = (f+g)(0) = \inf\{f(x)+g(x); x \in H\}$$

e assim $E_1 \cap E_2^0 = \emptyset$. Pelo Teorema de Hahn-Banach, ver [4], existe um hiperplano fechado em $H \times \mathbb{R}$ que separam E_1 e E_2. Em outras palavras existe $\langle w, \alpha\rangle \in H \times \mathbb{R}$ tal que

$$\langle x, w\rangle + \alpha\lambda \leqslant 0 \ \forall\ (x,\lambda) \in E_1$$

$$\langle x, w\rangle + \alpha\lambda \geqslant 0 \ \forall\ (x,\lambda) \in E_2.$$

O hiperplano é não vertical, isto é, $\alpha \neq 0$ pois, segue das últimas desigualdades que o hiperplano $\{x \in H; \langle x, w\rangle = 0\}$ separa $D(f)$ e $D(g)$. Para sermos mais específicos suponhamos que $\alpha = 1$. Então temos que

$$\langle x, w\rangle \leqslant -\lambda \leqslant -f(x) \ \forall\ x \in D(f)$$

e,

$$\langle x, w\rangle \geqslant -\lambda \leqslant g(x) \ \forall\ x \in D(g).$$

Assim $-w \in \partial f(0), w \in \partial g(0)$ como queríamos. □

Teorema 1.11. *(Condição de otimalidade para minimização de uma função convexa em um conjunto convexo). Sejam $f : H \to \mathbb{R}$ uma função convexa e $C \subset H$ um conjunto convexo. Então $\bar{x} \in H$ é um minimizador de f em C se, e somente se,*

$$\exists\ y \in \partial f(\bar{x}) \ \text{tal que}\ \langle y, x-\bar{x}\rangle \geqslant 0 \ \forall\ x \in C. \tag{1.12}$$

ou, de forma equivalente,

$$0 \in \partial f(\bar{x}) + N_C(\bar{x}). \tag{1.13}$$

Em particular, $\bar{x}$ é um minimizador de f em H se, e somente se,

$$0 \in \partial f(\bar{x}). \tag{1.14}$$

Demonstração. Como $\langle y, x - \bar{x}\rangle \geqslant 0$ para todo $x \in C$ significa que $-y \in N_C(\bar{x})$ (ver definição de cone normal), daí concluimos que (1.12) e (1.13 são equivalentes. Suponhamos que a hipótese (1.12) seja verdadeira. Pela definição de subgradinte, segue-se que

$$\forall\ x \in H,\ f(x) \geqslant f(\bar{x}) + \langle y, x - \bar{x}\rangle \geqslant f(\bar{x}),$$

isto é, $\bar{x}$ é um minimizador de f em C.

Suponhamos agora que $\bar{x}$ seja um minimizador de f em C. Escolhemos qualquer $d \in T_C(\bar{x})$, $d \neq 0$. Logo, existem sequências $\{t_k\}$ de números reais positivos e $\{d^k\} \subset H$ tais que $\{d^k\} \to d$, $\{t_k\} \to 0_+$, quando $k \to \infty$, $\bar{x} + t_k d^k \in C$ para todo k. Para todo k, temos que

$$0 \leqslant f(\bar{x} + t_k d^k) - f(\bar{x}) = t_k \frac{\partial f(\bar{x})}{\partial d^k} + o(t_k).$$

Dividindo os dois lados por $t_k > 0$, passando ao limite quando $k \to \infty$, e utilizando o Teorema 1.8, concluimos que

$$\frac{\partial f}{\partial d}(\bar{x}) \geqslant 0\ \forall\ d \in T_C(\bar{x}). \tag{1.15}$$

Suponhamos que (1.13) não seja válida, ou seja, que não exista $y \in H$ tal que $-y \in \partial f(\bar{x})$ e $y \in N_C(\bar{x})$. Ou seja,

$$(-\partial f(\bar{x})) \cap N_C(\bar{x}) = \emptyset.$$

Como $\partial f(\bar{x})$ é convexo, compacto e não vazio e $N_C(\bar{x})$ também é convexo, fechado e não vazio, pelo teorema da Separação Estrita, exitem $a \in H - \{0\}$ e $c \in \mathbb{R}$ tais que

$$-\langle a, y\rangle > c > \langle a, d\rangle\ \forall\ y \in \partial f(\bar{x}),\ \forall\ d \in N_C(\bar{x}). \tag{1.16}$$

Como $0 \in N_C(\bar{x})$, tem-se $c > 0$. Portanto, $\langle a, y\rangle < -c < 0$ para todo $y \in \partial f(\bar{x})$. Usando o Teorema 1.8, obtemos então que

$$\frac{\partial f}{\partial a}(\bar{x}) = \max_{y \in \partial f(\bar{x})} \langle a, y\rangle < 0. \tag{1.17}$$

Suponhamos que $\langle a, d\rangle > 0$ para algum $d \in N_C(\bar{x})$. Tomando $td \in N_C(\bar{x})$, $t > 0$, e passando ao limite quando $t \to \infty$, resulta numa contradição com (1.15), porque $c \in R$ é fixo. Portanto,$\langle a, d\rangle \leqslant 0$ para todo $d \in N_C(\bar{x})$. Concluimos então que

$$a \in (N_C(\bar{x}))^* = ((T_C(\bar{x}))^*)^* = T_C(\bar{x}), \tag{1.18}$$

onde as igualdades acima seguem da proposição 1.3. As relações (1.16) e (1.17) contradizem (1.15). Isto prova (1.12).

Além disso, é claro que quando $C = H$, as condições 1.12 e 1.13 são equivalentes a $0 \in \partial f(\bar{x})$. $\square$

Capítulo 2

Distâncias de Bregman

2.1 Funções e Distâncias de Bregman

Sejam $S \subset H$ um conjunto aberto e convexo e $g : H \to \mathbb{R}$ Gateaux diferenciável, defina $D_g : \overline{S} \times S \to \mathbb{R}_+$ como

$$D_g(x, y) = g(x) - g(y) - \langle \nabla g(y), x - y \rangle \tag{2.1}$$

onde $\nabla g(y)$ é a derivada de Gateaux de g em y.

Definição 2.1. *A função g é chamada função de Bregman e D_g a distância de Bregman induzida por g se satisfaz as seguintes condições:*

B_1) *g é fracamente semicontínua inferiormente em $\overline{S}$*

B_2) *g é estritamente convexa e continuamente diferenciável em S com respeito à topologia forte.*

B_3) *Os conjuntos $\Gamma_1(y, \delta) = \{x \in \overline{S}; D_g(x, y) \leqslant \delta\}$ e $\Gamma_2(x, \delta) = \{y \in S; D_g(x, y) \leqslant \delta\}$ são limitados para todo $\delta > 0$, $y \in S$, e $x \in \overline{S}$, respectivamente.*

B_4) *Se $\{x^k\}, \{y^k\} \subset S$, $x^k \rightharpoonup x$, $y^k \rightharpoonup x$, e $\lim_{k\to\infty} D_g(x^k, y^k) = 0$, então $\lim_{k\to\infty}[D_g(x, x^k) - D_g(x, y^k)] = 0$.*

B_5) *Se $\{x^k\} \subset \overline{S}$ é limitada, $\{y^k\} \subset S$ é tal que $y^k \rightharpoonup y$ e $\lim_{k\to\infty} D_g(x^k, y^k) = 0$, então $x^k \rightharpoonup y$.*

Uma função de Bregman g é dita compatível com a norma se, em adição com $B_1 - B_5$ satisfaz:

B_6) *Se $\{x^k\} \subset S$ e $\{y^k\} \subset S$ são tais que $x^k \rightharpoonup x$ e $y^k \rightharpoonup y$, com $x \neq y$, então $\liminf_{k\to\infty} |\langle \nabla g(x^k) - \nabla g(y^k), x - y \rangle| > 0$.*

A função de Bregman g *é chamada de zona coerciva se, em adição a* $B_1 - B_5$ *satisfaz a seguinte condição:*

B_7) *Para cada* $y \in H$, *existe um* $x \in S$ *tal que* $\nabla g(x) = y$.

Uma função de Bregman g *é dita ser fronteira coerciva se, em adição a* $B_1 - B_5$ *satisfaz:*

B_8) *Se* $\{y^k\} \subset S$ *é fracamente convergente para algum ponto da fronteira de* S *e* x *pertence* S, *então* $\lim_{k\to\infty} D_g(x, y^k) = \infty \ \forall\ x \in S$. *O conjunto* S *é chamado de zona da função* g.

Podemos verificar facilmente que $D_g(x, y) \geqslant 0$ para todo $x \in \overline{S}$, $y \in S$ e $D_g(x, y) = 0$ se, e somente se, $x = y$.

Agora vamos ver alguns exemplos de funções de Bregman:

Exemplo 2.1. *Seja* $g : H \to \mathbb{R}$, *tal que* $g(x) = ||x||^2$. *Neste caso* $D_g(x, y) = ||x - y||^2$.

Exemplo 2.2. $S = \mathbb{R}^n_{++}$, $g(x) = \sum_{i=1}^n x_i \log x_i$ *estendido com a continuidade até a fronteira de* $\mathbb{R}^n_{++}$ *com a convenção de que* $0 \log 0 = 0$. *Neste caso,*

$$D_g(x, y) = \sum_i^n (x_i \log \frac{x_i}{y_i} + y_i - x_i).$$

Proposição 2.1. *Se* g *é uma função de Bregman com zona* S *então:*

i) $D_g(x, y) - D_g(x, z) - D_g(z, y) = \langle \nabla g(y) - \nabla g(z), x - y \rangle$, $\forall x \in \overline{S}$, $\forall y, z \in S$.

ii) $\nabla_x D_g(x, y) = \nabla g(x) - \nabla g(y)$, $\forall\ x, y \in S$.

iii) $D_g(x, y)$ *é estritamente convexa para todo* $y \in S$.

Demonstração. i) Usando a igualdade (2.1), para $x \in \overline{S}$, $y, z \in S$, temos

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

ii) Sejam $x, z \in S$. Como $D_g(x, y) = g(x) - g(y) - \langle \nabla g(y), x - y \rangle$ então

$$\begin{aligned} \nabla_x D_g(x, y) &= \nabla g(x) - 0 - \langle \frac{\partial}{\partial x} \nabla g(y), x - y \rangle - \langle \nabla g(y), \frac{\partial}{\partial x}(x - y) \rangle \\ &= \nabla g(x) - \langle 0, x - y \rangle - \langle \nabla g(y), 1 \rangle \\ &= \nabla g(x) - \nabla g(y). \end{aligned}$$

iii) Para $y \in S$, defina

$$\begin{aligned} G : S &\longrightarrow \mathbb{R} \\ x &\longmapsto G(x) = D_g(x, y). \end{aligned}$$

Sejam $x_1, x_2 \in S$, com $x_1 \neq x_2$. Pelo item (ii), $\nabla G(x_1) = \nabla g(x_1) - \nabla g(y)$ e $\nabla G(x_2) = \nabla g(x_2) - \nabla g(y)$. Logo,

$$\begin{aligned} \langle \nabla G(x_1) - \nabla G(x_2), x_1 - x_2 \rangle &= \langle \nabla g(x_1) - \nabla g(y) - \nabla g(x_2) + \nabla g(y), x_1 - x_2 \rangle \\ &= \langle \nabla g(x_1) - \nabla g(x_2), x_1 - x_2 \rangle > 0, \end{aligned}$$

pois g é estritamente convexa em $\overline{S}$. Portanto, $G = D_g(\cdot, y)$ é estritamente covexa em $\overline{S}$, para cada $y \in S$. □

Para estudarmos as propriedades do subdiferencial de $D_g(., y)$ vamos assumir a partir de agora que as funções g e $D_g(., y)$ foram estendidas de forma usual, isto é, definimos

$$g(x) = D_g(x, y) = \infty, \forall\, x \notin \overline{S}. \tag{2.2}$$

Lema 2.1.1. *Se g é uma função fronteira coerciva ou zona coerciva com zona S então, para cada $y \in S$ fixado,*

$$\partial_x D_g(x, y) = \begin{cases} \{\nabla g(x) - \nabla g(y)\}, & \text{se} \quad x \in S \\ \emptyset, & \text{se} \quad x \notin S \end{cases}$$

Demonstração. O valor de $\partial_x D_g(x, y)$ para $x \in S$ segue diretamente da proposição 2.1(ii). Quando $x \notin \overline{S}$ temos que $\partial_x D_g(x, y) = \emptyset$ por (2.2). Portanto, com estas considerações, podemos estudar somente o caso em que x pertence à fronteira de S. Para simplificar a notação, usamos $f_y(\cdot) = D_g(\cdot, y)$. Tome $x \in \partial S$ e assuma que exista $\xi \in \partial f_y(\cdot)$. Note que f_y difere de g apenas em um termo linear e que f_y é contínua por B_1 e estritamente convexa em $\overline{S}$, pela proposição 2.1(iii).

Se g é uma função zona coerciva, então, por B_7 existe $z \in S$ tal que $\nabla f_y(z) = \nabla g(z) - \nabla g(y) = \xi$, implicando

$$0 = \langle \xi - \xi, z - x \rangle = \langle \nabla f_y(z) - \xi, z - x \rangle, \text{ com } \xi \in \partial f_y(x).$$

Pela convexidade de f_y, obtemos $x = z$, que é uma contradição, porque $z \in S$, $x \in \overline{S}$ e S é aberto. Segue que $\partial f_y(x) = \emptyset$.

Agora consideremos o caso de g ser fronteira coerciva. Vamos tomar uma sequência $\{\varepsilon_k\} \subset (0,1)$ tal que $\lim_{k\to\infty} \varepsilon_k = 0$ e seja $x^k = (1-\varepsilon_k)x + \varepsilon_k y$. Pela convexidade de S, $x^k \in S \ \forall\ k$ e usando a definição de subgradiente, temos

$$
\begin{aligned}
\varepsilon_k\langle \xi, y - x\rangle &= \langle \xi, x^k - x\rangle \\
&\leqslant f_y(x^k) - f_y(x) \\
&= g(x^k) - g(y) - \langle \nabla g(y), x^k - y\rangle - [g(x) - g(y) - \langle \nabla g(y), x - y\rangle] \\
&= g(x^k) - g(x) + \langle \nabla g(y), x - x^k\rangle \\
&\leqslant \langle -\nabla g(x^k), x - x^k\rangle + \langle \nabla g(y), x - x^k\rangle \\
&= \langle \nabla g(y) - \nabla g(x^k), x - x^k\rangle \\
&= \langle \nabla_g(x^k) - \nabla_g(y), x^k - x\rangle. \qquad (2.3)
\end{aligned}
$$

Como

$$x^k = (1-\varepsilon_k)x + \varepsilon_k y,$$

temos que

$$x^k - \varepsilon_k x^k = (1-\varepsilon_k)x + \varepsilon_k y - \varepsilon_k x^k$$

daí,

$$x^k - x = \frac{\varepsilon_k}{1-\varepsilon_k}(y - x^k). \qquad (2.4)$$

Segue de (2.1) que

$$\langle \nabla g(u) - \nabla g(v), u - v\rangle = Dg(u,v) + Dg(v,u) \qquad (2.5)$$

$\forall\ u, v \in S$. Então usando estes dois fatos em (2.3),

$$
\begin{aligned}
\varepsilon_k\langle \xi, y - x\rangle &\leqslant \langle \nabla g(x^k) - \nabla g(y), \frac{\varepsilon_k}{1-\varepsilon_k}(y - x^k)\rangle \\
&= \frac{\varepsilon_k}{1-\varepsilon_k}\langle \nabla g(x^k) - \nabla g(y), y - x^k\rangle \\
&= -\frac{\varepsilon_k}{1-\varepsilon_k}\langle \nabla g(x^k) - \nabla g(y), x^k - y\rangle \\
&= -\frac{\varepsilon_k}{1-\varepsilon_k}(D_g(x^k, y) + D_g(y, x^k)) \\
\varepsilon_k\langle \xi, y - x\rangle &\leqslant -\frac{\varepsilon_k}{1-\varepsilon_k}(D_g(x^k, y) + (D_g(y, x^k)) \qquad (2.6)
\end{aligned}
$$

Dividindo (2.6) por $\frac{\varepsilon_k}{1-\varepsilon_k} > 0 \ \forall\ k$, temos

$$(1-\varepsilon_k)\langle \xi, y - x\rangle \leqslant -(D_g(x^k, y) + D_g(y, x^K))$$

$$(1-\varepsilon_k)\langle \xi, y-x\rangle + D_g(x^k, y) \leqslant -D_g(y, x^k),$$

ou ainda,

$$(1-\varepsilon_k)\langle \xi, y-x\rangle + f_y(x^k) \leqslant -D_g(y, x^k). \tag{2.7}$$

Visto que $x^k \rightharpoonup x \in \partial S \subset \overline{S}$ e f_y é fracamente semicontínua em $\overline{S}$, então, tomando limites em 2.7, quando $k \to \infty$, temos que o lado esquerdo de (2.7) é maior ou igual a $\langle \xi, y-x\rangle + D_g(x, y)$. Por B_8, o limite do lado direito de (2.7) é $-\infty$, contradição. Esta contradição mostra que $\partial f_y(x) = \emptyset$. $\square$

Capítulo 3

Operadores Monótonos, Monótonos Maximais, Paramonótonos e Pseudomonótonos

Neste capítulo iremos trabalhar algumas noções de Operadores Monótonos Maximais que serão bastante utilizados no decorrer de nosso trabalho. Vejamos abaixo algumas definições.

3.1 Operadores Monótonos

Definição 3.1. *Um operador* $T : H \to H$ *é monótono se, e somente se,*

$$\langle x - y, T(x) - T(y) \rangle \geqslant 0 \ \forall \ x, y \in \ H.$$

O conjunto $D(T) = \{x \in H; \ T(x) \neq \emptyset\}$ *é o domínio de* T*. Já o conjunto* $\{(x, u); \ x \in H, \ \ u \in T(x)\}$ *é o gráfico de* T*,*

Exemplo 3.1. $T = \nabla f$ *com* f *convexa e diferenciável é um operador monótono.*
De fato, aplicando o Teorema 1.8*, temos que* $T = \nabla f$ *é um operador monótono.*

O operador $T : H \to H$, da definição 3.1, é chamado de operador ponto-a-ponto, pois o mesmo associa um ponto ou vetor do conjunto H a um outro ponto (ou vetor) de H. Esta definição de operadores monótonos pode ser estendida para operadores que associa cada $x \in H$ não apenas a um ponto (ou vetor), mas a um suconjunto de H. Assim, temos a noção de operadores ponto-conjunto. Seja $T : H \rightrightarrows H$ um operador ponto-conjunto.

Definição 3.2. *O operador* $T : H \rightrightarrows H$ *é monótono se, e somente se,*

$$\langle x - y, u - v \rangle \geqslant 0 \; \forall \; x, y \in H, \; u \in T(x), \; v \in T(y).$$

Exemplo 3.2. $T = \partial f$ *com* f *convexa é um operador monótono.*

. *Ver Teorema* 1.8.

Proposição 3.1. *Se* $T_1, T_2 : H \to H$ *são operadores monótonos tais que* $D(T_1) \cap D(T_2) \neq \emptyset$*, então* $T_1 + T_2$ *é um operdor monótono.*

Demonstração. Segue direto da definição 3.2. □

3.2 Operadores Monótonos Maximais

Definição 3.3. T *é monótono maximal se, e somente se,*

i) T *é monótono.*

ii) Para todo $\bar{T}$ *monótono tal que* $T(x) \subset \bar{T}(x)$ *para todo* x*, então* $T = \bar{T}$.

Neste caso, o gráfico de T não é subconjunto próprio de nenhum outro gráfico de outro operador monótono.

Exemplo 3.3. *Quando* $f : H \to \mathbb{R} \cup \{\infty\}$ *é convexa, própria e semicontínua inferiormente então* $T = \partial f$ *é um operador monótono maximal.*

Exemplo 3.4. *Seja* C *um subconjunto convexo e fechado de um espaço de Hilbert e defina a aplicação* $\mathsf{Proj}_C : H \to C$ *como sendo aquela que a cada ponto* $x \in H$ *associa o ponto de* C *mais próximo de* x*. Esta aplicação é conhecida como projeção de* H *sobre o conjunto* C*. A projeção é um exemplo de operdor monótono maximal que não é o subdiferencial de uma função convexa e semicontínua inferiormente.*

Proposição 3.2. *Sejam* $T_1, T_2 : H \rightrightarrows H$ *operadores monótonos maximais. Suponha que* $D(T_1) \cap (D(T_2))^0 \neq \emptyset$*. Então* $T_1 + T_2$ *é monótono maximal.*

Demonstração. ver [9] □

Agora vamos usar a noção de regularidade, que é uma propriedade de operadores monótonos maximais. Sejam $G(T) = \{(y, v); y \in H, v \in T(y)\}$ e $R(T) = \cup_{x \in H} T(x)$, respectivamente o gráfico e a imagem do operador monótono maximal T.

Definição 3.4. *Um operador monótono maximal* $T : H \rightrightarrows H$ *é regular se, e somente se,* $\forall\ u \in R(T)$ e $\forall\ x \in D(T)$, *temos que* $\sup_{(y,v) \in G(T)} \langle v - u, x - y \rangle < \infty$.

Proposição 3.3. *Se* $T = \partial f$, *com* f *convexa e semicontínua inferiormente, então* T *é regular.*

Demonstração. ver [3] □

Agora, vamos analizar a imagem da soma de dois operadores monótonos maximais, $T_1, T_2 : H \rightrightarrows H$, ou seja, vamos estudar os conjuntos

$$R(T_1 + T_2) = \cup_{u \in D(T_1) \cap D(T_2)} (T_1 u + T_2 u).$$

$$R(T_1) + R(T_2) = \cup_{u \in D(T_1), v \in D(T_2)} (T_1 u + T_2 v).$$

Para cada conjunto X, $conv(X)$, denota-se o convexo de X.

Lema 3.2.1. *Seja* $T_0 : H \rightrightarrows H$ *um operador monótono maximal e seja* $F \subset H$. *Se para cada* $u \in F$ *existe* $x \in H$ *tal que*

$$\sup_{(v,y) \in G(T_0)} \langle v - u, x - y \rangle < \infty \tag{3.1}$$

então,

i) $conv(F) \subseteq \overline{R(T_0)}$.

ii) $(conv(F))^0 \subseteq R(T_0)$.

Demonstração. Podemos verificar que (3.1) é válida para todo $u \in conv(F)$. De fato, seja $u = \sum \lambda_i u_i,\ u_i \in F,\ \lambda_i \geqslant 0$. Então, existem $x_i \in H$ e $c_i \in \mathbb{R}$, tais que

$$\langle v - u_i, x_i - y \rangle \leqslant c_i,\ \forall\ (y, v) \in G(T_0).$$

Então

$$\langle v, x_i \rangle - \langle v, y \rangle + \langle u_i, y \rangle \leqslant c_i + \langle u_i, x_i \rangle$$

e pondo $x = \sum \lambda_i x_i$ concluimos que (3.1) se verifica para $u \in conv(F)$. Então para provar (i) e (ii) basta mostrar, respectivamente, que

$$F \subset \overline{R(T_0)} \text{ e } F^0 \subset R(T_0).$$

Assim, dado $u \in F$, tome soluções y_ε de

$$\varepsilon y_\varepsilon + T_0 y_\varepsilon = u. \tag{3.2}$$

Usando (3.1) com $y = y_\varepsilon$ e $v = u - \varepsilon y_\varepsilon$, temos

$$\langle u - \varepsilon y_\varepsilon - u, x - y_\varepsilon \rangle \leqslant d$$

daí, $\sqrt{\varepsilon}\|y_\varepsilon\| \leqslant k$, onde d e k são constantes. Logo, por (3.2), $u \in \overline{R(T_0)}$. Portanto, $F \subset \overline{R(T_0)}$.

Agora seja $u \in F^o$ e $B(u, r) \subset F$. Então, para $\|w\| < r$ existe $x(w) \in H$ e $c(w) \in \mathbb{R}$ (x e c dependem de w) tais que

$$\langle v - u - w, x(w) - y \rangle < c(w), \ \forall \ (v, y) \in G(T_0). \tag{3.3}$$

Usando (3.3) para $(u - \varepsilon y_\varepsilon, y_\varepsilon) \in G(T_0)$, obtido de (3.2), temos

$$\langle -\varepsilon y_\varepsilon - w, x(w) - y_\varepsilon \rangle < c(w)$$

o que implica

$$\langle w, y_\varepsilon \rangle \leqslant c_1(w).$$

Logo, pelo princípio da limitação uniforme $|\langle w, y_\varepsilon \rangle| \leqslant T_0$, e daí $\|y_\varepsilon\| \leqslant T_0$. Passando a uma subsequência, se necessário, temos $y_\varepsilon \rightharpoonup y$ e $u - \varepsilon y_\varepsilon \to u$ e como $u - \varepsilon y_\varepsilon \in T_0(y_\varepsilon)$, concluimos que $u \in T_0(y)$. Portanto, $(conv(F))^0 \subset R(T_0)$.

□

Proposição 3.4. *Sejam* T_1, T_2 *operadores monótonos maximais. Suponha que*

a) T_1 *é regular,*

b) $D(T_1) \cap D(T_2) \neq \emptyset$ e $R(T_1) = H$,

c) $T_1 + T_2$ *é monótono maximal.*

Então $R(T_1 + T_2) = H$.

Demonstração. Vamos aplicar o lema 3.2.1 para $F = R(T_1) + R(T_2)$ e $T_0 = T_1 + T_2$. O resultado segue do lema 3.2.1(ii) e da afirmação (b), isto prova que F satisfaz (3.1). Para provar (3.1) vamos partir de $F = R(T_1) + R(T_2)$. Seja $u \in R(T_1) + R(T_2)$, e tome $x \in D(T_1) \cap D(T_2)$ e $w \in T_2(x)$. Então

$$u = w + (u - w)$$

para $w \in T_2(x)$.

Como $R(T_1) = H$, podemos encontrar $y \in H$ tal que $u - w \in T_1(y)$. Então, como T_1 é regular, dados $u - w \in R(T_1)$ e $x \in D(T_1)$, existe $\gamma \in \mathbb{R}$ tal que

$$\sup_{(z,s)\in G(T_1)} \langle s - (u - w), x - z\rangle \leqslant \gamma.$$

Então , para cada $(z, s) \in G(T_1)$, temos

$$\langle s - (u - w), x - z\rangle \leqslant \gamma. \tag{3.4}$$

Seja $v \in T_2(z)$. Para cada $z \in D(T_2) \cap D(T_1)$ e pela monotonicidade de T_2, temos que

$$\langle v - w, z - x\rangle \geqslant 0 \tag{3.5}$$

pois $w \in T_2(x)$. Adicionando (3.4) e (3.5), obtemos

$$\langle (s + v) - u, x - z\rangle \leqslant \gamma,$$

para cada $s \in T_1(z), v \in T_2(z)$, isto é, para cada $s + v \in (T_1 + T_2)(z)$. Além disso,

$$\sup_{(z,t)\in G(T_1+T_2)} \langle t - u, x - z\rangle \leqslant +\infty$$

e, desta forma, (3.1) é estabelecida para $F = R(T_1) + R(T_2)$ e $T_0 = T_1 + T_2$. Então $R(T_1 + T_2) = H$. □

Agora estamos em condições de provar o lema seguinte, onde para cada $y \in C^0$ fixo, $\partial_x D_g(\cdot, y)$ denota o subdiferencial de $D_g(\cdot, y)$ e $T : H \rightrightarrows H$ é um operador monótono maximal.

Lema 3.2.2. *Sejam* C *um conjunto convexo e fechado e* g *uma função de Bregman zona coerciva com zona* C^0*. Para cada* $y \in C^0$ *e* $\lambda > 0$ *defina* $\bar{T}(\cdot) = T(\cdot) + \lambda\partial_x D_g(\cdot, y)$ *e suponha que* $D(T) \cap C^0 \neq \emptyset$*. Então*
i) $\bar{T}$ *é monótono maximal,*
ii) $R(\bar{T}) = H$

Demonstração. i) O operador $\bar{T}$ é trivialmente monótono maximal, pois T é monótono maximal e $T_2 = \lambda\partial_x D_g(\cdot, y)$ também é monótono maximal. De fato, seja $T_1 = T$ e $T_2 = \lambda\partial_x D_g(\cdot, y)$. Pelo lema 2.1.1, $D(T_1) \cap D(T_2) \neq \emptyset$. Pela proposição 3.2, $\bar{T} = T_1 + T_2$ é monótono maximal.

ii) É uma consequêcia da proposição 3.4 onde são aplicadas as condições (a),(b) e (c) para $T_2 = T$ e $T_1 = \lambda\partial_x D_g(\cdot, y)$. Mais precisamente,

a) $\lambda\partial_x D_g(\cdot, y)$ é regular, pela proposição 3.3,

b) $D(T) \cap C^0 \neq \emptyset$ e $R(\partial_x D_g(\cdot, y)) = H$ por B_6, e

c) $\bar{T}$ é monótono maximal por (i).

Então $R(\bar{T}) = H$ pela proposição 3.4. □

Proposição 3.5. *Seja* $T : H \rightrightarrows H$ *um operador monótono maximal. Se* $D(T)$ *é limitado então* T *é sobrejetivo.*

Demonstração. ver [9] □

3.3 Operadores Paramonótonos

Definição 3.5. *Seja* $T : H \rightrightarrows H$ *um operador. Dizemos que* T *é paramonótono quando é monótono e* $\langle u - v, x - y\rangle = 0$*, com* $u \in T(x)$*,* $v \in T(y)$ *implica* $u \in T(y)$ *e* $v \in T(x)$.

Teorema 3.1. *Seja* $T = \partial f(x)$*, com* $f : H \to \mathbb{R}$ *uma função convexa. Então* T *é um operador paramonótono.*

Demonstração. Pelo exemplo (4.1) ∂f é um operador monótono. Dados $x, y \in H$, considere $\langle u - v, x - y\rangle = 0$, onde $u \in \partial f(x)$ e $v \in \partial f(y)$. Defina a função $\bar{f} : H \to \mathbb{R}$ por

$$\bar{f}(z) := f(z) + \langle u, x - z\rangle.$$

Note que, $\bar{f}$ é convexa, pois se $z, w \in H$ e $\alpha \in [0, 1]$, tem-se

$$\begin{aligned}
\bar{f}(\alpha z + (1-\alpha)w) &= f(\alpha z + (1-\alpha)w) + \langle u, x - \alpha z - (1-\alpha)w\rangle \\
&\leqslant \alpha f(z) + (1-\alpha)f(w) + \langle u, x + \alpha x - \alpha x - \alpha z - (1-\alpha)w\rangle \\
&= \alpha f(z) + (1-\alpha)f(w) + \alpha\langle u, x - z\rangle + (1-\alpha)\langle u, x - w\rangle \\
&= \alpha[f(z) + \langle u, x - z\rangle] + (1-\alpha)[f(w) + \langle u, x - w\rangle] \\
&= \alpha\bar{f}(z) + (1-\alpha)\bar{f}(w).
\end{aligned}$$

Portanto, $\bar{f}$ é uma função convexa.

Além disso, $\partial\bar{f}(z) = \partial f(z) - \{u\} = \{w - u;\ w \in \partial f(z)\}$. Para $z = x$, tem-se $\bar{f}(x) = f(x)$, e tomando $w = u \in \partial f(x)$ implica $0 \in \partial\bar{f}(x)$. Assim, x é minimizador de $\bar{f}$ em H pelo

Teorema 1.8.

Como

$$\langle u - v, x - y \rangle = 0,$$

tem-se

$$\langle u, x - y \rangle = \langle v, x - y \rangle.$$

Sendo $u \in \partial f(x)$ temos

$$\langle u, y - x \rangle \leqslant f(y) - f(x),$$

ou seja,

$$f(x) - f(y) \leqslant \langle u, x - y \rangle.$$

De maneira análoga, como $v \in \partial f(y)$, tem-se

$$\langle v, x - y \rangle \leqslant f(x) - f(y).$$

Portanto,

$$f(x) - f(y) \leqslant \langle u, x - y \rangle = \langle v, x - y \rangle \leqslant f(x) - f(y.) \tag{3.6}$$

Segue de ((3.6) que

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

Logo,

$$f(x) = f(y) + \langle u, x - y \rangle = \bar{f}(y).$$

Mas, $\bar{f}(x) = f(x)$. Então $\bar{f}(x) = \bar{f}(y)$. Assim, y é também um minimizador de $\bar{f}$. Pelo Teorema 1.8, tem-se $0 \in \partial\bar{f}(y)$.

Portanto, $0 = w - u$ para algum $w \in \partial f(y)$, isto é, $u \in \partial f(y)$.

Agora definamos, $\bar{g} : H \to \mathbb{R}$ por

$$\bar{g}(z) := f(z) + \langle v, y - z \rangle.$$

Analogamente a $\bar{f}$, mostra-se que $\bar{g}$ é convexa e $\partial\bar{g}(z) = \partial f(z) - \{v\} = \{w - v; w \in \partial f(z)\}$. Para $z = y$, tem-se $\bar{g}(y) = f(y)$. Tomando $w = v \in \partial f(y)$ implica $0 \in \partial\bar{g}(y)$. Potanto, pelo Teorema 1.8, tem-se que y é um minimizador de $\bar{g}$ em H.

De (3.6 segue que $f(y) = f(x) + \langle v, y - x \rangle = \bar{g}(x)$. Mas, $\bar{g}(y) = f(y)$. Logo, $\bar{g}(y) = \bar{g}(x)$. Assim, x é também um minimizador de $\bar{g}$ em H. Então, pelo Teorema 1.8 tem-se que $0 \in \partial\bar{g}(x)$.

Portanto, $0 = w - v$ para algum $w \in \partial f(x)$. Logo, $v \in \partial f(x)$. Com isso concluímos que, sendo ∂f um operador monótono e $\langle u - v, x - y\rangle = 0$, com $u \in \partial f(x)$,$v \in \partial f(y)$ implica $u \in \partial f(y)$ e $v \in \partial f(x)$. Portanto, $T = \partial f$ é um operador paramonótono. □

Proposição 3.6. *Se* $T_1 : H \rightrightarrows H$ *e* $T_2 : H \rightrightarrows H$ *são operadores paramonótonos, com* $D(T_1) \cap D(T_2) \neq \emptyset$. *Então* $T_1 + T_2$ *é um operador paramonótono.*

Demonstração. Sejam $x, y \in H$, $u \in (T_1 + T_2)(x)$ e $v \in (T_1 + T_2)(y)$. Então, para $u \in (T_1 + T_2)(x) = T_1(x) + T_2(x)$, existem $u_1 \in T_1(x)$ e $u_2 \in T_2(x)$ tais que $u = u_1 + u_2$. Analogamente, existem $v_1 \in T_1(y)$ e $v_2 \in T_2(y)$ tais que $v = v_1 + v_2$. Pela Proposição 3.1 $T_1 + T_2$ é um operador monótono. Ainda mais, se $\langle u - v, x - y\rangle = 0$, então $\langle u_1 - u_2 - v_1 - v_2, x - y\rangle = 0$, ou seja, $\langle u_1 - v_1, x - y\rangle + \langle u_2 - v_2, x - y\rangle = 0$. Sendo T_1 e T_2 monótonos, obtemos que $\langle u_1 - v_1, x - y\rangle = 0$ e $\langle u_2 - v_2, x - y\rangle = 0$. Mas, T_1 e T_2 são paramonótonos, então $u_1, \in T_1(y)$, $u_2 \in T_2(y)$, $v_1 \in T_1(x)$ e $v_2 \in T_2(x)$.
Portanto, $u = u_1 + u_2 \in T_1(y) + T_2(y) = (T_1 + T_2)(y)$ e $v = v_1 + v_2 \in T_1(x) + T_2(x) = (T_1 + T_2)(x)$.
Logo, $T_1 + T_2$ é um operador paramonótono. □

3.4 Operadores Pseudomonótonos

Definição 3.6. *Sejam* H *um espaço de Hilbert e* G *um subconjunto convexo e fechado de* $D(T) \subset H$. *O operador* $T : H \rightrightarrows H$ *é pseudomonótono em* G *se, e somente se, satisfaz a seguinte condição:*
Dada qualquer sequência $\{x^k\} \subset G$, *convergindo fracamente para um elemento* $x^0 \in G$, *e cada sequência* $\{w^k\} \subset H$, *com* $w^k \in T(x^k)\ \forall\ k$, *tal que*

$$\limsup_{k\to\infty} \langle w^k, x^k - x^0\rangle \leqslant 0.$$

Então, para cada $y \in G$ *existe um elemento* $w^0 \in T(x^0)$, *tal que*

$$\langle w^0, x^0 - y\rangle \leqslant \liminf_{k\to\infty} \langle w^k, x^k - y\rangle.$$

A próxima proposição enumera várias condições que garantem a pseudomonotonicidade, mas antes vamos relembrar a definição clássica de semicontinuidade superior.

Definição 3.7. *Sejam* S e S_1 *dois espaços topológicos e a aplicação* $T : S \rightrightarrows S_1$. *Dizemos que* T *é semicontínua superiormente no conjunto dos valores de* S *se para cada ponto* s_0

em S *e cada vizinhança aberta* V *de* $T(s_0)$ *em* S_1*, existe uma vizinhança* U *de* s_0 *em* S *(com* U *dependendo de* V*), tal que* $T(U) \subset V$.

Proposição 3.7. *Seja* $f : H \to \mathbb{R} \cup \infty$ *uma aplicação convexa e semicontínua inferiormente. Considere o operador* $T : H \rightrightarrows H$ *tal que* $T = \partial f$ *então* T *é pseudomonótono de* $D(T)$ *para* H.

Demonstração. Seja $\{x^k\}$ uma sequência em $D(\partial f)$ tal que,

$$\lim_{k\to\infty} x^k = \bar{x} \text{ e } \limsup_{k\to\infty}\langle u^k, x^k - \bar{x}\rangle \leqslant 0,$$

onde $u^k \in \partial f(x^k)$.

Pela definição de subdiferencial, temos $\langle u^k, y - x^k\rangle \leqslant f(y) - f(x^k)$ para todo $y \in D(\partial f)$ e todo $k \in H$. Assim, sendo f convexa e contínua. Então,

$$\liminf_{k\to\infty}\langle u^k, x^k - y\rangle \geqslant \liminf_{k\to\infty} f(x^k) - f(y) = f(\bar{x}) - f(y). \tag{3.7}$$

Da convexidade de f, existe $\bar{u} \in \partial f(\bar{x})$ tal que $f(y) \leqslant f(\bar{x}) + \langle \bar{u}, y - \bar{x}\rangle$. Assim,

$$\langle \bar{u}, \bar{x} - y\rangle \leqslant f(\bar{x}) - f(y). \tag{3.8}$$

Portanto de (3.7) e (3.4), existe $\bar{u} \in \partial f(\bar{x})$ tal que

$$\langle \bar{u}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u^k, x^k - y\rangle.$$

Logo, $T = \partial f$ é pseudomonótono.

□

Obs: Todo operador $T : H \to H$ contínuo é paramonótono.

Proposição 3.8. *Sejam* $T, S : H \rightrightarrows H$ *operadores pseudomonótonos tais que* $D(T) \cap D(S) \neq \emptyset$*. Então* $T + S$ *é pseudomonótono.*

Demonstração. Seja $\{x^k\}$ uma sequência em $D(T + S)$ tal que
$\lim_{k\to\infty} x^k = \bar{x} \in D(T + S)$ e

$$\limsup_{k\to\infty}\langle u^k, x^k - \bar{x}\rangle \leqslant 0, \tag{3.9}$$

onde $u^k \in (T + S)(x^k)$.

Seja $y \in D(T + S)(x^k)$. Então, existem $u_T^k \in T(x^k)$ e $u_S^k \in S(x^k)$ tais que

$$u^k = u_T^k + u_S^k.$$

De (3.9), temos que

$$\limsup_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle \leqslant 0.$$

Então,

$$\limsup_{k\to\infty}\langle u_T^k, x^k - \bar{x}\rangle \leqslant \limsup_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle \leqslant 0,$$

e

$$\limsup_{k\to\infty}\langle u_S^k, x^k - \bar{x}\rangle \leqslant \limsup_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle \leqslant 0.$$

Como T e S são pseudomonótonos, existem $\bar{u_T} \in T(\bar{x})$ e $\bar{u_S} \in S(\bar{x})$ tais que

$$\langle \bar{u_T}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_T^k, x^k - \bar{x}\rangle$$

e

$$\langle \bar{u_S}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_S^k, x^k - \bar{x}\rangle,$$

para todo $t \in D(T+S)$. Logo,

$$\langle \bar{u_T} + \bar{u_S}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle.$$

Seja

$$\bar{u} = \bar{u_T} + \bar{u_S} \in T(\bar{x}) + S(\bar{x}) = (T+S)(\bar{x}).$$

Então,

$$\langle \bar{u}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle,$$

para todo $y \in D(T+S)$. Portanto, $T+S$ é pseudomonótono. $\square$

Capítulo 4

Algoritmo do Ponto Proximal

4.1 Definição do Algoritmo

Sejam C um conjunto fechado e convexo de um espaço de Hilbert H, g uma função de Bregman com zona $C^0 \neq \emptyset$, $\{\lambda_k\}$ uma sequência de números reais positivos limitada superiormente por algum $\bar{\lambda}_k > 0$ e T um operador monótono maximal. Definimos o Algoritmo do Ponto Proximal Generalizado como segue:
i) Inicialização:

$$x^0 \in C^0 \tag{4.1}$$

ii)Etapa iterativa: Dado x^k, defina $T_k : H \rightrightarrows H$ com $T_k(\cdot) = T(\cdot) + \lambda_k D_g(\cdot, x^k)$ e seja $x^{k+1} \in H$ tal que

$$0 \in T_k(x^{k+1}) \tag{4.2}$$

Note que a exitência de x^{k+1} satisfazendo (4.2) não é totalmente imediata, e será assegurada somente com algumas hipóteses sobre T e g.

4.2 Resultados de Existência

Teorema 4.1. *Seja* $\{x^k\}$ *uma sequência dada por (4.1) e (4.2). Suponha que* $D(T) \cap C^0 \neq \emptyset$ *e que* g *é uma aplicação zona coerciva com zona* C^0*, então* $\{x^k\}$ *está bem definida e contida em* S.

Demonstração. A prova será feita por indução em k.
i) Por (4.1), temos que $x^0 \in C^0$.

ii) Suponhamos que $x^k \in C^0$ e devemos mostrar que $x^{k+1} \in C^0$.

De fato, sejam $B_k(\cdot) = \lambda_k \partial D_g(\cdot, x^k)$ e $T_k = T + B_k$. O operador $T_k = TB_k$ é estritamente monótono, pois T é monótono maximal e B_k é estritamente monótono, pela convexidade estrita de g. Daí, pelo lema 3.2.2(ii) $R(T_k) = H$, portanto, T_k tem um zero em $D(T_k)$, o qual é único pela monotonicidade estrita. Chamamos este zero de x^{k+1}. Devemos mostrar que este zero pertence ao interior de C. Pelo lema 2.1.1, $D(T_k) = D(T) \cap C^0$, e como $x^{k+1} \in D(T_k)$, obtemos que $x^{k+1} \in C^0$. Daí, concluimos que $\{x^k\} \subset C^0$. □

Teorema 4.2. *Seja* $\{x^k\}$ *uma sequência dada por* (4.1) *e* (4.2). *Suponha que* $D(T) \cap C^0 \neq \emptyset$. *Além disso, suponha que* $h_{T,C}(x) < \infty$ *(ver definição* A.2*) para todo* $x \in C \cap D(T)$ *e que* g *seja fronteira coerciva. Então* $\{x^k\}$ *está bem definida e contida em* C^0.

Demonstração. Agora não temos a propriedade $R(T_k) = H$, que é essencial no teorema 4.1 e é implicada pela zona coercividade de g. Esta hipótese é substituída pela fronteira coercividade de g e a finitude de $h_{T,C}(x)$ em $C \cap D(T)$. A prova também será feita por indução.

i) Para $k = 0$, temos que $x^0 \in C^0$ por (4.1).

ii) Agora, suponhamos que $x^k \in C^0$,então, existe $x^k \in D(T) \cap C^0$ com $0 \in T_{k-1}(x^k)$. Temos também que, em particular, $x^k \in D(T) \cap C$, e pela definição de $h_{T,C}$ segue-se que $0 \leqslant h_{T,C}(x^k) < \infty$. Para simplificar, usamos a notação $h := h_{T,C}$. Se $h(x^k) = 0$, então por (4.2) temos

$$\langle v, x^k - y \rangle \leqslant 0 \ \forall\, y \in C \text{ e } \forall\, v \in T(y).$$

Tome agora cada u em $T(x^k)$ e cada y C. Pela monotonicidade de T, temos que

$$\langle u - v, x^k - y \rangle \geqslant 0$$

logo,

$$\langle u, y - x^k \rangle \geqslant \langle v, y - x^k \rangle \geqslant 0$$

e então por (1) $x^k \in S^*$. Afirmamos que ao tomarmos $x^{k+1} := x^k$ obtemos $0 \in T_k(x^{k+1})$. De fato, usando o lema 2.1.1,

$$\begin{aligned} T_k(x^{k+1}) &= T(x^k) + \lambda_k \partial D_g(x^k, x^k) \\ &= T(x^k) + \lambda_k(\nabla g(x^k) - \nabla g(x^k)) \\ &= T(x^k) \end{aligned}$$

daí, verificamos que se $x \in D(T) \cap C^0 \cap S^*$ então x é um zero de T. Então neste caso $x^{k+1} = x^k$ também pertence a C^0 pela nossa hipótese de indução.

Então, podemos supor que $h(x^k) > 0$. Neste caso, definimos o conjunto

$$S_k := \{x \in C; D_g(x, x^k) \leqslant \frac{h(x^k)}{\lambda_k}\}.$$

Temos que S_k é fechado, convexo e limitado. Além disso, cada elemento x está no interior de S_k se, e somente se,

$$x \in C^0 \text{ e } D_g(x, x^k) < \frac{h(x^k)}{\lambda_k}. \tag{4.3}$$

Observe que $x^k \in S_k^0$ porque $D_g(x^k, x^k) = 0 < \frac{h(x^k)}{\lambda_k}$ e $x^k \in C^0$ pela hipótese de indução. Seja $N_k := N_{S_k}$, o perador normal de S_k. Ele está bem definido, de tal forma que $D(N_k) = S_k$, o qual, por B_3, é um conjunto limitado. Agora defina $B_k(\cdot) := N_k(\cdot) + \lambda_k \partial_x D_g(\cdot, x^k)$. Precisamos provar que B_k é monótono maximal. Para isto basta verificarmos que $D(N_k) \cap [D(\nabla_x D_g(\cdot, x^k))]^0 \neq \emptyset$. De fato, isto é verdade, pois pela nossa hipótese de indução e pelo lema 2.1.1, temos que $x^k \in D(N_k) \cap C^0 = S_k \cap D(\partial D_g(\cdot, x^k))$. Então $D(N_k) \cap [D(\nabla D_g(\cdot, x^k))]^0 \neq \emptyset$. Desta forma, concluimos que B_k é monótono maximal. Além disso, como $D(B_k)$ é limitado pois é, em particular, um subconjunto de S_k, pois $D(N_k) = S_k$, daí, usamos a proposição 3.5 para concluirmos que $R(B_k) = H$. Vamos considerar agora o operador $A_k := T + B_k$. Este operador também é monótono maximal. Para provarmos que A_k é monótono maximal, pela proposição 3.2, è suficinte mostrarmos que $D(T) \cap (D(B_k))^0 \neq \emptyset$. De fato, $D(T) \cap (D(B_k))^0 \neq \emptyset$, pois

$$x^k \in D(T) \cap C^0 \cap S_K^0 = D(T) \cap D(B_k)^0.$$

Logo A_k é monótono maximal, pela proposição 3.2. Agora usando a proposição (3.5) obtemos que A_k é sobrejetivo, porque o seu domínio é um subconjunto de $D(B_k)$, o qual é limitado. Então existe $y \in D(A_k) = D(T) \cap D(B_k) \subset D(T) \cap C^0$, tal que

$$0 \in T(y) + N_k(y) + \lambda_k D_g(y, x^k). \tag{4.4}$$

Sejam u^k, w^k e v^k elelmentos em H tais que

$$u^k \in T(y),\ w^k \in N_k(y),\ v^k \in \lambda_k \partial D_g(y, x^k)$$

e,

$$0 = u^k + w^k + v^k. \tag{4.5}$$

Afirmamos que $y \in S_k^0$. Além disso $y \in D(A_k) \subset D(B_k) \subset (\partial_x D_g(\cdot, x^k)) = C^0$ pelo lema 2.1.1. Como $x \in C^0 \Leftrightarrow D_g(x, x^k) < \frac{h(x^k)}{\lambda_k}$ é suficiente mostrarmos que

$$D_g(y, x^k) < \frac{h(x^k)}{\lambda_k}, \tag{4.6}$$

Vamos provar (4.6). Aplicando a definição de subdiferencial para a aplicação $D_g(\cdot, x^k)$, a qual é estritamente convexa por (2.1)(iii), obtemos

$$0 = D_g(x^k, x^k) > D_g(y, x^k) + \langle \frac{v^k}{\lambda_k}, x^k - v \rangle. \tag{4.7}$$

Usando (4.5) em (4.6) e reorganizando os termos, encontramos

$$\frac{1}{\lambda_k}(\langle u^k, x^k - y \rangle + \langle w^k, x^k - y \rangle) > D_g(y, x^k). \tag{4.8}$$

Como $w^k \in N_k(y)$, e $x^k \in S_k$, temos que

$$\langle w^k, x^k - y \rangle \leqslant 0 \tag{4.9}$$

Usando (4.9) em (4.8) obtemos

$$0 \leqslant \lambda_k D_g(y, x^k) < \langle u^k, x^k - y \rangle. \tag{4.10}$$

Como $\langle u^k, x^k - y \rangle \leqslant \sup_{z \in C \cap D(T), v \in Tz} \langle v, x^k - z \rangle = h(x^k)$, obtemos de (4.10) que $D_g(y, x^k) < \frac{h(x^k)}{\lambda_k}$, isto é, a desigualdade (4.6) está provada. Desta forma, provamos que $y \in S_k^0$ e além disso, $N_k(y) = 0$. Então $w^k = 0$ e $0 = u^k + v^k$ por (4.5). De (4.4), $0 \in T_k(y)$. Pela monotonicidade estrita de T_k, y é o único zero de T_k. Por (4.2), $y = x^{k+1}$. Desta forma, provamos que $y \in C^0$. Assim $x^{k+1} \in C^0$ e a etapa de indução está completa.

□

4.3 Resultados de Convergência

Nesta seção estabeleceremos propriedades de convergência da sequência gerada por (4.1) e (4.2).

Teorema 4.3. *Seja* $T : H \rightrightarrows H$ *um operador monótono maximal. Considere o PDV(T,C), onde* C *é um subconjunto convexo e fechado de* H*. Seja* g *uma função de Bregman com zona* C^0 *e suponha que sejam válidas as seguintes condições*

i) $D(T) \cap C^0 \neq \emptyset$

ii) $S^* \neq \emptyset$

iii) T *é pseudomonótono*

iv) $\lambda_k \in (0, \bar{\lambda})$ *para algum* $\bar{\lambda} > 0$

v) g *é zona coerciva, ou* g *é fronteira coerciva e* $h_{T,C} < \infty \ \forall\ x \in C \cap D(T)$.

*Então a sequência gerada por (*4.1) e (4.2) *satisfaz:*

a) A sequência $\{D_g(z, x^k)\}$ *é não crescente* $\forall\ z \in S^*$.

b) $\{x^k\}$ *é limitada e tem pontos fracos de acumulação.*

c) $\lim_{k\to\infty} D_g(x^{k+1}, x^k) = 0$.

d) Se $\bar{x}$ *é um ponto fraco de acumulação de* $\{x^k\}$*, então existe* $\bar{u} \in T(\bar{x})$ *tal que* $\langle \bar{u}, x^* - \bar{u}\rangle = 0$.

Demonstração. a) Devemos mostrar que

$$D_g(z, x^{k+1}) \leqslant D_g(z, x^k) - D_g(x^{k+1}, x^k)\ \forall\ z \in S^*. \tag{4.11}$$

Dado $z \in S^*$ $\exists v^* \in T(z)$ tal que

$$\langle v^*, x - z\rangle \geqslant 0 \text{ para todo } x \in C. \tag{4.12}$$

Pelos teoremas 4.1 e 4.2, usando o lema 2.1.1 e (4.2), existe $x^{k+1} \in C^0$ e $u^k \in T(x^{k+1})$ tal que

$$u^k = \lambda_k(\nabla g(x^k) - \nabla g(x^{k+1})). \tag{4.13}$$

Segue diretamente de (2.1)(i) que a igualdade abaixo vale para todo $y \in C$

$$\langle \nabla g(x^k) - \nabla g(x^{k+1}), x^{k+1} - y\rangle = D_g(y, x^k) - D_g(y, x^{k+1}) - D_g(x^{k+1}, x^k). \tag{4.14}$$

Por (4.13) e (4.14), temos

$$\langle u^k, x^{k+1} - y\rangle = \lambda_k \left(D_g(y, x^k) - D_g(y, x^{k+1}) - D_g(x^{k+1}, x^k)\right). \tag{4.15}$$

Agora tome $z \in S^*$. Por (4.12) e pela monotonicidade de T, encontramos

$$\langle u^k, x^{k+1} - z\rangle \geqslant \langle v^*, x^{k+1} - z\rangle \geqslant 0. \tag{4.16}$$

Assim (4.11) segue de (4.15) com $y = z$ e de (4.16). Por (4.12) e a não negatividade de D_g,

$$D_g(z, x^{k+1}) \leqslant D_g(z, x^k) - D_g(x^{k+1}, x^k). \tag{4.17}$$

e assim, provamos o item (a).

b) Iterando (4.17) obtemos $D_g(z, x^k) \leqslant D(z, x^0)\ \forall\ k$ e $\forall\ z \in S^*$. Agore fixe algum $z \in S^*$.

Logo, $\{x^k\}$ está contida no conjunto $\{x \in C; D_g(z,x) \leqslant D_g(z,x^0)\}$, o qual, por B_3, é limitado. Como $\{x^k\}$ é limitada, a existência de pontos fracos de acumulação segue do Teorema de Bourbaki-Alouglu.

c) Precisamos provar que

$$\lim_{k\to\infty} D_g(x^{k+1},x^k) = 0 \tag{4.18}$$

Pela desigualdade (4.11) temos que

$$0 \leqslant D_g(x^{k+1},x^k) \leqslant D_g(z,x^k) - D_g(z,x^{k+1}), \tag{4.19}$$

para cada $z \in S^*$. Pelo item (a), $\{D_g(z,x^k)\}$ é uma sequência não crescente, limitada inferiormente e por isso convergente. Este fato e (4.19) implicam que $\lim_{k\to\infty} D_g(x^{k+1},x^k) = 0$. Portanto o item (c) está provado.

d) Agora suponha que $\{x^k\}$ tenha um ponto fraco de acumulação. Primeiro, devemos observar que cada ponto fraco de acumulação de $\{x^k\}$ está em C. De fato, para cada conjunto convexo $A \subset H$, temos

$$\overline{A} = \overline{A}^w \tag{4.20}$$

onde $\overline{A}^w$ denota o fecho fraco de A (para mais detalhes ver[9]). Aplicando (4.20) em C e observando que C é fechado, temos que $C = \overline{C}^w$. Então, desta forma, cada ponto fraco de acumulação pertence ao fecho fraco de C. Assim, concluimos que este ponto fraco de acumulação também pertence a C. Agora seja $\bar{x}$ um ponto fraco de acumulação e $\{x^{k_j}\} \subset \{x^k\}$ tal que

$$x^{k_j} \rightharpoonup \bar{x}.$$

Da mesma forma que em (a), obtemos de (4.14), para $y = \bar{x}$ e $k = k_j$, que

$$\langle u^{k_j}, x^{k_j+1} - \bar{x}\rangle = \lambda_{k_j}\left(D_g(\bar{x},x^{k_j}) - D_g(\bar{x},x^{k_j+1}) - D_g(x^{k_j+1},x^{k_j})\right). \tag{4.21}$$

Agora vamos usar as condições B_4 e B_5 para provar que o lado direito de (4.21) tende a zero quando $j \to \infty$. Temos que

d_1) $\{x^{k_j}\}$ é limitada, pelo item (a),

d_2) $x^{k_j} \rightharpoonup \bar{x}$ por (4.7),

d_3) $\lim_{k\to\infty} D_g(x^{k_j+1},x^{k_j}) = 0$ pelo item (b). Então, aplicando B_5 para as sequências $\{x^{k_j+1}\}$ e $\{x^{k_j}\}$, obtemos

$$x^{k_j+1} \rightharpoonup \bar{x}. \tag{4.22}$$

Vemos que de (4.22), (d_2) e (d_3) podemos aplicar B_4 para as sequências $\{x^{k_j+1}\}$ e $\{x^{k_j}\}$ para obtermos

$$\lim_{j\to\infty}\left(D_g(\bar{x},x^{k_j+1})-D_g(\bar{x},x^{k_j})\right)=0. \tag{4.23}$$

Como, por (iv), $\{\lambda_k\}$ é limitada superiormente, então, se usarmos (d_3) e (4.23) em (4.21), obtemos

$$\lim_{j\to\infty}\langle u^{k_j},x^{k_j+1}-\bar{x}\rangle=0. \tag{4.24}$$

Agora podemos usar a hipótese da pseudomonotonicidade. De fato,

$$x^{k_j+1}\rightharpoonup\bar{x} \text{ e } u^{k_j}\in T(x^{k_j+1})\ \forall\, j \tag{4.25}$$

A hipótese da pseudomonotonicidade, (4.24) e (4.25) implicam que para cada $z\in S^*$ existe $\bar{u}\in T(\bar{x})$ tal que

$$\langle\bar{u},\bar{x}-z\rangle\leqslant\liminf_{j\to\infty}\langle u^{k_j},x^{k_j+1}-z\rangle. \tag{4.26}$$

Por (4.14) com $y=z\in S^*$, obtemos

$$\langle u^k,x^{k+1}-z\rangle=\lambda_k\left(D_g(z,x^k)-D_g(z,x^{k+1})-D_g(x^{k+1},x^k)\right). \tag{4.27}$$

Por (a) a sequência $D_g(z,x^k)$ é decrescente e então o lado direito de (4.27) tende a zero porque $\{\lambda_k\}$ é limitada superiormente, por (iv). Assim, o limite do lado direito de (4.26) é zero. Tomando limites em (4.26) com $j\to\infty$, temos

$$\langle\bar{u},\bar{x}-z\rangle\leqslant 0, \tag{4.28}$$

com $\bar{u}\in T(\bar{x})$. Como $z\in S^*$, existe $v^*\in T(z)$ tal que $\langle v^*,y-z\rangle\geqslant 0$ para todo $y\in C$. Além disso, pela monotonicidade de T,

$$\langle\bar{u},\bar{x}-z\rangle\geqslant\langle v^*,\bar{x}-z\rangle\geqslant 0. \tag{4.29}$$

As desigualdades (4.28) e (4.29) implicam que $\langle\bar{u},\bar{x}-z\rangle=0$ e assim, provamos o item (d). □

Observamos que a hipótese da pseudomonotonicidade no Teorema (4.3) implica, considerando a definição 3.6, que $D(T)$ deve ser fechado e convexo.

Agora, assumimos que a norma em C é compatível e que T é paramonótono em C. Provaremos que a sequência gerada pelo APPG é fracamente convergente para a solução do PDV(T,C).

Teorema 4.4. *Seja* $T : H \rightrightarrows H$ *monótono maximal. Considere o PDV(T,C), onde* C *é um subconjunto convexo e fechado de* H. *Seja* g *uma função de Bregman com zona* C^0, *isto é, satisfaz* $B_1 - B_5$ *e suponha que valem as seguintes condições:*

i) $D(T) \cap C^0 \neq \emptyset$.

ii) $S^* \neq \emptyset$.

iii) T *é pseudomonótono de* $D(T)$ *para* H.

iv) $\lambda_k \in (0, \bar{\lambda}]$ *para algum* $\bar{\lambda} > 0$.

v) g *é zona coerciva, ou* g *é fronteira coerciva e* $h_{T,C}(x) < \infty \ \forall\, x \in C \cap D(T)$.

vi) T *é paramonótono em* C.

Então a sequência $\{x^k\}$ *definida por* (4.1) e (4.2) *é fracamente convergente para uma solução* $\bar{x}$ *do PDV(T,C).*

Se além disso g *é norma compatível então a sequência* $\{x^k\}$ *tem um único ponto fraco de acumulação.*

Demonstração. Pelo Terema 4.3(b) a sequência $\{x^k\}$ tem pontos fracos de acumulação. Seja $\bar{x}$ um ponto fraco de acumulação da sequência $\{x^k\}$. Pelo Teorema 4.3(d), existe $\bar{u} \in T(\bar{x})$ tal que

$$\langle \bar{u}, x^* - \bar{x} \rangle = 0 \ \forall\, x^* \in S^*. \tag{4.30}$$

Sendo x^* uma solução do PDV(T,C), existe $u^* \in T(x^*)$ tal que

$$\langle u^*, x - x^* \rangle \geqslant 0 \ \forall\, x \in C. \tag{4.31}$$

Em particular para $\bar{x}$ tem-se

$$\langle u^*, \bar{x} - x^* \rangle \geqslant 0. \tag{4.32}$$

Pela monotonicidade de T, temos para $\bar{u} \in T(\bar{x})$ e $u^* \in T(x^*)$ que

$$\langle \bar{u} - u^*, \bar{x} - x^* \rangle \geqslant 0.$$

Logo,

$$0 = \langle \bar{u}, \bar{x} - x^* \rangle \geqslant \langle u^*, \bar{x} - x^* \rangle \geqslant 0, \tag{4.33}$$

daí,

$$\langle \bar{u}, \bar{x} - x^* \rangle \geqslant \langle u^*, \bar{x} - x^* \rangle = 0, \tag{4.34}$$

portanto,

$$\langle \bar{u} - u^*, \bar{x} - x^* \rangle = 0 \tag{4.35}$$

com $\bar{u} \in T(\bar{x})$ e $u^* \in T(x^*)$.

Pela paramonotonicidade de T, obtemos

$$\bar{u} \in T(x^*) \text{ e } u^* \in T(\bar{x}).$$

Portanto, usando 4.31 e 4.32 temos

$$\langle u^*, x - \bar{x}\rangle = \langle u^*, x - x^*\rangle + \langle u^*, x^* - \bar{x}\rangle \geqslant 0.$$

Logo $\bar{x}$ é uma solução do PDV(T,C). Como $\bar{x}$ é um ponto fraco de acumulação da sequência $\{x^k\}$, concluimos que todos os pontos de $\{x^k\}$ são soluções do PDV(T,C).

Provaremos que se g é norma compatível, então existe somente um ponto de acumulação fraco. Pelo Teorema 4.3(a), para cada $z \in S^*$ fixo, temos que a sequência $\beta_k(z) := \langle \nabla g(x^k), x^k - z\rangle - g(z)$ é decrescente e $\beta_k(z) \geqslant -g(z)$ porque $D_g(z, x^k) \geqslant 0$. Então a sequência $\{\beta_k(z)\}$ é convergente. Suponha que existam dois pontos mínimos diferentes, que são pontos fracos de acumulação, digamos $z_1, z_2 \in S^*$. Tome subsequências $\{x^{j_k}\}$ e $\{x^{i_k}\}$ de $\{x^k\}$ tais que $x^{j_k} \rightharpoonup z_1$ e $x^{i_k} \rightharpoonup z_2$. Pela convergência das sequências $\{\beta_k(z_1)\}$ e $\{\beta_k(z_2)\}$, existem $l_1, l_2 \in \mathbb{R}$ tais que

$$\lim_{k\to\infty} \beta_k(z_1) = l_1 \text{ e } \lim_{k\to\infty} \beta_k(z_2) = l_2.$$

Então

$$\lim_{k\to\infty} (\beta_k(z_1) - \beta_k(z_2)) = \lim_{k\to\infty} \langle \nabla g(x^k), z_2 - z_1\rangle = l_1 - l_2.$$

Consideremos esta última igualdade para $k = j_k$ e depois para $k = i_k$. Subtraindo os resultados destas expressões após as substituições, obtemos

$$\lim_{k\to\infty} \langle g(x^{j_k}) - \nabla g(x^{i_k}), z_2 - z_1\rangle = 0.$$

Visto que $z_1 \neq z_2$, esta última igualdade contradiz o fato da norma compatibilidade de g. Também existe um único ponto de acumulação da sequência $\{x^k\}$ na topologia fraca e desta forma a sequência inteira converge para ele. Portanto, este limite fraco e a sequência inteira resolvem o PDV(T,C). □

Vamos fechar esta seção mostrando que se o conjunto solução S^* do PDV(T,C) é vazio então a sequência $\{x^k\}$ gerada pelo APPG é não limitada inferiormente. Isto significa que a existência de soluções do PDV(T,C) é condição necessária para a convergência da sequência $\{x^k\}$. Mas antes, precizaremos de um resultado preliminar.

Proposição 4.1. *Se* V *é um conjunto convexo e fechado então o operador normal* N_V *é pseudomonótno em* $D(N_V)$.

Demonstração. Tome $\{x^k\}, \{w^k\}$,x^0 e y como na definição 3.6. Como $x^k, y \in D(N_V) = V$, segue da definição de N_V que $\langle w^k, x^k - y\rangle \geqslant 0$. Por outro lado, visto que $N_V(x^0)$ é um cone positivo, segue que $0 \in N_V(x^0)$. Tome $w^0 = 0$ para concluir que $\langle w^0, x^0 - y\rangle = 0 \leqslant \liminf_{k\to\infty}\langle w^k, x^k - y\rangle$. □

Lema 4.3.1. *Sob as hipóteses do Teorema* 4.3 *e assumindo que* T *é paramonótono em* C, *seja* $\{x^k\}$ *a sequência gerada pelo APPG. Se* $S^* = \emptyset$, *então* $\{x^k\}$ *é não limitada.*

Demonstração. Suponha que $\{x^k\}$ seja limitada. Assim, existe um conjunto limitado e convexo $B \subset H$ tal que $\overline{\{x^k\}} \subset B^0$. Considere o operador $\bar{T} = T + N_B$, onde N_B é o operador normal associado com B. Relembrando que $N_B(x) = \{0\}\ \forall\ x \in B^0$. Seja $\{\bar{x}^k\}$ a sequência gerada pelo APPG para o operador $\bar{T}$ com $\bar{x}^0 = x^0$. Provaremos por indução que $x^k = \bar{x}^k\ \forall\ k$.

i) Para $k = 0$ a afirmação segue.

ii) Agora, suponha, por hipótese que $\bar{x}^k = x^k$

Por (4.2) temos que

$$0 \in T(x^{k+1}) + \lambda_k \partial_x D_g(x^{k+1}, x^k) = \bar{T}(x^{k+1}) + \lambda_k \partial_x D_g(x^{k+1}, \bar{x}^k)$$

usando o fato de que $T = \bar{T}$ em B^0 e a hipótese de indução. Então x^{k+1} é uma solução da equação em y,

$$0 \in \bar{T}(y) + \lambda_k \partial_x D_g(y, x^k) \tag{4.36}$$

Pela monotonicidade estrita do operador $\bar{T}(y) + \lambda_k \partial_x D_g(\cdot, x^k)$, a solução y de (4.36) é única. Por (4.1) $\bar{x}^{k+1}$ é a solução de $0 \in \bar{T}(y) + \lambda_k D_g(y, x^k)$, daí, concluímos que $\bar{x}^{k+1} = x^{k+1}$. Desta forma, a etapa de indução está completa.

Vamos verificar agora as hipóteses do nosso lema para o PDV($\bar{T}$,C). Podemos verificar, usando o Teorema 3.2 que $\bar{T}$ é monótono maximal. Pelo Teorema 3.1 e pela proposição 3.6 o operador $\bar{T} = T + \partial(\delta_B)$ (ver definição é paramonótono e pela proposição 4.1 temos que N_B (ver definição 1.11) é pseudomonótono. Deste modo, $\bar{T}$ é a soma de dois operadores pseudomonótonos, e portanto ele é pseudomonótono, pela proposição 3.8. Finalmente, verificaremos que PDV($\bar{T}$,C) tem soluções. O conjunto das soluções do PDV($\bar{T}$,C) é o conjunto dos zeros de $\bar{T} + N_C = T + N_B + N_C$. Note que $D(\bar{T} + N_C) = D(T) \cap C \cap B \subset B$.

Como B é limitado, temos que $\bar{\mathsf{T}} + \mathsf{N}_C$ é sobrejetivo pela proposição 3.5 e portanto tem zeros. logo, podemos concluir que $\{\bar{x}^k\}$, e portanto $\{x^k\}$ possuem pontos fracos de acumulação e todos eles são soluções do PDV($\bar{\mathsf{T}}$,C). Seja x^* um ponto de acumulação de $\{x^k\}$. Segue da definição de $\{x^k\}$ que $x^* \in B^0$. Visto que $\mathsf{T}(x) = \bar{\mathsf{T}}(x)\ \forall\ x \in B^0$, obtemos $\mathsf{T}(x^*) = \bar{\mathsf{T}}(x^*)$ e então segue de (1) que x^* também é uma soluçaõ do PDV(T,C) em contradição com a nossa hipótese. Concluimos, desta forma, que $\{x^k\}$ é não limitada. □

Os nossos resultados podem ser resumidos no seguinte teorema:

Teorema 4.5. *Assumindo que sejam válidas as hipóteses do Teorema* 4.4. *Seja* $\{x^k\}$ *a sequência gerada pelo APPG. Então*

i) $\{x^k\}$ *é fracamente convergente se, e somente se,* $S^* \neq \emptyset$.

ii) Se $S^* \neq \emptyset$ *então o limite fraco* x^* *de* $\{x^k\}$ *pertence a* S^*.

iii) Se $S^* = \emptyset$ *então* $\{x^k\}$ *é não limitada.*

Demonstração. i) $S^* \neq \emptyset$, então o resultado segue do Teorema 4.4

Se a sequência converge então ela é limitada e segue do lema 4.3 que $S^* \neq \emptyset$.

ii) Segue diretamente do Teorema 4.4.

iii) Segue do lema 4.3. □

Apêndice A

Função Indicadora e Função Gap

A.1 Função Indicadora

Definição A.1. *Seja* $C \subset H$ *um conjunto convexo e fechado. A aplicação* $\delta_C : C \to \mathbb{R} \cup \{\infty\}$ *tal que*

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

é chamada de função indicadora

Proposição A.1. *A função indicadora definida acima é convexa se, e somente se,* C *é um conjunto convexo.*

Demonstração. De fato, se $x \notin V$, tem-se $\delta_C = \infty$. Assim, para todo $d \in \mathbb{R}$, tem-se que $d < \infty = \delta_C(x)$. Logo, $(x, d) \notin E_{\delta_C}$.

Se $x \in C$, tem-se $\delta_C(x) = 0$. Então, $\delta_C(x) \leqslant d$, $\forall\ d \in [0, , +\infty)$. Portanto, $(x, d) \in E_{\delta_C}$.

Logo, $E_{\delta_C} = C \times [0, +\infty)$.

Daí, se δ_C é convexa temos que E_{δ_C} é convexo. portanto $C \times [0, +\infty)$ é convexo. Logo, C é convexo.

Reciprocamente, se C é convexo, temos que E_{δ_C} é convexo. Logo, δ_C é convexa.

Se $x \in C$, tem-se $\delta_C(x) = 0$. Então, $\delta_C(x) \leqslant d$, $\forall\ d \in [0, +\infty)$. □

Agora vamos calcular o subidiferencial da função Indicadora.

Seja $s \in \partial f(x)$, então.

i)Se $x \notin C$, tem-se $\delta_C(x) = \infty$, daí

$$\delta_C(y) \geqslant \infty,\ \forall\ y \in H.$$

O que é um absurdo. Logo, $\partial\delta_C(x) = \emptyset$.

ii) Se $x \in C$, tem-se $\delta_C(x) = 0$, então

$$\delta_C(y) \geqslant \langle s, y - x\rangle, \ \forall \ y \in \ H.$$

Daí, considere os dois casos.

1) Se $y \notin C$, obtemos

$$\langle s, y - x\rangle \leqslant +\infty.$$

2) Se $y \in C$, tem-se $\delta_C(y) = 0$, assim

$$\langle s, y - x\rangle \leqslant 0.$$

Portanto, por 1) e 2), obtemos

$$\partial\delta_C(x) = \{s \in \ H; \langle s, y - x\rangle \leqslant 0, \ \forall \ y \in C\}.$$

Logo,

$$\partial\delta_C(x) = \begin{cases} \{s \in \ H; \langle s, y - x\rangle \leqslant 0, \ \forall \ y \in C\}, & \text{se} \quad x \in C \\ \emptyset, & \text{se} \quad x \notin C \end{cases}$$

Pela definição 1.11 e pelo cálculo do subdiferencial de δ_C vemos que

$$N_C = \partial\delta_C, \tag{A.1}$$

ou seja, o operador normal de C, $N_C : H \to H$, é igual ao conjunto dos subgradientes de δ_C em x, de forma mais precisa, temos $N_C = \emptyset$ se $x \notin C$, $N_C = 0$ se $x \in C^0$ e que $N_C(x)$ é um cone não vazio.

Como o conjunto

$$D(T) = \{x \in H; Tx \neq \emptyset\}$$

é o domínio do operador monótono maximal $T : H \rightrightarrows H$. Combinando estes resultados, temos que

$$D(N_C) = C.$$

e que N_C é um operador monótono maximal.

A.2 Função Gap

Agora vamos introduzir uma função que será utilizada a seguir.

Definição A.2. *Seja* $T : H \rightrightarrows H$ *um operador monótono maximal e* $C \neq \emptyset$ *um subconjunto convexo e fechado de* H *tal que* $D(T) \cap C \neq \emptyset$. *Definimos a função*

$$h_{T,C} : D(T) \cap C \longrightarrow \mathbb{R} \cup \{\infty\}$$

com

$$h_{T,C} := \sup_{(x,y) \in G_C(T)} \langle v, x - y \rangle, \tag{A.2}$$

onde $G_C(T) = \{(v, y); v \in T(y), y \in C\}$, denota o gráfico de T.
A função $h_{T,C}$ definida acima é chamada de função gap
Esta função tem uma propriedade interessante, para a qual descrevemos o seguinte lema:

Lema A.2.1. *Sejam* T, C e $h_{T,C}$ *dados pela definição acima. Então:*
i) $h_{T,C}$ *é convexa em* $D(T) \cap C$.
ii) $h_{T,C} \geqslant 0$ *em* $D(T) \cap H$.

Demonstração. i) $h_{T,C}$ é o supremo de uma família de transformações afins, que são, em particular, funções convexas, e o supremo de uma família de funções convexas é sempre uma função convexa.
ii) Basta tomarmos $x = y$ em (A.2). □

Lema A.2.2. *Considere* $h_{T,c}$ *como em (A.2). Suponha que* H *seja um espaço de Hilbert arbitrário e que* $D(T) \cap C^0 \neq \emptyset$. *Então* $h_{T,C}$ *tem zeros se, e somente se,* S^* *é não vazio e em tal caso* S^* *é precisamente o conjunto dos zeros de* $h_{T,C}$.

Demonstração. Primeiro iremos verificar que um zero de $h_{T,C}$ é uma solução do PDV(T,C). Por simplicidade denotaremos $h = h_{T,C}$. Seja $A := T + N_C$, tal que $D(A) = D(T) \cap C$. Se $h(x) = 0$, então por (A.2), temos que para cada $y \in D(T) \cap C,\ v \in Ty$,

$$\langle v, x - y \rangle \leqslant 0 \tag{A.3}$$

Por outro lado, se tomarmos $w \in N_C(x)$, então, para cada $y \in C \cap D(T)$,

$$\langle w, x - y \rangle \leqslant 0, \tag{A.4}$$

onde usamos a definiçao de N_C e assumimos que $x \in C$. Adicionando (A.3) e (A.4), obtemos

$$\langle 0 - (v + w), x - y \rangle \geqslant 0, \tag{A.5}$$

para cada $y \in D(T) \cap C = D(A)$ e cada $v + w \in A(y) = T(y) + N + C$. Pelo teorema 3.2 segue que A é monótono maximal. De (A.5) segue que

$$0 \in A(x) = T(x) + N_C(x), \tag{A.6}$$

e (A.7) mostra que $x \in S^*$.

Reciprocamente suponha que $x \in C \cap S^*$. Então, para cada $y \in C$, existe $v \in Tx$ com

$$\langle v, y - x \rangle \geqslant 0. \tag{A.7}$$

Usando esta última desigualdade e a monotonicidade, temos que para cada $w \in Ty$

$$\langle w, y - x \rangle \geqslant 0. \tag{A.8}$$

Segue de (A.8) que

$$h(x) = \sup_{(y,w) \in G_C(T)} \langle w, x - y \rangle \leqslant 0. \tag{A.9}$$

A equação (A.9) e o item (ii) do lema A.2.2 implicam que $h(x) = 0$, como queríamos mostrar. $\square$

Exemplo A.1. *Seja C um conjunto convexo fechado e com interior não vazio. Vemos que a função gap para $T = N_C$ é exatamente δ_C. Realmente, seja $x \in C = D(T)$. Então*

$$0 \leqslant h_{N_C,C}(x) = \sup_{(y,v) \in G_C(N_C)} \langle v, x - y \rangle \leqslant 0,$$

onde usamos a definição de N_C na última desigualdade. Assim, $h_{N_C}(x) = 0$ para todo $x \in C$. Agora seja $x \notin C$. Então, como $C^0 \neq \emptyset$, existe um hiperplano suporte não trivial dado para cada $\alpha \neq 0$ tal que

A) $\langle a, y - y_0 \rangle \leqslant 0$ para todo $y \in C$ e

B) $\langle a, x - y_0 \rangle > 0$. De (A) concluimos que $\lambda a \in N_C(y_0)$ para todo $\lambda > 0$. Então

$$h_{N_C,C}(x) = \sup_{(y,v) \in G_C(N_C)} \langle v, x - y \rangle \geqslant \lambda \langle a, x - y_0 \rangle.$$

Usando (B) e fazendo $\lambda \to \infty$ concluimos que $h_{N_C,C}(x) = \infty$ para $x \notin C$. Portanto, temos $h = \delta_C$.

Referências Bibliográficas

[1] Acker, F., Dickstein, F., Uma Introdução à análise Convexa. 14º Colóquio Brasileiro de Matemática. *Instituto de Matemática Pura e Aplicada*, Rio de Janeiro, RJ, (1983).

[2] Aubin, J.P, Applied Abstract Analysis. A Wiley Intercience Publication. França (1977).

[3] Brézis, H., Haraux, A., Image d'une som d'opérateurs monotones et applications. *Israel Journal of Mathematics* (1976)

[4] Brezis, H., Analyse Fonctionnelle Théorie et Applications. Masson, Paris (1983).

[5] Burachik, R.S, Generalized Proximal Point Methods for the Variational Inequality Problem. PhD Thesis, *Instituto de Matemática Pura e Aplicada*, Rio de Janeiro, RJ, (1995).

[6] Burachik, R.S., Iusem, A.N., A Generalized Proximal Point Algorithm for Variational Inequality Problem in a Hilbert Space. *Instituto de Matemática Pura e Aplicada*, Rio de Janeiro, RJ (1991).

[7] Cioranescu, I., Geometry of Banach Spaces, Duality Mappings an Nonlinear Problems. Kluwer Academic Publishers. Holanda (1990).

[8] Ekeland, I., Turnbull, T., Infinite Dimensional Otimization and Convexity. The Univerity of Chicago Press. EUA (1983).

[9] Figueiredo, D.G., Equações Elípticas não Lineares, 11º Colóquio Brasileiro de Matemática. *Instituto de Matemática Pura e Aplicada do C.N.Pq.*. IMPA. 1977.

[10] Friedman, Avner. Foundations of Modern Analysisy of Congress Cataloging in Publication Data. New York (1982).

[11] Iusem, A.N, Métodos de Ponto Proximal. 20º Colóquio Brasileiro de Matemática. Impa, RJ, (1995).

[12] Minty, G.J., Monotone Nonlinear Operators in Hilbert Space. Duke Mathematical Journal (1978).

[13] Moreau, J., Funtionnelles convexes, in Seminaire sur les equations aux derives partielles, Paris College de France, (1966).

[14] Rockafellar, R.T., Convex Analysis. *Princeton University Press*, New Jersey, (1970).

[15] Rockafellar, R.T., On the Maximality of Sums Nonlinear Monotone Operators. *Transations of the American Mathematical Society* 149 (1976) 887-898.

[16] Van Tiel, J, Wiley, J., Convex Analysis, an Introdutory Text. New York, (1984).