Universidade Federal do Piauí

Centro de Ciências da Natureza

Pós-Graduação em Matemática

Mestrado em Matemática

Algoritmo do Ponto Proximal Generalizado para o Problema de Desigualdade Variacional em $\mathbb{R}^n$

João Santos Andrade

Teresina - 2010

João Santos Andrade

Dissertação de Mestrado:

Algoritmo do Ponto Proximal Generalizado para o Problema de Desigualdade Variacional em $\mathbb{R}^n$

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

A553a

Andrade, João Santos.
Algoritmo do ponto proximal generalizado para o problema de desigualdade variacional R^n/ João Santos Andrade. Teresina: 2010.
68 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

Algoritmo do Ponto Proximal Generalizado para o Problema de Desigualdade Variacional em $\mathbb{R}^n$

Teresina - 2010

Dedico este trabalho a meus pais Nivaldo e Maria Francisca, a minha esposa Elcilene, a meus irmãos e irmãs (Franklivaldo, Liana, Walkiria, Niviane, Vânia, Liziane e Nivaldo Filho), a meu avô Raimundo e minha avó Angelita (In memorian).

Agradecimentos

Agradeço primeiramente a Deus pela vida e por ter me dado força para vencer as barreiras.

Aos meus pais Nivaldo e Maria Francisca pelo apoio e incentivo nessa conquista. A meus irmãos, irmãs, tios, primos e meu avô Raimundo que sempre acreditaram em mim.

A minha esposa Elcilene pelo apoio, carinho, amor, paciência e incentivo.

A todos os meus amigos em especial Luís Carlos (irmão), Aurineide (grande incentivadora), Claudio, Márcia, Julimar, Everaldo, Sulimery, Hélson, Jociel, Edivan, Fábio, Sandra, Diego (Parnaíba), Hudson, Anderson, Yuri, Rogério, Romeu, Diego (Barão), Déborah e Marciana pela amizade e incentivo. E todos os amigos do mestrado Gilberto, Daniel, Venâncio, Italo, Pedro, Domingos, João Carlos, Arimatéia, Renan e Cleyton pela amizade, solidariedade e companheirismos.

Ao professor Jurandir de Oliveira Lopes pela escolha do tema, orientação, paciência e compreensão.

Ao professor João Xavier pelas orientações, sugestões e incentivo. Aos professores Rolando Garciga e Sissy por terem participado da banca e pelas correções e sugestões.

A todos os professores e funcionários do departamento de matemática.

A CAPES pelo apoio financeiro.

"*O senhor é meu pastor: nada me faltará. Deitar-me faz me em verdes pastos, guia-me mansamente a águas tranquilas. ...* ".

Salmo 23.

Resumo

Nesta dissertação, inicialmente apresentamos o Algoritmo do Ponto Proximal sem restrições para encontrar zeros de Operadores Monótonos Maximais (com a norma euclideana) $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, que gera uma sequência $\{x^k\} \subset \mathbb{R}^n$, da seguinte forma: seja $x^0 \in \mathbb{R}^n$,

$$x^k \in (I + \frac{1}{\lambda_k}T)(x^{k+1}),$$

ou seja,

$$x^{k+1} \in (I + \frac{1}{\lambda_k}T)^{-1}(x^k),$$

onde $\{\lambda_k\}$ é uma sequência de números positivos satisfazendo

$$0 < \lambda_k \leqslant \bar{\lambda},$$

para algum $\bar{\lambda} > 0$.

Mostraremos que a sequência $\{x^k\}$ gerada por este algoritmo converge para um zero de T. Em seguida, substituindo a norma euclideana pela distância de Bregman apresentamos o Algoritmo do Ponto Proximal com Distância de Bregman (Generalizado) para resolver o Problema de Desigualdade Variacional $VIP(T, C)$ (Variational Inequality Problem), que é definido da seguinte forma: dado $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ monótono maximal e $C \subset \mathbb{R}^n$ um conjunto fechado e convexo, o $VIP(T, C)$ consiste de encontrar $z \in C$ tal que existe $u \in T(z)$ satisfazendo

$$\langle u, x - z\rangle \geqslant 0, \ \forall\ x \in C.$$

O algoritmo gera uma sequência $\{x^k\}$ da seguinte forma: seja $x^0 \in C^0$ e $x^k \in C^0$ defina $T_k : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ por

$$T_k(\cdot) := T(\cdot) + \lambda_k \partial_x D_h(\cdot, x^k)$$

e encontre $x^{k+1} \in C^0$ tal que

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

Equivalentemente,

$$x^{k+1} \in [\partial_x D_h(\cdot, x^k) + \frac{1}{\lambda_k} T]^{-1}(x^k),$$

ou ainda

$$\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1}),$$

onde C^0 é o interior de C.

Além disso, mostramos que a sequência $\{x^k\}$ gerada pelo algoritmo acima converge para uma solução do $\mathsf{VIP}(T, C)$.

Abstract

In this dissertation, we initially present the Proximal Point Algorithm without restrictions to find zeros of Maximal Monotone Operators (with the Euclidean norm) $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, that generates a sequence $\{x^k\} \subset \mathbb{R}^n$, in the following way: given $x^0 \in \mathbb{R}^n$,

$$x^k \in (I + \frac{1}{\lambda_k} T)(x^{k+1}),$$

i.e.

$$x^{k+1} \in (I + \frac{1}{\lambda_k} T)^{-1}(x^k),$$

where $\{\lambda_k\}$ is a sequence of positive numbers satisfying

$$0 < \lambda_k \leqslant \bar{\lambda}, \text{ for some } \bar{\lambda} > 0.$$

We show that the sequence $\{x^k\}$ generated by this algorithm converges to a zero of T. Next, by substituting the Euclidean norm by the Bregman is distance, we present the Proximal Point Algorithm with Bregman is distance (Generalized) to solve the Variational Inequality Problem $VIP(T, C)$, which is defined in the following way: given $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ maximal monotone and $C \subset \mathbb{R}^n$ a closed and convex set, the $VIP(T, C)$ consists of finding $z \in C$ such that there is $u \in T(z)$ satisfying

$$\langle u, x - z \rangle \geqslant 0, \ \forall\, x \in C.$$

The algorithm generates a sequence $\{x^k\}$ in the following way: given $x^0 \in C^0$ e $x^k \in C^0$ define
$T_k : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$T_k(\cdot) := T(\cdot) + \lambda_k \partial_x D_h(\cdot, x^k)$$

and find $x^{k+1} \in C^0$ such that

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

Equivalently,

$$x^{k+1} \in [\partial_x D_h(\cdot, x^k) + \frac{1}{\lambda_k} T]^{-1}(x^k),$$

or

$$\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1}),$$

where C^0 is the interior of C.

Moreover we show that the sequence $\{x^k\}$ generated by the above algorithm converges to a solution of $VIP(T, C)$.

Sumário

Introdução

Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$ convexa e limitada inferiormente, o problema de otimização convexa é definido por

$$\min_{x \in \mathbb{R}^n} f(x). \tag{1}$$

Quando x está restrito a um conjunto convexo e fechado $C \subset \mathbb{R}^n$, obtemos o problema de otimização convexa restrito

$$\min_{x \in C} f(x). \tag{2}$$

Utilizando o subdiferencial ∂f de f (ver Definição 1.9), podemos escrever o problema (1) e (2) respectivamente, da seguinte forma

$$\mathit{Encontrar}\ x^* \in \mathbb{R}^n,\ \mathit{tal\ que}\ 0 \in \partial f(x^*); \tag{3}$$

e

$$\mathit{Encontrar}\ x^* \in C,\ \mathit{tal\ que\ existe}\ u \in \partial f(x^*)\ \mathit{satisfazendo}$$

$$\langle u, y - x^* \rangle \geqslant 0, \tag{4}$$

para todo $y \in C$.

Quando f é uma função convexa e semicontínua inferiormente, temos que $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é um operador monótono maximal (ver Teorema 2.2). Assim, obtemos extensões naturais dos problemas (3) e (4), substituíndo o operador ∂f por qualquer outro operador monótono maximal $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ da seguinte maneira

$$\mathit{Encontrar}\ x^* \in \mathbb{R}^n,\ \mathit{tal\ que}\ 0 \in T(x^*); \tag{5}$$

e

$$\mathit{Encontrar}\ x^* \in C\ \mathit{tal\ que\ existe}\ u \in T(x^*)\ \mathit{satisfazendo}$$

$$\langle u, y - x^* \rangle \geqslant 0, \tag{6}$$

para todo $y \in C$, respectivamente.

Desta forma, o problema (5) associado a um operador monótono maximal, é o problema de encontrar zeros de T. O problema (6), é chamado o Problema de Desigualdade Variacional para o operador T no conjunto C, o qual denotamos por $VIP(T, C)$.

Faremos agora uma descrição do conteúdo dos capítulos desta dissertação.

No capítulo 1, apresentamos as definições e resultados de Análise Convexa, funções de Bregman e Distância de Bregman que são utilizados no decorrer da dissertação.

O capítulo 2 contém definições e resultados da teoria de Operadores Monótonos. Além disso, mostramos que o subdiferencial ∂f de uma função convexa é um operador Monótono Maximal, Paramonótono e Pseudomonótono.

No capítulo 3, apresentamos o Algoritmo do Ponto Proximal para Operadores Monótonos Maximais para resolver o problema (5). O qual gera uma sequência $\{x^k\}$, para $x^0 \in \mathbb{R}^n$, da seguinte maneira

$$0 \in T(x^{k+1}) + \lambda_k(x^{k+1} - x^k), \tag{7}$$

equivalentemente

$$x^k \in (I + \frac{1}{\lambda_k}T)(x^{k+1}), \tag{8}$$

ou seja

$$x^{k+1} \in (I + \frac{1}{\lambda_k}T)^{-1}(x^k), \tag{9}$$

onde $\{\lambda_k\}$ é uma sequência de números positivos satisfazendo

$$0 < \lambda_k \leqslant \bar{\lambda}, \tag{10}$$

para algum $\bar{\lambda} > 0$ e é utilizado a norma Euclideana.

Mostraremos que a sequência gerada por $(7) - (10)$ converge a uma solução do problema (5), ou seja, a um zero de T sempre que o problema possuir soluções.

Finalmente, no capítulo 4 apresentamos o Algorimo do Ponto Proximal com distância de Bregman para o $VIP(T, C)$, o qual gera uma sequência $\{x^k\}$ definida por

$$x^0 \in C^0. \tag{11}$$

Dado $x^k \in C^0$, defina $T_k : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ por $T_k(\cdot) := T(\cdot) + \lambda_k \partial_x D_h(\cdot, x^k)$ e encontre $x^{k+1} \in C^0$ tal que

$$0 \in T_k(x^{k+1}). \tag{12}$$

Equivalentemente,

$$x^{k+1} \in [\partial_x D_h(\cdot, x^k) + \frac{1}{\lambda_k} T]^{-1}(x^k), \tag{13}$$

e ainda pela Proposição 1.6 (ii)

$$\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1}). \tag{14}$$

Mostraremos que a sequência gerada por $(11) - (14)$ converge para a solução do problema (6), isto é, para a solução do $\mathsf{VIP}(T, C)$.

Capítulo 1

Noções Preliminares

1.1 Existência de Soluções

Sejam dados um conjunto $D \subset \mathbb{R}^n$ e uma função $f : D \to \mathbb{R}$. O problema de achar um minimizador de f no conjunto D é escrito como

$$\mathrm{min} f(x) \ \mathsf{sujeito\ a}\ x \in D. \tag{1.1}$$

O conjunto D é chamado de conjunto viável, os pontos de D são chamados de pontos viáveis e f será chamada função objetivo.

Definição 1.1. *Dizemos que um ponto* $x^* \in D$ *é*
(i) minimizador global de (1.1)*, se*

$$f(x^*) \leqslant f(x) \ \forall\ x \in D. \tag{1.2}$$

(ii) minimizador local de (1.1) *se existe uma vizihança V de* x^* *tal que*

$$f(x^*) \leqslant f(x) \ \forall\ x \in D \cap V. \tag{1.3}$$

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

Para $D = \mathbb{R}^n$ o problema (1.1) toma a seguinte forma:

$$\mathrm{min} f(x) \ \mathsf{sujeito\ a}\ x \in \mathbb{R}^n, \tag{1.4}$$

dizemos que o problema(1.4) é irrestrito.

Teorema 1.1. *(Teorema de Weierstrass) Sejam* D *um conjunto compacto não vazio e* $f : D \to \mathbb{R}$ *uma função contínua.*

Então, o problema de minimizar f *em* D *tem uma solução global.*

Demonstração. Ver [10]. □

Definição 1.2. *Dizemos que uma função* $f : D \to \mathbb{R}$ *é coerciva no conjunto* D *quando para toda sequência* $\{x^k\} \subset D$ *e* $\|x^k\| \to \infty$ *ou* $x^k \to x \in \overline{D}$*, quando* $k \to \infty$*, tem-se* $\limsup_{k\to\infty} f(x^k) = \infty$.

Exemplo 1.1. *Seja* $x_0 \in \mathbb{R}^n$*, suponha que o problema* (1.4) *tem solução e defina a função* $F : \mathbb{R}^n \to \mathbb{R}$ *por*

$$F(x) := f(x) + \frac{1}{2}\|x - x_0\|^2.$$

Então F *é coerciva em* $\mathbb{R}^n$.

De fato, considere $h(x) = \frac{1}{2}\|x - x_0\|^2, \ \forall\, x \in \mathbb{R}^n$, então $F(x) = f(x) + h(x), \forall\, x \in \mathbb{R}^n$. Como f é limitada inferiormente existe $\beta \in \mathbb{R}$ tal que $f(x) \geqslant \beta, \forall\, x \in \mathbb{R}^n$. Seja $\{x^k\}_{k\in\mathbb{N}}$ uma sequência em $\mathbb{R}^n$ tal que $\lim\limits_{k\to+\infty} \|x^k\| = +\infty$.

Assim, $\lim\limits_{k\to+\infty} h(x^k) = \lim\limits_{k\to+\infty} \frac{1}{2}\|x^k - x_0\|^2 = +\infty$.

Daí,

$$+\infty = \lim_{k\to+\infty} [\beta + h(x^k)] \leqslant \lim_{k\to+\infty} [f(x^k) + h(x^k)] = \lim_{k\to+\infty} F(x^k).$$

Portanto, $\lim\limits_{k\to+\infty} \sup F(x^k) = +\infty$, logo F é coerciva em $\mathbb{R}^n$.

Proposição 1.1. *Sejam* $D \subset \mathbb{R}^n$ *e* $f : D \to \mathbb{R}$ *uma função contínua coerciva em* D*. Então, o problema de minimizar* f *em* D *possui uma solução global.*

Demonstração. Ver [10]. □

Definição 1.3. *Dizemos que a função* $f : D \to \mathbb{R}$ *é semicontínua inferiormente no ponto* $x \in D \subset \mathbb{R}^n$*, quando para qualquer seqência* $\{x^k\} \subset \mathbb{R}^n$ *tal que* $\{x^k\} \to x$ *quando (*$k \to \infty$*), tem-se*

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

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

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

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

1.2 Conjuntos Convexos

Um conjunto convexo é caracterizado por conter todos os segmento de retas cujos extremos pertencem ao conjunto. Precisamente:

Definição 1.4. *Um conjunto* $D \subset \mathbb{R}^n$ *é convexo se para quaisquer* $x, y \in D$ *e* $\alpha \in [0,1]$, *tem-se*

$$\alpha x + (1-\alpha)y \in D. \tag{1.5}$$

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

Exemplo 1.2. *O conjunto vazio, o conjunto unitário e o próprio* $\mathbb{R}^n$ *são exemplos triviais de conjuntos convexos.*

Exemplo 1.3. *Um hiperplano do* $\mathbb{R}^n$ *é um conjunto da forma* $H = \{x \in \mathbb{R}^n | \langle a, x\rangle = c\}$, *onde* $a \in \mathbb{R}^n$ *e* $c \in \mathbb{R}$. *Segue da Definição* 1.4 *que H é um conjunto convexo.*

Exemplo 1.4. *Um semi-plano em* $\mathbb{R}^n$ *é um conjunto da forma* $\{x \in \mathbb{R}^n | \langle a, x\rangle \leqslant c\}$, *onde* $a \in \mathbb{R}^n$ *e* $c \in \mathbb{R}$. *Segue da Definição* 1.4 *que todo semi-plano é um conjunto convexo.*

Definição 1.5. *Dado* $x^i \in \mathbb{R}^n$, $\alpha_i \in [0,1]$, $i = 1, ..., p$, *tais que* $\sum_{i=1}^{p} \alpha_i = 1$, *o ponto* $\sum_{i=1}^{p} \alpha_i x^i$ *chama-se a combinação convexa de pontos* $x^i \in \mathbb{R}^n$ *com parâmetros* $\alpha_i, i = 1, ..., p$.

Teorema 1.2. *Um conjunto* $D \subset \mathbb{R}^n$ *é convexo se, e somente se, para quaisquer* $p \in \mathbb{N}$, $x^i \in D$ *e* $\alpha_i \in [0,1]$, $i = 1, ..., p$, *tais que* $\sum_{i=1}^{p} \alpha_i = 1$, *a combinação convexa* $\sum_{i=1}^{p} \alpha_i x^i$ *pertence a* D.

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

Dados $p \in \mathbb{N}$, $x^i \in D$ 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 $x = 1.x^1 \in D$.

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

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

$$x = \sum_{i=1}^{n+1} \alpha_i x^i = \sum_{i=1}^{n} \alpha_i x^i + \alpha_{n+1} x^{n+1} = (1-\alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{(1-\alpha_{n+1})} x^i + \alpha_{n+1} x^{n+1}. \quad (1.6)$$

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

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 $y \in D$.

Daí, substituindo y em (1.6), obtemos

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

Segue da convexidade de D que $x = \sum_{i=1}^{n+1} \alpha_i x^i \in D$. Logo, $x = \sum_{i=1}^{p} \alpha_i x^i \in D$,

$\forall p \in \mathbb{N}$, com $x^i \in D$ $\alpha_i \in [0,1]$ tais que $\sum_{i=1}^{p} \alpha_i = 1$.

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

Então, para $p = 2$, temos que $\sum_{i=1}^{2} \alpha_i x^i = \alpha_1 x^1 + \alpha_2 x^2 \in D$, com $x^1, x^2 \in D$ e $\alpha_1, \alpha_2 \in [0,1]$, tal que $\alpha_1 + \alpha_2 = 1$. Assim, $\alpha_1 = 1 - \alpha_2$, logo $(1-\alpha_2)x^1 + \alpha_2 x^2 = \alpha_1 x^1 + \alpha_2 x^2 \in D$.

Portanto, pela Definição 1.4, D é um conjunto convexo. $\square$

1.3 Funções Convexas

Nesta seção estudaremos a definição e algumas caracterizações de funções convexas, que serão de grande importância para estudarmos as seções seguintes e os próximos capítulos.

Definição 1.6. *Se $D \subset \mathbb{R}^n$ é um conjunto convexo, diz-se que $f : D \to \mathbb{R}$ é convexa em D quando para quaisquer $x, y \in D$ e $\alpha \in [0,1]$, tem-se*

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y). \quad (1.7)$$

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

Exemplo 1.5. *A função* $f : \mathbb{R}^n \to \mathbb{R}$ *dada por* $f(x) = \|x - x_0\|^2$ *para* $x_0 \in \mathbb{R}^n$ *é estritamente convexa.*

Proposição 1.2. *Seja* $D \subset \mathbb{R}^n$ *um conjunto convexo. Se* $f : D \to \mathbb{R}$ *é uma função convexa e* $h : D \to \mathbb{R}$ *é estritamente convexa, então* $f + g$ *é estritamente convexa.*

Demonstração. Sejam $x, y \in D$, então

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y), \text{ para } \alpha \in [0, 1] \tag{1.8}$$

e

$$h(\alpha x + (1-\alpha)y) < \alpha h(x) + (1-\alpha)h(y), \text{ para } \alpha \in (0, 1) \tag{1.9}$$

e $x \neq y$. Portanto, por (1.8) e (1.9), tem-se

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

Logo, $f + h$ é estritamente convexa. □

Definição 1.7. *O epígrafo da função* $f : D \to \mathbb{R}$ *é o conjunto*

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

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

Teorema 1.3. *Seja* $D \subset \mathbb{R}^n$ *um conjunto convexo. Uma função* $f : D \to \mathbb{R}$ *é convexa em* D *se, e somente se, o epígrafo de* f *é um conjunto convexo em* $\mathbb{R}^n \times \mathbb{R}$.

Demonstração. ($\Rightarrow$) Suponha que $f : D \to \mathbb{R}$ é convexa em D. Sejam $(x, c_1) \in E_f$ e $(y, c_2) \in E_f$. Então,

$$f(x) \leqslant c_1 \text{ e } f(y) \leqslant c_2.$$

Pela convexidade de f, para $\alpha \in [0, 1]$, segue que

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y) \leqslant \alpha c_1 + (1-\alpha)c_2.$$

Daí, pela Definição 1.7 , $(\alpha x + (1-\alpha)y, \alpha c_1 + (1-\alpha)c_2) \in E_f$.
Portanto, $\alpha(x, c_1) + (1-\alpha)(y, c_2) \in E_f$. Logo, E_f é um conjunto convexo.

($\Leftarrow$) Suponha agora que E_f é um conjunto convexo em $\mathbb{R}^n \times \mathbb{R}$. Sejam $x, y \in D$. Tem-se que $(x, f(x)) \in E_f$, pois $f(x) \leqslant f(x)$. De forma análoga, $(y, f(y)) \in E_f$.
Pela convexidade de E_f em $\mathbb{R}^n \times \mathbb{R}$, $\alpha(x, f(x)) + (1-\alpha)(y, f(y)) \in E_f$, $\alpha \in [0,1]$. Daí $(\alpha x + (1-\alpha)y, \alpha f(x) + (1-\alpha)f(y)) \in E_f$. Pela Definição 1.7,

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

Portanto, f é uma função convexa em D. □

Dizemos que

$$\min f(x) \text{ sujeito a } x \in D \tag{1.10}$$

é um problema de minimização convexa quando $D \subset \mathbb{R}^n$ é um conjunto convexo e $f : D \to \mathbb{R}$ é uma função convexa no conjunto D. A importância da convexidade já pode ser vista no resultado seguinte.

Teorema 1.4. *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : D \to \mathbb{R}$ uma função convexa em D. Então todo minimizador local do problema (1.10) é global.*
Além disso, o conjunto de minimizadores é convexo. Se f é estritamente convexa, não pode haver mais de um minimizador.

Demonstração. Ver [10]. □

Definição 1.8. *Seja $D \subset \mathbb{R}^n$. O conjunto de nível da função $f : D \to \mathbb{R}$ associado a $c \in \mathbb{R}$, é o conjunto dado por*

$$L_{f,D}(c) = \{x \in D | f(x) \leqslant c\}.$$

Quando uma função é diferenciável, a convexidade admite caracterizações que são muito úteis para determinar se uma função é convexa ou não.

Teorema 1.5. ***(Caracterizações de funções convexas diferenciáveis)***
Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e aberto e $f : D \to \mathbb{R}$ uma função diferenciável em D. Então as propriedades seguintes são equivalentes:

1) A função f *é convexa em* D.

2) Para todo $x \in D$ *e todo* $y \in D$,

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

3)Para todo $x \in D$ *e todo* $y \in D$,

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

Demonstração. 1) $\Rightarrow$ 2) Suponha que f seja convexa, então para $x, y \in D$ e $\alpha \in (0, 1]$, tem-se que

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

Considere $d = y - x$. Assim, $x + \alpha d = x + \alpha(y - x) = (1-\alpha)x + \alpha y$. Com isso temos que $f(x + \alpha d) \leqslant \alpha f(y) + (1-\alpha) f(x)$. Então,

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

Isto implica que

$$f(y) - f(x) \geqslant \frac{f(x + \alpha d) - f(x)}{\alpha},$$

pois $\alpha > 0$.

Daí, passando ao limite quando $\alpha \to 0^+$, obtemos

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

Portanto, para todo $x, y \in D$

$$f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle. \tag{1.11}$$

2) $\Rightarrow$ 3) Trocando x por y em (1.11), obtemos

$$f(x) \geqslant f(y) + \langle \nabla f(y), x - y \rangle \tag{1.12}$$

Somando (1.11) e (1.12), tem-se

$$\begin{aligned}
\langle \nabla f(x), y - x \rangle + \langle \nabla f(y), x - y \rangle &\leqslant 0 \\
\langle \nabla f(x), y - x \rangle - \langle \nabla f(y), y - x \rangle &\leqslant 0 \\
\langle \nabla f(x) - \nabla f(y), y - x \rangle &\leqslant 0 \\
\langle \nabla f(y) - \nabla f(x), y - x \rangle &\geqslant 0, \ \forall x, y \in D.
\end{aligned}$$

3) $\Rightarrow$ 2) Suponha que 3) seja verdadeiro. Sejam $x, y \in D$. Pelo teorema do valor médio (ver [11]), existe $\alpha \in (0,1)$ tal que

$$f(y) - f(x) = \frac{\partial f}{\partial d}(x + \alpha d) = \langle \nabla f(x + \alpha d), d \rangle \tag{1.13}$$

onde $d = x - y$.

Aplicando 3) para $x + \alpha d$ e x, obtemos

$$0 \leqslant \langle \nabla f(x + \alpha d) - \nabla f(x), x + \alpha d - x \rangle = \langle \nabla f(x + \alpha d) - \nabla f(x), \alpha d \rangle.$$

Então

$$\langle \nabla f(x + \alpha d), \alpha d \rangle \geqslant \langle \nabla f(x), \alpha d \rangle \Rightarrow \langle \nabla f(x + \alpha d), d \rangle \geqslant \langle \nabla f(x) d \rangle. \tag{1.14}$$

Combinando (1.13) e (1.14) e usando $d = y - x$, obtemos

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

Portanto, para todo $x, y \in D$

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

2) $\Rightarrow$ 1) Suponha agora que 2) seja verdadeiro. Aplicando 2) para x e $x + \alpha d$, com $d = y - x$, segue que

$$f(x) \geqslant f(x + \alpha d) + \langle \nabla f(x + \alpha d), x - (x + \alpha d) \rangle = f(x + \alpha d) + \langle \nabla f(x + \alpha d), -\alpha d \rangle.$$

Portanto,

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

Daí,

$$(1 - \alpha) f(x) \geqslant (1 - \alpha) f(x + \alpha d) - (1 - \alpha)\alpha \langle \nabla f(x + \alpha d), d \rangle. \tag{1.15}$$

Analogamente, aplicando 2) para y e $x + \alpha d$, obtemos

$$\alpha f(y) \geqslant \alpha f(x + \alpha d) + \alpha(1 - \alpha) \langle \nabla f(x + \alpha d), d \rangle. \tag{1.16}$$

Somando (1.15) e (1.16), tem-se

$$(1 - \alpha) f(x) + \alpha f(y) \geqslant (1 - \alpha) f(x + \alpha d) + \alpha f(x + \alpha d) = f(x + \alpha d). \tag{1.17}$$

Subtituindo $d = y - x$ em (1.17), temos

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

Portanto, f é convexa em D. $\square$

Corolário 1.3.1. *Sejam* $D \subset \mathbb{R}^n$ *um conjunto convexo e aberto e* $f : D \to \mathbb{R}$ *uma função diferenciável em* D. *Então as propriedades seguintes são equivalentes:*

1) A função f *é estritamente convexa em* D.

2) Para todo $x \in D$ *e todo* $y \in D$ *tal que* $x \neq y$,

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

3) Para todo $x \in D$ *e todo* $y \in D$ *tal que* $x \neq y$,

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

Demonstração. Segue da Definição 1.6 e do Teorema 1.5 □

Teorema 1.6. *Sejam* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função diferenciável e* $\bar{x}$ *minimizador local de* f. *Então* $\nabla f(\bar{x}) = 0$.

Demonstração. Seja $d \in \mathbb{R}^n$ arbitrário, porém fixo. Como $\bar{x}$ é mínimo de f, existe $\varepsilon > 0$ tal que

$$f(\bar{x}) \leqslant f(\bar{x} + td), \ \forall t \in [0, \varepsilon]. \tag{1.18}$$

Da diferenciabilidade de f, tem-se

$$f(\bar{x} + td) - f(\bar{x}) = \langle \nabla f(\bar{x}), td \rangle + r(t), \ \text{com} \lim_{t \to 0^+} \frac{r(t)}{t} = 0. \tag{1.19}$$

Daí, usando (1.18) em (1.19), obtemos

$$t\langle \nabla f(\bar{x}), d \rangle + r(t) \geqslant 0. \tag{1.20}$$

Dividindo (1.20) por $t > 0$, tem-se

$$\langle \nabla f(\bar{x}), d \rangle + \frac{r(t)}{t} \geqslant 0, \tag{1.21}$$

e tomando o limite em (1.21) quando $t \to 0^+$, obtemos

$$\langle \nabla f(\bar{x}), d \rangle \geqslant 0. \tag{1.22}$$

Como $d \in \mathbb{R}^n$ é arbitrário, podemos escolher $d = -\nabla f(\bar{x})$ e substituindo em (1.22), tem-se

$$\langle \nabla f(\bar{x}), -\nabla f(\bar{x}) \rangle \geqslant 0.$$

Portanto,

$$\langle \nabla f(\bar{x}), \nabla f(\bar{x}) \rangle \leqslant 0.$$

Logo, $\nabla f(\bar{x}) = 0$. □

1.4 Subgradiente de uma Função

A seguinte noção é fundamental em Análise Convexa.

Definição 1.9. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa. Dizemos que* $y \in \mathbb{R}^n$ *é um subgradiente de* f *no ponto* $x \in \mathbb{R}^n$ *se*

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

O conjunto de todos os subgradientes de f em x, é chamado o Subdiferencial de f em x; e denotado por $\partial f(x)$. Isto é,

$$\partial f(x) := \{y \in \mathbb{R}^n \mid f(z) \geqslant f(x) + \langle y, z - x\rangle\}.$$

Observação 1.1. *Considerando* $z \in \mathbb{R}^n$ *como* $z = x + \alpha d$*, onde* $d \in \mathbb{R}^n$ *e* $\alpha > 0$. *Seja* $s \in \partial f(x)$*, pela Definição* 1.9*, tem-se*

$$\begin{aligned}
s \in \partial f(x) \quad &\Leftrightarrow \quad f(z) - f(x) \geqslant \langle s, z - x\rangle, \ \forall \ z \in \mathbb{R}^n \\
&\Leftrightarrow \quad f(x + \alpha d) - f(x) \geqslant \langle s, x + \alpha d - x\rangle \\
&\Leftrightarrow \quad f(x + \alpha d) - f(x) \geqslant \langle s, \alpha d\rangle \\
&\Leftrightarrow \quad \frac{f(x + \alpha d) - f(x)}{\alpha} \geqslant \langle s, d\rangle.
\end{aligned}$$

Então,

$$\lim_{\alpha \to 0+} \frac{f(x + \alpha d) - f(x)}{\alpha} \geqslant \langle s, d\rangle.$$

Portanto,

$$\frac{\partial f}{\partial d}(x) \geqslant \langle s, d\rangle, \ \forall \ d \in \mathbb{R}^n.$$

Assim, temos que a Definição 1.9 *é equivalente a seguinte definição*

$$\partial f(x) := \{y \in \mathbb{R}^n \mid \frac{\partial f}{\partial d}(x) \geqslant \ \langle y, d\rangle, \ \forall \ d \in \mathbb{R}^n\}.$$

Notação: $\dfrac{\partial f}{\partial d}(x) = f'(x, d)$.

Teorema 1.7. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa. Então para todo* $x \in \mathbb{R}^n$*, o conjunto* $\partial f(x)$ *é convexo, compacto e não vazio.*

Demonstração. Ver [10]. □

Proposição 1.3. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa. Então, para todo* $d \in \mathbb{R}^n$*, temos*

$$\frac{\partial f}{\partial d}(x) = \max_{s \in \partial f(x)} \langle s, d\rangle.$$

Demonstração. Ver [10]. □

Proposição 1.4. *Uma função convexa* $f : \mathbb{R}^n \to \mathbb{R}$ *é diferenciável se, e somente se, o conjunto* $\partial f(x)$ *contém um só elemento. Neste caso,* $\partial f(x) = \{\nabla f(x)\}$.

Demonstração. $(\Rightarrow)$ Suponha que f seja diferenciável em $x \in \mathbb{R}^n$. Então, sendo f convexa, tem-se pelo Teorema 1.5 que $f(z) \geqslant f(x) + \langle \nabla f(x), z - x\rangle$, $z \in \mathbb{R}^n$. Assim, pela Definição 1.9 $\nabla f(x) \in \partial f(x)$. Seja $s \in \partial f(x)$. Logo, $f(z) \geqslant f(x) + \langle s, z - x\rangle$, $z \in \mathbb{R}^n$. Considere $\lambda d = z - x$, com $\lambda > 0$. Então,

$$f(x + \lambda d) \geqslant f(x) + \langle s, \lambda d\rangle. \tag{1.23}$$

Como f é diferenciável em x,

$$f(x + \lambda d) = f(x) + \langle \nabla f(x), \lambda d\rangle + \|\lambda d\| r(\lambda d), \text{ com } \lim_{\lambda \to 0} r(\lambda d) = 0. \tag{1.24}$$

Então, juntando (1.23) e (1.24), obtemos

$$\langle s, \lambda d\rangle \leqslant \langle \nabla f(x), \lambda d\rangle + \|\lambda d\| r(\lambda d)$$

$$\langle s, d\rangle \leqslant \langle \nabla f(x), d\rangle + \|d\| r(\lambda d) \tag{1.25}$$

Passando ao limite em (1.25) quando $\lambda \to 0$, obtemos

$$\langle s, d\rangle \leqslant \langle \nabla f(x), d\rangle + \|d\| \lim_{\lambda \to 0} r(\lambda d) = \langle \nabla f(x), d\rangle.$$

Então,

$$\langle s - \nabla f(x), d\rangle \leqslant 0, \ \forall d \in R^n. \tag{1.26}$$

Assim, tomando $d = s - \nabla f(x)$ em (1.26), obtemos

$$\langle s - \nabla f(x), s - \nabla f(x)\rangle \leqslant 0. \text{Logo}, s = \nabla f(x).$$

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

$(\Leftarrow)$ Reciprocamente suponha que $\partial f(x) = \{s\}$. Pela Proposição 1.3, $\dfrac{\partial f}{\partial d}(x) = \langle s, d\rangle$ $\forall d \in \mathbb{R}^n$. Escolhendo d como os elementos da base canônica de $\mathbb{R}^n$, tem-se que $s_i = \dfrac{\partial f}{\partial x_i}(x), i = 1, ..., n$. Assim, $\dfrac{\partial f}{\partial d}(x) = \langle \nabla f(x), d\rangle, \ \forall \ d \in \mathbb{R}^n$.

Portanto, f é diferenciável em x. □

Teorema 1.8. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *um função convexa. O ponto* $\bar{x} \in \mathbb{R}^n$ *é mínimo de* f *se, e somente se,* $0 \in \partial f(\bar{x})$.

Demonstração. ($\Rightarrow$) Suponha que $\bar{x} \in \mathbb{R}^n$ seja mínimo de f. Então

$$f(\bar{x}) \leqslant f(x), \ \forall x \in \mathbb{R}^n.$$

Daí,

$$f(x) \geqslant f(\bar{x}) + \langle 0, x - \bar{x} \rangle, \ \forall \ x \in \mathbb{R}^n.$$

Portanto, pela Definição 1.9, $0 \in \partial f(\bar{x})$.

($\Leftarrow$) Reciprocamente, se $0 \in \partial f(\bar{x})$, pela Definição 1.9,

$$f(x) \geqslant f(\bar{x}) + \langle 0, x - \bar{x} \rangle = f(\bar{x}), \quad \forall \ x \in \mathbb{R}^n.$$

Logo, $f(x) \geqslant f(\bar{x}) \ \ \forall \ x \in \mathbb{R}^n$. Portanto, $\bar{x}$ é mínimo de f. □

Exemplo 1.6. *Seja* $f(x) = |x|$ *para todo* $x \in \mathbb{R}$. *Note que* f *é uma função convexa. Encontremos o seu subdiferencial que é dado pela definição por:*

$\partial f(x) = \{s \in \mathbb{R}; s(y - x) \leqslant |y| - |x|, \ \forall \ y \in \mathbb{R}\}$. *Então,*

i) Se $x > 0$, *tem-se que* f *é diferenciável, portanto* $\partial f(x) = \{1\}$, *que é o gradiente de* f.

ii) Se $x < 0$, f *é diferenciável, assim* $\partial f(x) = \{1\}$, *que é o gradiente de* f.

iii) Se $x = 0$, *tem-se* $f(x) = 0$, *daí para todo* $y \in \mathbb{R}$, *obtemos* $|y| \geqslant sy$ *então, se* $y \geqslant 0$, $s \leqslant 1$ *e se* $y < 0$, *implica* $s \geqslant -1$.

Portanto, $\partial f(x) = [-1, 1]$.

Logo, por i),ii) e iii),

$$\partial f(x) = \begin{cases} \{1\}, & \text{se} \quad x > 0 \\ [-1, 1], & \text{se} \quad x = 0 \\ \{-1\}, & \text{se} \quad x < 0 \end{cases}.$$

1.5 Funções Convexas com valores em $\mathbb{R} \cup \{+\infty\}$

Nesta seção trataremos de estender as funções convexas definidas na seção 1.3.

Definição 1.10. *Uma função* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, *não identicamente* $+\infty$ *é dita convexa quando, para todo* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ *e todo* $\alpha \in (0, 1)$, *tem-se*

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

como uma desigualdade em $\mathbb{R} \cup \{+\infty\}$.

Podemos observar que de acordo com a Definição 1.6, para $D \subset \mathbb{R}^n$ um conjunto convexo, não vazio e $f : D \to \mathbb{R}$ uma função convexa, podemos estender a função f a uma função convexa como na Definição 1.10 da seguinte forma $\bar{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$,

$$\bar{f}(x) = \begin{cases} f(x), & \text{se} \quad x \in D \\ +\infty, & \text{se} \quad x \notin D \end{cases}$$

Então, obtemos uma função $\bar{f}$, a qual chamamos de função convexa própria.

Definição 1.11. *O domínio efetivo de uma função convexa* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é o conjunto denotado por* $\mathrm{dom}(f) = \{x \in \mathbb{R}^n; f(x) < +\infty\}$.

Definição 1.12. *Uma função* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é chamada fechada quando seu epígrafo é fechado em* $\mathbb{R}^n \times \mathbb{R}$ *ou equivalentemente se seus conjuntos de nível são fechados.*

Definição 1.13. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa. O subdiferencial de* f *em* $x \in \mathbb{R}^n$ *é o conjunto* $\partial f(x)$ *definido por*

$$\partial f(x) = \begin{cases} \{s \in \mathbb{R}^n; f(y) \geqslant f(x) + \langle s, y - x \rangle, \ \forall\, y \in \mathbb{R}^n\}, & \textit{se} \quad x \in \mathrm{dom}(f) \\ \emptyset, & \textit{se} \quad x \notin \mathrm{dom}(f) \end{cases}$$

Exemplo 1.7. *Seja* $C \subset \mathbb{R}^n$ *fechado e convexo. A função* $\delta_C : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *definida por:*

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

é chamada a função indicadora de C.

A qual é convexa se, e somente se, C é um conjunto convexo.
De fato, se $x \notin C$ tem-se $\delta_C = +\infty$. Assim, para todo $d \in \mathbb{R}$, $d < +\infty = \delta_C(x)$. 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.
Agora vamos calcular o subdiferencial da função indicadora.
Seja $s \in \partial\delta_C(x)$, então.
i)Se $x \notin C$, tem-se $\delta_C(x) = +\infty$, daí

$$\delta_C(y) \geqslant +\infty, \ \forall\, y \in \mathbb{R}^n.$$

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

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

$$\delta_C(y) \geqslant \langle s, y - x\rangle, \ \forall\ y \in \mathbb{R}^n.$$

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

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

Logo,

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

Teorema 1.9. *Sejam* $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *funções convexas.*

i)Tem-se $\partial(f + g)(x) \supset \partial f(x) + \partial g(x)$*, para* $x \in \mathbb{R}^n$*.*

ii) Seja $x \in \mathrm{ri}(\mathrm{dom}(f)) \cap \mathrm{ri}(\mathrm{dom}(g))$ *tal que* $\partial f(x)$ *e* $\partial g(x)$ *sejam não vazios. Então,*

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

Demonstração. i) Se $s \in \partial f(x)$, $z \in \partial g(x)$, então $s + z \in \partial f(x) + \partial g(x)$.

Daí, para todo $y \in \mathbb{R}^n$

$$f(y) \geqslant f(x) + \langle s, y - x\rangle \ \text{e}\ g(y) \geqslant g(x) + \langle z, y - x\rangle.$$

Portanto,

$$(f + g)(y) \geqslant (f + g)(x) + \langle s + z, y - x\rangle.$$

Segue que, $s + z \in \partial(f + g)(x)$. Logo, $\partial(f + g)(x) \supset \partial f(x) + \partial g(x)$.

ii) Suponha que $\partial(f + g)(x) \not\subset (\partial f(x) + \partial g(x))$, isto é, existe $\omega \in \partial(f + g)(x)$, tal que $\omega \notin (\partial f(x) + \partial g(x))$. Pela compacidade de $\partial f(x) + \partial g(x)$ e pelo Teorema da Separação Estrita existem $a \in \mathbb{R}^n$ e $c \in \mathbb{R}$ tal que

$$\langle a, s + z\rangle < c < \langle a, \omega\rangle \leqslant (f + g)'(x, a) = f'(x, a) + g'(x, a), \ \forall\ s \in \partial f(x), \ \forall\ z \in \partial g(x),$$

que é um absurdo pois, pela Proposição 1.3, existe um $s_0 + z_0 \in (\partial f(x) + \partial g(x))$ tal que

$$\langle a, s_0\rangle = f'(x, a) \text{ e } \langle a, z_0\rangle = g'(x, a).$$

Portanto, $\partial(f + g)(x) = \partial f(x) + \partial g(x)$. □

1.6 Função Conjugada de uma Função Convexa

Para demonstrar que o subdiferencial de uma função convexa é um operador monótono maximal precisaremos da definição de função conjugada de uma função convexa e de alguns resultados.

Definição 1.14. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa própria. A função* $f^* : \mathbb{R}^n \to R \cup \{+\infty\}$ *definida por*

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

para todo $s \in \mathbb{R}^n$*, é chamada a função conjugada da função* f*.*

Exemplo 1.8. *Considere a função convexa* $f(x) = |x|$ *para todo* $x \in \mathbb{R}$*, encontremos sua função conjugada.*

Pela definição de conjugada, para todo $s \in \mathbb{R}^n$*,* $f^*(s) = \sup_{x \in \mathbb{R}^n} \{\langle x, s\rangle - f(x)\}$*. Portanto,* $f^*(s) = \sup_{x \in \mathbb{R}^n} \{(s - 1)x\}$*, se* $x \geqslant 0$ *e* $f^*(s) = \sup_{x \in \mathbb{R}^n} \{(s + 1)x\}$*, se* $x < 0$*. Logo,*

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

Observação 1.2. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$*. Segue da Definição* 1.14 *que para cada* $x \in \mathbb{R}^n$ *e cada* $s \in \mathbb{R}^n$*, tem-se* $f^*(s) \geqslant \langle x, s\rangle - f(x)$*, isto é,*

$$f^*(s) + f(x) \geqslant \langle x, s\rangle \tag{1.27}$$

sempre que o lado esquerdo da desigualdade (1.27) *estiver definido. A relação* (1.27) *é chamada a desigualdade de Fenchel.*

Teorema 1.10. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa própria semicontínua inferiormente , onde* f *é finita em* $x \in \mathbb{R}^n$*. Então,* $s \in \partial f(x)$ *se, e somente se,* $f^*(s) = \langle x, s\rangle - f(x)$*.*

Demonstração. ($\Rightarrow$) Seja $s \in \partial f(x)$. Então

$$\langle s, z - x\rangle \leqslant f(z) - f(x)$$

isto é,

$$\langle s, z\rangle - f(z) \leqslant \langle s, x\rangle - f(x),$$

para todo $z \in \mathbb{R}^n$.

Portanto,

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

Logo, $f^*(s) = \langle s, x\rangle - f(x)$.

($\Leftarrow$) Se $f^*(s) = \sup_{z \in \mathbb{R}^n}\{\langle s, z\rangle - f(z)\} = \langle s, x\rangle - f(x)$, tem-se $f^*(s) < +\infty$. Assim,

$$\langle s, z\rangle - f(z) \leqslant \langle s, x\rangle - f(x).$$

Portanto,

$$\langle s, z - x\rangle \leqslant f(z) - f(x), \forall z \in \mathbb{R}^n.$$

Logo, $s \in \partial f(x)$. □

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

Teorema 1.11. *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função convexa própria semicontínua inferiormente e $x \in \mathbb{R}^n$. Então, a função conjugada $f^* : \mathbb{R}^n \to R \cup \{+\infty\}$ é uma função convexa.*

Demonstração. Dados $s_1, s_2 \in \mathbb{R}^n$. Faremos a prova em dois casos:

i) Se $s_1, s_2 \in \mathbb{R}^n$, tais que $f^*(s_1) < +\infty$ e $f^*(s_2) < +\infty$. Então para $t \in [0, 1]$, tem-se,

$$f^*((1-t)s_1 + ts_2) = \sup_{x \in \mathbb{R}^n}\{\langle x, (1-t)s_1 + ts_2\rangle - f(x)\}. \tag{1.28}$$

Daí, somando e subtraindo $tf(x)$ a (1.28) obtemos

$$f^*((1-t)s_1 + ts_2) = \sup_{x \in \mathbb{R}^n}\{\langle x, (1-t)s_1 + ts_2\rangle - f(x) + tf(x) - tf(x)\}.$$

Assim,

$$f^*((1-t)s_1 + ts_2) = \sup_{x \in \mathbb{R}^n}\{(1-t)[\langle x, s_1\rangle - f(x)] + t[\langle x, s_2\rangle - f(x)]\}.$$

Segue que,

$$f^*((1-t)s_1 + ts_2) \leqslant (1-t)\sup_{x\in\mathbb{R}^n}\{\langle x, s_1\rangle - f(x)\} + t\sup_{x\in\mathbb{R}^n}\{\langle x, s_2\rangle - f(x)\}.$$

Poratnto, pela Definição1.14

$$f^*((1-t)s_1 + ts_2) \leqslant (1-t)f^*(s_1) + tf^*(s_2).$$

Logo, f^* é uma função convexa.

ii) Se para $s_1, s_2 \in \mathbb{R}^n$, $f^*(s_1) = +\infty$ ou $f^*(s_2) = +\infty$, então temos trivialmente que

$$f^*((1-t)s_1 + ts_2) \leqslant (1-t)f^*(s_1) + tf^*(s_2).$$

Portanto, f^* é uma função convexa. □

Definição 1.15. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa. Dizemos que a função* $f^{**} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é a função conjugada* $(f^*)^*$ *da conjugada de* f.

Lema 1.6.1. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa. Então* f^{**} *é o supremo do conjunto de todas as funções afins que minoram* f.

Demonstração. Seja A o conjunto de todas as funções afins que minoram f. Considere $\mu = \sup\{g; g \in A\}$, provaremos que $\mu = f^{**}$.

De fato, para cada $s \in \mathbb{R}^n$, a função $g(x) = \langle x, s\rangle - f^*(s)$ é uma função afim. Pela desigualdade de Fenchel,

$$f(x) + f^*(s) \geqslant \langle x, s\rangle \ x \in \mathbb{R}^n.$$

Segue que,

$$g(x) = \langle x, s\rangle - f^*(s) \leqslant f(x) + f^*(s) - f^*(s) = f(x).$$

Assim, g é uma minorante de f. Então, $g \in A$.

Portanto,

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

Logo,

$$f^{**}(x) \leqslant \mu(x),\ \forall\, x \in \mathbb{R}^n. \tag{1.29}$$

Agora, se $h \in A$, então h é da forma $h(x) = \langle x, s\rangle - \alpha$, onde $s \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, que implica que, para todo $x \in \mathbb{R}^n$, tem-se $\langle x, s\rangle - \alpha \leqslant f(x)$.

Isto é,

$$\langle x, s\rangle - f(x) \leqslant \alpha.$$

Então,

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

Assim,

$$-\alpha \leqslant -f^*(s).$$

Segue que,

$$h(x) = \langle x, s\rangle - \alpha \leqslant \langle x, s\rangle - f^*(s).$$

Portanto, para cada $x \in \mathbb{R}^n$

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

Logo,

$$\mu(x) \leqslant f^{**}(x), \ \forall\, x \in \mathbb{R}^n. \tag{1.30}$$

De (1.29) e (1.30), obtemos $\mu = f^{**}$. □

Proposição 1.5. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função convexa própria semicontínua inferiormente. Então* $f = f^{**}$.

Demonstração. i)Mostraremos que $f^{**} \leqslant f$. De fato, por definição

$$f^{**}(\omega) := \sup_{y\in\mathbb{R}^n}\{\langle \omega, y\rangle - f^*(y)\}. \tag{1.31}$$

Por outro lado,

$$f^*(y) = \sup_{\omega\in\mathbb{R}^n}\{\langle y, \omega\rangle - f(\omega)\} \geqslant \langle y, \omega\rangle - f(\omega),$$

para todo $\omega \in \mathbb{R}^n$. Segue que,

$$\langle y, \omega\rangle - f^*(y)\} \leqslant f(\omega), \ \forall\ \omega \ \in \mathbb{R}^n. \tag{1.32}$$

Portanto, de (1.31) e (1.32), tem-se $f^{**}(\omega) \leqslant f(\omega)$, $\forall\ \omega \ \in \mathbb{R}^n$.
ii) Mostraremos que $f \leqslant f^{**}$. Suponha que não ocorra $f \leqslant f^{**}$, isto é, existe $\omega_0 \in \mathbb{R}^n$ tal que $f^{**}(\omega_0) < f(\omega_0)$. Assim, existe $\alpha \in \mathbb{R}$ tal que

$$f^{**}(\omega_0) < \alpha < f(\omega_0). \tag{1.33}$$

Da convexidade de f, existe uma função afim h minorando f dada por
$h(x) = \alpha - \langle s, x - \omega_0\rangle$, onde $h(\omega_0) = \alpha$.

Como $h < f$, temos pelo Lema 1.6.1 que $h(x) \leqslant f^{**}(x)$, em particular para $x = \omega_0$, tem-se $h(\omega_0) \leqslant f^{**}(\omega_0)$ que implica $\alpha \leqslant f^{**}(\omega_0)$, contrariando (1.33).

Portanto, $f < f^{**}$.

Logo, de i) e ii), obtemos $f = f^{**}$. □

1.7 Funções e Distâncias de Bregman

Seja $S \subset \mathbb{R}^n$ um conjunto aberto e convexo, e $\overline{S}$ o fecho de S. Considere a função $h : \overline{S} \to \mathbb{R}$ convexa e diferenciável em S. Defina $D_h : \overline{S} \times S \to \mathbb{R}$ por

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \tag{1.34}$$

induzida por h.

Definição 1.16. *Dizemos que* h *é uma função de Bregman e* D_h *a distância de Bregman induzida por* h *se as seguintes condições são satisfeitas:*

B_1) h *é continuamente diferenciável em* S.

B_2) h *é estritamente convexa e contínua em* $\overline{S}$.

B_3) *Para todo* $\delta > 0$ *os conjuntos* $\Gamma_1(y, \delta) = \{x \in \overline{S}; D_h(x, y) \leqslant \delta\}$ *e*

$\Gamma_2(x, \delta) = \{y \in S; D_h(x, y) \leqslant \delta\}$ *são limitados para todo* $y \in S$, *e* $x \in \overline{S}$, *respectivamente.*

B_4) *Se a sequência* $\{y^k\} \subset S$, *converge para* y^*, *então* $\lim_{k\to\infty} D_h(y^*, y^k) = 0$.

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

Dizemos que uma função de Bregman h *é coerciva na fronteira se satisfaz a seguinte condição:*

B_6) *Se* $\{y^k\} \subset S$ *é tal que* $\lim_{k\to\infty} y^k = y \in \partial S$, *então*

$$\lim_{k\to\infty} \langle \nabla h(y^k), y - y^k \rangle = -\infty, \ \forall\, y \in S. \tag{1.35}$$

Dizemos que h *é zona coerciva se satisfaz a seguinte condição:*

B_7) *Para todo* $y \in \mathbb{R}^n$ *existe* $x \in S$ *tal que* $\nabla h(x) = y$.

S é chamado de zona de h.

Vejamos agora alguns exemplos.

Observação 1.3. *A definição de funções e distância de Bregman acima pode ser definida de outra maneira* (ver[*19*]), *sendo suficiente considerar apenas* $(B_1) - (B_4)$, *onde em* (B_3) *é utilizado apenas um dos conjuntos de níveis e* (B_5) *segue de* (B_1) *e* (B_2).

Exemplo 1.9. *Seja* $h : \mathbb{R}^n \to \mathbb{R}$, *tal que* $h(x) = ||x||^2$. *Neste caso* $D_h(x, y) = ||x - y||^2$.

Exemplo 1.10. $S = \mathbb{R}^n_{++}$, $h(x) = \sum_{i=1}^n x_i \log x_i$ *estendido com a continuidade até a fronteira de* $\mathbb{R}^n_{++}$ *com a convenção de que* $0 \log 0 = 0$. *Neste caso,*

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

que é Kulblback-Leibeler divergente, bastante usada em Estatística.

Observação 1.4. *Podemos verificar facilmente que* $D_h(x, y) \geqslant 0$ *para todo* $x \in \overline{S}$, $y \in S$ *e* $D_g(x, y) = 0$ *se, e somente se,* $x = y$.

Proposição 1.6. *Se* h *é uma função de Bregman com zona* S *então:*

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

ii) $\partial_x D_h(x, y) = \nabla h(x) - \nabla h(y)$, $\forall x, y \in S$.

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

Demonstração. i) De (1.34), para $x \in \overline{S}$, $y, z \in S$, temos

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

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

$$\begin{aligned} \partial_x D_h(x, y) &= \nabla h(x) - 0 - \frac{\partial}{\partial x} \langle \nabla h(y), x - y \rangle \\ &= \nabla h(x) - \nabla h(y). \end{aligned}$$

iii) Para $y \in S$, defina

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

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

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

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

A importância de distância de Begman pode ser vista em [18].

Capítulo 2

Operadores Monótonos

2.1 Operadores Monótonos e Monótonos Maximais

Nesta seção apresentaremos os principais resultados e definições de operadores monótonos. Seja $A : \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear, diz-se que A é semidefinida positiva se $\langle x, Ax \rangle \geqslant 0$, para todo $x \in \mathbb{R}^n$.
Então para quaisquer $x, y \in \mathbb{R}^n$, tem-se

$$0 \leqslant \langle x - y, A(x - y) \rangle = \langle x - y, Ax - Ay \rangle.$$

Um operador monótono é uma generalização de uma transformação linear semidefinida positiva (não necessariamente linear).

Definição 2.1. *Um operador* $T : \mathbb{R}^n \to \mathbb{R}^n$ *(não necessariamente linear) diz-se monótono se para todo* $x, y \in \mathbb{R}^n$

$$\langle T(x) - T(y), x - y \rangle \geqslant 0. \tag{2.1}$$

Observação 2.1. T *é chamado estritamente monótono quando a desigualdade* (2.1) *é estrita para* $x \neq y$.

Exemplo 2.1. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa (resp. estritamente convexa) e diferenciável. Então* $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ *é um operador monótono (resp. estritamente monótono). Segue do Teorema* 1.5.

Como ∂f associa a cada vetor $x \in \mathbb{R}^n$, não exatamente a cada vetor de $\mathbb{R}^n$, mas a um subconjunto do $\mathbb{R}^n$. Estenderemos a noção de operador monótono ponto a ponto a operador ponto-conjunto. Notação: $T : \mathbb{R}^n \to P(\mathbb{R}^n)$ ou $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

Definição 2.2. *Seja* $\mathsf{T}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador.*

i) O domínio de T *é o conjunto* $\mathrm{dom}(\mathsf{T})=\{x\in\mathbb{R}^n;\mathsf{T}(x)\neq\emptyset\}$.

ii) A imagem de T *é o conjunto*

$$\mathsf{R}(\mathsf{T})=\bigcup_{x\in\mathrm{dom}(\mathsf{T})}\mathsf{T}(x).$$

Definição 2.3. *Dado* $\mathsf{T}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$*, o operador* $\mathsf{T}^{-1}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é definido pela seguinte relação*

$$x\in\mathsf{T}^{-1}(y)\Leftrightarrow y\in\mathsf{T}(x).$$

Definição 2.4. *Um ponto* $\bar{x}\in\mathbb{R}^n$ *é chamado zero de* $\mathsf{T}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *quando* $0\in\mathsf{T}(\bar{x})$.

Definição 2.5. *Seja* $\mathsf{T}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$*. Dizemos que T é monótono quando para quaisquer* $x,y\in\mathbb{R}^n, u\in\mathsf{T}(x)$ e $v\in\mathsf{T}(y)$ *tem-se*

$$\langle x-y,u-v\rangle\geqslant 0. \tag{2.2}$$

Observação 2.2. T *é estritamente monótono se a desigualdade* (2.2) *é estrita para* $x\neq y$.

Proposição 2.1. *Sejam* $\mathsf{T}_1:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *e* $\mathsf{T}_2:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *operadores monótonos, tais que* $\mathrm{dom}(\mathsf{T}_1)\cap\mathrm{dom}(\mathsf{T}_2)\neq\emptyset$*. Então* $\mathsf{T}_1+\mathsf{T}_2$ *é um operador monótono.*

Demonstração. Sejam $x,y\in\mathbb{R}^n$, $u\in(\mathsf{T}_1+\mathsf{T}_2)(x)$ e $v\in(\mathsf{T}_1+\mathsf{T}_2)(y)$. Então, para $u\in(\mathsf{T}_1+\mathsf{T}_2)(x)=\mathsf{T}_1(x)+\mathsf{T}_2(x)$, existem $u_1\in\mathsf{T}_1(x)$ e $u_2\in\mathsf{T}_2(x)$ tais que $u=u_1+u_2$. Analogamente, existem $v_1\in\mathsf{T}_1(y)$ e $v_2\in\mathsf{T}_2(y)$ tais que $v=v_1+v_2$. Como T_1 e T_2 são operadores monótonos, tem-se $\langle u_1-v_1,x-y\rangle\geqslant 0$ e $\langle u_2-v_2,x-y\rangle\geqslant 0$.
Portanto,

$$\begin{aligned}\langle u-v,x-y\rangle &= \langle u_1+u_2-v_1-v_2,x-y\rangle\\ &= \langle u_1-v_1,x-y\rangle+\langle u_2-v_2,x-y\rangle\geqslant 0.\end{aligned}$$

Logo,$\mathsf{T}_1+\mathsf{T}_2$ é um operador monótono. □

Exemplo 2.2. *Seja* $\partial f:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$*, onde* f *é uma função convexa própria semicontínua inferiormente. Então* ∂f *é um operador monótono.*

De fato, sejam $x,y\in\mathbb{R}^n$, $\xi\in\partial f(x)$ e $\eta\in\partial f(y)$. Pela definição de ∂f, obtemos

$$\langle\xi,y-x\rangle\leqslant f(y)-f(x) \tag{2.3}$$

$$-\langle \eta, y - x\rangle = \langle \eta, x - y\rangle \leqslant f(x) - f(y). \tag{2.4}$$

Então, somando (2.3) e (2.4) obtemos

$$\langle \xi, y - x\rangle - \langle \eta, y - x\rangle \leqslant 0$$

Logo,

$$\langle \xi - \eta, y - x\rangle \leqslant 0 \Rightarrow \langle \xi - \eta, x - y\rangle \geqslant 0.$$

Portanto, ∂f é um operador monótono.

Definição 2.6. *Um operador* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é chamado monótono maximal quando*
i) T *é monótono*
ii) Para todo T' *operador monótono tal que* $T(x) \subset T'(x)$ *para todo* $x \in \mathbb{R}^n$*, tem-se* $T = T'$*.*

Teorema 2.1. *Sejam* $T_1 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *e* $T_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *operadores monótonos maximais tais que* $(\mathrm{dom} T_1) \cap (\mathrm{dom} T_2)^0 \neq \emptyset$ *.*
i) Então $T_1 + T_2$ *é um operador monótono maximal.*
Em adição, temos.
ii) Se T_2 *é o subdiferencial de uma função convexa própria fechada e* T_2 *é sobrejetivo, então* $T_1 + T_2$ *é sobrejetivo.*

Demonstração. Ver [3]. □

A seguir provaremos que $T = \partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é um operador monótono maximal.
Sejam $f : \mathbb{R}^n \to \mathbb{R}$ e $x_0 \in \mathbb{R}^n$, defina a função $F : \mathbb{R}^n \to \mathbb{R}$ por

$$F(x) := f(x) + \frac{1}{2}\|x - x_0\|^2. \tag{2.5}$$

Mostraremos agora que F possui um único minimizador.
De fato, considere $h(x) = \frac{1}{2}\|x - x_0\|^2$, $\forall\ x \in \mathbb{R}^n$, então $F(x) = f(x) + h(x), \forall\ x \in \mathbb{R}^n$.
Como f convexa e h é estritamente convexa, temos pela Proposição 1.1 que $F = f + h$ é estritamente convexa.
Temos que F é coerciva em $\mathbb{R}^n$ pelo exemplo 1.1. Com isso, temos que F possui um minimizador global. Sendo F estritamente convexa, esse minimizador é único.
Assim, podemos definir $\mathrm{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$, dado por

$$\mathrm{prox}_f(x_0) := \mathrm{argmin}\{f(x) + \frac{1}{2}\|x - x_0\|^2\}. \tag{2.6}$$

Lema 2.1.1. *Sejam* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa semicontínua inferiormente e* $x, y, z \in \mathbb{R}^n$. *As seguintes condições são equivalentes:*

i) $z = x + y$ *e* $f(x) + f^*(y) = \langle x, y\rangle$;

ii) $x = \text{prox}_f(z)$ *e* $y = \text{prox}_{f^*}(z)$.

Demonstração. Faremos algumas considerações preliminares antes de provarmos as equivalências. Sendo $h(x) = \frac{1}{2}\|x - x_0\|^2$, temos que $F = f + h$. Fazendo $x = \text{prox}_f(z)$, temos que x é um minimizador da função F se, e somente se, $0 \in \partial F(x)$, o que é equivalente a $0 \in (\partial f(x) + \partial h(x))$. Mas, $\nabla h(x) = x - z$, isto implica

$$x = \text{prox}_f(z) \Leftrightarrow 0(\in \partial f(x) + \{x - z\}), \tag{2.7}$$

ou seja, existe $y \in \partial f(x)$ tal que $z = x + y$. Analogamente, podemos mostrar que

$$y = \text{prox}_f^*(z). \tag{2.8}$$

Isto é, existe $x \in \partial f^*(y)$ tal que $z = x + y$. O Teorema 1.10 e a Proposição 1.5 implicam

$$y \in \partial f(x) \Leftrightarrow f(x) + f^*(y) = \langle x, y\rangle \Leftrightarrow x \in \partial f^*(y). \tag{2.9}$$

$(i) \Rightarrow (ii)$ Por hipótese $z = x + y$ e $f(x) + f^*(y) = \langle x, y\rangle$. Então, da equação (2.9) obtemos

$$y \in \partial f(x) \text{ e } x \in \partial f^*(y).$$

Combinando estes resultados com (i),(2.7) e (2.8) obtemos

$$x = \text{prox}_f(z) \text{ e } y = \text{prox}_f^*(z).$$

$(i) \Leftarrow (ii)$ Seja $x = \text{prox}_f(z)$. De (2.7) obtemos a existência de um $y_0 \in \partial f(x)$ tal que $z = x + y_0$. Observemos a equação (2.9) implica

$$x \in \partial f(y_0) \text{ e } f(x) + f^*(y_0) = \langle x, y_0\rangle.$$

Daí, combinando esse resultado com (2.8) obtemos $y_0 = \text{prox}_f^*(z) = y$. Segue que

$$z = x + y \text{ e } f(x) + f^*(y) = \langle x, y\rangle.$$

$\square$

Teorema 2.2. *Se* $f : \mathbb{R}^n \to \mathbb{R}$ *é um função convexa semicontínua inferiormente, então* $\partial f : \mathbb{R}^n \rightrightarrows R^n$ *é um operador monótono maximal.*

Demonstração. Pelo exemplo 2.2, ∂f é um operador monótono. Seja $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ um operador monótono tal que $\partial f(x) \subset T(x)$, para todo $x \in \mathbb{R}^n$. Provemos agora que $T(x) \subset \partial f(x)$, para todo $x \in \mathbb{R}^n$.

Dados $x_0, y_0 \in \mathbb{R}^n$, com $y_0 \in T(x_0)$ e $y \in T(x)$. Como T é monótono, obtemos

$$\langle x - x_0, y - y_0 \rangle \geqslant 0, \tag{2.10}$$

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

Considere $x_1 = \mathrm{prox}_f(x_0 + y_0)$ e $y_1 = \mathrm{prox}_{f^*}(x_0 + y_0)$. Pelo Lema 2.1.1, $x_0 + y_0 = x_1 + y_1$ e $f(x_1) + f^*(y_1) = \langle x_1, y_1 \rangle$.

Então, pelo Teorema 1.10, $y_1 \in \partial f(x_1) \subset T(x_1)$. Daí, tomando $x = x_1$ e $y = y_1$ em (2.10), obtemos

$$\langle x_1 - x_0, y_1 - y_0 \rangle \geqslant 0. \tag{2.11}$$

Assim,

$$0 \leqslant \langle x_1 - x_0, x_0 - x_1 \rangle = -\langle x_1 - x_0, x_1 - x_0 \rangle = -\|x_1 - x_0\|^2.$$

Segue que, $x_1 = x_0$. Portanto, $y_0 = y_1 \in \partial f(x_1) = \partial f(x_0)$.

Como x_0 e y_0 são arbitrários, tem-se que $T(x) \subset \partial f(x),\ \forall\, x \in \mathbb{R}^n$.

Então, $\partial f(x) = T(x),\ \forall\, x \in \mathbb{R}^n$. Logo, $\partial f : \mathbb{R}^n \to P(\mathbb{R}^n)$ é um operador monótono maximal. □

Exemplo 2.3. *Seja* C *um subconjunto convexo e fechado do* $\mathbb{R}^n$*. A aplicação* $P_C(x) : R^n \to C$ *que associa cada* $x \in \mathbb{R}^n$ *o ponto de* C *mais próximo de* x *é chamada de projeção de* x *sobre* C*. Essa aplicação é um exemplo de operador monótono maximal que não é o subdiferencial de uma função convexa semicontínua inferiormente.*

2.2 Operadores Paramonótonos

Definição 2.7. *Seja* $T : \mathbb{R}^n \to \mathbb{R}^n$ *um operador. Diz-se que* T *é paramonótono quando é monótono e* $\langle T(x) - T(y), x - y \rangle = 0$ *implica* $T(x) = T(y)$.

Para operadores ponto-conjunto temos a seguinte definição.

Definição 2.8. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *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 2.3. *Seja* $T = \partial f(x)$, *com* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa semicontínua inferiormente. Então* T *é um operador paramonótono.*

Demonstração. Pelo Exemplo 2.2 ∂f é um operador monótono. Dados $x, y \in \mathbb{R}^n$, considere $\langle u - v, x - y \rangle = 0$, onde $u \in \partial f(x)$ e $v \in \partial f(y)$. Defina a função $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ por

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

Note que, $\bar{f}$ é convexa, pois é soma de funções convexas.
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 $\mathbb{R}^n$ pelo Teorema 1.8.
Como $\langle u - v, x - y \rangle = 0$, temos

$$\langle u, x - y \rangle = \langle v, x - y \rangle. \tag{2.12}$$

Sendo $u \in \partial f(x)$ temos $\langle u, y - x \rangle \leqslant f(y) - f(x)$, ou seja,

$$f(x) - f(y) \leqslant \langle u, x - y \rangle. \tag{2.13}$$

De maneira análoga, como $v \in \partial f(y)$, temos

$$\langle v, x - y \rangle \leqslant f(x) - f(y). \tag{2.14}$$

Segue de (2.12), (2.13) e (2.14), 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, temos que $0 \in \partial \bar{f}(y)$.
Portanto, $0 = w - u$ para algum $w \in \partial f(y)$, isto é, $u \in \partial f(y)$.
Agora definamos, $\bar{g} : \mathbb{R}^n \to \mathbb{R}$ 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 $0 \in \partial \bar{g}(y)$. Portanto, pelo Teorema 1.8, temos que y é um minimizador de $\bar{g}$ em $\mathbb{R}^n$.
De (2.12), (2.13) e (2.14), 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 $\mathbb{R}^n$. Então, pelo Teorema 1.8, temos 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 2.2. *Se $T_1 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ e $T_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ são operadores paramonótonos, tais que $\mathrm{dom}(T_1) \cap \mathrm{dom}(T_2) \neq \emptyset$. Então $T_1 + T_2$ é um operador paramonótono.*

Demonstração. Sejam $x, y \in \mathbb{R}^n$, $u \in (T_1 + T_2)(x)$ 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 2.1, $T_1 + T_2$ 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. □

2.3 Operadores Pseudomonótonos

Definição 2.9. *Seja $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ um operador tal que $\mathrm{dom}(T)$ é convexo e fechado. Dizemos que T é pseudomonótono se satisfaz a seguinte condição:*

Se para toda sequência $\{x^k\} \subset \mathrm{dom}(T)$ convergindo para o ponto $\bar{x} \in \mathrm{dom}(T)$ e todo $u^k \in T(x^k)$ para todo $k \in \mathbb{N}$ tem-se

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

então, existe $\bar{u} \in T(\bar{x})$ tal que

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

para todo $y \in \mathrm{dom}(T)$.

Proposição 2.3. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa semicontínua inferiormente. Então $T = \partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é um operador pseudomonótono.*

Demonstração. Seja $\{x^k\}$ uma sequência em $\mathrm{dom}(\partial f)$ tal que
$\lim_{k\to\infty} x^k = \bar{x} \in \mathrm{dom}(\partial f)$ 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 $y \in \mathrm{dom}(\partial f)$ e todo $k \in \mathbb{N}$. Então,

$$\liminf_{k\to\infty}\langle u^k, x^k - y\rangle \geqslant \liminf_{k\to\infty}[f(x^k) - f(y)] = f(\bar{x}) - f(y). \tag{2.15}$$

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,

$$\langle \bar{u}, \bar{x} - y\rangle \leqslant f(\bar{x}) - f(y). \tag{2.16}$$

Portanto de (2.15) e (2.3), 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. □

Exemplo 2.4. *Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador monótono contínuo. Então T é um exemplo de operador pseudomonótono que não é o subdiferencial de uma função convexa semicontínua inferiormente.*

Proposição 2.4. *Sejam $T, S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ operadores pseudomonótonos tais que $\mathrm{dom}(T) \cap \mathrm{dom}(S) \neq \emptyset$. Então $T + S$ é pseudomonótono.*

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

$$\limsup_{k\to\infty}\langle u^k, x^k - \bar{x}\rangle \leqslant 0, \tag{2.17}$$

onde $u^k \in (T + S)(x^k)$.
Seja $y \in \mathrm{dom}(T+S)(x^k)$. Então, existem $u_T^k \in T(x^k)$ e $u_S^k \in S(x^k)$ tal que $u^k = u_T^k + u_S^k$.
De (2.17), $\limsup_{k\to\infty}\langle u_T^k + u_S^k, x^k - \bar{x}\rangle \leqslant 0$. Então

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

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

Como T e S são pseudomonótonos, existem $\bar{u_T} \in T(\bar{x})$ e $\bar{u_S} \in S(\bar{x})$ tais que
$\langle \bar{u_T}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_T^k, x^k - y\rangle$, $\langle \bar{u_S}, \bar{x} - y\rangle \leqslant \liminf_{k\to\infty}\langle u_S^k, x^k - y\rangle$,
para todo $y \in \mathrm{dom}(T + S)$.

Logo, $\langle \bar{u_T} + \bar{u_S}, \bar{x} - y \rangle \leqslant \liminf_{k \to \infty} \langle u_T^k + u_S^k, x^k - y \rangle$.
Seja $\bar{u} = \bar{u_T} + \bar{u_S} \in T(\bar{x}) + S(\bar{x}) = (T + S)(\bar{x})$. Então,

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

para todo $y \in \mathsf{dom}(T + S)$. Portanto, $T + S$ é pseudomonótono. $\square$

Capítulo 3

O Algoritmo do Ponto Proximal

3.1 O Algoritmo do Ponto Proximal para Operadores Monótonos Maximais(APPOMM)

Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa semicontínua inferiormente. O Algoritmo do Ponto Proximal (ver Apêndice) gera, para um ponto de inicial $x^0 \in \mathbb{R}^n$, a sequência $\{x^k\} \subset \mathbb{R}^n$ pela iteração

$$x^{k+1} := \arg\min_{x \in \mathbb{R}^n} \{f(x) + \lambda_k \|x - x^k\|^2\},$$

onde $\{\lambda_k\}$ é uma sequência de números positivos satifazendo,

$$0 < \lambda_k \leqslant \bar{\lambda},$$

para algum $\bar{\lambda} > 0$ e $\| \cdot \|$ é a norma euclideana. A sequência das iteradas converge para solução do problema

$$\min_{x \in \mathbb{R}^n} f(x). \tag{3.1}$$

Equivalentemente, temos pelo Teorema1.8

$$0 \in \partial f(x^{k+1}) + 2\lambda_k(x^{k+1} - x^k),$$

ou seja

$$2\lambda_k(x^k - x^{k+1}) \in \partial f(x^{k+1}),$$

assim

$$x^k \in (I + \frac{1}{2\lambda_k}\partial f)(x^{k+1}). \tag{3.2}$$

Motivados por essa observação e lembrando que o subdiferencial ∂f de uma função convexa semicontínua inferiormente , é um operador monótono maximal, podemos extender o Algoritmo do Ponto Proximal para encontrar zeros de Operadores Monótonos Maximais $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

O Algoritmo do Ponto Proximal para Operadores Monótonos Maximais (APPOMM) gera uma sequência $\{x^k\} \subset \mathbb{R}^n$, da seguinte forma

$$x^0 \in \mathbb{R}^n,$$

$$x^k \in (I + \frac{1}{\lambda_k} T)(x^{k+1}),$$

ou seja

$$x^{k+1} \in (I + \frac{1}{\lambda_k} T)^{-1}(x^k),$$

onde λ_k é uma sequência de números positivos satisfazendo

$$0 < \lambda_k \leqslant \bar{\lambda},$$

para algum $\bar{\lambda} > 0$.

Observação 3.1. *Para mostrarmos a bem definição e a convergência desse algoritmo utilizaremos a norma euclideana.*

Agora, veremos algumas definições e resultados que serão úteis na análise de convergência do (APPOMM).

Definição 3.1. *Uma sequência $\{y^k\} \subset \mathbb{R}^n$ é dita Fejér convergente a um conjunto $U \subset \mathbb{R}^n$, com respeito a norma euclidiana, se*

$$\|y^{k+1} - u\| \leqslant \|y^k - u\|, \ \forall \ k \geqslant 0 \text{ e } \forall \ u \in U. \tag{3.3}$$

Proposição 3.1. *Se $\{y^k\}$ é uma sequência Fejér convergente a um conjunto $U \neq \emptyset$, então $\{y^k\}$ é limitada. Além disso se um ponto de acumulação $\bar{y}$ de $\{y^k\}$ pertence a U, então $\lim\limits_{k\to\infty} y^k = \bar{y}$.*

Demonstração. Seja $u \in U$, um ponto fixado. Como $\{y^k\}$ é Fejér convergente a U, temos que

$$\|y^{k+1} - u\| \leqslant \|y^k - u\| \leqslant ... \leqslant \|y^0 - u\|.$$

Portanto, $\{y^k\}$ está contida na bola $B = B(u, \|y^0 - u\|)$. Logo, $\{y^k\}$ é limitada.

Para a segunda parte considere uma subsequência $\{y^{k_j}\}$ de $\{y^k\}$, tal que $\lim\limits_{k\to\infty} y^{k_j} = \bar{y}$.

Como $\bar{y} \in U$ e $\{y^k\}$ é Fejér convergente, pela Definição 3.1 a sequência $\{\|y^k - \bar{y}\|\}$ é não-crescente e não-negativa, portanto convergente.

Como, a subsequência $\{\|y^{k_j} - \bar{y}\|\}$ converge a zero, segue que $\{\|y^k - \bar{y}\|\}$ converge a zero. Isto é, $\lim\limits_{k\to\infty} \|y^k - \bar{y}\| = 0$. Logo, $\lim\limits_{k\to\infty} y^k = \bar{y}$. □

Definição 3.2. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador ponto-conjunto, dizemos que*

i) T *é sobrejetivo quando, para todo* $y \in \mathbb{R}^n$*, existe* $x \in \mathbb{R}^n$*, tal que* $y \in T(x)$.

ii) T *é injetivo se, para* $x \neq y$ *tivermos* $T(x) \cap T(y) = \emptyset$.

Teorema 3.1. *(Teorema de Minty)*

Se $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é um operador monótono maximal e* $\lambda > 0$*, então o operador* $I + \lambda T$ *é injetivo e sobrejetivo.*

Demonstração. Ver [13]. □

O Teorema 3.1 garante que (APPOMM) está bem definido, como veremos na Proposição seguinte.

Proposição 3.2. *A sequência* $\{x^k\}$ *gerada pelo (APPOMM) está bem definida.*

Demonstração. De fato, seja $x^0 \in \mathbb{R}^n$, e seja $x^k \in \mathbb{R}^n$, sendo $I + \lambda T$ sobrejetivo, temos que existe $x^{k+1} \in \mathbb{R}^n$ para todo $k \geqslant 0$ tal que

$$x^k \in (I + \frac{1}{\lambda_k} T)(x^{k+1}),$$

portanto

$$x^{k+1} \in (I + \frac{1}{\lambda_k} T)^{-1}(x^k).$$

A injetividade de $(I + \lambda T)$ garante a unicidade de x^{k+1}. □

Lema 3.1.1. *Se* $\lim\limits_{k\to\infty} y^k = \bar{y}$*,* $\lim\limits_{k\to\infty} z^k = \bar{z}$*,* T *é um operador monótono maximal e* $y^k \in T(z^k)$*, então* $\bar{y} \in T(\bar{z})$.

Demonstração. Defina,

$$T'(z) = \begin{cases} T(z), & \text{se} \quad z \neq \bar{z} \\ T(\bar{z}) \cup \{\bar{y}\}, & \text{se} \quad z = \bar{z} \end{cases} . \tag{3.4}$$

Afirmação: T' é monótono. Então, precisamos verificar que:

$$\langle y - y', z - z' \rangle \geqslant 0,\ \forall\ z, z',\ \forall\ y \in T'(z) \text{ e } \forall\ y' \in\ T'(z'). \tag{3.5}$$

Da definição de T' e T sendo monótono, basta verificarmos (3.5) para $y' = \bar{y}$ e $z' = \bar{z}$. Por hipótese temos, $\{y^k\} \in T(z^k)$, portanto

$$\langle y - y^k, z - z^k \rangle \geqslant 0,\ \forall\ z,\ \forall\ y \in T(z). \tag{3.6}$$

Tomando o limite em (3.6) quando $k \to \infty$, obtemos

$$\langle y - \bar{y}, z - \bar{z} \rangle \geqslant 0,\ \forall\ z,\ \forall\ y \in T(z).$$

Logo, T' é um operador monótono.
Como $T(x) \subset T'(x),\ \forall\ x \in \mathbb{R}^n$ e por hipótese T é monótono maximal, concluímos que $T = T'$.
Em particular, $T(\bar{z}) = T'(\bar{z}) = T(\bar{z}) \cup \{\bar{y}\}$. Logo, $\bar{y} \in T(\bar{z})$. □

O lema acima mostra que os Operadores Monótonos Maximais são fechados. Veremos agora resultados que garantem a convergência do (APPOMM).

Lema 3.1.2. *Sejam* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e o conjunto* $U = \{x \in \mathbb{R}^n;\ 0 \in T(x)\}$ *não-vazio. Então, a sequência* $\{x^k\}$ *gerada pelo (APPOMM) satisfaz a seguinte desigualdade*

$$\|x^{k+1} - \bar{x}\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2,\ \forall\ k \geqslant 0 \text{ e } \forall\ \bar{x} \text{ tal que } 0 \in T(\bar{x}).$$

Demonstração. De fato, seja $\bar{x} \in U$, então

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

Assim,

$$\|x^k - \bar{x}\|^2 - \|x^k - x^{k+1}\|^2 - \|x^{k+1} - \bar{x}\|^2 = 2\frac{1}{\lambda_k}\langle \lambda_k(x^k - x^{k+1}), x^{k+1} - \bar{x} \rangle.$$

Como,

$$x^{k+1} \in (I + \frac{1}{\lambda_k}T)^{-1}(x^k),$$

temos

$$\lambda_k(x^k - x^{k+1}) \in T(x^{k+1}).$$

Então, como $0 \in T(\bar{x})$, temos pela monotonicidade de T

$$\langle \lambda_k(x^k - x^{k+1}), x^{k+1} - \bar{x}\rangle = \langle \lambda_k(x^k - x^{k+1}) - 0, x^{k+1} - \bar{x}\rangle \geqslant 0.$$

Portanto,

$$\|x^k - \bar{x}\|^2 - \|x^k - x^{k+1}\|^2 - \|x^{k+1} - \bar{x}\|^2 \geqslant 0.$$

Logo,

$$\|x^{k+1} - \bar{x}\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^k - x^{k+1}\|^2, \ \forall\, k \geqslant 0 \text{ e } \forall\, \bar{x} \text{ tal que } 0 \in T(\bar{x}).$$

□

Lema 3.1.3. *Sejam* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e o conjunto* $U = \{x \in \mathbb{R}^n;\ 0 \in T(x)\}$ *não-vazio. Então,* $\lim_{k\to\infty}(x^{k+1} - x^k) = 0$*, onde* $\{x^k\}$ *é a sequência gerada pela (APPOMM).*

Demonstração. Do Lema 3.1.2, temos

$$0 \leqslant \|x^{k+1} - x^k\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2, \ \forall\, k \geqslant 0. \tag{3.7}$$

Mas, $0 \leqslant \|x^{k+1} - \bar{x}\|^2 \leqslant \|x^k - \bar{x}\|^2$, implicando que a sequência $\{\|x^k - \bar{x}\|\}$ é não-crescente e não-negativa. Portanto, convergente.

Assim,

$$\lim_{k\to\infty}(\|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2) = 0.$$

Logo, passando o limite em (3.7) quando $k \to \infty$, obtemos

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

□

Lema 3.1.4. *Sejam* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e o conjunto* $U = \{x \in \mathbb{R}^n;\ 0 \in T(x)\}$ *não-vazio. Então, a sequência* $\{x^k\}$ *gerada por (APPOMM) possui pontos de acumulação e todos pertencem a* U*.*

Demonstração. Pelo Lema 3.1.2, $\{x^k\}$ é Fejér convergente a U, portanto é limitada pela Proposição 3.1, então $\{x^k\}$ possui pontos de acumulação. Sejam $\hat{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim_{k\to\infty} x^{k_j} = \hat{x}$.

Por hipótese, temos que

$$x^{k+1} \in (I + \frac{1}{\lambda_k}T)^{-1}(x^k),$$

então

$$\lambda_{k_j}(x^{k_j} - x^{k_j+1}) \in T(x^{k_j+1}).$$

Pelo Lema 3.1.3 $\lim_{k\to\infty}(x^{k_j} - x^{k_j+1}) = 0$ e $\lambda_{k_j} \leqslant \bar{\lambda}$, obtemos

$$\lim_{k\to\infty} \lambda_{k_j}(x^{k_j} - x^{k_j+1}) = 0.$$

Daí, aplicando o Lema 3.1.1, para $\bar{y} = 0$, $y^k = \lambda_{k_j}(x^{k_j} - x^{k_j+1})$, $z^k = x^{k_j+1}$ e $\bar{z} = \hat{x}$, temos que $0 \in T(\hat{x})$. Logo, $\hat{x} \in U$. □

Teorema 3.2. *Sejam* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e o conjunto* $U = \{x \in \mathbb{R}^n;\ 0 \in T(x)\}$ *não-vazio. Então, a sequência* $\{x^k\}$*, gerada por (APPOMM), converge a um elemento de* U*.*

Demonstração. Pelo Lema 3.1.2, $\{x^k\}$ é Fejér convergente a U. Assim, pela Proposição 3.1 $\{x^k\}$ é limitada. Seja $\hat{x}$ ponto de acumulação de $\{x^k\}$. Pelo Lema 3.1.4, $\hat{x} \in U$. Portanto $\lim_{k\to\infty} x^k = \hat{x}$ pela Proposição 3.1. Logo, $\{x^k\}$ converge a um elemento de U. □

Capítulo 4

Algoritmo do Ponto Proximal com Distâncias de Bregman para o Problema de Desigualdade Variacional

4.1 Problema de Desigualdade Variacional

A extensão natural do problema

$$\min \mathsf{f}(x) \text{ sujeito a } x \in \mathsf{C}, \tag{4.1}$$

(onde $\mathsf{C} \subset \mathbb{R}^n$ é fechado e convexo) a operadores monótonos é chamado o problema de desigualdade variacional, definido segundo a definição abaixo.

Definição 4.1. *Dados* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador monótono maximal e* $\mathsf{C} \subset \mathbb{R}^n$ *um conjunto convexo e fechado o Problema de Desigualdade Variacional* $\mathsf{VIP}(\mathsf{T}, \mathsf{C})$ *consiste em encontrar* $z \in \mathsf{C}$ *tal que existe* $u \in \mathsf{T}(z)$ *satisfazendo*

$$\langle u, x - z \rangle \geqslant 0 \tag{4.2}$$

para todo $x \in \mathsf{C}$.

Observação 4.1. *Se* $\mathsf{T} = \partial \mathsf{f}$*, para* $\mathsf{f} : \mathbb{R}^n \to \mathbb{R}$ *convexa, temos que* $u \in \mathsf{T}(z) = \partial \mathsf{f}(z)$ *se, e somente se,* $0 \leqslant \langle u, x - z \rangle \leqslant \mathsf{f}(x) - \mathsf{f}(z)$ *para todo* $x \in \mathsf{C}$*, e portanto* z *minimiza* f *em* C*. Assim, o problema* $\mathsf{VIP}(\mathsf{T}, \mathsf{C})$ *generaliza o problema de minimizar funções convexas em conjuntos convexos.*

4.2 Algoritmo do Ponto Proximal Generalizado para o $\mathsf{VIP}(T, C)$

Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa. O Algoritmo do Ponto Proximal com Distância de Bregman gera(Ver apêndice), para um ponto inicial $x^0 \in S$, a sequência $\{x^k\} \subset S$ pela iteração

$$x^{k+1} := \arg\min_{x \in \overline{S}} \{f(x) + \lambda_k D_h(x, x^k)\},$$

onde h é uma função de Bregman com zona S e $\{\lambda_k\}$ é uma sequência de números positivos satisfazendo,

$$0 < \lambda_k \leqslant \bar{\lambda},$$

para algum $\bar{\lambda} > 0$. A sequência das iteradas converge para solução do problema

$$\min f(x) \ \mathit{sujeito\ a} \ x \in \overline{S}, \tag{4.3}$$

com $S \subset \mathbb{R}^n$ aberto e convexo, $\overline{S}$ o fecho de S, e f contínua em $\overline{S}$.

Equivalentemente, temos pelo Teorema 1.8

$$0 \in \partial[f(\cdot) + \lambda_k D_h(\cdot, x^k)](x^{k+1}),$$

ou seja

$$0 \in [\partial f(\cdot) + \lambda_k \partial_x D_h(\cdot, x^k)](x^{k+1}).$$

Com essa observação e lembrando que o subdiferencial de uma função convexa semicontínua inferiormente ∂f é um operador monótono maximal, podemos estender o Algoritmo do Ponto Proximal com Distância de Bregman para resolver o $\mathsf{VIP}(T, C)$, onde $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é um operador monótono maximal.

O algoritmo do ponto proximal generalizado para o $\mathsf{VIP}(T, C)$ é definido da seguinte forma.

Considere uma função de Bregman h com zona C^0 e uma sequência de números positivos $\{\lambda_k\}$ limitada acima por $\bar{\lambda} > 0$. Seja $\{x^k\}$ a sequência definida por

Inicialização:

$$x^0 \in C^0. \tag{4.4}$$

Iteração: Dado $x^k \in C^0$, defina $T_k : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ por $T_k(\cdot) = T(\cdot) + \lambda_k \partial_x D_h(\cdot, x^k)$ e encontre $x^{k+1} \in C^0$ tal que

$$0 \in T_k(x^{k+1}). \tag{4.5}$$

Equivalentemente a (4.5) temos

$$x^{k+1} \in [\partial_x D_h(\cdot, x^k) + \frac{1}{\lambda_k} T]^{-1}(x^k) \tag{4.6}$$

e também equivalente a (4.5) temos

$$\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1}) \tag{4.7}$$

pela Proposição 1.6 (ii).

Apresentaremos agora alguns resultados que utilizaremos na análise de convergência do algoritmo do ponto proximal generalizado para o $VIP(T, C)$, para isso suponhamos que o $VIP(T, C)$ possua soluções.

Teorema 4.1. *Seja* $\{x^k\}$ *a sequência gerada por* (4.4) *e* (4.5)*. Suponha que* $dom(T) \cap C^0 \neq \emptyset$ *e a função de Bregman* h *é zona coerciva com zona* C^0*. Então* $\{x^k\}$ *está bem definida e contida em* C^0*.*

Demonstração. Faremos a prova por indução sobre k. De (4.4) $x^0 \in C^0$. Suponha que $x^k \in C^0$. Seja $B_k(\cdot) := \lambda_k \partial_x D_h(\cdot, x^k)$, temos $dom(B_k) = C^0$ (ver Apêndice Lema A.2.1 seção A.2).

Assim, $T_k = T + B_k$ e $dom(T) \cap dom(B_k) = dom(T) \cap C^0 \neq \emptyset$. Então, T_k é monótono ma-

ximal pela Teorema 2.1 (i). Sendo h estritamente convexa por (B_2), temos que B_k é estritamente monótono, então T_k é estritamente monótono.

Como h é zona coerciva, B_k é sobrejetivo. Então, pelo Teorema 2.1 (ii) T_k é sobrejetivo.

Logo, T_k possui um zero em $dom(T_k)$, que é único pela monotonicidade estrita de T_k .

Denotemos este zero por x^{k+1}.

Como $dom(T_k) = dom(T) \cap dom(B_k) = dom(T) \cap C^0$ e x^{k+1} pertence a $dom(T_k)$, obtemos que $x^{k+1} \in C^0$ para todo k.

Portanto, $\{x^k\}$ está bem definida e contida em C^0. □

Em posse das hipóteses do Teorema 4.1, temos os seguintes lemas.

Lema 4.2.1. *A sequência* $\{x^k\}$ *gerada por* (4.4) *e* (4.5) *satisfaz a seguinte desigualdade*

$$D_h(\bar{x}, x^{k+1}) \leqslant D_h(\bar{x}, x^k) - D_h(x^{k+1}, x^k)$$

para todo k e cada solução $\bar{x}$ *do* $VIP(T, C)$*. Consequentemente a sequência* $\{D_h(\bar{x}, x^k)\}$ *é não crescente.*

Demonstração. De (4.7), temos $\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1})$. Então, existe $u^k \in T(x^{k+1})$ tal que

$$u^k = \lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in T(x^{k+1}). \tag{4.8}$$

Pelo item (i) da Proposição 1.6, $D_h(x, y) - D_h(x, z) - D_h(z, y) = \langle \nabla h(y) - \nabla h(z), z - x \rangle$, para cada $x \in C$ e todo $y, z \in C^0$. Tomando $y = x^k$, $z = x^{k+1}$ e $x = y$, obtemos

$$\langle \nabla h(x^k) - \nabla h(x^{k+1}), x^{k+1} - y \rangle = D_h(y, x^k) - D_h(y, x^{k+1}) - D_h(x^{k+1}, x^k). \tag{4.9}$$

De (4.8) e (4.9), tem-se

$$\langle u^k, x^{k+1} - y \rangle = \lambda_k[D_h(y, x^k) - D_h(y, x^{k+1}) - D_h(x^{k+1}, x^k)]. \tag{4.10}$$

Seja $\bar{x}$ solução do $VIP(T, C)$ e tome $v \in T(\bar{x})$ tal que

$$\langle v, x - \bar{x} \rangle \geqslant 0, \text{ para todo } x \in C. \tag{4.11}$$

Como T é monótono, $\langle u^k - v, x^{k+1} - \bar{x} \rangle \geqslant 0$. Assim de (4.11), temos

$$\langle u^k, x^{k+1} - \bar{x} \rangle \geqslant \langle v, x^{k+1} - \bar{x} \rangle \geqslant 0. \tag{4.12}$$

Tomando $y = \bar{x}$ em (4.10) e usando (4.12), obtemos

$$D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1}) - D_h(x^{k+1}, x^k) \geqslant 0.$$

Daí, sendo D_h não negativa, temos

$$D_h(\bar{x}, x^{k+1}) \leqslant D_h(\bar{x}, x^k).$$

□

Lema 4.2.2. *A sequência* $\{x^k\}$ *gerada por* (4.4) *e* (4.5) *é limitada e possui pontos de acumulação.*

Demonstração. Pelo Lema 4.2.1 temos que a sequência $\{D_h(\bar{x}, x^k)\}$ é não crescente e não negativa. Então, $D_h(\bar{x}, x^{k+1}) \leqslant D_h(\bar{x}, x^k) \leqslant ... \leqslant D_h(\bar{x}, x^0)$ para todo k e todo $\bar{x}$ solução do $VIP(T, C)$. Portanto, $\{x^k\}$ está contida no conjunto $\{x \in C; D_h(\bar{x}, x) \leqslant D_h(\bar{x}, x^0)\}$ que é limitado por (B_3). Logo, $\{x^k\}$ é limitada e possui pontos de acumulação. □

Lema 4.2.3. $\lim\limits_{k \to \infty} D_h(x^{k+1}, x^k) = 0.$

Demonstração. Pelo Lema 4.2.1

$$,0 \leqslant D_h(x^{k+1}, x^k) \leqslant D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1}), \tag{4.13}$$

para todo $\bar{x}$ solução do $VIP(T, C)$. Como $\{D_h(\bar{x}, x^k)\}$ é não crescente pelo Lema 4.2.1 e limitada por baixo ($D_h \geqslant 0$), temos que $\{D_h(\bar{x}, x^k)\}$ é convergente. Então,

$$\lim_{k\to\infty} [D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1})] = 0. \tag{4.14}$$

Assim, de (4.13) e (4.14), segue que $\lim_{k\to\infty} D_h(x^{k+1}, x^k) = 0$.

□

Lema 4.2.4. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *pseudomonótono. Se* $\tilde{x}$ *é um ponto de acumulação de* $\{x^k\}$*, então existe* $\tilde{u} \in T(\tilde{x})$ *tal que* $\langle \tilde{u}, x^* - \tilde{x}\rangle = 0$*, para todo* x^* *solução do* $VIP(T, C)$*.*

Demonstração. Seja $\tilde{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim_{k\to\infty} x^{k_j} = \tilde{x} \in C$, pois C é fechado.
Tomando $y = \tilde{x}$ e $k = k_j$ em (4.10), obtemos

$$\langle u^{k_j}, x^{k_j+1} - \tilde{x}\rangle = \lambda_{k_j}[D_h(\tilde{x}, x^{k_j}) - D_h(\tilde{x}, x^{k_j+1}) - D_h(x^{k_j+1}, x^{k_j})]. \tag{4.15}$$

Observe que

1) $\{x^{k_j+1}\}$ é limitada pelo Lema 4.2.2.

2) $\lim_{k\to\infty} x^{k_j} = \tilde{x}$.

3) $\lim_{k\to\infty} D_h(x^{k_j+1}, x^{k_j}) = 0$ pelo Lema 4.2.3.

Então, aplicando (B_5) para as sequências $\{x^{k_j+1}\}$ e $\{x^{k_j}\}$, temos

$$\lim_{k\to\infty} x^{k_j+1} = \tilde{x}. \tag{4.16}$$

Em posse de (4.16) e (2), podemos aplicar (B_4) as sequências $\{x^{k_j+1}\}$ e $\{x^{k_j}\}$, então

$$\lim_{k\to\infty} D_h(\tilde{x}, x^{k_j+1}) = 0 \text{ e } \lim_{k\to\infty} D_h(\tilde{x}, x^{k_j}) = 0.$$

Assim,

$$\lim_{k\to\infty} [D_h(\tilde{x}, x^{k_j+1}) - D_h(\tilde{x}, x^{k_j})] = 0. \tag{4.17}$$

Como $\{\lambda_k\}$ é limitada. Segue de (3) e (4.17) e fazendo k tender a ∞ em (4.15) que

$$\limsup_{k\to\infty} \langle u^{k_j}, x^{k_j+1} - \tilde{x}\rangle = \lim_{k\to\infty} \langle u^{k_j}, x^{k_j+1} - \tilde{x}\rangle = 0. \tag{4.18}$$

onde $u^{k_j} \in T(x^{k_j+1})$, para todo $j \in \mathbb{N}$.

Sendo T pseudomonótono, então para todo z solução do $\mathrm{VIP}(T, C)$ existe $\tilde{u} \in T(\tilde{x})$ tal que

$$\langle \tilde{u}, \tilde{x} - z \rangle \leqslant \liminf_{k\to\infty} \langle u^{k_j}, x^{k_j+1} - z \rangle. \tag{4.19}$$

Tomando $y = z$ em (4.10)

$$\langle u^k, x^{k+1} - z \rangle = \lambda_k [D_h(z, x^k) - D_h(z, x^{k+1}) - D_h(x^{k+1}, x^k)]. \tag{4.20}$$

Pelo Lema 4.2.1, temos que a sequência $\{D_h(z, x^k)\}$ é não crescente. Sendo $\{\lambda_k\}$ limitada, então em (4.20), temos

$$\lim_{k\to\infty} \langle u^k, x^{k+1} - z \rangle = \lim_{k\to\infty} \lambda_k [D_h(z, x^k) - D_h(z, x^{k+1}) - D_h(x^{k+1}, x^k)] = 0.$$

Assim, $\liminf_{k\to\infty} \langle u^{k_j}, x^{k_j+1} - z \rangle = 0$. Então, de (4.19), temos

$$\langle \tilde{u}, \tilde{x} - z \rangle \leqslant 0, \tag{4.21}$$

com $\tilde{u} \in T(\tilde{x})$.

Por outro lado, como z é solução do $\mathrm{VIP}(T, C)$, existe $v \in T(z)$ tal que $\langle v, y - z \rangle \geqslant 0$ para todo $y \in C$. Além disso, sendo T monótono , $\langle \tilde{u} - v, \tilde{x} - z \rangle \geqslant 0$. Portanto,

$$\langle \tilde{u}, \tilde{x} - z \rangle \geqslant \langle v, \tilde{x} - z \rangle \geqslant 0. \tag{4.22}$$

Logo, de (4.21) e (4.22) segue que $\langle \tilde{u}, \tilde{x} - z \rangle = 0$, para todo z solução do $\mathrm{VIP}(T, C)$. □

Os teorema que veremos a seguir garantem que a sequência gerada por (4.4) e (4.5) converge para uma solução do VIP(T,C).

Teorema 4.2. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador monótono maximal. Considere o problema* $\mathrm{VIP}(T, C)$, *onde* $C \subset \mathbb{R}^n$ *é convexo e fechado. Seja* h *uma função de Bregman com zona* C^0. *Suponha que as seguintes condições são válidas:*

i) $\mathrm{dom}(T) \cap C^0 \neq \emptyset$.

ii) $\mathrm{VIP}(T, C)$ *possui soluções.*

iii) T *é pseudomonótono.*

iv) $0 < \lambda_k \leqslant \bar{\lambda}$, *para algum* $\bar{\lambda} > 0$.

v) Ou, h *é zona coerciva, ou* h *é coerciva na fronteira.*

Então a sequência $\{x^k\}$, *gerada por* (4.4) *e* (4.5), *satisfaz:*

a) A sequência $\{D_h(\bar{x}, x^k)\}$ *é não crescente, para toda solução* $\bar{x}$ *de* $\mathsf{VIP}(T, C)$.

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

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

d) Se $\hat{x}$ *é um ponto de acumulação de* $\{x^k\}$*, então existe* $\hat{u} \in T(\hat{x})$ *tal que* $\langle \hat{u}, x^* - \hat{x}\rangle = 0$*, para toda solução* x^* *do* $\mathsf{VIP}(T, C)$.

Demonstração. a) Lema 4.2.1

b) Lema 4.2.2

c) Lema 4.2.3

d) Lema 4.2.4. □

Teorema 4.3. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador monótono maximal. Considere o problema* $\mathsf{VIP}(T, C)$*, onde* $C \subset \mathbb{R}^n$ *é convexo e fechado. Seja* h *uma função de Bregman com zona* C^0*. Suponha que as seguintes condições são válidas:*

i) $\mathsf{dom}(T) \cap C^0 \neq \emptyset$.

ii) $\mathsf{VIP}(T, C)$ *possui soluções.*

iii) T *é pseudomonótono.*

iv) $0 < \lambda_k \leqslant \bar{\lambda}$*, para algum* $\bar{\lambda} > 0$.

v) Ou, h *é zona coerciva, ou* h *é coerciva na fronteira.*

vi) T *é paramonótono em* C.

Então a sequência $\{x^k\}$*, gerada por* (4.4) *e* (4.5)*, converge para uma solução* $\bar{x}$ *do* $\mathsf{VIP}(T, C)$.

Demonstração. Pelo Lema 4.2.2, $\{x^k\}$ possui pontos de acumulação. Seja $\bar{x}$ um ponto de acumulação de $\{x^k\}$. Pelo Lema 4.2.4 existe $\bar{u} \in T(\bar{x})$ tal que

$$\langle \bar{u}, x^* - \bar{x}\rangle = 0, \tag{4.23}$$

para todo x^* solução do $\mathsf{VIP}(T, C)$.

Como x^* é solução do $\mathsf{VIP}(T, C)$ existe $u \in T(x^*)$ tal que

$$\langle u, x - x^*\rangle \geqslant 0 \tag{4.24}$$

para todo $x \in C$.

Daí, pela monotonicidade de T, $\langle u - \bar{u}, x^* - \bar{x}\rangle \geqslant 0$, ou seja, por (4.23) e (4.24) $0 = \langle \bar{u}, x^* - \bar{x}\rangle \leqslant \langle u, x^* - \bar{x}\rangle \leqslant 0$, segue que

$$\langle \bar{u}, x^* - \bar{x}\rangle = \langle u, x^* - \bar{x}\rangle = 0. \tag{4.25}$$

De (4.25), $\langle u - \bar{u}, x^* - \bar{x}\rangle = 0$ com $u \in T(x^*)$, $\bar{u} \in T(\bar{x})$, então pela paramonotonicidade de T, obtemos

$$u \in T(\bar{x}). \tag{4.26}$$

Portanto, usando (4.25) e em seguida (4.24), temos

$$\langle u, x - \bar{x}\rangle = \langle u, x - x^*\rangle + \langle u, x^* - \bar{x}\rangle \geqslant 0. \tag{4.27}$$

Logo, por (4.26) e (4.27) concluímos que $\bar{x}$ é solução do $VIP(T, C)$.

Seja $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim\limits_{k\to\infty} x^{k_j} = \bar{x}$. Então, por (B_4) $\lim\limits_{k\to\infty} D_h(\bar{x}, x^{k_j}) = 0$.

Pelo Lema 4.2.1 a sequência $\{D_h(\bar{x}, x^k)\}$ é não crescente e não negativa. Assim $\{D_h(\bar{x}, x^k)\}$ é convergente e com subsequência convergindo a zero. Logo $\{D_h(\bar{x}, x^k)\}$ converge a zero. Então segue de (B_4) que $\lim\limits_{k\to\infty} x^k = \bar{x}$.

Portanto, a sequência $\{x^k\}$ gerada por (4.4) e (4.5), converge para uma solução $\bar{x}$ do $VIP(T, C)$. $\square$

Observação 4.2. *A convergência do nosso algoritmo pode ser estabelecida retirando a hipótese de o operador* T *ser pseudomonótono* (ver[*19*]).

Apêndice A

O Algoritmo do Ponto Proximal

A.1 O Algoritmo do Ponto Proximal em $\mathbb{R}^n$

Neste apêndice , é apresentado o algoritmo do ponto proximal para resolver o problema de otimização convexa sem restrições (seção A.1) e com restrições (seção A.2). Apresentaremos nesta seção o algoritmo do ponto proximal em $\mathbb{R}^n$ que foi proposto por Martinet (1970) e desenvolvido por Rockafellar, para resolver o problema de otimização convexa em $\mathbb{R}^n$ como definido em (1).

Com as hipóteses do problema (1) sobre f o Algoritmo do Ponto Proximal gera, para um ponto de inicial $x^0 \in \mathbb{R}^n$, a sequência $\{x^k\} \subset \mathbb{R}^n$, pela iteração

$$x^{k+1} = \arg\min_{x \in \mathbb{R}^n} \{f(x) + \lambda_k \|x - x^k\|^2\}, \tag{A.1}$$

onde $\{\lambda_k\}$ é uma sequência de números positivos satifazendo,

$$0 < \lambda_k \leqslant \bar{\lambda}, \tag{A.2}$$

para algum $\bar{\lambda} > 0$ e $\|\cdot\|$ é a norma euclideana.

O resultado que veremos agora garante que a sequência gerada por (A.1) e (A.2) está bem definida.

Proposição A.1. *A sequência* $\{x^k\}$ *gerada por* (A.1) *e* (A.2) *está bem definida.*

Demonstração. Defina $f_k(x) := f(x) + \lambda_k \|x - x^k\|^2$, então $\lim_{\|x\| \to \infty} f_k(x) = \infty$, pois f é limitada inferiormente. Portanto, f_k é coerciva. Assim f_k possui um mínimo, denotemos este mínimo por x^{k+1}.

Mas, f convexa e $\lambda_k\|x - x^k\|^2$ é estritamente convexa, então f_k é estritamente convexa. Logo, x^{k+1} é único. □

Mostraremos que a sequência gerada pelo Algoritmo do Ponto Proximal converge a uma solução do problema (1)

Lema A.1.1. *Suponhamos que* $f : \mathbb{R}^n \to \mathbb{R}$ *é uma função convexa e continuamente diferenciável e o conjunto de minimizadores de* f*,* $U \subset \mathbb{R}^n$*, é não vazio. Então a sequência* $\{x^k\}$ *gerada por* (A.1) *e* (A.2) *satisfaz a seguinte desigualdade*

$$\|x^{k+1} - \bar{x}\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2, \ \forall\ k \geqslant 0 \text{ e } \forall\ \bar{x} \in U.$$

Demonstração. Seja $\bar{x} \in U$, então

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

Como x^{k+1} satisfaz (A.1), temos

$$0 = \nabla f_k(x^{k+1}) = \nabla f(x^{k+1}) + 2\lambda(x^{k+1} - x^k).$$

Assim,

$$\begin{aligned} \|x^k - \bar{x}\|^2 - \|(x^k - x^{k+1}) - (x^{k+1} - \bar{x})\|^2 &= 2\langle x^k - x^{k+1}, x^{k+1} - \bar{x}\rangle \\ &= \frac{1}{\lambda_k}\langle \nabla f(x^{k+1}), x^{k+1} - \bar{x}\rangle \\ &\geqslant \frac{1}{\lambda_k}\langle f(x^{k+1}), x^{k+1} - \bar{x}\rangle \geqslant 0, \end{aligned}$$

onde a primeira desigualdade segue da convexidade de f e a segunda do fato de $\bar{x}$ ser minimizador do problema (1). □

Lema A.1.2. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa e contínuamente diferenciável e o conjunto de minimizadores de* f*,* $U \subset \mathbb{R}^n$*, é não vazio e a sequência* $\{x^k\}$ *gerada por* (A.1) *e* (A.2)*. Então,* $\lim_{k\to\infty}(x^{k+1} - x^k) = 0$.

Demonstração. Pelo Lema A.1.1, temos

$$0 \leqslant \|x^{k+1} - x^k\|^2 \leqslant \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2.$$

Como $\{x^k\}$ é uma sequência não crescente e não negativa, é convergente. Logo o lado direito da desigualdade acima converge a zero, e segue o resultado. □

Lema A.1.3. *Sejam* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa e contínuamente diferenciável e o conjunto de minimizadores de* f*,* $U \subset \mathbb{R}^n$*, não vazio. Então a sequência* $\{x^k\}$ *gerada por* $(A.1)$ *e* $(A.2)$ *possui pontos de acumulação e todos pertencem a* U*.*

Demonstração. Pelo Lema $A.1.1$, temos que $\{x^k\}$ é uma sequência Fejér convergente a U, assim pela Proposição 3.1, ela é limitada.
Seja $\bar{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$, tal que $\lim\limits_{k\to\infty} x^{k_j} = \bar{x}$. Por $(A.1)$, temos que

$$0 = \nabla f_k(x^{k+1}) = \nabla f(x^{k+1}) + 2\lambda(x^{k+1} - x^k),$$

Assim,

$$\nabla f(x^{k_j+1}) = 2\lambda(x^{k_j} - x^{k_j+1}). \tag{A.3}$$

Pelo Lema $A.1.2$, segue que

$$\lim_{k\to\infty} x^{k_j+1} = \lim_{k\to\infty} x^{k_j} = \bar{x}.$$

Usando o fato que $\lambda_k \leqslant \bar{\lambda}$ e que f é continuamente diferenciável e aplicando ao limite em $(A.3)$, obtemos que $\nabla f(\bar{x}) = 0$. Portanto pela convexidade de f, $\bar{x} \in U$. □

Teorema A.1. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa e contínuamente diferenciável e o conjunto de minimizadores de* f*,* $U \subset \mathbb{R}^n$*, é não vazio. Então a sequência* $\{x^k\}$ *gerada por* $(A.1)$ *e* $(A.2)$*, converge para um ponto* $x^* \in U$*.*

Demonstração. O Lema $A.1.2$ garante que a sequência $\{x^k\}$ gerada por $(A.1)$ e $(A.2)$ é Fejér convergente a U e o Lema $A.1.3$ garante que todos os seus pontos de acumulação pertencem a U. Portanto, pela Proposição 3.1, a sequência $\{x^k\}$ converge a $x^* \in U$. □

A.2 O Algoritmo do Ponto Proximal com Distâncias de Bregman(APPDB)

Nesta seção apresentaremos o Algoritmo do Ponto Proximal com Distâncias de Bregman que é uma generalização do algoritmo apresentado na seção anterior, onde é substituida a norma Euclideana pela distância de Bregman.
O problema de interesse é

$$\min f(x) \text{ sujeito a } x \in \overline{S}, \tag{A.4}$$

onde $S \subset \mathbb{R}^n$ é aberto e convexo, com $\overline{S}$ o fecho de S, e f contínua em $\overline{S}$. O Algoritmo do Ponto Proximal com Distância de Bregman (APPD) é definido como $x_0 \in S$,

$$x^{k+1} := \arg\min_{x \in \overline{S}}\{f(x) + \lambda_k D_h(x, x^k)\} \tag{A.5}$$

onde h é uma função de Bregman com zona S e $\{\lambda_k\}$ uma sequência de números positivos satisfazendo

$$0 < \lambda_k \leqslant \bar{\lambda}, \text{ para algum } \bar{\lambda} > 0. \tag{A.6}$$

Prosseguimos com a análise de convergência. Assumimos a partir de agora que (A.4) tenha soluções de modo que f seja limitada inferiormente em $\overline{S}$.

Lema A.2.1. *A sequência $\{x^k\}$ gerada pelo (APPD) está bem definida e contida em S.*

Demonstração. Seja β um limite inferior para f em $\overline{S}$ e $f_k(x) = f(x) + \lambda_k D_h(x, x^k)$. Então, $f_k(x) \geqslant \beta + \lambda_k D_h(x, x^k)$ e decorre de (B_3) que os conjuntos de nível de f_k são limitados de modo que a minimização em (A.5) se reduz a um compacto e o mínimo é atingido. Note que f_k é estritamente convexa pela convexidade de h e
pela Proposição 1.6 (iii), de modo que o mínimo é único e assim x^{k+1} é unicamente determinado. Em seguida provaremos que $x^{k+1} \in S$.
É fácil verificar a partir de (A.5) que x^{k+1} é único $x \in \overline{S}$ tal que

$$\lambda_k \nabla h(x^k) \in \partial(f + \lambda_k h)(x). \tag{A.7}$$

Mostraremos que de acordo com (B_6), $\partial(f + \lambda_k h)(x) = \emptyset$ para todo $x \in \partial S$, assim, tendo em conta (A.7) temos $x^{k+1} \in S$.
De fato, tome $x \in \partial S$ e suponha que exista $\xi \in \partial(f + \lambda_k h)(x)$. Tome $z \in S$ e defina

$$y^t = (1 - \varepsilon_t)x + \varepsilon_t z \tag{A.8}$$

com $\lim_{t\to\infty} \varepsilon_t = 0$. Então $y^t \in S$ pela convexidade de S e $\lim_{t\to\infty} y^t = x$.
Assim, de (A.8), da definição de $\partial(f + \lambda_k h)$, B_2 e da convexidade de h e f, tem-se

$$\begin{aligned}
\varepsilon_t\langle \xi, z - x\rangle &= \langle \xi, y^t - x\rangle \\
&\leqslant (f + \lambda_k h)(y^t) - (f + \lambda_k h)(x) \\
&= f(y^t) - f(x) + \lambda_k(h(y^t) - h(x)) \\
&\leqslant f(y^t) - f(x) + \lambda_k\langle \nabla h(y^t), y^t - x\rangle.
\end{aligned}$$

Daí, usando (A.8) e da convexidade de f, obtemos

$$\varepsilon_t\langle \xi, z-x\rangle \leqslant \varepsilon_t(f(z)-f(x)) + \lambda_k \frac{\varepsilon_t}{(1-\varepsilon_t)}\langle \nabla h(y^t), z-y^t\rangle. \tag{A.9}$$

Então, de (A.9)

$$\frac{1-\varepsilon_t}{\lambda_k}[f(x)-f(z)+\langle \xi, z-x\rangle] \leqslant \langle \nabla h(y^t), z-y^t\rangle. \tag{A.10}$$

Como $\lim_{t\to\infty} y^t = x \in \partial S$, temos por (B_6) que o lado direito de (A.10) tende a $-\infty$ quando t tende a ∞, enquanto o lado esquerdo tem limite finito. Que é uma contradição, implicando que $\partial(f+\lambda_k h) = \emptyset$ para todo $x \in \partial S$ e que $x^{k+1} \in S$. □

Lema A.2.2. *A sequência $\{x^k\}$ gerada pelo (APPD)satisfaz a seguinte*

$$D_h(\bar{x}, x^{k+1}) \leqslant D_h(\bar{x}, x^k) - D_h(x^{k+1}, x^k)$$

para todo k e cada solução $\bar{x}$ de (A.4).

Demonstração. Usando a Proposição 1.6 (i) com $x=\bar{x}$, $y=x^k$, $z=x^{k+1}$, obtemos

$$D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1}) - D_h(x^{k+1}, x^k) = \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^{k+1}-\bar{x}\rangle. \tag{A.11}$$

De (A.5)

$$0 \in \partial[f+\lambda_k D_h(\cdot, x^k)](x^{k+1}). \tag{A.12}$$

De (A.12) e da Proposição 1.6 (ii), temos

$$\lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})] \in \partial f(x^{k+1}). \tag{A.13}$$

Seja $y^k = \lambda_k[\nabla h(x^k) - \nabla h(x^{k+1})]$. De (A.11) e da definição de subgradiente, temos

$$D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1}) - D_h(x^{k+1}, x^k) = \frac{1}{\lambda_k}\langle y^k, x^{k+1}-\bar{x}\rangle \geqslant \frac{1}{\lambda_k}(f(x^{k+1})-f(\bar{x})). \tag{A.14}$$

Como $\bar{x}$ minimiza f em $\overline{S}$, temos $f(\bar{x}) \leqslant f(x^{k+1})$, ou seja, $0 \leqslant f(x^{k+1}) - f(\bar{x})$. Logo,

$$D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1}) - D_h(x^{k+1}, x^k) \geqslant 0.$$

□

Lema A.2.3. *Se a sequência $\{x^k\}$ gerada pelo (APPD) é limitada e $\lim_{k\to\infty} x^{k_j} = \hat{x}$, então $\lim_{k\to\infty} x^{k_j+1} = \hat{x}$.*

Demonstração. Do Lema A.2.2 a sequência $\{D_h(\bar{x}, x^k)\}$ é decrescente e não-negativa e $D_h(x^{k+1}, x^k) \leqslant D_h(\bar{x}, x^k) - D_h(\bar{x}, x^{k+1})$ de modo que

$$\lim_{k\to\infty} D_h(x^{k+1}, x^k) = 0. \tag{A.15}$$

Sendo $\{D_h(\bar{x}, x^k)\}$ decrescente, temos $D_h(\bar{x}, x^k) \leqslant D_h(\bar{x}, x^0)$. Portanto, por (B_3) a sequência $\{x^k\}$ é limitada.

Seja $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim\limits_{k\to\infty} x^{k_j} = \hat{x}$. Então, por (A.15) temos $\lim\limits_{k\to\infty} D_h(x^{k_j+1}, x^{k_j}) = 0$. Logo, por (B_5) $\lim\limits_{k\to\infty} x^{k_j+1} = \hat{x}$. □

Lema A.2.4. *A sequência $\{x^k\}$ gerada pelo (APPD) possui pontos de acumulação, os quais são soluções de* (A.4)*.*

Demonstração. Tome uma solução $\bar{x}$ de (A.4). Pelo Lema A.2.3 existe $\hat{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim\limits_{k\to\infty} x^{k_j} = \hat{x}$. E ainda $\lim\limits_{k\to\infty} x^{k_j+1} = \hat{x}$. De (A.14) e de

$$0 \leqslant \frac{1}{\bar{\lambda}}[f(x^{k_j+1}) - f(\bar{x})] \leqslant \frac{1}{\lambda_{k_j}}[f(x^{k_j+1}) - f(\bar{x})],$$

temos

$$0 \leqslant \frac{1}{\bar{\lambda}}[f(x^{k_j+1}) - f(\bar{x})] \leqslant D_h(\bar{x}, x^{k_j}) - D_h(\bar{x}, x^{k_j+1}) - D_h(x^{k_j+1}, x^{k_j}). \tag{A.16}$$

Usando a convergência de $\{D_h(\bar{x}, x^k)\}$ e B_5, segue que

$$\lim_{k\to\infty} [D_h(\bar{x}, x^{k_j}) - D_h(\bar{x}, x^{k_j+1}) - D_h(x^{k_j+1}, x^{k_j})] = 0. \tag{A.17}$$

Daí, usando a contínuidade de f e (A.17) e fazendo k tender a ∞ em (A.16), obtemos $f(\hat{x}) = f(\bar{x})$.

Portanto, sendo que $\overline{S}$ é fechado e $\{x^k\} \subset \overline{S}$, temos que $\hat{x} \in \overline{S}$ e $\hat{x}$ é solução de (A.4). □

Teorema A.2. *Se o problema* (A.4) *tem solução e h é froneira coecirva com respeito a S, então a sequência gerada por* (A.5) *converge para a solução x^* do problema* (A.4)*.*

Demonstração. Seja $\hat{x}$ um ponto de acumulação de $\{x^k\}$ e tome a subsequência $\{x^{k_j}\}$ de $\{x^k\}$ tal que $\lim\limits_{k\to\infty} x^{k_j} = \hat{x}$.

Então por (B_4) $\lim\limits_{k\to\infty} D_h(\hat{x}, x^{k_j}) = 0$. Pelo Lema A.2.4 $\hat{x}$ é solução de (A.4) e pelo Lema A.2.2 $\{D_h(\hat{x}, x^k)\}$ é uma sequência não negativa e decrescente com subsequência convergindo a 0.

Assim a sequência $\{D_h(\hat{x}, x^k)\}$ converge a zero, logo por (B_4) temos $\lim\limits_{k\to\infty} x^k = \hat{x}$. □

Referências Bibliográficas

[1] AUSLENDER, A., TEBOULLE, M. Asymptotic Cones Functions in Optimization and Variational Inequalities. Springer Monographs in Mathemathics.

[2] BERTSEKAS, D. P., Nonlinear Programming, Athena Scientific, Belmont. Massachusetts, 1995.

[3] BRÉZIS H., Opérateurs Monotones Maximaux et Semigrups de Contraction dans les Espaces de Hilbert, Mathematics Studies 5, North-Holland, New York, 1973.

[4] BURACHIK, R. S., Generalized Proximal Point Methods for the Variational Inequality Problem, Tese do Grau de Doutor em Matemática, IMPA, Rio de Janeiro, 1995.

[5] BURACHIK, R. S., IUSEM, A. N., A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Space, SIAM J. Optim., 1991.

[6] DA SILVA, S. R. P., Algoritmo de Ponto Proximal para Otimização em $\mathbb{R}^n$, Disertação de Mestrado em Matemática, UFG, Goiania, 1999.

[7] FIGUEIREDO, D. G., Equações Elípticas não Lineares, Impa, Rio de Janeiro, 1977.

[8] IUSEM, A. N., Métodos de Ponto Proximal em Otimização, 20^0 Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 1995.

[9] IUSEM, A. N., On some Properties of Paramonotone Operators, Jurnal of Convex Analysis, 1998.

[10] IZMAILOV, A., SOLODOV, M., Otimização – volume 1, IMPA, Rio de Janeiro, 2005.

[11] LIMA, E. L., Curso de Análise - volume 2, Sexta Edição, IMPA, Rio de Janeiro, 2006.

[12] MACEDO, A. C., Método de Ponto Proximal Para Otimização, Dissetação de Mestrado em Matemática, UFG, Goiânia, 2009.

[13] Minty, G., Monotone (nonlinear) operators in Hilbert space, Duke Math. J., 29 (1976), 341–346.

[14] NETO, A. M., Método do Ponto Proximal Usando Distâncias Generalizadas Separáveis - Reescala e Seleção do Coprimente do Passo, Dissetação de Mestrado em Matemática, UFG, Goiânia, 2008.

[15] POLYAK, B. T., Introduction to Optimization. Optimization Software, New York (1987).

[16] ROCKAFELLAR, R. T., Monotone Operators and the Proximal Point Algorithm, SIAM Journal on Control and Optimization 14, pp. 877-898, 1976.

[17] ROCKAFELLAR, R. T., On the Maximality of Sums of Nonlinear monotone Operators, Transactions of the American Mathematical Society 149, p.75-88, 1970.

[18] Sissy da S. Souza., P.R. Oliveira., J.X. da Cruz Neto., A. Soubeyran., A proximal method with separable Bregman distances for quasiconvex minimization over the nonnegative orthant. European Journal of Operational Research

[19] SOLODOV. M. V., SVAITER. B. F.,AN INEXACT HYBRID GENERALIZED PROXIMAL POINT ALGORITHM AND SOME NEW RESULTS ON THE THEORY OF BREGMAN FUNCTIONS.

[20] TIEL, J. V., Convex analysis: an introductory text, Royal Netherlands Meteorological Institute, 1984.