Universidade Federal do Piauí
Centro de Ciências da Natureza
Pós-Graduação em Matemática
Mestrado em Matemática

Algoritmo do Ponto Proximal para Encontrar Zeros de Operadores Quase-Monótonos

Joel Conceição Rabelo

Teresina - 2014

Joel Conceição Rabelo

Dissertação de Mestrado:

Algoritmo do Ponto Proximal para Encontrar Zeros de Operadores Quase-Monótonos

Dissertação submetida à Coordenação do Programa 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 - 2014

Rabelo, Joel Conceição.

Algoritmos do Ponto Proximal para encontrar
zeros de operadores quase-monotónos

Joel Conceição Rabelo – Teresina: 2014.

Orientador: Prof. Dr. Jurandir de Oliveira Lopes

1. Área de Concentração: Otmização

CDD 516.36

- Aos meus pais Raimundo Nonato e Maria de Jesus.

-Aos meus irmãos. Gerson Rabelo (in memoriam).

-Aos meus amigos.

Agradecimentos

• A Deus pela vida, por ter me presenteado com pessoas maravilhosas que me ajudaram a chegar até aqui e pela permissão de realizar este trabalho.

• A minha família. Ao meu pai Raimundo Nonato pelos sermões e consequentemente pelos ensinamentos. A minha mãe Maria de Jesus pelo carinho e compreenção, o que contribuição a pessoa que sou hoje. Aos meus irmãos Iolanda, José Orlando, Manoel, Maria Aguida, Jesmo, todos fazem parte desta vitória.

• Obrigada a todos os famíliares pelo carinho, pela torcida e por todas as ajudas dadas em tantos momentos.

• A minha noiva e amiga, Renata , pelo respeito, amor, e pelos bons momentos que passo ao seu lado. Você tornou minha vida muito mais feliz. E também pela leitura e sugestões a este trabalho.

• Aos professores e amigos Ricardo Araújo, Eliziran e Amsterdã Lopes os quais me incentivaram e deram a maior força.

• A todos os meus colegas de infância e aos amigos da Universidade pela amizade e pelo ombro amigo. Em especial Alexandre lima (xandre), Carlos Adriano (chorrão), Cleidinaldo (Big Cleidi), Rui Marques (chefe), Sandoel de Brito (san), Weslay Vieira, Gilson Nascimento (irmão), Emerson de Matos (baiano), Edvalter, Vitaliano Amaral e também a Atécio Alves (técio), Andressa Gomes (dressa) e Lucas Vidal (Banach space) (representando a turma 2013.1 do mestrado) obrigado a todos pela amizade, convivência e pelas conversas, mais especificamente sobre matemática as quais contribuiram para meu aprendizado. Muito obrigado.

• A todos os professores que estiveram comigo durante a graduação e mestrado pela amizade, dicas e aprendizado. Em especial Jurandir de Oliveira (orientador e amigo), Mário Gomes, Vicente Lima, Gilvan Lima, João Xavier, João Benício, Jucelino Silva, Paulo Sérgio, Sissy Souza, Roger Peres, Liane Mendes, Paulo Alexandre, Newton Luis, Barnabé Pessoa e Jefferson Leite.

• Ao professor Jurandir de Oliveira Lopes pela atenção e paciência pelo aprendizado, dicas e conversas sobre matemática e também sobre este trabalho

• Aos professores João Xavier da Cruz Neto e Pedro Antônio Soraes Júnior por aceitarem participar da banca examinadora e também pela leitura e sugestões deste trabalho.

• A CAPES pelo apoio financeiro.

"Adoramos a perfeição, porque não a podemos ter, repugna-la-íamos se a tivéssemos."

Fernando Pessoa.

Resumo

Neste trabalho será apresentado um algoritmo que tem como objetivo resolver o problema da desigualdade variacional associadado a um operador quase-monótono e a um conjunto não vazio, convexo e fechado de $\mathbb{R}^n$. Mostraremos a boa definição do algoritmo, no sentido de que o método gera um sequência $\{x^k\}$ onde cada x^k resolve o subproblema proposto. Faremos a análise de convergência do algoritmo com cada uma das regularizações que utilizaremos, tais como as funções tipo-distâncias φ-divergente, Bregman e Log-Quadrática provando que a sequência $\{x^k\}$ gerada pelo algoritmo converge a um ponto solução do problema.

Palavras chave: Problema da desigualdade variacional, Operadores quase-monótonos, Método do Ponto Proximal

Abstract

In this paper an algorithm that aims to solve the problem of variational inequality will be presented associate a quasi-monotone operator and a non-empty, convex and closed set $\mathbb{R}^n$. We will show good definition of the algorithm, in the sense that the method generates a sequence $\{x^k\}$ where each x^k solves the subproblem proposed. We do the analysis of convergence of the algorithm with each of the regularization that we use, such as type-distances functions φ-divergent, Bregman and log-quadratic, proving that the sequence $\{x^k\}$ generated by the algorithm converges to a point solution of the problem.
Keywords: Problem of variational inequality, quasi-monotone operators, the Proximal Point Method.

Sumário

Introdução

Sejam T um operador e C um conjunto convexo, fechado e não vazio de $\mathbb{R}^n$, associamos a estes, o Problema da Desigualdade Variacional associado a T e C, PDV(T;C), definido por:

$$\mathsf{PDV}(T;C) = \begin{cases} \text{obter } x^* \in C \text{ tal que exista } u^* \in T(x^*), \text{com} \\ \langle u^*, x - x^* \rangle \geqslant 0, \forall x \in C. \end{cases} \tag{1}$$

Quando T é o subdiferencial de uma função $f : U \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convexa, própria e semicontínua inferiormente, o PDV(T;C) reduz-se ao problema de minimização não suave com restrições:

$$(P_0) \quad min_{x \in C} f(x).$$

O PDV(T;C) é equivalente a:

$$\text{Encontrar } x^* \in C \text{ tal que } 0 \in (T + N_C)(x^*),$$

onde N_C é o operador normalizador de C, o qual será definido adiante, e portanto P_0 pode ser reapresentado por:

$$\text{Encontrar } x^* \in C \text{ tal que } 0 \in (\partial f + N_C)(x^*).$$

A abordagem clássica para determinar zeros de operadores, quando T é monótono maximal, pode ser vista em Rockafellar, [17], de forma que dado x^{k-1}, o método define x^k tal que

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

onde λ_k é uma sequência de números reais tal que $\lambda_k > \lambda > 0$ e $\widehat{T} = T + N_C$. Este problema está retrito ao conjunto C, para eleminar isso, varios estudos consideram generalizações do tipo encontrada em [3, 4, 6, 10], em que a regularização usual é substituida por outras, tais como: φ-divergente, Bregman e Log-quadrática, as quais incorporam a restrição de C, de forma que cada solução de cada subproblema permaneça no interior do conjunto C.

Abdellah em [1], usando a função φ-divergente e inspirado no método Log-quadrático propôs método para resolver problema da desigualdade variacional em que o conjunto viável é o $\mathbb{R}^n_{++}$ e T um operador pseudo-monótono. Em [16], Langenberg propôs um algoritmo para resolver $\mathsf{PDV}(T; K)$ usando a regularização de Bregman, onde T é um operador Pseudo-monotóno e C um conjunto convexo não vazio.

Este trabalho é inspirado em [9] Brito, Lopes, Cruz Neto et al, no qual apresenta um algoritmo proximal interior usando além da regularização log-quadrática as funções tipo distancia como, Bregman e φ-divergente, que visam solucionar (1), com T quase-monótono.

Este trabalho está divido em 4 capítulos. No capítulo 1, apresentaremos alguns resultados clássicos e conceitos sobre as funções regularizadoras que sustentarão a parte teórica básica para os capítulos seguintes. No capítulo 2, é feita uma reunião de definições e resultados sobre operadores, por exemplo, operadores monótonos, pseudo-monótonos e quase-monótonos. No capítulo 3, propomos um algoritmo com o objetivo de resolver o problema da desigualdade variacional com restrições lineares para um operador quase-monótono, provaremos a boa definição do nosso algoritmo, faremos a análise de convergência para cada uma das funções tipo-distâncias e, sob certas condições, mostraremos que a sequência gerada por esse algoritmo converge para solução do problema proposto. No capítulo 4, apresentamos as considerações finais.

Capítulo 1

Noções Preliminares

O objetivo deste capítulo é fornecer as principais definições e resultados já conhecidos na literatura sobre certos tipos de funções regularizadoras que permitirão que a sequência gerada pelo algoritmo permaneça sempre no conjunto viável.

1.1 A Distância de Bregman

A definição da distância de Bregman foi introduzida por Lev M. Bregman em 1967. Destacamos aqui os estudos feitos por Teboulle em [20] e por Attouch e Teboulle em [2].

Sejam S um subconjunto aberto e convexo de $\mathbb{R}^n$ e $\overline{S}$ seu fecho. Consideremos uma função real $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ estritamente convexa, a distância de Bregman é, para cada $y \in \mathbb{R}^n$ fixo, a aplicação $D_h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, definida da seguinte forma:

$$D_h(x, y) = \begin{cases} h(x) - h(y) - \langle \nabla h(y), x - y \rangle, & \text{se } x \in \overline{S}, \\ +\infty & , \text{ caso contrário.} \end{cases} \tag{1.1}$$

$D_h(x, y)$ pode ser vista como a diferença entre $h(x)$ e a aproximação linear de h numa vizinhaça de y.

Definição 1. *A função* h *é chamada uma função de Bregman com zona* S *se:*

(B1) h *é estritamente convexa e contínua em* $\overline{S}$*;*

(B2) h *é continuamente diferenciável em* S*;*

(B3) *Dado qualquer* $x \in \overline{S}$ *e* $\delta \in \mathbb{R}$*, o conjunto de nível parcial à direita* $L_{D_h}(x, \delta) = \{y \in S; D_h(x, y) \leqslant \delta\}$ *é limitado;*

(B4) *Se* $\{y^k\} \subset S$ *converge para* y *então* $D_h(y, y^k)$ *converge para* 0.

A definição clássica de função de Bregman requer duas condições adicionais:

(B5) O conjunto de nível parcial à esquerda $L_{D_h}(\delta, y) = \{x \in \overline{S}; D_h(x,y) \leqslant \delta\}$ *é limitado para todo* $y \in S$.

(B6) Se $\{z^k\} \subset \overline{S}$ *é limitada,* $\{y^k\} \subset S$ *converge para* y *e* $\lim_{k\to+\infty} D_h(z^k, y^k) = 0$*, então* $\lim_{k\to+\infty} z^k = y$.

Sob certas condições adicionais é possível assegurar (B5) e (B6)(Ver [18]).

Apresentaremos a seguir uma propriedade que algumas funções de Bregman h satisfazem.

(B7) Para todo $y \in \mathbb{R}^n$, existe $x \in S$ tal que $\nabla h(x) = y$.

(B8) Se $\{x^k\}, \{y^k\} \in \text{int dom}(h)$ são sequências limitadas tal que $\lim_{k\to\infty} \|x^k - y^k\| = 0$, então $\lim_{k\to\infty} \big(\nabla h(x^k) - \nabla h(y^k)\big) = 0$.

Definição 2. *Uma função de Bregman satisfazendo* (B7) *é chamada coerciva, com zona* S.

Quando a função de Bregman h tem a forma

$$h(x) = \sum_{i=1}^{n} h_i(x_i),$$

dizemos que h e sua associada D_h são separáveis.

É interessante notar que, quando $h_i(t) = t^2, \ \forall\, i = 1, ...n$, $h(x) = \sum_{i=1}^{n} h_i(x_i)$ é uma função de Bregman separável, a distância de Bregman associada a ela é o quadrado da distância euclidiana usual, como verificaremos a seguir:

$$\begin{aligned}
D_h(x,y) &= \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} {y_i}^2 - \langle (2y_1, ..., 2y_n), (x_1 - y_1, ..., x_n - y_n) \rangle \\
&= \|x\|^2 - \|y\|^2 - \sum_{i=1}^{n} 2x_i y_i + \sum_{i=1}^{n} 2y_i^2 \\
&= \|x\|^2 - \|y\|^2 - 2\sum_{i=1}^{n} x_i y_i + 2\|y\|^2 \\
&= \|x\|^2 - 2\langle x, y \rangle + \|y\|^2 \\
&= \|x - y\|^2.
\end{aligned}$$

Proposição 1. *Seja h 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) $\nabla_x D_h(x,y) = \nabla h(x) - \nabla h(y), \forall x, y \in S$;

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

Demonstração. i) Da definição de D_h

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

ii) É imediato da definição de D_h.

iii) De fato,

$$\begin{aligned} D_h(tx_1 + (1-t)x_2, y) &= h(tx_1 + (1-t)x_2) - h(y) - \langle \nabla h(y), tx_1 + (1-t)x_2 - y\rangle \\ &< th(x_1) + (1-t)h(x_2) - h(y) - th(y) + th(y) - \\ &\quad \langle \nabla h(y), tx_1 + (1-t)x_2 - y - ty + ty\rangle \\ &= t(h(x_1) - h(y)) + (1-t)(h(x_2) - h(y)) - \\ &\quad \langle \nabla h(y), t(x_1 - y)\rangle - \langle \nabla h(y), (1-t)(x_2 - y)\rangle \\ &= t(h(x_1) - h(y) - \langle \nabla h(y), x_1 - y\rangle) + \\ &\quad (1-t)(h(x_2) - h(y) - \langle \nabla h(y), x_2 - y\rangle) \\ &= tD_h(x_1, y) + (1-t)D_h(x_2, y). \end{aligned}$$

□

O item *i*) deixa claro que a distância de Bregman não satisfaz a desigualdade triangular.

1.2 A Distância φ-divergente

A utilização dessa função tipo-distância pode ser encontrada em Teboulle [19] e Auslender e Teboulle [5].

Nesta seção apresentaremos uma classe de funções tipo distância, adequada para o ortante não negativo, os elementos desta classe serão denotados por $d_\varphi(\cdot,\cdot)$.
Seja $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ uma função convexa própria que satisfaz as seguintes propriedades:

i) φ é duas vezes continuamente diferenciável em int (dom φ)$= (0,+\infty)$;

ii) φ é estritamente convexa no seu domínio;

iii) $\varphi(1) = \varphi'(1) = 0$ e $\varphi''(1) > 0$;

iv) $\lim_{t\to 0^+} \varphi'(t) = -\infty$.

Vamos denotar por Φ a classe de funções satisfazendo i) - iv).

Definição 3. *Se $\varphi \in \Phi$, então a função $d_\varphi : \mathbb{R}^n_+ \times \mathbb{R}^n_{++} \to \mathbb{R} \cup \{+\infty\}$ definida por*

$$d_\varphi(x,y) = \begin{cases} \sum_{i=1}^{n} y_i \varphi\left(\frac{x_i}{y_i}\right), & \text{se } x, y \in \mathbb{R}^n_{++}, \\ +\infty, & \text{caso contrário.} \end{cases}$$

é chamada φ-divergente.

Exemplo 1. *Estes são alguns exemplos de funções da classe Φ:*

1. $\varphi_1(t) = t - \log t - 1$,
2. $\varphi_2(t) = t\log t - t + 1$,
3. $\varphi_3(t) = at - t^a + (1-a) \quad (0 < a < 1)$,
4. $\varphi_4(t) = at + t^{-a} - (a+1) \quad (a > 0)$.

A análise de convergência do algoritmo proximal com a φ-divergente, em geral, requer algumas condições de regularidade adicionais sobre φ. Desse modo, é usual considerar duas subclasses de Φ, a saber

$$\Phi_1 := \{\varphi \in \Phi : \varphi'(t) \leqslant \varphi''(1)\log t, \quad \forall\ t > 0\}$$
$$\Phi_2 := \left\{\varphi \in \Phi : \varphi''(1)\left(1 - \frac{1}{t}\right) \leqslant \varphi'(t) \leqslant \varphi''(1)\log t, \quad \forall\ t > 0\right\}.$$

Então $\Phi_2 \subset \Phi_1 \subset \Phi$. É possível checar que as funções φ_1, φ_2 e φ_3 estão em Φ_2, enquanto φ_4 está em $\Phi_1 - \Phi_2$.

Exemplo 2. *Seja* $\varphi\colon \mathbb{R}_{++} \to \mathbb{R}$ *definida por,* $\varphi(t) := \varphi_1(t) = t - \log t - 1$. *Então*

$$d_\varphi(x,y) := \sum_{i=1}^{n} y_i \varphi\left(\frac{x_i}{y_i}\right) = \sum_{i=1}^{n}\left(y_i \log \frac{y_i}{x_i} + x_i - y_i\right), \quad \forall\ (x,y) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}.$$

A mesma pode ser extendida para $\mathbb{R}^n_{++} \times \mathbb{R}^n_{+}$ *se aceitarmos a convenção* $0 \log 0 = 0$. *Além disso,*

$$(\nabla_1 d_\varphi(x,y))_i = \varphi'\left(\frac{x_i}{y_i}\right) = 1 - \frac{y_i}{x_i}.$$

Exemplo 3. *Seja* $\varphi\colon \mathbb{R}_{++} \to \mathbb{R}$ *definida por* $\varphi(t) = (\sqrt{t} - 1)^2$.*Um cálculo rápido nos dá* $\varphi'(t) = 1 - \frac{1}{\sqrt{t}}$, $\varphi''(t) = \frac{1}{2}\frac{1}{\sqrt{t^3}}$, $\varphi''(1) = \frac{1}{2}$.
Agora faça o seguinte artifício
$(\sqrt{t}-1)^2 = t - 2\sqrt{t} + 1 \geqslant 0 \Rightarrow 1 - 2\frac{\sqrt{t}}{t} + \frac{1}{t} \geqslant 0 \Rightarrow 2 - 2\frac{\sqrt{t}}{t} \geqslant 1 - \frac{1}{t} \Rightarrow \varphi'(t) = 1 - \frac{1}{\sqrt{t}} \geqslant \frac{1}{2}\left(1 - \frac{1}{t}\right) = \varphi''(1)\left(1 - \frac{1}{t}\right)$.
E ainda, para todo $x > 0$ *temos* $ex \leqslant e^x$, *em particular quando* $x = \frac{1}{\sqrt{t}}$, *portanto*
$\frac{e}{\sqrt{t}} \leqslant e^{\frac{1}{\sqrt{t}}} \Rightarrow e^{1-\frac{1}{\sqrt{t}}} \leqslant \sqrt{t} \Rightarrow \varphi'(t) = 1 - \frac{1}{\sqrt{t}} \leqslant \log \sqrt{t} = \frac{\log t}{2} = \varphi''(1) \log t$.
Assim temos que $\varphi \in \Phi_1$, *e mais ainda*

$$d_\varphi(x,y) = \sum_{j=1}^{n} (\sqrt{x_j} - \sqrt{y_j})^2$$

pois sendo $x = (x_1, ..., x_n)$ *e* $y = (y_1, ..., y_n)$ *com cada* $x_j, y_j > 0, j = \{1, ..., n\}$ *temos*

$$\begin{aligned}
d_\varphi(x,y) = \sum_{j=i}^{n} y_j \varphi\left(\frac{x_j}{y_j}\right) &= \sum_{j=1}^{n} y_j \left(\sqrt{\frac{x_j}{y_j}} - 1\right)^2 \\
&= y_1\left(\frac{x_1}{y_1} - 2\sqrt{\frac{x_1}{y_1}} + 1\right) + \cdots + y_n\left(\frac{x_n}{y_n} - 2\sqrt{\frac{x_n}{y_n}} + 1\right) \\
&= (x_1 - 2\sqrt{x_1}\sqrt{y_1} + y_1) + \cdots + (x_n - 2\sqrt{x_n}\sqrt{y_n} + y_n) \\
&= (\sqrt{x_1} - \sqrt{y_1})^2 + \cdots + (\sqrt{x_n} - \sqrt{y_n})^2 \\
&= \sum_{j=1}^{n} (\sqrt{x_j} - \sqrt{y_j})^2.
\end{aligned}$$

Proposição 2. *Se* $\varphi \in \Phi$, *então:*

a) $\varphi(t) \geqslant 0$ *e* $\varphi(t) = 0$ *se, e somente se,* $t = 1$*;*

b) $\varphi(t)$ *é decrescente em* $(0,1)$ *e crescente em* $(1, +\infty)$*;*

c) $\lim_{t \to +\infty} \varphi(t) = +\infty$.

Demonstração. a) Pela propriedade iii), segue que 1 é ponto crítico de φ e que $\varphi(1) = 0$, logo, pela propriedade ii), 1 é ponto de mínimo e, consequentemente, $\varphi(t) \geqslant 0$ e $\varphi(t) = 0$ se, e somente, se $t = 1$.

b) Segue da convexidade estrita de φ e do fato de 1 ser ponto de mínimo de φ.

c) Pela Propriedade ii) da função φ segue que

$$\varphi(t) \geqslant \varphi(2) + \varphi'(2)(t-2).$$

Como $\varphi'(2) > 0$ então $\lim_{t\to+\infty} \varphi(t) \geqslant \lim_{t\to+\infty} \varphi'(2)(t-2) = +\infty$.
Logo, $\lim_{t\to+\infty} \varphi(t) = +\infty$. □

Do item a) da Proposição 2, verificamos que

$$d_\varphi(x,y) \geqslant 0 \ e \ d_\varphi(x,y) = 0 \Longleftrightarrow x = y \quad \forall\ (x,y) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}.$$

Além disso, d_φ *também satisfaz as seguintes propriedades:*

1. *Os conjuntos de nível de* $d_\varphi(\cdot, y)$ *são limitados para cada* $y \in \mathbb{R}^n_{++}$*;*

2. *Se* $\{y^k\} \subset \mathbb{R}^n_{++}$ *converge para* $y \in \mathbb{R}^n_+$*, então* $\lim_{k\to+\infty} d_\varphi(y^k, y) = 0$*;*

3. *Se* $\{y^k\} \subset \mathbb{R}^n_{++}$ *satisfaz* $\lim_{k\to+\infty} y^k = y \in \mathbb{R}^n_+$*,* $\{z^k\} \subset \mathbb{R}^n_+$ *é limitada e* $\lim_{k\to+\infty} d_\varphi(y^k, z^k) = 0$*, então* $\lim_{k\to+\infty} z^k = y$*.*

Os resultados acima são obtidos diretamente das propriedades da φ*.*

Sendo o domínio da d_φ *o* $\mathbb{R}^n_+ \times \mathbb{R}^n_{++}$ *podemos fazer uma composição de* d_φ *com uma função afim do seguinte modo:*
seja C *um politopo de* $\mathbb{R}^n$ *dado por:*

$$C = \{x \in \mathbb{R}^n; Ax \leqslant b\}, \tag{1.2}$$

com $b \in \mathbb{R}^m$*,* $A_{(m\times n)}$ *tendo posto máximo e* $m \geqslant n$*. Defina*

$$f_i(x) = b_i - \langle a_i, x\rangle, \text{com} \quad i = 1, \cdots, m;$$
$$F(x) = (f_i(x), \cdots, f_m(x)) = (b - Ax)^T;$$

Considere agora

$$D_\varphi(x,y) = d_\varphi(F(x), F(y)). \tag{1.3}$$

É possível averiquar que D_φ *possui as mesmas propriedades de* d_φ*. Além disso, conclui-se a partir de* (1.3)*, que para todo* $x, y \in \text{int}C$ *vale:*

$$\nabla_x D_\varphi(x,y) = -\sum_{i=1}^{m} a_i \varphi^{'}\left(\frac{f_i(x)}{f_i(y)}\right),$$

$$\nabla_y D_\varphi(x,y) = -\sum_{i=1}^{m} a_i \left(1 - \frac{f_i(x)}{f_i(y)}\right). \tag{1.4}$$

Lema 1. *Seja* $x \in C$ *e* $x^n \in \text{int}C$*, então para todo* $n \in \mathbb{N}$*, temos:*

$$D_\varphi(x, x^n) - D_\varphi(x, x^{n+1}) \geqslant \sum_{i=1}^{m} \langle a_i, x^{n+1} - x\rangle \left(1 - \frac{f_i(x^n)}{f_i(x^{n+1})}\right).$$

Demonstração. Ver lema 2.3 de Auslender, [3]. □

1.3 A Distância log-Quadrática

Definiremos abaixo, uma família de regularizações que será de muita importância para nossos estudos. Este t ipo de regularização já é conhecida na literatura, esta foi indroduzida Auslender, Teboulle e Ben-Tiba em [6]. As funções nas quais estamos interessados será construida a partir de uma família de funções dada assim,

$$\varphi(t) = \mu h(t) + \frac{\nu}{2}(t-1)^2 \tag{1.5}$$

onde $h : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *é uma função convexa, própria e semicontínua inferiormente satisfazendo as seguintes propriedades adicionais:*

1. h *é duas vezes continuamente diferenciável sobre* int(dom φ)= $(0, +\infty)$*;*

2. h *é estritamente convexa sobre seu domínio;*

3. $\lim_{t \to 0^+} h^{'}(t) = -\infty$*;*

4. $h(1) = h^{'}(1) = 0$ *e* $h^{''}(1) > 0$*;*

5. $h \in \Phi_1$*, e os* μ *e* ν *são tais que* $\nu > \mu h^{''}(1) > 0$*.*

Agora, definiremos a distância log-quadrática $d_\varphi(x, y)$*, como sendo*

$$d_\varphi(x,y) = \begin{cases} \sum_{i=1}^n y_i^2 \varphi\left(\frac{x_i}{y_i}\right) & , \textit{se } x, y \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}, \\ +\infty & , \textit{caso contrário.} \end{cases} \tag{1.6}$$

De forma análoga ao que fizemos em (1.3), *podemos para cada* $x, z \in \text{int}C$ *definir:* $D_\varphi(x,z) = d_\varphi(F(x), F(z))$.

Observação 1.

$$\nabla_1 D_\varphi(x,z) = -\sum_{i=1}^n a_i y_i(z) \varphi'\left(\frac{y_i(x)}{y_i(z)}\right) = -A^T \nabla_1 d_\varphi(y(x), y(z)).$$

Dada a matriz A *definida em* (1.2) *a função* $\langle u, v\rangle_A : (u,v) \mapsto \langle A^T A u, v\rangle$ *define um produto interno em* $\mathbb{R}^n \times \mathbb{R}^n$ *com norma* $\|u\|_A := \|Au\| = \sqrt{\langle Au, Au\rangle}$.

Observação 2.

$$\begin{aligned} D_\varphi(x,z) &= d_\varphi(y(x), y(z)) \\ &= \sum_{i=1}^m y_i^2(z) \cdot \varphi\left(\frac{y_i(x)}{y_i(z)}\right) \\ &= \sum_{i=1}^m y_i^2(z) \left[\mu \cdot h\left(\frac{y_i(x)}{y_i(z)}\right) + \frac{\nu}{2}\left(\frac{y_i(x)}{y_i(z)} - 1\right)^2\right] \\ &= \sum_{i=1}^m y_i^2(z) \cdot \mu \cdot h\left(\frac{y_i(x)}{y_i(z)}\right) + \sum_{i=1}^m y_i^2(z) \cdot \frac{\nu}{2}\left(\frac{y_i^2(x)}{y_i^2(z)} - 2 \cdot \frac{y_i(x)}{y_i(z)} + 1\right) \\ &= \mu \cdot \sum_{i=1}^m y_i^2(z) \cdot h\left(\frac{y_i(x)}{y_i(z)}\right) + \frac{\nu}{2} \cdot \sum_{i=1}^m \left(y_i^2(x) - 2y_i(x)y_i(z) + y_i^2(z)\right) \\ &= \mu \cdot D_h(x,z) + \frac{\nu}{2} \cdot \sum_{i=1}^m (y_i(z) - y_i(x))^2 \\ &= \mu \cdot D_h(x,z) + \frac{\nu}{2} \cdot \sum_{i=1}^m (b_i - \langle a_i, z\rangle - b_i + \langle a_i, x\rangle)^2 \\ &= \mu \cdot D_h(x,z) + \frac{\nu}{2} \cdot \sum_{i=1}^m \langle a_i, x - z\rangle^2 \\ &= \mu \cdot D_h(x,z) + \frac{\nu}{2} \cdot \langle A(x-z), A(x-z)\rangle \\ &= \mu \cdot D_h(x,z) + \frac{\nu}{2} \cdot \|A(x-z)\|^2. \end{aligned}$$

Portanto,

$$D_\varphi(x,z) = \mu D_h(x,z) + \frac{\nu}{2}\|x - z\|_A^2, \forall \ x, z \in \text{int}C.$$

Por praticidade, denotaremos

$$\theta := \frac{\nu + h''(1)\mu}{2}.$$

O lema a seguir é de grande importância para a análise de convergência do algoritmo.

Lema 2. *Seja* D_φ *definida como em (1.3), então para quaisquer* $x, z \in \operatorname{int} C$ *e* $w \in C$ *vale:*

i) $D_\varphi(x,z) \geqslant 0$ *e* $D_\varphi(x,z) = 0$ *se, e somente se,* $x = z$*;*

ii) $D_\varphi(\cdot, z)$ *é fortemente convexa com módulo* ν*, e portanto*

$$\langle \nabla_1 D_\varphi(x,p) - \nabla_1 D_\varphi(z,p), x - z\rangle \geqslant \nu \|x - z\|_A^2,\ \forall\ p \in \operatorname{int} C;$$

iii) $\langle \nabla_1 D_\varphi(x,z), w - x\rangle \leqslant \theta(\|w - z\|_A^2 - \|w - x\|_A^2),\ \forall\ w \in C$*;*

iv) $\|x - z\|_A^2 \geqslant \lambda_{min}(A^T A)\|x - z\|^2$*, onde* $\lambda_{min}(A^T A)$ *é o autovalor mínimo da matriz simétrica definida positiva* $A^T A$*.*

Demonstração. i) $D_\varphi(x,z) \geqslant 0$ é claro, e $D_\varphi(x,z) = 0 \Leftrightarrow d_\varphi(x,z) = 0 \Leftrightarrow \sum_{i=1}^n z_i^2 \varphi(\frac{x_i}{z_i}) = 0 \Leftrightarrow \varphi(\frac{x_i}{z_i}) = 0 \Leftrightarrow \frac{x_i}{z_i} = 1$, portanto $x = z$.

ii) Uma consequência das exigências sobre h, dada em (1.5) é que, φ é fortemente convexa, e como veremos aposteriori, $\nabla_1 d_\varphi(\cdot\,, z)$ é fortemente monótono, e é plausível que $\nabla_1 D_\varphi(\cdot\,, z) = -A^T \nabla_1 d_\varphi(\cdot\,, z)$ seja fortemente monótono, e portanto a desigualdade $\langle \nabla_1 D_\varphi(x,p) - \nabla_1 D_\varphi(z,p), x - z\rangle \geqslant \nu\|x - z\|_A^2,\ \forall\ p \in \operatorname{int} C$ é verificada.

iii) Para provar este item, usaremos as seguintes desigualdades:

$$\langle b - a, c - a\rangle = \frac{1}{2}(\|a - c\|^2 - \|b - c\|^2 + \|a - b\|^2) \quad \text{e}$$

$$\langle b - a, c - b\rangle = \frac{1}{2}(\|a - c\|^2 - \|b - c\|^2 - \|a - b\|^2).$$

Inicialmente, notemos que, como $x, z \in \operatorname{int} C$, temos $y_j(x) > 0$ e $y_j(z) > 0$. Seja então $t_j = \frac{y_j(x)}{y_j(z)}$, $j = 1, ..., m$. Pela propriedade v) da seção 1.1, para todo $t > 0$, temos $h'(t) \leqslant h''(1)(t-1)$. Logo, multiplicando por $y_j(z)y_j(w) \geqslant 0$, obtemos para cada $j = 1, ..., m$

$$\begin{aligned} y_j(z)y_j(w)h'\left(\frac{y_j(x)}{y_j(z)}\right) &\leqslant h''(1)y_j(z)y_j(w)\left(\frac{y_j(x)}{y_j(z)} - 1\right) \\ &= h''(1)y_j(w)(y_j(x) - y_j(z)). \end{aligned} \tag{1.7}$$

Também temos $-h'(t) \leqslant -h''(1)\left(1 - \dfrac{1}{t}\right)$, para todo $t > 0$. Daí, multiplicando por $y_j(z)y_j(x) \geqslant 0$ obtemos, para cada $j = 1, ..., m$

$$\begin{aligned} -y_j(z)y_j(x)h'\left(\frac{y_j(x)}{y_j(z)}\right) &\leqslant -h''(1)y_j(z)y_j(x)\left(1 - \frac{y_j(z)}{y_j(x)}\right) \\ &= -h''(1)y_j(z)(y_j(x) - y_j(z)). \end{aligned} \tag{1.8}$$

Somando (1.7) e (1.8), encontramos

$$y_j(z)h'\left(\frac{y_j(x)}{y_j(z)}\right)(y_j(w) - y_j(x)) \leqslant h''(1)(y_j(w) - y_j(z))(y_j(x) - y_j(z)).$$

Multiplicando ambos os lados por μ e somando sobre $j = 1, ..., m$ também em ambos os lados da desigualdade, obtemos

$$\begin{aligned} \sum_{j=1}^{m}\left[(y_j(w) - y_j(x))y_j(z)\mu h'\left(\frac{y_j(x)}{y_j(z)}\right)\right] &\leqslant \sum_{j=1}^{m}\left[\mu h''(1)(y_j(w) - y_j(z))(y_j(x) - y_j(z))\right] \\ &= \mu h''(1)\langle y(x) - y(z), y(w) - y(z)\rangle. \end{aligned} \tag{1.9}$$

Somando então $\sum_{j=1}^{m}(y_j(w) - y_j(x))y_j(z)\nu\left(\dfrac{y_j(x)}{y_j(z)} - 1\right)$ no primeiro membro da desigualdade acima, obtemos

$$\begin{aligned} &\sum_{j=1}^{m}(y_j(w) - y_j(x))y_j(z)\left[\mu h'\left(\frac{y_j(x)}{y_j(z)}\right) + \nu\left(\frac{y_j(x)}{y_j(z)} - 1\right)\right] \\ &= \sum_{j=1}^{m}(y_j(w) - y_j(x))y_j(z)\varphi'\left(\frac{y_j(x)}{y_j(z)}\right) \\ &= \left\langle\left(y_1(z)\varphi'\left(\frac{y_1(x)}{y_1(z)}\right), ..., y_m(z)\varphi'\left(\frac{y_m(x)}{y_m(z)}\right)\right), y(w) - y(x)\right\rangle \\ &= \langle\nabla_1 d_\varphi(y(x), y(z)), b - Aw - (b - Ax)\rangle \\ &= \langle\nabla_1 d_\varphi(y(x), y(z)), A(x - w)\rangle \\ &= \langle A^T\nabla_1 d_\varphi(y(x), y(z)), x - w\rangle \\ &= \langle -A^T\nabla_1 d_\varphi(y(x), y(z)), w - x\rangle \\ &= \langle\nabla_1 D_\varphi(x, z), w - x\rangle. \end{aligned}$$

Somando agora o mesmo termo no segundo membro da desigualdade (1.9),usando (1.3) e a definição de θ dada acima, encontramos

$$\mu h''(1)\langle y(x)-y(z),y(w)-y(z)\rangle+\nu\sum_{j=1}^{m}(y_j(w)-y_j(x))(y_j(x)-y_j(z))$$
$$=\mu h''(1)\langle y(x)-y(z),y(w)-y(z)\rangle+\nu\langle y(x)-y(z),y(w)-y(x)\rangle$$
$$=(2\theta-\nu)\langle y(x)-y(z),y(w)-y(z)\rangle+\nu\langle y(x)-y(z),y(w)-y(x)\rangle$$
$$=(2\theta-\nu)\frac{1}{2}(\|y(z)-y(w)\|^2-\|y(x)-y(w)\|^2+\|y(z)-y(x)\|^2)+\nu\frac{1}{2}(\|y(z)-y(w)\|^2-\|y(x)-y(w)\|^2-\|y(z)-y(x)\|^2)$$
$$=\theta(\|y(z)-y(w)\|^2-\|y(x)-y(w)\|^2+\|y(z)-y(x)\|^2)-\nu\|y(z)-y(x)\|^2$$
$$=\theta(\|y(z)-y(w)\|^2-\|y(x)-y(w)\|^2)+(\theta-\nu)\|y(z)-y(x)\|^2$$
$$\leqslant\theta(\|y(z)-y(w)\|^2-\|y(x)-y(w)\|^2)$$
$$=\theta\left(\|(b-Az)-(b-Aw)\|^2-\|(b-Ax)-(b-Aw)\|^2\right)$$
$$=\theta\left(\|Aw-Az\|^2-\|Aw-Ax\|^2\right)=\theta\left(\|w-z\|_A^2-\|w-x\|_A^2\right).$$

Portanto, $\langle\nabla_1 D_\varphi(x,z),w-x\rangle\leqslant\theta(\|w-z\|_A^2-\|w-x\|_A^2)$.

iv) Considere $\varphi: S^n\to\mathbb{R}$ dada por $\varphi(u)=\langle A^TAu,u\rangle$, note que φ é contínua, portanto admite um mínimo em $u_0\in S^n$, já que S^n é compacta e pelo método dos multiplicadores de Lagrange existe λ tal que vale

$$\nabla\varphi(u_0)=\lambda\nabla f(u_0) \tag{1.10}$$

onde $f:\mathbb{R}^n\to\mathbb{R}$ é o vínculo $f(x)=||x||^2$. Assim, (1.10) diz que $A^TAu_0=\lambda u_0$ portanto

$$\langle A^TAu,u\rangle\geqslant\langle A^TAu_0,u_0\rangle=\langle\lambda u_0,u_0\rangle=\lambda \tag{1.11}$$

e ainda mais, $\lambda=\lambda_{min}(A^TA)$, pois se não fosse, por (1.11), para qualquer outro autovalor λ_k de A^TA teriamos $\lambda_k\geqslant\lambda_{min}(A^TA)$. Portanto

$$||x-z||_A^2=||x-z||^2\left\langle A^TA\left(\frac{x-z}{||x-z||}\right),\frac{x-z}{||x-z||}\right\rangle\geqslant\lambda_{min}(A^TA)||x-z||^2.$$

□

Capítulo 2

Operadores

Apresentaremos neste capítulo as definições e os principais resultados sobre operadores os quais podem ser encontrados em Rockafellar, [17] e [18].

2.1 Operadores

Definição 4. *Um operador ponto-conjunto* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é um operador que associa a cada ponto* $x \in \mathbb{R}^n$ *a um conjunto (possivelmente vazio)* $\mathsf{T}(x) \subset \mathbb{R}^n$.

O domínio, a imagem e gráfico de T *é definido como:*

Definição 5. *Seja* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador, então:*

(i) *O domínio de* $\mathsf{T}, \mathsf{D}(\mathsf{T})$, *é definido por* $\mathsf{D}(\mathsf{T}) := \{x \in \mathbb{R}^n : \mathsf{T}(x) \neq \emptyset\}$;

(ii) *A imagem de* $\mathsf{T}, \mathrm{Im}(\mathsf{T})$, *é definido por*
$\mathrm{Im}(\mathsf{T}) := \{u \in \mathbb{R}^n : u \in \mathsf{T}(x)$, *para algum* $x \in \mathbb{R}^n\}$;

(iii) *O gráfico de* $\mathsf{T}, \mathsf{G}(\mathsf{T})$, *é definido por* $\mathsf{G}(\mathsf{T}) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in \mathsf{T}(x)\}$.

Definição 6. *Um operador* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é monótono se*

$$\langle u - v, x - y \rangle \geqslant 0, \text{ para todo } x, y \in \mathsf{D}(\mathsf{T}) \text{ e para todo } u \in \mathsf{T}(x),\ v \in \mathsf{T}(y).$$

Definição 7. *Um operador monótono* T *é dito maximal se para qualquer outro operador* T_1 *tal que* $\mathsf{T}_1(x) \supseteq \mathsf{T}(x), \forall x \in \mathbb{R}^n$, *tem-se que* $\mathsf{T}_1 = \mathsf{T}$, *isto é,*

$$\langle u - v, x - y \rangle \geqslant 0,\ \forall v \in \mathsf{T}(y), y \in \mathsf{D}(\mathsf{T}) \Rightarrow u \in \mathsf{T}(x).$$

Definição 8. *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*

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

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

$$\partial f(x) := \{y \in \mathbb{R}^n \mid f(z) \geqslant f(x) + \langle y, z - x\rangle\}. \tag{2.1}$$

Exemplo 4. *Se f é uma função convexa, própria e semicontínua inferiormente o* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é monótono maximal.*

De fato, sejam $x, y \in \mathbb{R}^n$, $\theta \in \partial f(x)$ *e* $\omega \in \partial f(y)$*. Por* (2.1)

$$\langle \theta, y - x\rangle \leqslant f(y) - f(x) \tag{2.2}$$

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

Somando (2.2) *e* (2.3) *obtemos*

$$\langle \theta - \omega, x - y\rangle \geqslant 0.$$

E portanto, ∂f *é monótono. A demonstração sobre maximalidade de* ∂f *pode ser encontrada em Rockafellar, [17].*

Exemplo 5. *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),

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

Observação 3. *Um operator* T *é dito ponto-ponto de para cada* $x \in \mathsf{D}(\mathsf{T})$ *o conjunto* $\mathsf{T}(x)$ *é unitário.*

Exemplo 6. *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa (estritamente convexa) e diferenciável. Então* $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ *é um operador ponto-ponto monótono (estritamente monótono).*

Como o nosso principal objetivo neste trabalho é um relaxamento das propriedades de monotocidade do operador, somos motivados a apresentar algumas definições sobre classes de monotonicidade.

Definição 9. *Dado um operador* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$*. Sejam* $(x, u), (y, v) \in \mathsf{G}(\mathsf{T})$*, então* T *é dito:*

1- Quase-monótono em $\mathsf{D}(\mathsf{T})$*, quando* $\langle v, x - y\rangle > 0$ *implica* $\langle u, x - y\rangle \geqslant 0$*;*

2- Fortemente monótono em $\mathsf{D}(\mathsf{T})$*, quando existir* $\lambda > 0$ *tal que,* $\langle u - v, x - y\rangle \geqslant \lambda\|x - y\|^2$ *dizemos também que* T *é fortemente monótono com módulo* λ*;*

3- Fracamente monótono em $\mathsf{D}(\mathsf{T})$*, quando existe* $\mathsf{L} > 0$ *tal que,* $\langle u - v, x - y\rangle \geqslant -\mathsf{L}\|x - y\|^2$*, dizemos também que* T *é fracamente monótono com módulo L;*

4- Pseudo-monótono em $\mathsf{D}(\mathsf{T})$*, quando* $\langle v, x - y\rangle \geqslant 0$ *implica* $\langle u, x - y\rangle \geqslant 0$*;*

5- Coercivo, se $\mathsf{D}(\mathsf{T})$ *é limitado ou existir* $x^\circ \in \mathsf{X}$ *tal que* $\displaystyle\lim_{||x||\to\infty} \frac{\langle u, x - x^\circ\rangle}{||x||} = \infty.$

A definição de operador pseudo-monótono citada acima é no sentido de Karamardian, [15], a qual é diferente da definição de operador pseudo-monótono introduzida por Brézis em [8].

Observação 4. $\langle u - v, x - y\rangle \geqslant 0 \Rightarrow \langle u, x - y\rangle \geqslant \langle v, x - y\rangle$*, portanto* $\langle v, x-y\rangle \geqslant 0 \Rightarrow \langle u, x-y\rangle \geqslant 0$*, isto é, monotonicidade implica em pseudo-monotonicidade, e mais ainda, sem muita dificuldade, podemos observar o seguinte:*
monotonicidade forte $\Rightarrow$ *monotonicidade* $\Rightarrow$ *pseudo-monotonicidade* $\Rightarrow$ *quase-monotonicidade.*

Entretando a recíproca da afirmação acima não é verídica:

1. *$T : [0,1] \to [0,1]$ dado por $T(x) = 1 - x$, é pseudo-monótono mas não é monótono. De fato para $x, y \in [0,1]$, $(x-y)(x-y) \geqslant 0 \Rightarrow ((1-y)-(1-x))(x-y) \geqslant 0 \Rightarrow (1-y)(x-y) \geqslant (1-x)(x-y)$, então $(1-x)(x-y) \geqslant 0$ implica $(1-y)(x-y) \geqslant 0$ e assim T é pseudo. Porém não é verdade que $\langle T(u) - T(v), x - y\rangle \geqslant 0$, já que $((1-x)-(1-y))(x-y) \leqslant 0$.*

2. *Agora $T : [-1,0] \to \mathbb{R}$ sendo $T(x) = -x$ é quase-monótono, já que $-y(x-y) > 0 \Rightarrow x - y > 0 \Rightarrow -x(x-y) \geqslant 0$, sendo que podemos ter $x = 0$. Mas T não é pseudo pois, quando $-y(x-y) \geqslant 0$ consideremos o caso em que $x \neq y$ e $y = 0$, implicando $-x^2 > 0$, T ainda é fracamente monótono com módulo -1, observe, por Cauchy-Schwarz, $\langle -x+y, x-y\rangle \geqslant -|x-y|^2$.*

Proposição 3. *Se T é um operador fracamente monótono com módulo $L > 0$, então para $\overline{L} > L$ dado, o operador $T + \overline{L}I$ é um operador fortemente monotóno com módulo $\overline{L} - L$.*

Demonstração. Ver El Farouq, [12]. □

Proposição 4. *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 $D(T_1) \cap D(T_2) \neq \emptyset$. Então $T_1 + T_2$ é monótono maximal.*

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 cada $u \in (T_1 + T_2)(x)$ existem $u_1 \in T_1(x)$ e $u_2 \in T_2(x)$ tais que $u = u_1 + u_2$ pelo mesmo argumento exitem $v_1 \in T_1(y)$ e $v_2 \in T_2(y)$, com $v = v_1 + v_2$. Como T_1 e T_2 são monótonos

$$\langle u_1 - v_1, x - y\rangle \geqslant 0$$
$$\langle u_2 - v_2, x - y\rangle \geqslant 0$$

com isso temos

$$\langle u - v, x - y\rangle = \langle u_1 + u_2 - v_1 - v_2, x - y\rangle = \langle u_1 - v_1, x - y\rangle + \langle u_2 - v_2, x - y\rangle \geqslant 0$$

e portanto $T_1 + T_2$ é monótono. A demonstração da maximalidade de $T_1 + T_2$ pode encontrada em Rockafellar, [17]. □

Proposição 5. *Sejam $T_1 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ fortemente monótono e $T_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ monótono, tais que $D(T_1) \cap D(T_2) \neq \emptyset$. Então $T_1 + T_2$ é fortemente monótono .*

Demonstração. Segue-se de imediato das definições de monótono e fortemente monótono. □

Definição 10. *Seja* C *convexo, fechado e não vazio em* $\mathbb{R}^n$. *A função* $\delta_C : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *dada por*

$$\delta_C(x) = \begin{cases} 0 \quad , & \text{se } x \in C \\ +\infty, & \text{caso contrário.} \end{cases} \tag{2.4}$$

é dita função indicadora de C, *onde* δ_C *é convexa, própria e semicontínua inferiormente.*

Definição 11. *Seja* $C \subset \mathbb{R}^n$ *convexo, fechado e não vazio. O operador* $N_C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *definido assim*

$$N_C(x) = \begin{cases} \{w \in \mathbb{R}^n; \langle w, y - x\rangle \leqslant 0 \;\; \forall y \in C\}, & \text{se } x \in C \\ \emptyset, & \text{caso contrário.} \end{cases} \tag{2.5}$$

Para $x \in C$, *seja* $u \in \partial\delta_C(x)$ *então devemos ter* $\langle u, y-x\rangle \leqslant \delta_C(y) - \delta_C(x) = 0$ *isto é,* $\partial\delta_C(x) \subset N_C(x)$. *reciprocamente se* $w \in N_C(x)$, *então* $\langle w, y - x\rangle \leqslant 0$ *para todo* $y \in C$, *assim* $N_C(x) = \partial\delta_C(x)$, *e portanto* $N_C(x)$ *é monotóno maximal, ver Rockafellar, [17]. Além disso,* $D(N_C) = C$, *e* $N_C(x) = 0$ *se, e somente se,* $x \in \operatorname{int} C$.

Observação 5. *Se* x^* *resolve o* $PDV(T, C)$ *se, e somente se* $0 \in (T + N_C)(x^*)$.

Proposição 6. *Se* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *for fortemente monotóno em* C *então* $PDV(T; C)$ *admite uma única solução.*

Demonstração. Ver Corolário 3.2 em Harker, [14]. □

Exemplo 7. *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função fortemente convexa. Então* $\partial f : \mathbb{R}^n \to \mathbb{R}^n$ *é um operador fortemente monótono.*

No que segue relebremos a definição de subdiferencial de Clarke, uma noção bastante usada na literatura, mais resultados e exemplos podem ser encontrados em [11].

Seja $f : X \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, *uma função localmente lipschitz, denotaremos por* $f^\circ(x, v)$ *a derivada direcional generalizada de f no ponto x na direção de v, a qual será determinada por:*
$f^\circ(x, v) := \limsup_{\substack{y \to x \\ t \to 0^+}} \dfrac{f(y + tv) - f(y)}{t}$. *O subdiferencial de Clarke,* $\partial^C f$, *é definido para todo* $x \in \operatorname{dom}(f)$ *como sendo:*

$$\partial^C f(x) := \{x^* \in \mathbb{R}^n; f^\circ(x, v) \geqslant \langle x^*, v\rangle, \text{para todo } v \in \mathbb{R}^n\}.$$

Quando f *for de classe* C^1, $\partial^C f = \{\nabla f\}$, *e se* f *é uma função convexa diferenciável então o subdiferencial de clarke coincide com o subdiferencial clássico em análise funcional.*

Observação 6. *Em geral, não é verdade que* $\partial^C f = -\partial^C(-f)$*, a menos que f seja localmente lipschitz. E nada impede que* $\partial^C f = \emptyset$*.*

Exemplo 8. *Para* $f : \mathbb{R}_+ \to \mathbb{R}$*, dada por* $f(x) = \sqrt{x}$*, temos* $\partial^C f(0) = \mathbb{R}$ *e* $\partial^C(-f)(0) = \emptyset$

De fato, como

$$\begin{aligned} f^\circ(0,v) &= \limsup_{\substack{y\to 0\\ t\to 0^+}} \frac{\sqrt{y+tv}-\sqrt{y}}{t} = \limsup_{\substack{y\to 0\\ t\to 0^+}} \frac{tv}{t(\sqrt{y+ty}+\sqrt{y})} \\ &= \limsup_{y\to 0} \frac{v}{\sqrt{y+ty}+\sqrt{y}} = +\infty. \end{aligned}$$

Logo, para qualquer x^* *real* $f^\circ(0,v) \geqslant x^* v$*, para todo* $v \in \mathbb{R}$*, assim* $\partial^C f(0) = \mathbb{R}$*. A mesma linha de raciocínio justifica que* $\partial^C(-f)(0) = \emptyset$*. Para melhor clareza na demonstração do teorema que está por vir, usaremos as seguintes notações:*
$[u,v] = \{x \in X / x = \lambda u - (1-\lambda)v$*, para algum* $\lambda \in [0,1]\}$.
$]u,v[= [u,v] \setminus \{u,v\}$.
$B_\lambda([u,v]) = \{x \in X / \|x - y\| < \lambda$*, para algum* $y \in [u,v]\}$.

Lema 3. *Seja* $f : X \subset \mathbb{R}^n \to \mathbb{R}$ *uma função semicontínua inferiormente. Seja* $u, v, w \in X$*, com* $v \in [u,w]$*,* $f(v) > f(u)$ *e* $\lambda > 0$*. Então existe* $\overline{x} \in X$ *e* $\overline{x}^* \in \partial^C f(\overline{x})$ *tal que*

$$\overline{x} \in B_\lambda([u,v]) \quad \text{e} \quad \langle \overline{x}^*, w - \overline{x} \rangle$$

Teorema 1. *Seja* $f : X \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função semicontínua inferiormente. Então:*

1. f *é quase-convexa se, e somente se,* $\partial^C f$ *é quase-monótono;*

2. f *é pseudo-convexa, então,* $\partial^C f$ *é pseudo-monótono.*

Demonstração. 1. Se f não fosse quase-convexa, existiria $v \in]u,w[$ tal que $f(v) >$máx$\{f(u), f(w)\}$, seja $\lambda > 0$ de modo que $f(x) >$máx$\{f(u), f(v)\}$ para todo $x \in B_\lambda(v)$. Pelo Lema 3 existe $\overline{x} \in X$ e $\overline{x}^* \in \partial^C f(\overline{x})$ tal que

$$\overline{x} \in B_\lambda([u,v]) \quad \text{e} \quad \langle \overline{x}^*, w - \overline{x} \rangle \tag{2.6}$$

Afirmamos que $]\overline{x}, w] \cap B_\lambda(v) \neq \emptyset$. Isso é verdade se $\overline{x}$ estiver próximo de v, isto é, ($\|\overline{x} - v\| < \lambda$), se não, considere $P\overline{x}$ próximo de $\overline{x} \in [u,v[$, então escrevendo

$v = tP\bar{x} + (1-t)w$, com $0 \leqslant t < 1$ o ponto $\bar{z} = t\bar{x} + (1-t)w$ pertence à $]\bar{x}, w] \cap B_\lambda(v)$. Fixado $\bar{z} \in]\bar{x}, w] \cap B_\lambda(v)$, por (2.6), para algum $y \in [\bar{z}, w]$, temos

$$\begin{aligned}
\langle \bar{x}^*, y - \bar{x} \rangle &= \langle \bar{x}^*, \mu\bar{z} + (1-\mu)w - \bar{x} \rangle \\
&= \langle \bar{x}^*, \mu(\bar{z} - w) \rangle + \langle \bar{x}^*, w - \bar{x} \rangle \\
&= \langle \bar{x}^*, t\mu(\bar{x} - w) \rangle + \langle \bar{x}^*, w - \bar{x} \rangle \\
&= (1 - t\mu)\langle \bar{x}^*, w - \bar{x} \rangle
\end{aligned}$$

assim existe $\lambda_1 > 0$ tal que

$$\langle \bar{x}^*, y - \bar{x} \rangle > 0, \qquad \forall\ y \in B_{\lambda_1}([\bar{z}, w]) \tag{2.7}$$

Desde que $f(\bar{z}) > f(w)$,aplicando o Lema 3 novamente, temos que existe $\overline{y} \in X$ e $\overline{y}* \in \partial^C f(\overline{y})$ tal que

$$\overline{y} \in B_{\lambda_1}([\bar{z}, w]) \text{ e } \langle \overline{y}^*, \bar{x} - y \rangle > 0.$$

Já que $\overline{y} \in [\bar{z}, w]$, temos $\langle \bar{x}^*, \overline{y} - \bar{x} \rangle \geqslant 0$ por (2.7), e daí $\partial^C f$ não seria quase-monotóno. Mostraremos que se f for quase-convexa então $\partial^C f$ será quase-monotóno. Seja $u^* \in \partial^C f(u)$ e $v^* \in \partial^C f(v)$ com $\langle u^*, v - u \rangle > 0$. A demonstração acaba quando vericarmos que $f^\circ(v, u - v) \leqslant 0$ pois, isso implica $\langle v^*, v - u \rangle \geqslant 0$.
Então fixe $\varepsilon > 0$ e $\gamma \in (0, \varepsilon)$ tal que

$$\langle u^*, y - u \rangle > 0 \quad \forall y \in B_\gamma(v).$$

Agora considere $y \in B_\gamma(v)$ fixo. Como $x^* \in \partial^C f(u)$, temos
$f^\circ(u, y - u) \geqslant \langle u^*, y - u \rangle > 0$ o que torna possível encontrar $\varepsilon' \in (0, \varepsilon - \gamma)$, $x \in B_{\varepsilon'}(u)$ e $\tau \in (0,1)$ tal que $f(x + \tau(y - x)) > f(x)$. Pela quase-covexidade de f, temos $f(x) < f(y)$ e portanto temos

$$f(y + t(x - y)) \leqslant f(y) \text{ sempre que } t \in (0,1)$$

dai $f^\circ(v, u - v) = \limsup\limits_{\substack{y \to v \\ t \to 0^+}} \dfrac{f(y + t(x - y)) - f(y)}{t} \leqslant 0.$

2. A demonstração pode ser encontrada em Hadijisavvas, [13].

□

Teorema 2. *Fixado* $x \in \mathrm{int}C$. *Sejam* T *um operador ponto-ponto e fracamente monótono com módulo* $L_1 > 0$, ∇D *um operador ponto-ponto fortemente monótono com módulo* $L_2 > 0$ *e* ω_k *uma sequencia de números reais positivos satisfazendo* $\omega_k \geqslant \omega > \frac{L_1}{L_2}$. *Então o operador*

$$F(y) = \begin{cases} T(y) + \omega_k \nabla_1 D(y, x), & \textit{se } x, y \in \mathrm{int}C \\ \emptyset, & \textit{caso contrário.} \end{cases}$$

é fortemente monótono em C com constante $[\omega L_2 - L_1]$.

Demonstração. Da definição de operador fracamente e fortemente monótono, segue que:

$$\begin{aligned} \langle F(y_1) - F(y_2), y_1 - y_2 \rangle &= \langle T(y_1) - T(y_2), y_1 - y_2 \rangle \\ &+ \omega_k \langle \nabla_1 D(y_1, x) - \nabla_1 D(y_2, x), y_1 - y_2 \rangle \\ &\geqslant -L_1 ||y_1 - y_2||^2 + \omega L_2 ||y_1 - y_2||^2 \\ &= [\omega L_2 - L_1] ||y_1 - y_2||^2. \end{aligned}$$

e portanto, pelo item 2 da Definição 9, F é um operador fortemente monótono com módulo $\omega L_2 - L_1$.

□

Capítulo 3

Problema da Desigualdade Variacional

Neste capítulo apresentaremos um algoritmo para resolver o problema da desigualdade variacional associado a um operador quase-monótono e a um conjunto convexo, fechado e não vazio. Mostraremos sua boa definição e que o ponto limite da sequência gerada por esse algoritmo é uma solução do $\mathsf{PDV}(\mathsf{T};\mathsf{C})$.

3.1 Algoritmo

Sejam $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador ponto-conjunto e C um conjunto não vazio, convexo e fechado de* $\mathbb{R}^n$. *Neste trabalho consideremos o Problema da Desigualdade Variacional associado a T e C, PDV(T;C), definido por:*

$$\mathsf{PDV}(\mathsf{T};\mathsf{C}) = \begin{cases} \textit{obter } x^* \in \mathsf{C} \textit{ tal que exista } u^* \in \mathsf{T}(x^*), \textit{com} \\ \langle u^*, x - x^* \rangle \geqslant 0, \forall x \in \mathsf{C} \end{cases} \tag{3.1}$$

O conjunto solução do PDV(T;C) será denotado por SOL(T;C). De agora em diante, vamos considerar as seguintes hipóteses:

(H1) $\mathsf{D}(\mathsf{T}) \cap \operatorname{int}\mathsf{C} \neq \emptyset$.

(H2) $\mathsf{SOL}(\mathsf{T};\mathsf{C}) \neq \emptyset$.

Observação 7. *O interior ao qual nos referimos acima é o interior topológico.*

Considere D *uma função tipo-distância. O algoritmo para resolver (3.1) é definido da seguinte maneira:*

ALGORITMO

Inicialização. *Escolha um* $x^0 \in \mathrm{int}C$. $k := 0$

Passo 1. *Determine a iteração* $x^{k+1} \in \mathrm{int}C$ *e* $u^{k+1} \in T(x^{k+1})$ *tal que*

$$u^{k+1} + \omega_k \nabla_1 D(x^{k+1}, x^k) = 0, \tag{3.2}$$

onde ω_k *é uma sequência de números reais positivos limitada superiormente.*

Passo 2. *Se* $x^{k+1} = x^k$, *pare. Caso contrário.*

Passo 3 *Faça* $k := k + 1$ *e retorne ao passo 1.*

Observação 8. *Se ocorrer* $x^{k+1} = x^k$, *para algum* k, *concluiremos da propriedade da* D_φ, *que* $\nabla_1 D_\varphi(x^{k+1}, x^k) = 0$. *Daí,* $0 = u^{k+1} \in T(x^{k+1})$, *consequentemente* x^{k+1} *resolve o PDV(T;C).*

3.2 Boa Definição

Mostraremos sob algumas hipóteses adicionais ao operator T, *a boa definição do nosso algoritmo. O teorema a seguir garante a boa definição da sequência* $\{x^k\}$ *gerada pelo algoritmo.*

Teorema 3. *Em cada um dos casos abaixo,* (3.2) *tem uma única solução:*

(a) T *é um operador ponto-ponto, contínuo e fracamente monótono com módulo* $L > 0$ *e* $\{\omega_k\}$ *uma sequência como no Teorema 2;*

(b) $T = \partial^C f$, *fracamente monótono com módulo* $L > 0$, f *uma função semicontínua inferiormente e* $\{\omega_k\}$ *uma sequência como no Teorema 2.*

Demonstração. (a) Pelo Teorema 2, o operador $F = T + \omega_k \nabla D$ é ponto-ponto, contínuo e fortemente monótono em todo o $\mathrm{int}C$. Portanto pela Proposição 6, a equação (3.2) tem uma única solução.

(b) Por (H1), $D(T) \cap \mathrm{int}C \neq \emptyset$ e pelo item (ii) do Lema 2, $\nabla_1 D\varphi$ é fortemente monótono.

Por outro lado, $T + \beta_k \nabla_1 D_\varphi = \partial^C(f + \beta_k D_\varphi)$ é o subdiferencial de clarke de uma função própria, semicontínua inferior e fortemente convexa. Neste caso, o subdiferencial de Clarke coincide com o subdiferencial clássico (ver [11]). Logo, $T+\beta_k \nabla_1 D$ é um operador fortemente monótono e maximal, assim existe uma única solução para (3.2). □

Em seguinda introduziremos a noção de Fejér convergência de uma sequência. Esse conceito será de grande importância para a análise de convergência do nosso algoritmo que está por vir.

Definição 12. *Uma sequência* $\{x^k\} \subset \mathbb{R}^n$ *é dita Fejér convergente para uma conjunto não vazio* $U \subset \mathbb{R}^n$ *com respeito à uma função tipo distância* $d(\cdot,\cdot)$*, se para cada* $u \in U$ *acontecer:*

$$d(u, x^{k+1}) \leqslant d(u, x^k) \quad , \ \forall \ k \in \mathbb{N} \tag{3.3}$$

Proposição 7. *Se* $\{x^k\} \subset \mathbb{R}^n$ *é Fejér convergente para uma conjunto não vazio* $U \subset \mathbb{R}^n$ *com respeito à uma função tipo distância* $d(\cdot,\cdot)$*, então:*

1. $\{x^k\}$ *é limitada;*

2. *Se um ponto de acumulação* x*, da sequência* $\{x^k\}$ *pertence a* U*, então* $\lim_{n\to\infty} x^k = x$.

Demonstração.

1. Por (3.3), $d(u, x^{k+1}) \leqslant d(u, x^k) \leqslant d(u, x^{k-1}) \leqslant \ldots \leqslant d(u, x^0)$, isto é, a sequência $\{x^k\}$ está contida na bola $B[u, d(u, x^0)]$, e portanto ela é limitada.

2. Agora considere $\{x^{k_j}\}$ uma subsequênica de $\{x^k\}$ tal que $\lim_{n\to\infty} x^{k_j} = x$. Desde que $x \in U$, novamente por (3.3) a sequência $\{d(x, x^k)\}$ é decrescente e não negativa, e no mais esta possui uma subsequênica $d(x, x^{k_j}) \to 0$, então a sequência toda converge a zero, isto é, $\lim_{n\to\infty} d(x, x^k) = 0$, implicando $\lim_{n\to\infty} x^k = x$.

□

3.3 Análise de Convergência

Provaremos para cada uma das tipos-distâncias a convergência do método. Vamos supor doravante que para todo k*, (3.2) tem solução e* $x^{k+1} \neq x^k$*. De início provaremos a* *análise de convergência para* T*, um operador pseudomonotono. Depois sob certa hipótese adcional mostraremos a análise de convergência quando* T *for quase-monótono. Vamos também supor:*

(H3) T *é um operador localmente limitado e* $\mathsf{G(T)}$ *é fechado.*

Observação 9. *Se* T *é um operador monótono maximal, então* (H3) *é válida.*

3.3.1 Análise de Convergência Via Distância de Bregman

Nosso objetivo se reduz a averiquação da convergência da sequência gerada pelo algoritmo segundo a distância de Bregman. Então do **Passo 1**, *do nosso algoritmo , temos:*

$$u^{k+1} + \omega_k \nabla_1 D_h(x^{k+1}, x^k) = 0, \tag{3.4}$$

onde D_h *representa a distância de Bregman.*

Teorema 4. *Seja* T *um operador pseudomonótono e assuma que* (3.4) *seja solúvel. Supondo* (H1) − (H3)*, então a sequência* $\{x^k\}$ *gerada pelo algoritmo converge à um ponto de* $\mathsf{SOL(T; C)}$.

Demonstração. Seja $\overline{x} \in \mathsf{SOL(T;C)}$, então $\langle \overline{u}, x - \overline{x}\rangle \geqslant 0$, para todo $x \in \mathsf{C}$ e para $\overline{u} \in \mathsf{T}(\overline{x})$, em particular $\langle \overline{u}, x^k - \overline{x}\rangle \geqslant 0$. Usando a pseudomonotonicidade de T, $\langle u^k, x^k - \overline{x}\rangle \geqslant 0$, com $u^k \in \mathsf{T}(x^k)$ assim, por (3.4) e pela Proposição 1, temos

$$\begin{aligned}
0 \leqslant \langle u^k, x^k - \overline{x}\rangle &= \omega_{k-1}\langle -\nabla_1 D_h(x^k, x^{k-1}), x^k - \overline{x}\rangle \\
&= \omega_{k-1}\langle \nabla h(x^{k-1}) - \nabla h(x^k), x^k - \overline{x}\rangle \\
&= \omega_{k-1}\left(D_h(\overline{x}, x^{k-1}) - D_h(\overline{x}, x^k) - D_h(x^k, x^{k-1})\right).
\end{aligned}$$

Portanto,

$$D_h(\overline{x}, x^k) \leqslant D_h(\overline{x}, x^{k-1}) - D_h(x^k, x^{k-1}) \leqslant D_h(\overline{x}, x^{k-1}).$$

Logo $\{x^k\}$ é Fejér convergente ao conjunto $\mathsf{SOL(T;C)}$ com respeito a distância $D_h(\overline{x}, \cdot)$, assim $\{x^k\}$ é limitada, então esta possui uma subsequênca convergente, digamos $x^{k_j} \to x^*$. Sendo T localmente limitado, $\{u^{k_j}\}$ também é limitada, portanto existe uma subsequênica $\{u^{k_{j_i}}\}$ de $\{u^{k_j}\}$ convergindo para u^*, sem perda de generalidade, vamos redefinir $\{u^{k_{j_i}}\}$ por $\{u^{k_j}\}$. Desde que $\mathsf{G(T)}$ é fechado $u^* \in \mathsf{T}(x^*)$, novamente podemos assumir que

$u^{k_j} \to u^* \in T(x^*)$. Agora veja que, para cada $j \in \mathbb{N}$

$$\begin{aligned} \langle u^{k_j}, x - x^{k_j} \rangle &= \omega_{k_{j-1}} \langle -\nabla_1 D_h(x^{k_j}, x^{k_{j-1}}), x - x^{k_j} \rangle \\ &= \omega_{k_{j-1}} \langle \nabla h(x^{k_{j-1}}) - \nabla h(x^{k_j}), x - x^{k_j} \rangle \\ &\geqslant -\omega_{k_{j-1}} \|\nabla h(x^{k_{j-1}}) - \nabla h(x^{k_j})\| \, \|x - x^{k_j}\| \end{aligned}$$

como $\lim\limits_{j\to\infty} \|x^{k_j} - x^{k_{j-1}}\| = 0$, por (B8) da definição 1, $\lim\limits_{j\to\infty} \|\nabla h(x^{k_{j-1}}) - \nabla h(x^{k_j})\| = 0$, assim

$$\lim_{j\to\infty} \langle u^{k_j}, x - x^{k_j} \rangle \geqslant 0, \quad \forall x \in C, u^{k_j} \in T(x^{k_j}).$$

Logo

$$\langle u^*, x - x^* \rangle \geqslant 0, \quad \forall\, x \in C, u^* \in T(x^*).$$

Dessa forma $x^* \in SOL(T; C)$, e pela Proposição 7, $\{x^k\}$ converge para x^*.

□

Para o caso em que T *quase-monótono, vamos considerar o seguinte subconjunto de* $SOL(T; C)$.

$$S(T; C) := \{x^* \in SOL(T; C); \exists\, u^* \neq 0, u^* \in T(x^*)\}.$$

Observação 10. *Se a solução do* $PDV(T; C)$ *for um ponto do interior de C, então o problema da desigualdade variacional se reduz a encontrar os zeros do operador T, o que é um caso irrestrito. O interessante aqui é considerar o caso em que* $S(T; C) \neq \emptyset$*, isto é,* $SOL(T; C) \cap \partial C \neq \emptyset$.

Motivados pela última observação assumiremos:

(H4) $S(T; C) \neq \emptyset$.

Agora considere o cone normal de C em x^*

$$N_C(x^*) := \{s \in \mathbb{R}^n; \langle s, x - x^* \rangle \leqslant 0, \forall x \in C\,\}$$

Lema 4. *Suponha* (H4). *Se* $x^* \in S(T; C)$ *e* $w \in \text{int}C$, *então vale que* $\langle u^*, w - x^* \rangle > 0$.

Demonstração. Seja $x^* \in S(T : C)$, então existe $u^* \neq 0$ em $T(x^*)$ tal que $\langle u^*, x - x^* \rangle \geqslant 0$ $\forall x \in C$. Portanto

1) $-u^* \in N_C(x^*)$, e em particular;

2) $\langle u^*, w - x^*\rangle \geqslant 0$, para todo $w \in \text{int}C$.

Suponha que para algum $w \in \text{int}C$, ocorra $\langle u^*, w - x^*\rangle = 0$. Desde que $w \in \text{int}C$, existe uma bola de centro w e raio $r > 0$, $B(w, r) \subset \text{int}C$. Como $u^* \neq 0$, existe $\varepsilon > 0(\varepsilon < r\|u^*\|)$, de modo que $\overline{x} = w - \frac{\varepsilon}{\|u^*\|^2}u^* \in B(w, r)$. Portanto encontramos um $\overline{x} \in C$ tal que $\langle -u^*, \overline{x} - x^*\rangle = \langle -u^*, w - \frac{\varepsilon}{\|u^*\|^2}u^* - x^*\rangle = \langle -u^*, w - x^*\rangle + \varepsilon = \varepsilon > 0$. O que é uma contradição pois por 1), $-u^* \in N_C(x^*)$. □

O teorema a seguir nos diz que a sequência $\{x^k\}$ *converge para uma solução do PDV(T;C). Denotaremos agora os pontos de acumulação da sequência* $\{x^k\}$ *por* $\text{Acc}(x^k)$.

Teorema 5. *Considere T um operador quase-monotóno tal que* (3.2) *tenha solução. Se* $(H1) - (H4)$ *forem verdadeiras, então:*

1) $\text{Acc}(x^k)$ *é não vazio e todo ponto de* $\text{Acc}(x^k)$ *é solução do PDV(T;C);*

2) Se $\text{Acc}(x^k) \cap S(T; C) \neq \emptyset$ *então* $\{x^k\}$ *converge para um ponto de SOL(T;C).*

Demonstração. (1) Dado $x \in S(T; C)$, pelo Lema 4, existe $u \in T(x)$ com $\langle u, x^k - x\rangle > 0$, para todo $k \in \mathbb{N}$, desde que T é quase monotóno temos $\langle u^k, x^k - x\rangle \geqslant 0$ para todo $k \in \mathbb{N}$ e $u^k \in T(x^k)$. Similarmente ao Teorema 4, a sequência $\{x^k\}$ é Féjer convergente ao conjunto $S(T; C)$, portanto esta é limitada e como consequência admite uma subsequênica convergente, dai $\text{Acc}(x^k) \neq \emptyset$. E também pelo Teorema 4, todo ponto de $\text{Acc}(x^k)$ é solução do PDV(T;C).
(2) Dado $x^* \in \text{Acc}(x^k) \cap S(T; C)$ pela Proposição 7, $\{x^k\}$ converge para x^*.

□

3.3.2 Análise de Convergência Via Distância φ-divergente

Nosso objetivo se reduz a averiguação da convergência da sequência gerada pelo algoritmo segundo a distância ϕ*-divergência. Então do* **Passo 1***, do nosso algoritmo , temos:*

$$u^{k+1} + \omega_k \nabla_y D_\varphi(x^k, x^{k+1}) = 0, \tag{3.5}$$

onde D_φ *é distância* φ*-divergente.*

Vamos considerar a seguinte hipótese:

$(\mathsf{H5})$ *existe* $y^* \in \mathsf{SOL}(\mathsf{T};\mathsf{C})$ *de forma que para todo* $\varepsilon > 0$ *e para todo conjunto limitado* $\mathsf{K} \subset \mathsf{SOL}(\mathsf{T};\mathsf{C})^c_\varepsilon \cap \mathrm{int}\mathsf{C}$, *tem-se*

$$\inf\{\langle \nu, y - y^*\rangle; y \in \mathsf{K}, \nu \in \mathsf{T}(y)\} > 0.$$

onde $\mathsf{SOL}(\mathsf{T};\mathsf{C})_\varepsilon = \{s \in \mathsf{C}; \exists\, y \in \mathsf{SOL}(\mathsf{T};\mathsf{C})$ *com* $\|s - y\| \leqslant \varepsilon\}$ *e* $\mathsf{SOL}(\mathsf{T};\mathsf{C})^c_\varepsilon$ *denota o seu complementar.*

Os lemas seguintes serão de grande importância para a nossa próxima análise de convergência.

Lema 5. *Seja* $\{x^k\} \in \mathrm{int}\mathsf{C}$. *Se para um ponto de acumulação x de* $\{x^k\}$ *a sequência* $\{D_\varphi(x, x^k)\}$ *converge, então a sequência* $\{x^k\}$ *converge para x, onde* D_φ *é dada em* (1.3).

Demonstração. Ver Lema 2.2 de [3]. □

O próximo lema tira proveito da pseudo-monótonicidade do operador T.

Lema 6. *Seja A uma matriz de posto máximo, suponha* $(\mathsf{H4})$ *e* T *um operador pseudo-monótono. Então*

i) A sequência $\{x^k\}$ *é limitada;*

ii) Se for adicionado $(\mathsf{H5})$ *então a sequência* $\{x^k\}$ *converge para uma solução de* $\mathsf{SOL}(\mathsf{T};\mathsf{C})$.

Demonstração. i) Seja $u^{k+1} \in \mathsf{T}(x^{k+1})$ de forma que

$$u^{k+1} + \omega_k \nabla_y D_\varphi(x^k, x^{k+1}) = 0 \tag{3.6}$$

Seja $x^* \in \mathsf{SOL}(\mathsf{T};\mathsf{C})$ com $u^* \in \mathsf{T}(x^*)$, assim, $\langle u^*, x^{k+1} - x^*\rangle \geqslant 0$, sendo T pseudo-monótono temos,

$$\langle u^{k+1}, x^{k+1} - x^*\rangle \geqslant 0,$$

então por (3.6), (1.4) e pelo Lema 2, temos

$$\begin{aligned} 0 \leqslant \langle u^{k+1}, x^{k+1} - x^* \rangle &= \langle -\omega_k \nabla_y D_\varphi(x^k, x^{k+1}), x^{k+1} - x^* \rangle \\ &= \langle \omega_k \sum_{i=1}^{m} a_i \left(1 - \frac{f_i(x^k)}{f_i(x^{k+1})}\right), x^{k+1} - x^* \rangle \\ &= \omega_k \sum_{i=1}^{m} \langle a_i, x^{k+1} - x^* \rangle \left(1 - \frac{f_i(x^k)}{f_i(x^{k+1})}\right) \\ &\leqslant \omega_k (D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1})) \end{aligned}$$

desde que cada termo ω_k é positivo temos $D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1}) \geqslant 0$, o que implica que a sequência $\{x^k\}$ é Fejér convergente com respeito a distância $D_\varphi(x^*, \cdot)$ e portanto $\{x^k\}$ é limitada.

ii) Basta mostrar que $\mathsf{Acc}(x^k) \cap \mathsf{SOL}(T; C) \neq \emptyset$, pelo Teorema 4, a sequência $\{x^k\}$ converge para um ponto de $\mathsf{Acc}(x^k) \cap \mathsf{SOL}(T; C)$. Suponha por absurdo que possa ocorrer $\mathsf{Acc}(x^k) \cap \mathsf{SOL}(T; C) = \emptyset$, então existe $\varepsilon > 0$ tal que $\mathsf{Acc}(x^k) \subset \mathsf{SOL}(T : C)^c_\varepsilon$, e pelo item acima $\forall\ x^* \in \mathsf{SOL}(T; C)$,

$$D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1}) \geqslant \frac{1}{\omega_k} \langle u^{k+1}, x^{k+1} - x^* \rangle \geqslant 0.$$

desde que $\{D_\varphi(x^*, x^k) - D_\varphi(x^*, x^{k+1})\}$ converge a zero, o que é uma contradição de (H5), assim $\mathsf{Acc}(x^k) \cap \mathsf{SOL}(T; C) \neq \emptyset$.

□

Teorema 6. *Seja* T *um operador pseudomonótono e assuma que* (3.7) *seja solúvel. Supondo* $(H1) - (H5)$, *então a sequência* $\{x^k\}$ *gerada pelo algoritmo converge a um ponto de* $\mathsf{SOL}(T; C)$.

Demonstração. Segue direto do Lema 6. □

Teorema 7. *Considere T um operador quase-monotóno tal que* (3.7) *tenha solução. Se* $(H1) - (H4)$ *for verdade então:*

1) $\mathsf{Acc}(x^k)$ *é não vazio e todo ponto de* $\mathsf{Acc}(x^k)$ *é solução do PDV(T;C);*

2) Se $\mathsf{Acc}(x^k) \cap S(T; C) \neq \emptyset$ *então* $\{x^k\}$ *converge para um ponto de SOL(T;C).*

Demonstração. Análoga ao Teorema 5. □

3.3.3 Análise de Convergência Via Distância Log-Quadrática

Nosso objetivo se reduz a averiguação da convergência da sequência gerada pelo algoritmo segundo a distância Log-Quadrática. Então do **Passo 1**, *do nosso algoritmo , temos:*

$$u^{k+1} + \omega_k \nabla_1 D_{\varphi}(x^{k+1}, x^k) = 0, \tag{3.7}$$

onde D_{φ} *é distância Log-Quadrática.*

Teorema 8. *Seja* T *um operador pseudomonótono. Supondo* $(H1) - (H3)$, *então a sequência* $\{x^k\}$ *gerada pelo algoritmo converge à um ponto de* $SOL(T; C)$.

Demonstração. Seja $\overline{x} \in SOL(T; C)$ e $\overline{u} \in T(\overline{x})$, tal que $\langle \overline{u}, x - \overline{x} \rangle \geqslant 0$, $\forall x \in C$. Em particular $\langle \overline{u}, x^k - \overline{x} \rangle \geqslant 0$ e por T ser pseudomonótono $\langle u^k, x^k - \overline{x} \rangle \geqslant 0$, desde que $u^k \in T(x^k)$. Por (3.7) e pelo item iii) do Lema 2, temos

$$0 \leqslant \langle u^k, x^k - \overline{x} \rangle = \omega_{k-1} \langle \nabla_1 D_{\varphi}(x^k, x^{k-1}), \overline{x} - x^k \rangle \leqslant \theta \omega_{k-1} \left(\|\overline{x} - x^{k-1}\|_A^2 - \|\overline{x} - x^k\|_A^2 \right),$$

onde θ é dado em (1.3) e desde que ω_k e θ são estritamente positivos, ganhamos a Fejér convergência da sequência $\{x^k\}$ ao conjunto $SOL(T; C)$, assim pela Proposição 7, $\{x^k\}$ é limitada.

Seja x^* um ponto de acumulação para $\{x^k\}$ e uma subsequênica $\{x^{k_j}\}$ de $\{x^k\}$ com $x^{k_j} \to x^*$, de modo que $u^{k_j} \in T(x^{k_j})$. Por T ser localmente limitado $\{u^{k_j}\}$ também é limitada, portanto existe uma $\{u^{k_{j_i}}\}$ de $\{u^{k_j}\}$ convergindo para u^*. Mas $u^* \in T(x^*)$ pois $G(T)$ é fechado, sem perda de generalidade podemos assumir que $u^{k_j} \to u^* \in T(x^*)$. Fazendo a adaptação para operadores da proposição 42 de [9], obtemos

$$\lim_{j \to \infty} \langle u^{k_j}, x - x^{k_j} \rangle \geqslant 0, \quad \forall x \in C, u^{k_j} \in T(x^{k_j})$$

logo

$$\langle u^*, x - x^* \rangle \geqslant 0, \quad \forall\, x \in C, u^* \in T(x^*)$$

implicando que $x^* \in SOL(T; C)$, novamente pela Proposição 7, $\{x^k\}$ converge para x^*. □

Teorema 9. *Considere T um operador quase-monotóno. Se* $(H1) - (H4)$ *for verdade então:*

1) $Acc(x^k)$ *é não vazio e todo ponto de* $Acc(x^k)$ *é solução do PDV(T;C);*

2) *Se* $Acc(x^k) \cap S(T; C) \neq \emptyset$ *então* $\{x^k\}$ *converge para um ponto de SOL(T;C).*

Demonstração. (1) Dado $x \in S(T;C)$, pelo Lema 4, existe $u \in T(x)$ com $\langle u, x^k - x\rangle > 0$, para todo $k \in \mathbb{N}$, sendo T quase-monotóno temos $\langle u^k, x^k - x\rangle \geqslant 0$ para todo $k \in \mathbb{N}$ e $u^k \in T(x^k)$. Aos passos do Teorema 4, temos que a sequência $\{x^k\}$ é Fejér convergente em $S(T;C)$ com respeito a norma $\|\cdot\|_A$, portanto esta é limitada e como consequência admite uma subsequênica convergente, dai $Acc(x^k) \neq \emptyset$. E também pelo Teorema 4 todo ponto de $Acc(x^k)$ é solução do PDV(T;C).

(2) Dado $x^* \in Acc(x^k) \cap S(T;C)$ pela proposição 7, $\{x^k\}$ converge para x^*. $\square$

Capítulo 4

Conclusão

Neste trabalho apresentamos um algoritmo, com objetivo de resolver o problema da desigualdade variational aasociado a um operador quase-monótono T *e* C *um conjunto convexo, fechado e não vazio* C*, denominado de* $\mathsf{PDV}(\mathsf{T};\mathsf{C})$*. Mostramos sua análise de convergência para cada uma das funções tipo-distância: Bregman,* φ*-divergente e log-quadrática. Sob as hipóteses de que operador* T *é ponto-ponto, contínuo e fracamente monótono ou* T *é o subdiferencial de uma função semicontínua inferiormente e fracamente monótono garantimos a boa definição do algoritmo. Supondo a boa definição do algoritmo e sob a hipótese adicinal razoável,* $(\mathsf{H}4)$ *obtevemos a convergência da sequência a um ponto solução do* $\mathsf{PDV}(\mathsf{T};\mathsf{C})$ *quando* T *é um operador quase-monótono.*

Referências Bibliográficas

[1] *Abdellah, B.: An LQP method for pseudomonotone variational inequalities. J. Glob. Optim. 36, 351-363 (2006).*

[2] *Attouch,H., Teboulle, M.: Regularized Lotka-Volterra Dynamical System as continuous Proximal-Like Method in Optimization. Journal of Optimization Theory and aplications, 121,3. 541-570 (2004).6*

[3] *Auslender, A., Haddou, M.: An interior-proximal method for convex linearly constrained problems and its extension to variational inequalities. Math. Programming 71 no.1, Ser. A, 77-100 (1995).*

[4] *Auslender, A.,Teboulle, M., Ben-Tiba, S.: Interior proximal and multiplier methods based on second order homogeneous kernels. Math. Oper.Res. 24, 645-668 (1999).*

[5] *Auslender, A., Teboulle, M.: Interior Gradient and Proximal Methods for Convex and Conic Optimization. Siam j. Optim 16, 693-725 (2006).*

[6] *Auslender, A.,Teboulle, M., Ben-Tiba, S.: A Logarithmic-quadrática proximal method for variational inequalities. Comput. Optim. Appl.12, 31-40 (1990).*

[7] *Aussel, D., Corvellec, J.N.: Subdifferential characterization of quasiconvexity and convexity: J. of convex Analysis V.1, No.2 195-201 (1994).*

[8] *Brezis, H.: Equations et inequalitions nonlinear dans les espaces vestoriels en dualité. Ann. Inst. Fourier. 18, 115-175 (1968).*

[9] *Brito, Arnaldo S. ; Lopes, J. O. ; Cruz Neto, J. X. ; Oliveira, P. Roberto.: Interior Proximal Algorithm for Quasiconvex Programming Problems and Variational Inequalities with Linear Constraints. Journal of Optimization Theory and Applications, v. 154, p. 217-234, 2012.*

[10] *Censor, Y. Zenios, S.A.: The proximal minimization algorithm with D-funtions. J.optim. Theory Appl. 73. 451-464 (1992)*

[11] *Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley Intercience, New york (1983).*

[12] *El Farouq, L.: Pseudomonotone Variational Inequalities: Convergence of Proximal Methods. J. Optim. Theory and Appl, Vol. 109(2), 311-326, 2001.*

[13] *Hadijisavvas, N.: Continuity and maximality properties of pseudomonotone operators. J. Convex Anal. 10 459-469 (2003)*

[14] *Harker, P. T., Pang, J. S.: Finite-dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms, and applications. Math. Program. 48, 161-220 (1990).*

[15] *Karamardian, S.: Complementarity problems over cones with monotone and pseudomonotone maps. J. Optim. Theory and Appl. 18, 445-455 (1976).*

[16] *Langenberg, N.: Pseudomonotone operators and the bregman proximal point algorithm. J.Glob. Optim. 47, 535-555 (2010).*

[17] *Rockafellar, R.T.: On the maximality of sums of nonlinear monotone operators, Trans. Amer. Math. Soc., 149, 75-88 (1970).*

[18] *Rockafellar. R. T.: Monotone operators and the proximal point algorithm. SAIM J. cont. optim. 14 877-898 (1976).*

[19] *Teboulle, M.: Entropic Proximal Mapping With Applications to Nonlinear Programming. Mrath. Oper. Res. 17, 670-690 (1992).*

[20] *Teboulle, M.: Convergence of Proximal-Like algorithms. Siam J. Optim. 7 1069-1083 (1997).*