COPPE UFRJ

Instituto Alberto Luiz Coimbra de
Pós-Graduação e Pesquisa de Engenharia

MÉTODO DE PONTO PROXIMAL PARA O PROBLEMA DE OTIMIZAÇÃO QUASE-CONVEXA E DESIGUALDADE VARIACIONAL COM RESTRIÇÕES LINEARES

Arnaldo Silva Brito

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientadores: Paulo Roberto Oliveira
Jurandir de Oliveira Lopes

Rio de Janeiro
Maio de 2012

MÉTODO DE PONTO PROXIMAL PARA O PROBLEMA DE OTIMIZAÇÃO QUASE-CONVEXA E DESIGUALDADE VARIACIONAL COM RESTRIÇÕES LINEARES

Arnaldo Silva Brito

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Paulo Roberto Oliveira, D.Ing.

Prof. Jurandir de Oliveira Lopes, D.Sc.

Prof. João Xavier da Cruz Neto, D.Sc.

Prof. Nelson Maculan Filho, Dr.

Prof. Orizon Pereira Ferreira, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
MAIO DE 2012

Brito, Arnaldo Silva

Método de Ponto Proximal para o Problema de Otimização Quase-Convexa e Desigualdade Variacional com Restrições Lineares/Arnaldo Silva Brito. – Rio de Janeiro: UFRJ/COPPE, 2012.

VIII, 51 p. $29,7$cm.

Orientadores: Paulo Roberto Oliveira

Jurandir de Oliveira Lopes

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2012.

Referências Bibliográficas: p. 47 – 51.

1. método de ponto proximal. 2. função quase-convexa. 3. restrições lineares. 4. método inviável. 5. operador quase-monótono. 6. operador pseudo-monótono. 7. operador monótono maximal. 8. quasi-Fejér convergência. I. Oliveira, Paulo Roberto *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Aos meus pais Joaquim e Antonia.
À minha esposa Coracide e ao meu filho Marcelo.

Agradecimentos

Primeiramente a Deus, Senhor da vida, por me permitir conquistar algo tão almejado por mim e minha família, pela força e coragem que me deu para enfrentar os obstáculos que sugiram no decorrer desta caminhada.

Ao Prof. Paulo Roberto Oliveira pela orientação, apoio, confiança e especialmente pelo empenho na realização do Dinter UFRJ/UESPI/UFPI.

Ao Prof. Jurandir de Oliveira Lopes por ter aceitado me orientar e pelo apoio, amizade e incentivo, com que pude contar em todos momentos da realização deste trabalho e de outros também.

Aos Professores João Xavier da Cruz Neto, Nelson Maculan Filho e Orizon Pereira Ferreira por terem aceito participar da banca de defesa desta tese de doutorado, pelo tempo que disponibilizaram à leitura da mesma e pelas consideráveis observações e sugestões.

À minha querida esposa Coracide e meu amado filho Marcelo, em especial, por sua compreensão, incentivo, paciência e principalmente pelo apoio incondicional durante todo o tempo.

Aos meus pais Joaquim e Antonia, irmãos Fernando e Joaquim Filho e demais familiares, em especial tia Valquiria, tio Paulinho, Julyanno, Débora, Denise e a tia Giselda que me acolheram durante toda a minha vida acadêmica.

Aos amigos de doutorado, em especial Afonso, Pedro e Jesus Ossian por todo o apoio, amizade e pela ajuda em todas as ocasiões.

Aos meus irmãos de fé Antonio Carlos, Alda, Augusto, Adaylto e Coralia pela amizade, pelo apoio direto e indireto desde a graduação até hoje.

Aos amigos da UESPI e da UFPI por toda a estrutura disponibilizada e pela amizade de longos anos.

A todos os professores e funcionários do PESC/COPPE pelos conhecimentos transmitidos e pela ajuda em todas as ocasiões.

À CAPES e UESPI, pelo apoio financeiro e institucional.

A todos que de alguma forma contribuíram para a realização deste trabalho. Muito obrigado.

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

MÉTODO DE PONTO PROXIMAL PARA O PROBLEMA DE OTIMIZAÇÃO QUASE-CONVEXA E DESIGUALDADE VARIACIONAL COM RESTRIÇÕES LINEARES

Arnaldo Silva Brito

Maio/2012

Orientadores: Paulo Roberto Oliveira
Jurandir de Oliveira Lopes

Programa: Engenharia de Sistemas e Computação

Neste trabalho, propomos dois algoritmos proximais baseados na busca global para duas importantes classes de problemas: problema de otimização quase-convexa com restrições lineares, e problema da desigualdade variacional associado a um operador quase-monótono. No primeiro método, obtemos convergência global quando a sequência dos parâmetros de regularização converge a zero. Esta condição pode ser substituida por limitação quando a função objetivo for pseudo-convexa. No segundo método, assumindo que o problema variacional admite pelo menos uma solução no interior da região viável, provamos convergência global a uma solução do problema.

Em outro enfoque, usando a distância proximal introduzida por Auslender e Teboulle em [3], apresentamos dois algoritmos inviáveis para resolver o problema da desigualdade variacional associado a um operador monótono maximal. Diferentemente do que ocorre no método proposto em [3], estes algoritmos podem ser aplicados em problemas cujo interior topológico da região viável seja vazio. Além disso, podem ser inicializados a partir de um ponto arbitrário do $\mathbb{R}^n$. Para estes métodos estabelecemos convergência global sob pressupostos razoáveis.

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

INTERIOR PROXIMAL METHOD FOR QUASICONVEX PROGRAMMING PROBLEMS AND VARIATIONAL INEQUALITY WITH LINEAR CONSTRAINTS

Arnaldo Silva Brito

May/2012

Advisors: Paulo Roberto Oliveira
Jurandir de Oliveira Lopes

Department: Systems Engineering and Computer Science

In this work, we propose two proximal algorithms based on the global search for two important classes of problems: quasiconvex minimization problem with linear constraints, and variational inequality problem associated to a quasimonotone operator. In the first method, we obtain global convergence when the sequence of regularization parameters converges to zero. This condition can be replaced by boundededness when the objective function is pseudoconvex. In the second method, assuming that the variational problem admits at least one solution inside the feasible region, we prove global convergence to a solution of the problem.

In another approach, using the proximal distance introduced by Auslender and Teboulle in [3], we present two infeasible algorithms to solve the variational inequality problem associated to a maximal monotone operator. Differently from what occurs in the method proposed in [3], these algorithms can be applied to problems whose topological interior of the feasible region is empty. Furthermore, they can be initialized from an arbitrary point in $\mathbb{R}^n$. For these methods we establish global convergence under reasonable assumptions.

Sumário

Introdução

Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função quase-convexa, própria, fechada. Sejam $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ um operador ponto-conjunto e C um subconjunto convexo, não vazio do $\mathbb{R}^n$, definido por

$$C := \{x \in \mathbb{R}^n : Ax \leq b\}, \tag{1}$$

onde A é a matriz $m \times n$ com $m \geq n$ e $b \in \mathbb{R}^m$.

Neste trabalho consideramos os seguintes problemas: primeiro, o Problema de Otimização Quase-Convexa, denotado (P), definido por

$$\text{(P)} \quad \begin{cases} \min \ f(x), \\ \text{s.a. } x \in C. \end{cases} \tag{2}$$

e segundo, o Problema da Desigualdade Variacional associado a T e C, $PDV(T;C)$, definido por

$$\text{(PDV)} \quad \begin{cases} \text{Obter } x^* \in C \text{ tal que exista } u^* \in T(x^*) \text{ com} \\ \langle u^*, x - x^* \rangle \geq 0 \text{ para todo } x \in C. \end{cases}$$

Nos últimos anos foram obtidos progressos na teoria do método de ponto proximal para o problema de otimização convexa (quando a função objetivo f dada em (2) é convexa), baseados em distâncias generalizadas, sobre os quais damos uma breve descrição a seguir.

O algoritmo de ponto proximal clássico, APP, desenvolvido para o problema de encontrar zeros de um operador T, ver [39, 49], é um método iterativo que inicia em um ponto $x^0 \in \mathbb{R}^n$ e gera, iterativamente, uma sequência de pontos $\{x^{k+1}\}$ tais que $0 \in T_k(x^{k+1})$, onde $T_k(x^{k+1}) := T(x^{k+1}) + \lambda_k(x^{k+1} - x^k)$ e $\{\lambda_k\}$ é uma sequência de números reais positivos convenientemente escolhida. Observamos que o termo $(x - x^k)$ é o gradiente de $\frac{1}{2}\|x - x^k\|^2$.

Diversos trabalhos posteriores propuseram generalizações de APP, substituindo o termo quadrático por funcionais tipo-distância, tais como: distância de Bregman, ϕ-divergência e a distância log-quadrática, veja por exemplo [4, 5, 7, 16, 27, 37, 44, 52, 53].

O problema de otimização quase-convexa possui um largo domínio de aplicações em vários campos das ciências e engenharia, tais como: economia [24], teoria da localização [32], teoria da demanda [45], teoria do controle [9] e teoria da aproximação [8]. Entretanto, ainda há poucos trabalhos em otimização quase-convexa. A seguir, destacamos os mais recentes [25, 44, 45, 51].

Cunha, Cruz Neto e Oliveira [25], considerando uma função especifica ϕ e inspirado no método proximal proposto por Teboulle [52], propuseram um método interior-proximal com ϕ-divergência para programação quase-convexa em R^n_{++}. Sob a hipótese de que os parâmetros de regularização convergem a zero, foi provado que a sequência gerada converge a uma solução do problema proposto. Papa e Oliveira [44] apresentaram um método de ponto proximal com distância de Bregman em Variedade de Hadamard. Em [45] os autores trabalharam numa extensão do método de ponto proximal para minimização quase-convexa irrestrita e restrita ao ortante não negativo. Em [51] Souza et al propuseram um algoritmo interior-proximal com distância de Bregman sob o ortante não negativo. Sob as mesmas hipóteses do trabalho [25], foram obtidos os mesmos resultados de convergência.

Motivados por esses trabalhos, pelo vasto campo de aplicações da programação quase-convexa e inspirados no método introduzido por Auslender, Teboulle e Ben-Tiba [4] apresentamos, na primeira parte desta tese, um algoritmo interior-proximal com distância log-quadrática para resolver o problema (P) com restrições lineares. Além disso, diferente do que ocorre em [25, 45, 51], generalizamos os nossos resultados, na medida em que desenvolvemos um segundo algoritmo proximal com a finalidade de resolver o problema da desigualdade variacional PDV$(T;C)$ associado a um operador quase-monótono, que em geral é menos restritivo do que a monotonicidade do operador. Além dos trabalhos já mencionados, destacamos outros que foram importantes no desenvolvimento do segundo algoritmo.

Abdellah [1], considerando uma função especifica φ e inspirado no método proximal log-quadrático, desenvolveu um método corretor-preditor para resolver o $PDV(T;\mathbb{R}^n_{++})$, onde T é um operador pseudo-monótono. Langenberg [40] apresentou um algoritmo proximal com distância de Bregman para resolver $PDV(T;K)$, onde T é um operador pseudo-monótono e K é um conjunto convexo. Vale destacar que os resultados obtidos nesta parte da tese foram publicados na revista *Journal of Optimization Theory and Applications* [14].

Outro objeto de interesse desta tese, foi desenvolver métodos proximais inviáveis para resolver o $PDV(T;C)$, onde T é um operador monótono maximal. Nesse novo enfoque não será exigido que o interior topológico do conjunto viável C seja diferente do vazio. Auslender, Teboulle e Ben-Tiba [4, 5] propuseram um método de ponto proximal com distância log-quadrática para resolver o problema de otimização convexa. A hipótese de que o conjunto tenha interior topológico não vazio é crucial para este médoto, pois do contrário ele pode não estar definido. Burachik e Svaiter [17] apresentaram um método de ponto proximal usando esta distância log-quadrática para resolver PDV($T; R^n_+$), onde T é um operador monótono maximal. Os autores provaram sob hipóteses razoáveis convergência global. Yamashita et al [55] propuseram um método usando distância log-quadrática para resolver o problema de otimização convexa com restrições lineares sem supor interior topológico não vazio. Esse método consiste em substituir o problema original por uma sequência de problemas (P^k), cujo conjunto de restrições tem sempre interior não vazio, onde em cada problema (P^k) eles aplicam uma regularização do tipo log-quadrática. Sob hipóteses de somabilidade dos parâmetros, foi provado que a sequência gerada converge a uma solução do problema original. Burachik, Lopes e Silva [15] generalizam o método proposto por Yamashita et al [55] para o problema PDV($T; C$), onde T é um operador monótono maximal. Sob hipóteses similares às utilizadas em [55] os autores obtiveram convergência global a uma solução do PDV($T; C$).

Na segunda etapa deste estudo, apresentamos dois algoritmos proximais inviáveis com distância proximal generalizada cujo objetivo é resolver o $PDV(T; C)$. A ideia básica desses métodos é substituir o $PDV(T; C)$ por uma sequência de problemas $PDV(T; C^k)$, nos quais o conjunto de restrições C^k contém o conjunto original C. No primeiro método, os resultados de existência e convergência são estabelecidos sob a hipótese de coercividade do operador. No segundo método não exigimos esta hipótese. Entretanto, definimos cada subproblema de tal maneira que os iterados sempre existam. Mais especificamente, em cada subproblema usaremos uma regularização do "tipo Tikhonov", ver [33, 54]. Todas estas hipóteses serão formalizadas no Capítulo 4.

Este trabalho está dividido em 5 capítulos. No Capítulo 1, apresentamos vários resultados clássicos e conceitos que formarão a parte teórica básica para os capítulos subsequentes. No Capítulo 2, apresentamos o Algoritmo 1, o qual é aplicado para resolver o problema de otimização quase-convexa (P). Obteremos convergência a uma solução do problema proposto. No Capítulo 3, propomos o Algoritmo 2, com a finalidade de resolver o problema da desigualdade variacional quase-monótona com restrições lineares. Sob hipóteses razoáveis, provamos a sua boa definição e obteremos convergência a uma solução do problema proposto. No Capítulo 4, definimos o algoritmos 3, o qual é aplicado para resolver o problema da desigualdade variacional com restrições lineares. Mostramos a sua boa definição e otimalidade dos pontos de acumulação sob a hipótese de coercividade do operador monótono maximal T. Usando uma regularização do tipo Tikhonov definimos também o Algoritmo 4, que é aplicado para resolver o mesmo problema abordado pelo Algoritmo 3.

Mostramos a boa definição do método e, além disso, obtemos a convergência global a uma solução do problema proposto. Finalmente, no Capítulo 5, são feitas as considerações finais e apresentadas algumas possibilidades de pesquisa futura.

Capítulo 1

Notações e Resultados Preliminares

Este capítulo tem como finalidade facilitar a leitura deste trabalho, fornecendo as definições e resultados que são essenciais para o seu desenvolvimento. Iniciamos com as funções definidas sobre $\mathbb{R}^n$ a valores estendidos:

Definição 1.1 *Seja $f : X \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função:*

(i) *[47, p. 23] O domínio efetivo de f, $\operatorname{dom} f$, é dado por*

$$\operatorname{dom} f := \{x \in X : f(x) < +\infty\}.$$

(ii) *[47, p. 23] O epígrafo de f, $\operatorname{epi} f$, é definido por*

$$\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}.$$

Uma função estendida $f : X \to \mathbb{R} \cup \{+\infty\}$ é dita própria se, e somente se, o seu domínio efetivo, $\operatorname{dom} f$, é não vazio, ou seja $f(x) < +\infty$ para algum $x \in X$. Caso contrário, diremos que f é imprópria.

Definição 1.2 *[47, p. 51] Diremos que $f : X \to \mathbb{R} \cup \{+\infty\}$ é semicontínua inferiormente no ponto $x \in X \subset \mathbb{R}^n$, quando para qualquer sequência $\{x^k\} \subset X$ tal que $x^k \to x$ $(k \to \infty)$, tem-se*

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

A função f é semicontínua inferiormente (sci) no conjunto X, quando ela é semicontínua inferiormente em todos os pontos de X. A continuidade inferior de uma função f é equivalente a dizer que o seu epígrafo é um conjunto fechado. Para mais detalhes veja [47, Teorema 7.1].

Definição 1.3 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ uma função própria, onde X é um subconjunto convexo não vazio do $\mathbb{R}^n$:*

(i) *[38, p. 55] f é denominada convexa se para todo $x, y \in \operatorname{dom} f$ e $\alpha \in (0, 1)$, vale a desigualdade*

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

(ii) *[38, p. 132] f é dita quase-convexa se para todo $x, y \in \operatorname{dom} f$ e $\alpha \in (0, 1)$, vale a desigualdade*

$$f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}.$$

Quando ocorre a desigualdade estrita nos itens acima, as funções são ditas: *estritamente* convexa e *estritamente quase-convexa*, respectivamente. A função f diz-se *fortemente (fracamente) convexa* com módulo $\gamma > 0$ ($\gamma < 0$), quando para quaisquer $x, y \in \operatorname{dom} f$ e $\alpha \in (0, 1)$, tem-se $f(\alpha x+(1-\alpha)y) \leq \alpha f(x)+(1-\alpha)f(y)-\gamma\alpha(1-\alpha)\|x-y\|^2$.

É imediato ver que convexidade implica em quase-convexidade. Entretanto, a recíproca não é verdadeira. Uma função f é dita quase-côncava se $-f$ é quase-convexa.

Exemplo 1.1 *São quase-convexas as seguintes funções:*

(i) $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^3$,

(ii) $f : \mathbb{R} \to \mathbb{R},\ f(x) = \arctan x$,

(iii) $f : \mathbb{R}_{++} \to \mathbb{R},\ f(x) = \log x$,

(iv) *[23, Seção 6, p. 16]* $f : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$

$$f(x,y) := \begin{cases} 0 & \text{se } x < 0 \text{ ou } x = 0, y \leq 0, \\ \sqrt{y} & \text{se } x = 0, y \geq 0, \\ +\infty & \text{caso contrário.} \end{cases}$$

Quase-convexidade tem uma interpretação geométrica. Enquanto uma função convexa pode ser caracterizada pela convexidade de seu epígrafo, uma função quase-convexa pode ser caracterizada pela convexidade de seu conjunto de nível.

Teorema 1.1 *[38, p. 133] Seja $f : X \to \mathbb{R} \cup \{+\infty\}$, onde X é um subconjunto convexo não vazio em $\mathbb{R}^n$. A função f é quase-convexa se, e somente se,*

$$L_f(\alpha) := \{x \in \operatorname{dom} f : f(x) \leq \alpha\},$$

é convexo para todo número real α.

Para funções quase-convexas contínuas em $\mathbb{R}$, tem-se a seguinte caracterização: Uma função contínua $f : \mathbb{R} \to \mathbb{R}$ é quase-convexa se, e somente se, pelo menos uma das seguintes condições acontece:

- f é não-decrescente.
- f é não-crescente.
- Existe um ponto $x_0 \in \operatorname{dom} f$ tal que para $x \leq x_0$ (e $x \in \operatorname{dom} f$), f é não-crescente, e para $x \geq x_0$ ($x \in \operatorname{dom} f$), f é não-decrescente.

No teorema abaixo, apresentamos uma caracterização para funções quase-convexas diferenciáveis.

Teorema 1.2 *[38, p. 134] Considere $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria e diferenciável. Então, f é quase-convexa se, e somente se, para todo $x, y \in \operatorname{dom} f$*

$$f(x) \leq f(y) \Longrightarrow \langle \nabla f(y), x - y \rangle \leq 0. \tag{1.1}$$

A condição (1.1) pode ser interpretada geometricamente. Com efeito, quando $\nabla f(y) \neq 0$, $\nabla f(y)$ define um hiperplano suporte para o conjunto de nível $\{x \in \operatorname{dom} f : f(x) \leq f(y)\}$, no ponto y.

Observação 1.1 *Se $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é uma função quase-convexa, própria e diferenciável onde $\nabla f(\overline{x}) = 0$ para algum $\overline{x}$, então $\overline{x}$ não é necessariamente um ponto de mínimo da função f, como podemos ver por exemplo, considerando $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x^3$ e o ponto $\overline{x} = 0$. Quando f é uma função convexa, a condição $\nabla f(\overline{x}) = 0$ é conhecida como uma condição necessária e suficiente para que $\overline{x}$ seja um ponto de mínimo da função f.*

No Capítulo 2 vamos estudar o problema de minimização quase-convexa, onde a função objetivo é diferenciável. Por isso, somos motivados a apresentar a seguinte definição.

Definição 1.4 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ uma função própria e diferenciável. Um ponto $\overline{x}$ é dito um ponto estacionário de um problema de otimização restrito $\min_{x \in X} f(x)$, onde X é um subconjunto convexo não vazio em $\mathbb{R}^n$, se $\langle \nabla f(\overline{x}), x - \overline{x} \rangle \geq 0$ para todo $x \in X$.*

Como as funções quase-convexas diferenciáveis não gozam da propriedade particular das funções convexas, conforme Observação 1.1, apresentamos outra classe de funções f que satisfazem esta importante propriedade, as funções pseudo-convexas.

Definição 1.5 *[38, p. 140] Uma função própria e diferenciável* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é dita pseudo-convexa, se para todo* $x, y \in \operatorname{dom} f$*, com* $f(x) < f(y)$ *implica* $\langle \nabla f(y), x - y \rangle < 0$.

Para as funções pseudo-convexas temos o seguinte resultado que é bastante conhecido na literatura.

Proposição 1.1 *[38, Teorema 3. p. 141] Se* $f : X \to \mathbb{R} \cup \{+\infty\}$ *é uma função pseudo-convexa, própria e diferenciável, então todo ponto estacionário de* f *é um ponto de mínimo global.*

Introduzimos a seguir, dois resultados técnicos que serão usados no Capítulo 4.

Lema 1.1 *[46, Lema 2.2.2] Sejam* $\{\sigma_k\}$ *e* $\{\beta_k\}$ *sequências de números reais não negativas, satisfazendo:*

(i) $\sigma_{k+1} \leq \sigma_k + \beta_k$;

(ii) $\sum_{k=1}^{+\infty} \beta_k < +\infty$.

Então a sequência $\{\sigma_k\}$ *é convergente.*

Lema 1.2 *Sejam* $C = \{x \in \mathbb{R}^n : Ax \leq b\}$ *e* $C^k = \{x \in \mathbb{R}^n : Ax \leq b + \delta^k\}$ *onde* A *é uma matriz* $m \times n$ *com* $m \geq n$ *e* $b, \delta^k \in \mathbb{R}^m$*. Dado* $x^k \in C^k$ *existe uma constante* $\alpha > 0$ *tal que*

$$dist(x^k, C) := \inf_{y \in C} \|y - x^k\| = \|p^k - x^k\| \leq \alpha \|\delta^k\|,$$

onde p^k *é a projeção de* x^k *em* C.

Prova: Ver Lema de Hoffman [31]. □

1.1 Regularização Log-quadrática

Definiremos agora uma família de regularizações que será de muita importância na primeira etapa deste trabalho. Este tipo de regularização foi introduzida por Auslender, Teboulle e Ben-Tiba em [4, 5]. Daqui em diante denotaremos por: $\operatorname{int} C$, $\operatorname{ir} C$ e $\operatorname{cl} C$ o interior topológico, interior relativo e o fecho do conjunto C, respectivamente. Assumindo que o interior do conjunto C é não vazio a função quase-distância, denotada por D_φ é construida a partir de uma classe de funções $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, definida da forma:

$$\varphi(t) := \mu h(t) + \frac{\nu}{2}(t-1)^2 \tag{1.2}$$

onde h é uma função convexa, própria, fechada satisfazendo as seguintes condições adicionais:

(i) h é duas vezes continuamente diferenciável sobre $\mathrm{int}(\mathrm{dom}\, h) = (0, +\infty)$;

(ii) h é estritamente convexa sobre seu domínio;

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

(iv) $h(1) = h'(1) = 0$ e $h''(1) > 0$;

(v) Para todo $t > 0$,

$$h''(1)\left(1 - \frac{1}{t}\right) \leq h'(t) \leq h''(1)(t-1). \tag{1.3}$$

Onde os parâmetros μ, ν são tais que

$$\nu > \mu h''(1) > 0.$$

Os itens **(i)-(iv)** e **(i)-(v)** usados acima, definem em [4] as famílias Φ e Φ_2, respectivamente.

Observação 1.2 *A função φ definida em (1.2) também satisfaz as mesmas condições* **(i)-(v)** *da função h. De fato, os itens* **(i)-(iv)** *decorrem diretamente da definição da função φ e das propriedades da função h. Por isso, mostraremos apenas o item* **(v)***, ou seja, para todo $t > 0$,*

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t-1). \tag{1.4}$$

Pela definição de φ e (1.3), obtemos

$$(\varphi''(1) - \nu)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) - \nu(t-1) \leq (\varphi''(1) - \nu)(t-1). \tag{1.5}$$

(i) *Para $0 < t < 1$, temos que $\frac{t-1}{t} < t - 1$, assim*

$$-\nu(t-1) < -\nu\left(\frac{t-1}{t}\right), \tag{1.6}$$

combinando o lado esquerdo da desigualdade (1.5) com (1.6), obtemos

$$\varphi''(1)\left(1 - \frac{1}{t}\right) - \nu\left(1 - \frac{1}{t}\right) \leq \varphi'(t) - \nu(t-1) \leq \varphi'(t) - \nu\left(1 - \frac{1}{t}\right),$$

consequentemente,

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t).$$

(ii) *Para $t > 1$, temos $\frac{t-1}{t} < t - 1$, e de forma semelhante ao que foi feito em* **(i)**, *obtemos*

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t).$$

(iii) *Para $t = 1$, obtemos trivialmente a última desigualdade. Portanto, de* **(i)-(iii)** *segue que*

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t), \quad \text{para todo} \quad t > 0. \tag{1.7}$$

Por outro lado, segue-se diretamente do lado direito da desigualdade dada em (1.5) que

$$\varphi'(t) \leq \varphi''(1)(t-1), \quad \text{para todo} \quad t > 0. \tag{1.8}$$

De (1.7) e (1.8), obtemos o resultado desejado.

Dentre todas as funções h que satisfazem as propriedades **(i)-(v)**, destacamos:

Exemplo 1.2 *As funções $h_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, $i \in \{1, 2, 3, 4\}$ definidas por:*

$$h_1(t) = \begin{cases} t - \ln t - 1, & se \;\; t > 0, \\ +\infty, & \text{caso contrário.} \end{cases}$$

$$h_2(t) = \begin{cases} t \ln t - t + 1, & se \;\; t > 0, \\ +\infty, & \text{caso contrário.} \end{cases}$$

$$h_3(t) = t^\alpha - t^\beta + t(\beta - \alpha) + \alpha - \beta, \qquad \alpha \geq 1 \;\; and \;\; 0 < \beta < 1.$$

$$h_4(t) = \alpha t - t^\alpha + (1 - \alpha), \qquad 0 < \alpha < 1,$$

satisfazem as condições **(i)-(v)**.

Considerando a função h_1 com parâmetros $\nu > \mu > 0$ fixados, a correspondente função φ é denominada função log-quadrática. Esta goza de várias propriedades importantes para o desenvolvimento de algoritmos eficientes na resolução de problemas de otimização restrita, assim como nos problemas de programação matemática, ver por exemplo [4, 5, 7]. A função h_2 também é bastante conhecida na literatura, veja [10, 20, 25].

Associando a função φ, dada em (1.2), definimos a função tipo-distância $d_\varphi(u, v)$, dada por

$$d_\varphi(u, v) := \begin{cases} \sum\limits_{i=1}^{n} v_i^2 \varphi(u_i/v_i), & u, v \in R^n_{++}, \\ +\infty, & \text{caso contrário.} \end{cases} \tag{1.9}$$

Exemplo 1.3 *Considerando $u, v \in \mathbb{R}^n_{++}$ e as funções $h_i, i \in \{1, 2, 3, 4\}$, do Exemplo 1.2, as correspondentes funções distância generalizada d_{φ_i} são dadas por:*

$$d_{\varphi_1}(u, v) := \begin{cases} \sum\limits_{i=1}^{n} \mu \left[v_i^2 \ln(\frac{v_i}{u_i}) + u_i v_i - v_i^2 \right] + \frac{\nu}{2}(u_i - v_i)^2, \\ +\infty, & \text{caso contrário.} \end{cases}$$

$$d_{\varphi_2}(u, v) := \begin{cases} \sum\limits_{i=1}^{n} \mu \left[u_i v_i \ln(\frac{u_i}{v_i}) - u_i v_i + v_i^2 \right] + \frac{\nu}{2}(u_i - v_i)^2, \\ +\infty, & \text{caso contrário.} \end{cases}$$

$$d_{\varphi_3}(u,v) := \begin{cases} \sum_{i=1}^{n} \mu \left[u_i^{\alpha} v_i^{2-\alpha} - u_i^{\beta} v_i^{2-\beta} + u_i v_i(\beta-\alpha) + v_i^2(\alpha-\beta)\right] + \frac{\nu}{2}(u_i - v_i)^2, \\ +\infty, \qquad\qquad\qquad\qquad \text{caso contrário.} \end{cases}$$

$$d_{\varphi_4}(u,v) := \begin{cases} \sum_{i=1}^{n} \mu \left[\alpha u_i v_i - u_i^{\alpha} v_i^{2-\alpha} + v_i^2(1-\alpha)\right] + \frac{\nu}{2}(u_i - v_i)^2, \\ +\infty, \qquad\qquad\qquad\qquad \text{caso contrário.} \end{cases}$$

1.2 Distância Proximal

Nesta seção, relembramos as definições de distância proximal e distância proximal induzida, apresentadas por Auslender e Teboulle em [3].

Definição 1.6 *[3, Definição 2.1] Uma função $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é dita distância proximal com respeito a um conjunto aberto, convexo e não vazio $K \subset \mathbb{R}^n$ se, para cada $y \in K$, d verifica as seguintes propriedades:*

(P1) *$d(\cdot, y)$ é própria, sci, convexa e continuamente diferenciável sobre K;*

(P2) *$\operatorname{dom} d(\cdot, y) \subset \operatorname{cl} K$ e $\operatorname{dom} \partial_1 d(\cdot, y) = K$, onde $\operatorname{dom} \partial_1 d(\cdot, y)$ denota o subdiferencial da função $d(\cdot, y)$ em relação à primeira variável;*

(P3) *$d(\cdot, y)$ é coerciva em $\mathbb{R}^n$, isto é, $\lim_{\|x\| \to +\infty} d(x, y) = +\infty$;*

(P4) *$d(y, y) = 0$.*

Denotamos por $\mathcal{D}(K)$ a família das funções d que satisfazem a Definição 1.6. A definição abaixo associa a cada distância proximal $d \in \mathcal{D}(K)$ uma outra função satisfazendo certas propriedades.

Definição 1.7 *[3, Definição 2.2] Sejam $K \subset \mathbb{R}^n$ um conjunto aberto, não vazio e $d \in \mathcal{D}(K)$. Uma função $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n_+ \cup \{+\infty\}$ é dita distância proximal induzida por d, se H é finita em $K \times K$ e para cada $u, v \in K$, temos*

$$H(u, u) = 0,$$

$$\langle w - v, \nabla_1 d(v, u)\rangle \leq H(w, u) - H(w, v) \;\; \forall\, w \in K.$$

Denotamos por $(d, H) \in \mathcal{F}(K)$ a distância proximal e a distância proximal induzida que satisfaz as condições da Definição 1.7. Escreveremos $(d, H) \in \mathcal{F}(\operatorname{cl} K)$ para a tripla $[\operatorname{cl} K, d, H]$ sempre que existir H com valores finitos em $\operatorname{cl} K \times K$ que satisfaça as seguintes condições: $H(u, u) = 0$ e $\langle w - v, \nabla_1 d(v, u)\rangle \leq H(w, u) - H(w, v) \;\; \forall\, w \in \operatorname{cl} K$, com $H(w, \cdot)$ nível limitada em K.

Para este trabalho, estamos interessados no caso $K = \mathbb{R}^n_{++}$, conforme veremos no Capítulo 4.

A seguir, relembramos alguns exemplos de distância proximal e distância proximal induzida.

Exemplo 1.4 (Distância Proximal de Bregman) *Seja S um subconjunto convexo e aberto do $\mathbb{R}^n$. Considerando uma função convexa $h : \operatorname{cl} S \to \mathbb{R}$. Seja $D_h : \operatorname{cl} S \times S \to \mathbb{R}$ tal que,*

$$D_h(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle. \tag{1.10}$$

Definição 1.8 *A função h é chamada uma função de Bregman com zona S se:*

(B1) *h é estritamente convexa e contínua sobre $\operatorname{cl} S$;*

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

(B3) *Dados quaisquer $x \in \operatorname{cl} S$ e $\delta \in \mathbb{R}$, o conjunto de nível parcial à direita $L_{D_h}(x, \delta) := \{y \in S : D_h(x, y) \leq \delta\}$ é limitado;*

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

Apresentamos a seguir uma propriedade sobre distâncias de Bregman que é bastante conhecida na literatura, esta segue diretamente de (1.10) e das condições (B1)-(B4).

Proposição 1.2 *Seja h uma função de Bregman com zona S. Então,*

(i) $D_h(x, y) - D_h(x, z) - D_h(z, y) = \langle \nabla h(y) - \nabla h(z), z - x \rangle$ *para todo* $x \in \operatorname{cl} S; y, z \in S$;

(ii) $\nabla_1 D_h(x, y) = \nabla h(x) - \nabla h(y)$ *para todo* $x, y \in S$;

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

Observação 1.3 *Conforme pode ser observado em [3], no caso da Distância de Bregman, a distância proximal induzida coincide com a função D_h, ou seja, $H \equiv D_h$.*

Para mais detalhes sobre a função de Bregman ver [11, 19, 21, 27, 36].

Exemplo 1.5 (Distância ϕ-Divergente) *Seja $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e fechada com $\operatorname{dom} \phi \subseteq \mathbb{R}_+$ e $\operatorname{dom} \partial\phi = \mathbb{R}_{++}$, que satisfaz as seguintes propriedades:*

(i) *ϕ é duas vezes continuamente diferenciável no $\operatorname{int}(\operatorname{dom} \phi) = (0, +\infty)$;*

(ii) *ϕ é estritamente convexa sobre seu domínio;*

(iii) $\lim_{t \to 0^+} \phi'(t) = -\infty$;

(iv) $\phi(1) = \phi'(1) = 0$ *e* $\phi''(1) > 0$.

Denotamos por Φ *a classe das funções satisfazendo* (i)-(iv) *e por* Φ_1 *a subclasse:*

$$\Phi_1 := \{\phi \in \Phi : \phi''(1)\left(1 - \frac{1}{t}\right) \leq \phi'(t) \leq \phi''(1)\log t, \ \forall t > 0\}. \tag{1.11}$$

Dado $\phi \in \Phi_1$*, a distância proximal* ϕ*-divergente é definida por:*

$$d_\phi(x, y) := \sum_{i=1}^{n} y_i \phi\left(\frac{x_i}{y_i}\right). \tag{1.12}$$

Observação 1.4 *As funções* h_1 *e* h_2 *do Exemplo 1.2, pertencem a subclasse* Φ_1*, ver [53].*

Observação 1.5 *Conforme pode ser observado em [3], dada* $\phi \in \Phi_1$ *a distância proximal induzida* H*, associada a distância* d_ϕ*, é dada pela função*

$$H(x, y) = d_\phi(x, y) := K(x, y) = \sum_{i=1}^{n} x_i \log\frac{x_i}{y_i} + y_i - x_i \quad \forall x \in \mathbb{R}^n_+ \ , \forall y \in \mathbb{R}^n_{++},$$

que é conhecida na literatura como a função entropia de Kullback-Leibler [52].

Observação 1.6 *Considerando o contexto de sistema dinâmico, Attouch e Teboulle em [3] definiram a distância proximal* $d(x, y) := d_\phi(x, y) + \frac{\nu}{2}\|x - y\|^2$*, onde* $\phi = h_1$ *do Exemplo 1.2 e* d_ϕ *é dada em (1.12). Nesse caso, a distância proximal induzida* H *é dada por:* $H(x, y) = K(x, y) + \frac{\nu}{2}\|x - y\|^2$*.*

Para mais detalhes sobre a distância proximal ϕ-divergente ver [37, 52, 53].

Exemplo 1.6 (Distância proximal homogênea de segunda ordem) *A distância proximal homogênea de segunda ordem é dada pela função* d_φ *definida em (1.9), onde* φ *é definida em (1.2). Nesse caso, tem-se a distância proximal induzida:*

$$H(x, y) = \frac{\mu + \nu}{2}\|x - y\|^2.$$

Observação 1.7 *Verifica-se que a distância log-quadrática é um caso particular de uma distância proximal homogênea de segunda ordem.*

1.3 Fejér Convergente

Nesta seção recordamos a definição de Fejér convergência para uma sequência. Além disso, apresentamos um resultado que será fundamental na análise de convergência dos algoritmos propostos nos capítulos subsequentes. Antes, consideremos B uma matriz $m \times n$, $m \geq n$, com posto completo. Como B tem posto máximo, a função $(x, y) \to \langle B^T Bx, y\rangle$ define um produto interno em $\mathbb{R}^n$, que denotamos por $\langle x, y\rangle_B$ com $\|x\|_B := \|Bx\| = \langle Bx, Bx\rangle^{\frac{1}{2}}$.

Definição 1.9 *Uma sequência $\{z^k\} \subset \mathbb{R}^n$ é dita Fejér Convergente para um conjunto não vazio $U \subset \mathbb{R}^n$ com relação à função $\|\cdot\|_B$, se para cada $u \in U$, verifica-se*

$$\|u - z^{k+1}\|_B \leq \|u - z^k\|_B, \text{ para todo } k \in \mathbb{N}. \tag{1.13}$$

Definição 1.10 *Uma sequência $\{z^k\} \subset \mathbb{R}^n$ é dita quase-Fejér Convergente para um conjunto não vazio $U \subset \mathbb{R}^n$ com relação à função $\|\cdot\|_B$, se para cada $u \in U$, existe uma sequência $\{\zeta_k\}$ tal que*

$$\|u - z^{k+1}\|_B \leq \|u - z^k\|_B + \zeta_k \text{ para todo } k \in \mathbb{N}, \tag{1.14}$$

com $\sum_{k=1}^{+\infty} \zeta_k < +\infty$.

Proposição 1.3 *Se $\{z^k\} \subset \mathbb{R}^n$ é Fejér Convergente para um conjunto não vazio $U \subset \mathbb{R}^n$ com relação à função $\|\cdot\|_B$, então $\{z^k\}$ é limitada. Além disso, se um ponto de acumulação z de $\{z^k\}$ pertence a U, então $\lim_{k\to\infty} z^k = z$.*

Prova: Segue diretamente de (1.13) que $\|u - z^k\|_B \leq \|u - z^0\|_B$ para todo $u \in U$, logo a sequência $\{z^k\}$ está contida numa bola de centro u e raio $\|u - z^0\|_B$, e portanto, ela é limitada. Seja $\{z^{k_j}\}$ uma subsequência de $\{z^k\}$ tal que $\lim_{k\to\infty} z^{k_j} = z$. Se $z \in U$ por (1.13) a sequência $\{\|z - z^k\|\}$ é decrescente e não negativa, e além disso ela possui uma subsequência ($\{\|z - z^{k_j}\|_B\}$) que converge para zero. Então a sequência converge para zero, isto é, $0 = \lim_{k\to\infty} \|z - z^k\|_B$ implicando $\lim_{k\to\infty} z^k = z$. $\square$

O resultado da proposição acima também é verificado quando a sequência $\{z^k\}$ é quase-Fejér convergente.

1.4 Operadores

Apresentaremos agora alguns conceitos e propriedades sobre operador ponto-conjunto.

Definição 1.11 *[6, Seção 1.4 p.20] Sejam X e Y dois subconjuntos não vazios do $\mathbb{R}^n$. Diremos que $T : X \to \mathcal{P}(Y)$ é um operador ponto-conjunto se a cada elemento $x \in X$ associamos um subconjunto $T(x) \subseteq Y$, ou seja, $T(x)$ é um elemento do conjunto de partes de Y, $\mathcal{P}(Y)$.*

Quando $T(x)$ é apenas um ponto, diremos que o operador é ponto-ponto. A seguir, listamos algumas definições relativas a operatores ponto-conjunto veja, por exemplo, em [50].

Definição 1.12 *Seja $T : X \to \mathcal{P}(X)$ um operador ponto-conjunto.*

(i) *O gráfico de $T, G(T)$, é definido por*

$$G(T) := \{(x, y) \in X \times X \ : \ y \in T(x)\}.$$

(ii) *O domínio de $T, D(T)$, é definido por*

$$D(T) := \{x \in X \ : \ T(x) \neq \emptyset\}.$$

(iii) *A imagem de T,* $\mathrm{Im}(A)$, *é definida por*

$$\mathrm{Im}(T) := \{u \in X : u \in T(x) \text{ para algum } x \in X\}.$$

A seguir, apresentamos caracterização para duas propriedades do operador ponto-conjunto: fecho e limitação local. Veja, por exemplo, [6] e [50], respectivamente.

Definição 1.13 *Um operador ponto-conjunto $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ é fechado em $\overline{x}$ se, para quaisquer sequências $\{x^k\} \subset \mathbb{R}^n$ e $\{y^k\} \subset \mathbb{R}^n$, com $(x^k, y^k) \in G(T), (x^k, y^k) \to (\overline{x}, \overline{y})$, implica que $\overline{y} \in T(\overline{x})$.*

Proposição 1.4 *Um operador ponto-conjunto $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ é localmente limitado se, e somente se, $T(X)$ é limitado para todo conjunto limitado X. Isto é equivalente à propriedade: quando $y^k \in T(x^k)$ e a sequência $\{x^k\} \subset \mathbb{R}^n$ é limitada, então a sequência $\{y^k\} \subset \mathbb{R}^n$ também é limitada.*

Como um dos nossos principais objetivos na primeira etapa desta tese é essencialmente um relaxamento das propriedades de monotonicidade do operador (ver Capítulo 3), somos motivados a apresentar a definição de algumas classes de monotonicidade. Esta definição pode ser encontrada em [28].

Definição 1.14 *Sejam X um subconjunto não vazio em $\mathbb{R}^n$ e $T : X \to \mathcal{P}(X)$ um operador ponto-conjunto, com $(x, u), (y, v) \in G(T)$. T é dito:*

(i) *Fortemente monótono em X, se existe um número real $\alpha > 0$ tal que* $\langle u - v, x - y\rangle \geq \alpha \|x - y\|^2$;

(ii) *Monótono em X, se* $\langle u - v, x - y\rangle \geq 0$;

(iii) *Pseudo-monótono em* X*, quando* $\langle v, x-y\rangle \geq 0$ implica $\langle u, x-y\rangle \geq 0$;

(iv) *Quase-monótono em* X*, quando* $\langle v, x-y\rangle > 0$ implica $\langle u, x-y\rangle \geq 0$;

(v) *Fracamente monótono em* X*, se existe um número real* $L > 0$*, tal que* $\langle u-v, x-y\rangle \geq -L\|x-y\|^2$.

A definição de operador Pseudo-monótono dada acima, é no sentido de Karamardian (ver [42]), que é diferente da definição de pseudomonotonicidade introduzida por Brezis em [12]. A seguir, apresentamos a definição de operador monótono maximal conforme [49].

Definição 1.15 *Um operador* $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ *é dito monótono maximal se seu gráfico não está inteiramente contido no gráfico de qualquer outro operador monótono, o que é equivalente a*

$$\langle u-v, x-y\rangle \geq 0, \forall v \in T(y), y \in \operatorname{dom} T \Longrightarrow u \in T(x).$$

A propriedade de maximalidade assume para operadores monótonos o mesmo papel que a continuidade assume para funções.

Observação 1.8 *Se* T *é um operador fracamente monótono com módulo* $L > 0$ *e* $\tilde{L} > L$ *é dado, então* $(T + \tilde{L}I)$ *é um operador fortemente monótono com módulo* $\tilde{L} - L$*, onde* I *denota o operador identidade. De fato, sejam* $(x,u),(y,v) \in G(T)$*, assim* $(x, u + \tilde{L}x), (y, v + \tilde{L}y) \in G(T + \tilde{L}I)$*, logo*

$$\begin{aligned} \langle (u+\tilde{L}x) - (v+\tilde{L}y), x-y\rangle &= \langle u-v, x-y\rangle + \tilde{L}\langle x-y, x-y\rangle \\ &\geq -L\|x-y\|^2 + \tilde{L}\|x-y\|^2 \\ &= (\tilde{L}-L)\|x-y\|^2. \end{aligned}$$

Observação 1.9 *Segue-se da Definição 1.14 que*
Monotonicidade forte $\Rightarrow$ *Monotonicidade* $\Rightarrow$ *Pseudomonotonicidade* $\Rightarrow$ *Quase-monotonocidade.*
Entretanto, a recíproca da relação acima não é verdadeira em geral, pois conforme [28, Exemplo 2.1], tem-se.
(i) *O operador* $T(x) = 1-x$ *é pseudo-monótono em* $X = [0,1]$*, mas não é monótono.*
(ii) *O operador* $T(x) = -x$ *é quase-monótono em* $X = [0, +\infty[$*, mas não é pseudo-monótono.*

O resultado que apresentamos a seguir é bastante conhecido na literatura, e nos mostra em que condição a maximalidade é preservada quando somamos dois operadores monótonos maximais.

Proposição 1.5 *[48, Teorema 2] Sejam T_1, T_2 operadores monótonos maximais. Se* $\mathrm{ir}D(T_1) \cap \mathrm{ir}D(T_2) \neq \emptyset$, *então $T_1 + T_2$ é um operador monótono maximal.*

A seguir, relembramos a definição do subdiferencial de Clarke de uma função f no ponto x, esta definição pode ser encontrada em [22].

Dada uma função localmente Lipschitz $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, o subdiferencial de Clarke, ∂f, é definido para todo $x \in \mathrm{dom}\,(f)$ por

$$\partial f(x) := \{x^* \in \mathbb{R}^n : f^0(x; d) \geq \langle x^*, d\rangle \ \text{ para todo } \ d \in \mathbb{R}^n\},$$

onde $f^0(x, d) := \lim\limits_{t \downarrow 0} \sup\limits_{y \to x} \left[\frac{f(y+td)-f(y)}{t}\right]$.

Quando f é uma função continuamente diferenciável, isto é, f é de classe C^1, $\partial f(x)$ é somente $\{\nabla f(x)\}$, e quando f é uma função convexa o subdiferencial de Clarke coincide com o subdiferencial clássico da análise convexa. No próximo resultado apresentamos dois exemplos clássicos de operadores ponto-conjunto.

Teorema 1.3 *Sejam $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função localmente Lipschitz e ∂f o seu subdiferencial de Clarke:*

(i) *[26, Teorema 2.1] f é quase-convexa se, e somente se, ∂f é quase-monótono;*

(ii) *[34, Corolário 3.2] Se f é pseudo-convexa, então ∂f é pseudo-monótono.*

Definição 1.16 *Seja T um operador ponto-conjunto.*

(i) *[43, p. 166] T é dito coercivo se, $D(T)$ é limitado, ou para todo $y_0 \in D(T)$*

$$\lim_{\|y\| \to +\infty} \frac{\langle v, y - y_0\rangle}{\|y\|} = +\infty \ \ \forall (v, y) \in G(T).$$

(ii) *[2] T é dito regular se*

$$\forall u \in \mathrm{Im}(T) \text{ e } \forall y \in D(T), \ \sup_{(z,v) \in G(T)} \langle v - u, y - z\rangle < \infty.$$

Os próximos resultados podem ser encontratos em [13] e [18], respectivamente.

Teorema 1.4 *Seja T um operador monótono maximal. Se T é um operador coercivo, então T é sobrejetivo e regular.*

Proposição 1.6 *Sejam $T_1, T_2 : \mathbb{R}^n \to \mathcal{P}(R^n)$ operadores monótonos maximais, tais que:*

(i) *T_1 é regular e sobrejetivo,*

(ii) $T_1 + T_2$ *é monótono maximal.*

Então $T_1 + T_2$ *é sobrejetivo.*

Para finalizar este capítulo, apresentamos o seguinte lema.

Lema 1.3 *[30, Lema 6.5] Se* T *é um operador fortemente monótono, então* T *é coercivo.*

Capítulo 2

Algoritmo Interior-Proximal para Otimização Quase-Convexa com Restrições Lineares

Neste capítulo apresentamos um algoritmo de ponto interior-proximal, denotado Algoritmo 1, para resolver o problema de otimização quase-convexa (P). Para tanto, estabelecemos a boa-definição da sequência gerada pelo Algoritmo 1. Mostramos convergência global a um ponto solução do problema proposto, quando a sequência dos parâmetros de regularização tende a zero. Esta condição pode ser retirada quando a função for pseudo-convexa.

Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria, fechada e C um conjunto poliedral em $\mathbb{R}^n$ definido por:

$$C := \{x \in \mathbb{R}^n : Ax \leq b\}, \tag{2.1}$$

onde A é uma matriz $m \times n, b \in R^m$ e $m \geq n$.

O Problema de Otimização Quase-Convexa, denotado (P), é dado por

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

No decorrer deste capítulo serão consideradas as seguintes hipóteses:

(H1) f é uma função continuamente diferenciável e limitada inferiormente;

(H2) $\mathrm{dom} f \cap \mathrm{int}\, C \neq \emptyset$;

(H3) $\mathrm{Posto}(A) = n$ (e portanto, A injetiva);

(H4) f é uma função quase-convexa.

Seja a_i a i-ésima linha da matriz A, para cada $x \in C$ defina:

$$y_i(x) := b_i - \langle a_i, x \rangle.$$

$$y(x) := (y_1(x), y_2(x), ..., y_m(x))^T.$$

$$D_\varphi(x, z) := d_\varphi(y(x), y(z)). \qquad (2.2)$$

Onde d_φ é dada por (1.9).

Observação 2.1 *Para cada $x, z \in \operatorname{int} C$, temos*

$$\begin{aligned} \nabla_1 D_\varphi(x, z) &= \sum_{i=1}^{n} (y_i(z))^2 \varphi' \left(\frac{y_i(x)}{y_i(z)} \right) \left(\frac{y_i(x)}{y_i(z)} \right)' \\ &= -\sum_{i=1}^{n} a_i y_i(z) \varphi' \left(\frac{y_i(x)}{y_i(z)} \right) \\ &= -A^T \nabla_1 d_\varphi(y(x), z(x)). \end{aligned}$$

Segue-se diretamente de (2.2) que,

$$D_\varphi(x, z) = \mu D_h(x, z) + \frac{\nu}{2} \parallel A(x - z) \parallel^2 \quad \text{para todo} \quad x, z \in \operatorname{int} C.$$

Sendo A uma matriz de posto máximo, (H3), a função $(x, y) \to \langle A^T A x, y \rangle$ define um produto interno em $\mathbb{R}^n$, que denotamos por $\langle x, y \rangle_A$ com $\|x\|_A := \|Ax\| = \langle Ax, Ax \rangle^{\frac{1}{2}}$. Assim,

$$D_\varphi(x, z) = \mu D_h(x, z) + \tfrac{\nu}{2} \parallel x - z \parallel_A^2 \qquad \text{para todo} \quad x, z \in \operatorname{int} C.$$

O lema a seguir é útil na análise de convergência. Para facilitar a notação defina

$$\theta := \frac{\nu + \mu}{2} \qquad (2.3)$$

onde ν, μ são dados em (1.2).

Lema 2.1 *Seja D_φ definida em (2.2). Então, para todo $x, z \in \operatorname{int} C$ e $w \in C$, valem as seguintes afirmações:*

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

(ii) $D_\varphi(\cdot, z)$ *é fortemente convexa com módulo ν, ou seja*
$\langle \nabla_1 D(x, p) - \nabla_1 D(z, p), x - z \rangle \geq \nu \parallel x - z \parallel_A^2,$ *para todo* $p \in \operatorname{int} C$;

(iii) $\langle \nabla_1 D_\varphi(x, z), w - x \rangle \leq \theta (\parallel w - z \parallel_A^2 - \parallel w - x \parallel_A^2)$*;*

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

Prova: Veja [7, Proposição 2.1]. □

Observação 2.2 *Como consequência do lema acima segue que os conjuntos de nível* $L_{D_\varphi(\cdot,z)}(\alpha) = \{x \in \operatorname{int} C : D_\varphi(x,z) \leq \alpha\}$ *são limitados para todo* $z \in \operatorname{int} C$ *e todo* $\alpha \in \mathbb{R}$*. De fato, seja* $\alpha = 0$*, então pelo item (i) do lema acima,* $L_{D_\varphi(\cdot,z)}(0) = \{x \in \operatorname{int} C : D_\varphi(x,z) \leq 0\} = \{x\}$*, o qual é limitado. Pelo item (ii) do mesmo lema* $D_\varphi(\cdot,z)$ *é convexa para todo* $z \in \operatorname{int} C$*. Portanto, de [47, Corolário 8.7.1], segue que* $L_{D_\varphi(\cdot,z)}(\alpha)$ *é limitado para todo* $\alpha \geq 0$*.*

2.1 Algoritmo 1

Agora, podemos formalizar o nosso algoritmo de ponto interior-proximal com a função definida em (2.2) para resolver o problema (P). Para isso escolhemos uma sequência de números reais positivos $\{\lambda_k\}$ satisfazendo:
(R1) $0 < \lambda_k \leq \overline{\lambda}$ para algum $\overline{\lambda} > 0$.

Algoritmo 1

Inicialização. Escolha algum $x^0 \in \operatorname{int} C$*.* $k := 0$*.*
Passo 1. Encontre a iteração $x^{k+1} \in \operatorname{int} C$ *tal que:*

$$x^{k+1} \in \arg\min\{f(x) + \lambda_k D_\varphi(x, x^k)\}. \tag{2.4}$$

Passo 2. Se $x^{k+1} = x^k$*, então pare. Caso contrário,*
Passo 3. Faça $k := k+1$ *e retorne ao passo 1.*

Observação 2.3 *Se* $x^{k+1} = x^k$ *para algum* k*, então* $\nabla f(x^{k+1}) = 0$*. De fato, pelo Lema 2.1 obtemos que* $\nabla_1 D_\varphi(x^{k+1}, x^k) = 0$*, o que implica* $\nabla f(x^{k+1}) = 0$*. Portanto,* x^{k+1} *é um ponto estacionário do problema (P).*

Observação 2.4 *Se* $x^k \neq x^{k-1}$*, então* x^k *não é uma solução do problema (P). De fato, suponha por absurdo que* x^k *seja uma solução do problema proposto, então* $\nabla f(x^k) = 0$*. Segue-se por (2.4) que* $\nabla_1 D_\varphi(x^k, x^{k-1}) = 0$*, mas pelo Lema 2.1 obtemos* $x^{k+1} = x^k$ *o que gera uma contradição.*

O processo iterativo gerado pelo Algoritmo 1 termina se para algum valor de k, digamos k_0, resulta $x^{k_0+1} = x^{k_0}$. Neste caso x^{k_0} é um ponto estacionário. Essa condição é necessária mas não é suficiente para deduzir que x^{k_0} seja uma solução do problema proposto. Na verdade, este processo nos leva a detectar "candidatos" à solução. No que segue vamos assumir que o Algoritmo 1 não tenha terminação finita, isto é, $x^{k+1} \neq x^k$ para todo k.

2.2 Existência

A seguir mostramos a boa definição da sequência gerada pelo Algoritmo 1, através da garantia da existência de cada iterado $x^k \in \operatorname{int} C$, para cada $k \in \mathbb{N}$.

Proposição 2.1 *Seja $\{x^k\}$ a sequência gerada pelo Algoritmo 1. Supondo (H1)-(H3), temos:*

(i) *A sequência $\{x^k\}$ está bem definida;*

(ii) *A sequência $\{f(x^k)\}$ é estritamente decrescente e convergente.*

Prova: **(i)** Provaremos por indução sobre k. Pela inicialização do algoritmo admitimos a existência de x^0. Suponha que x^k está bem definido. Seja $f_k(x) : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ definida por

$$f_k(x) := f(x) + \lambda_k D_\varphi(x, x^k)$$

A hipótese (H2) implica $\operatorname{dom} f_k \neq \emptyset$. Mostraremos que S_k é não vazio, onde $S_k := \arg\min f_k(x)$. Segue-se da Observação 2.2 que $D_\varphi(x, x^k)$ é nível limitada e como f é limitada inferiormente, (H1), obtemos que a função f_k também é nível limitada. Agora, pela continuidade de f_k, segue que os conjuntos de nível de f_k são compactos. Usando, novamente, a continuidade de f_k e a compacidade de seus conjuntos de níveis concluimos que f_k possui um mínimo global, o qual pode não ser único devido à não-convexidade de f. Assim, S_k é não vazio.

(ii) Pela definição de x^k, dada no Algoritmo 1, temos

$$f(x^k) + \lambda_k D_\varphi(x^k, x^{k-1}) \leq f(x) + \lambda_k D_\varphi(x, x^{k-1}) \quad \text{para todo} \quad x \in C.$$

Tomando $x = x^{k-1}$ na última desigualdade e aplicando o Lema 2.1 (i), segue que

$$f(x^k) + \lambda_k D_\varphi(x^k, x^{k-1}) \leq f(x^{k-1}).$$

Como $x^k \neq x^{k-1}$ temos que $D_\varphi(x^k, x^{k-1}) > 0$ e desde que $\lambda_k > 0$, obtemos

$$0 < \lambda_k D_\varphi(x^k, x^{k-1}) \leq f(x^{k-1}) - f(x^k).$$

Portanto, $\{f(x^k)\}$ é uma sequência estritamente decrescente e por (H1) segue-se o resultado desejado. $\square$

Observação 2.5 *Desde que S_k é não vazio obtemos*

$$\nabla f(x^{k+1}) = -\lambda_k \nabla_1 D_\varphi(x^{k+1}, x^k). \tag{2.5}$$

Agora estamos aptos a apresentar o primeiro resultado de convergência deste capítulo.

2.3 Análise de Convergência

Nesta seção provamos a convergência do Algoritmo 1 para uma solução do problema (P). Inicialmente, estabelecemos convergência a um ponto estacionário do problema proposto, quando a sequência gerada pelo Algoritmo 1, $\{x^k\}$, for convergente.

Proposição 2.2 *Assuma (H1)-(H3). Se a sequência $\{x^k\}$ for convergente, então ela converge para um ponto estacionário do problema (P).*

Prova: Seja $\overline{x} = \lim\limits_{k\to\infty} x^k$. Mostraremos que $\langle \nabla f(\overline{x}), x - \overline{x}\rangle \geq 0$ para todo $x \in C$, ou seja, $\overline{x}$ é um ponto estacionário do problema (P).
Para todo $x \in C$, seja $y = y(x) = b - Ax$ e $y^k = y(x^k) = b - Ax^k$ para todo $k \in \mathbb{N}$. Da Observação 2.1 e de (2.5) temos

$$\begin{aligned}
\langle \nabla f(x^k), x - x^k\rangle &= \langle -\lambda_k \nabla_1 D_\varphi(x^k, x^{k-1}), x - x^k\rangle \\
&= \lambda_k \langle A^T \nabla_1 d_\varphi(y^k, y^{k-1}), x - x^k\rangle \\
&= \lambda_k \langle \nabla_1 d_\varphi(y^k, y^{k-1}), A(x - x^k)\rangle \\
&= \lambda_k \langle \nabla_1 d_\varphi(y^k, y^{k-1}), y^k - y\rangle \\
&= \lambda_k \langle y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right), y^k - y\rangle \\
&= \lambda_k \langle y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right), y^k\rangle - \lambda_k \langle y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right), y\rangle, \qquad (2.6)
\end{aligned}$$

onde $y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right) := \left(y_1^{k-1}\varphi'\left(\frac{y_1^k}{y_1^{k-1}}\right), ..., y_m^{k-1}\varphi'\left(\frac{y_m^k}{y_m^{k-1}}\right)\right)^T$.
Como $x \in C$ e $x^k \in \operatorname{int} C$ temos que $y \in \mathbb{R}^n_+$ e $y^k \in \mathbb{R}_{++}$ para todo $k \in \mathbb{N}$.
Tomando $t = \frac{y_j^k}{y_j^{k-1}} > 0$ para algum $j \in \{1, ..., m\}$ em (1.4), obtemos

$$\varphi''(1)\left(1 - \frac{y_j^{k-1}}{y_j^k}\right) \leq \varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right) \leq \varphi''(1)\left(\frac{y_j^k}{y_j^{k-1}} - 1\right). \qquad (2.7)$$

Multiplicando a desigualdade acima por $y_j^{k-1}y_j^k$, segue que

$$\varphi''(1)y_j^{k-1}\left(y_j^k - y_j^{k-1}\right) \leq y_j^{k-1}y_j^k\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right) \leq \varphi''(1)y_j^k\left(y_j^k - y_j^{k-1}\right). \qquad (2.8)$$

Usando a convergência da sequência $\{x^k\}$, o Teorema do Confronto e o fato que $\varphi''(1) > 0$, obtemos

$$\lim_{k\to\infty}\left[y_j^{k-1}y_j^k\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right)\right] = 0, \quad \text{para todo} \quad j = 1, ..., m.$$

Consequentemente

$$\lim_{k\to\infty}\left\langle y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right), y^k\right\rangle = 0. \tag{2.9}$$

Considerando, agora, o lado direito da desigualdade (2.7) multiplicado por $y_j y_j^{k-1}$, obtemos

$$y_j y_j^{k-1}\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right) \leq \varphi''(1)y_j\left(y_j^k - y_j^{k-1}\right). \tag{2.10}$$

Usando novamente a convergência da sequência $\{x^k\}, \varphi''(1) > 0, y_j \geq 0$, obtemos que

$$\lim_{k\to\infty}\left[y_j y_j^{k-1}\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right)\right] \leq 0, \quad \text{para todo} \quad j = 1, ..., m.$$

Assim,

$$\lim_{k\to\infty}\left[-\left\langle y^{k-1}\varphi'\left(\frac{y^k}{y^{k-1}}\right), y\right\rangle\right] \geq 0. \tag{2.11}$$

Passando o limite em (2.6) combinado com (2.9) e (2.11), obtemos

$$\lim_{k\to\infty}\left\langle \nabla f(x^k), x - x^k\right\rangle \geq 0. \tag{2.12}$$

Como f é continuamente diferenciável, (H1), concluimos

$$\langle \nabla f(\overline{x}), x - \overline{x}\rangle \geq 0 \quad \text{para todo} \quad x \in C.$$

$\square$

Para provarmos o próximo resultado de convergência, assumiremos que a sequência do parâmetros $\{\lambda_k\}$ congerve a zero.

Teorema 2.1 *Se as hipóteses da Proposição 2.2 são verificadas e* $\lim_{k\to\infty}\lambda_k = 0$, *então a sequência gerada pelo Algoritmo 1 converge para uma solução do problema (P).*

Prova: Seja $\overline{x} = \lim_{k\to\infty} x^k$. Como x^{k+1} é um minimizador de $f(x) + \lambda_k D_\varphi(x, x^k)$, temos

$$f(x^{k+1}) + \lambda_k D_\varphi(x^{k+1}, x^k) \leq f(x) + \lambda_k D_\varphi(x, x^k) \quad \forall x \in \operatorname{int} C.$$

Pela não negatividade de D_φ e λ_k, segue que

$$f(x^{k+1}) \leq f(x) + \lambda_k D_\varphi(x, x^k) \quad \forall x \in \operatorname{int} C.$$

Tomando o limite na desigualdade acima, com $k \to \infty$, combinado com a continuidade da função f, $\lim_{k\to\infty} \lambda_k = 0$ e o Lema 2.1, obtemos

$$f(\overline{x}) \leq f(x) \quad \forall x \in \operatorname{int} C. \tag{2.13}$$

Consideremos agora uma sequência $\{z^k\} \subset \operatorname{int} C$ e $z \in C$, tal que $\lim_{k\to\infty} z^k = z$. Segue diretamente de (2.13) que

$$f(\overline{x}) \leq f(z^k) \quad \forall k \in \mathbb{N}.$$

Passando o limite na última desigualdade com $k \to \infty$, sendo f contínua, obtemos

$$f(\overline{x}) \leq f(z) \quad \forall z \in C.$$

Portanto $\overline{x}$ é uma solução do problema (P). □

Com o objetivo de estabelecer uma condição suficiente para a convergência da sequência gerada pelo Algoritmo 1, definimos o seguinte subconjunto de C.

$$U := \{x \in C : f(x) < f(x^k), \quad k = 0, 1, ...\}.$$

Observação 2.6 *Segue diretamente da Proposição 2.1 que x^k não resolve o problema (P). Todavia, se o problema (P) tem uma solução então $U \neq \emptyset$.*

Proposição 2.3 *Assuma que (H1)-(H4) são verificadas. Se $U \neq \emptyset$, então a sequência $\{x^k\}$ é Fejér convergente para U com relação à função $\|\cdot\|_A$.*

Prova: Dado $x \in U$, temos $f(x) < f(x^{k+1})$ para todo $k \in \mathbb{N}$, onde $\{x^k\}$ é gerada pelo Algoritmo 1. Aplicando a hipótese de quase-convexidade da função f (H4), (H1) e o Teorema 1.2, obtemos que

$$\langle \nabla f(x^{k+1}), x - x^{k+1} \rangle \leq 0, \quad \forall x \in U. \tag{2.14}$$

Combinando (2.5), (2.14) e Lema 2.1 (iii) com $x = x^{k+1}, z = x^k$ e $w = x$, segue que

$$\begin{aligned} 0 \leq \langle \nabla f(x^{k+1}), x^{k+1} - x \rangle &= \langle -\lambda_k \nabla_1 D_\varphi(x^{k+1}, x^k), x^{k+1} - x \rangle \\ &= \lambda_k \langle \nabla_1 D_\varphi(x^{k+1}, x^k), x - x^{k+1} \rangle \\ &\leq \lambda_k \theta \left[\|x - x^k\|_A^2 - \|x - x^{k+1}\|_A^2 \right], \end{aligned}$$

onde θ é dado em (2.3). Como $\lambda_k > 0$ e $\theta > 0$, obtemos

$$\|x - x^{k+1}\|_A^2 \quad \leq \quad \|x - x^k\|_A^2.$$

Pela Definição 1.9, isto completa a prova. □

No próximo resultado provamos que a sequência $\{x^k\}$ é convergente. Além disso, se a sequência dos parâmetros converge a zero, então como consequência direta do teorema anterior obteremos convergência a uma solução do problema (P).

Teorema 2.2 *Se (H1)-(H4) são verificadas e $U \neq \emptyset$, então a sequência $\{x^k\}$ gerada pelo Algoritmo 1 é convergente. Além disso, se $\lim_{k\to\infty} \lambda_k = 0$, então $\{x^k\}$ converge para uma solução de (P).*

Prova: Segue das Proposições 2.3 e 1.3 que a sequência $\{x^k\}$ é limitada. Seja $\overline{x}$ um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência convergente para $\overline{x}$. Pela continuidade da função f, temos que

$$\lim_{j\to\infty} f(x^{k_j}) = f(\overline{x}).$$

Como $\{f(x^k)\}$ é uma sequência estritamente decrescente e convergente, obtem-se que $\overline{x} \in U$. Agora, usando novamente a Proposição 1.3 concluimos que a sequência $\{x^k\}$ é convergente. A conclusão da prova segue diretamente do Teorema 2.1. □

Quando a função objetivo no problema (P) é uma função pseudo-convexa, a hipótese $\lim_{k\to\infty} \lambda_k = 0$ pode ser retirada.

Corolário 2.1 *Se (H1)-(H3), $U \neq \emptyset$ são satisfeitas e f é uma função pseudo-convexa, então a sequência $\{x^k\}$ converge para um ponto ótimo do problema* $\min_{x\in C} f(x)$.

Prova: O resultado segue diretamente do Teorema 2.2 e das Proposições 1.1 e 2.2. □

Finalmente, vamos considerar o caso onde $U = \emptyset$. Mostraremos que a sequência gerada pelo Algoritmo 1 é ilimitada. Entretando, a sequência $\{f(x^k)\}$ converge para o ínfimo de f em C.

Proposição 2.4 *Se (H1)-(H3) são satisfeitas e $U = \emptyset$, então*

(i) $\lim_{k\to\infty} f(x^k) = \inf_{x\in C} f(x)$;

(ii) *A sequência $\{x^k\}$ é ilimitada.*

Prova: **(i)** Sabemos que a sequência $\{f(x^k)\}$ é estritamente decrescente e convergente (Proposição 2.1). Assim, existe $\lim_{k\to\infty} f(x^k)$ e $\inf_{x\in C} f(x) \leq \lim_{k\to\infty} f(x^k) < f(x^k)$. Suponha, por absurdo, que $\inf_{x\in C} f(x) < \lim_{k\to\infty} f(x^k)$. Então, existe $\hat{x} \in C$ tal que $f(\hat{x}) < \lim_{k\to\infty} f(x^k) < f(x^k)$. Consequentemente, $f(\hat{x}) < f(x^k)$ para todo $k \in \mathbb{N}$, portanto $\hat{x} \in U$, o que é uma contradição, pois $U = \emptyset$, logo

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

(ii) Suponha, por absurdo, que $\{x^k\}$ seja limitada. Então existe uma subsequência $\{x^{k_j}\}$ de $\{x^k\}$ tal que,

$$\lim_{j\to\infty} x^{k_j} = w,$$

com $w \in C$. Pela continuidade da função f, obtemos

$$\lim_{j\to\infty} f(x^{k_j}) = f(w).$$

Como $\{f(x^k)\}$ é uma sequência estritamente decrescente e convergente, segue que

$$f(w) = \lim_{j\to\infty} f(x^{k_j}) = \lim_{k\to\infty} f(x^k) < f(x^k), \quad \forall k \in \mathbb{N}.$$

Portanto $w \in U$, o que é uma contradição. Assim, $\{x^k\}$ é ilimitada. $\square$

Capítulo 3

Problema da Desigualdade Variacional com Restrições Lineares

Neste capítulo apresentamos um algoritmo, denotado Algoritmo 2, para resolver o problema da desigualdade variacional associado a um operador quase-monótono. Mostramos sua boa definição e que o ponto limite da sequência gerada por este algoritmo é uma solução do $PDV(T;C)$.

Seja $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ um operador ponto-conjunto e C o conjunto definido em (2.1). O Problema da Desigualdade Variacional associado a T e C, chamado $PDV(T;C)$ é definido por

$$\text{(PDV)} \quad \begin{cases} \text{Obter } x^* \in C \text{ tal que exista } u^* \in T(x^*) \text{ com} \\ \langle u^*, x - x^* \rangle \geq 0 \text{ para todo } x \in C. \end{cases}$$

O conjunto solução do $PDV(T;C)$ será denotado por $SOL(T;C)$. De agora em diante vamos assumir que

(H5) $D(T) \cap \operatorname{int} C \neq \emptyset$;
(H6) O conjunto solução $SOL(T;C)$ do $VIP(T;C)$ é não vazio.

O próximo lema será fundamental para garantir a boa definição do Algoritmo 2. Lembramos que $\lambda_{\min}(A^T A)$ denota o autovalor mínimo da matriz simétrica positiva definida $(A^T A)$ e ν é a constante real dada na definição da função φ em (1.2).

Lema 3.1 *Fixado $x \in \operatorname{int} C$. Se T é um operador fracamente monótono com módulo $L > 0$ e $\{\beta_k\}$ é uma sequência de números reais positivos satisfazendo $\beta_k \geq \beta > \frac{L}{\nu \lambda_{\min}(A^T A)}$, então o operador $F(y) = \begin{cases} T(y) + \beta_k \nabla_1 D_\varphi(y, x) & \text{se } x, y \in \operatorname{int} C \\ \emptyset & \text{caso contrário} \end{cases}$ é* *fortemente monótono em C com constante $\left[\beta\nu\lambda_{\min}(A^TA) - L\right]$.*

Prova: Pelo Lema 2.1 (ii), (iv) e pela definição de operador fracamente monótono, temos

$$\begin{aligned}
\langle F(y_1) - F(y_2), y_1 - y_2\rangle &= \langle T(y_1) - T(y_2), y_1 - y_2\rangle \\
&+ \beta_k\langle \nabla_1 D_\varphi(y_1, x) - \nabla_1 D_\varphi(y_2, x), y_1 - y_2\rangle \\
&\geq -L\|y_1 - y_2\|^2 + \beta\nu\|y_1 - y_2\|_A^2 \\
&\geq -L\|y_1 - y_2\|^2 + \beta\nu\lambda_{\min}(A^TA)\|y_1 - y_2\|^2 \\
&= \left[\beta\nu\lambda_{\min}(A^TA) - L\right]\|y_1 - y_2\|^2.
\end{aligned}$$

□

3.1 Algoritmo 2

Para formalizar o nosso segundo algoritmo, consideramos
(R2) Uma sequência de números reais positivos $\{\beta_k\}$, satisfazendo o Lema 3.1.

Algoritmo 2

Inicialização. Escolha algum $x^0 \in \operatorname{int} C$. $k := 0$
Passo 2. Encontre a iteração $x^{k+1} \in \operatorname{int} C$ e $u^{k+1} \in T(x^{k+1})$ tais que:

$$u^{k+1} + \beta_k \nabla_1 D_\varphi(x^{k+1}, x^k) = 0. \tag{3.1}$$

Passo 2. Se $x^{k+1} = x^k$, então pare. Caso contrário,
Passo 3. Faça $k := k + 1$ e retorne ao passo 1.

Observação 3.1 *Se $x^{k+1} = x^k$ para algum k, então análogo à Observação 2.3, com u^k no lugar de $\nabla f(x^k)$, concluimos que $0 \in T(x^k)$. Consequentemente, x^k resolve o $PDV(T;C)$. Como estamos interessados na análise de convergência do Algoritmo 2, assumiremos que a iteração $x^{k+1} \neq x^k$ para todo k. Além disso, assumiremos até o fim deste capítulo que $\{x^k\}$ denota a sequência gerada pelo Algoritmo 2.*

3.2 Boa Definição

O teorema a seguir garante a boa definição da sequência $\{x^k\}$, considerando T um operador ponto-ponto.

Teorema 3.1 *Se T é um operador ponto-ponto contínuo fracamente monótono com módulo $L > 0$ e $\{\beta_k\}$ satisfaz (R2), então (3.1) tem solução única.*

Prova: O operador F definido no Lema 3.1 é contínuo e fortemente monótono no $\operatorname{int} C$. Então de acordo com [35, Corolário 3.2] x^{k+1} existe e é único. $\square$

A seguir garantimos a boa definição para uma classe de operadores ponto-conjunto.

Teorema 3.2 *Se $T = \partial f$ é o subdiferencial de Clarke de uma função semi-contínua inferior e fracamente monótono com módulo $L > 0$, então (3.1) tem solução única.*

Prova: $\nabla_1 D_\varphi$ é fortemente monótono e $D(T) \cap \operatorname{int} C \neq \emptyset$, por Lema 2.1 (ii) e (H5), respectivamente. Obtemos então que $T + \beta_k \nabla_1 D_\varphi = \partial(f + \beta_k D_\varphi)$ é o subdiferencial de Clarke de uma função própria, semi-contínua inferior e fortemente convexa. Neste caso, (convexo) o subdiferencial de Clarke coincide com o subdiferencial convexo [22]. Portanto, $T + \beta_k \nabla_1 D_\varphi$ é um operador monótono maximal e fortemente monótono, assim existe uma única solução para (3.1). $\square$

3.3 Análise de Convergência

Para mostrarmos a convergência do Algoritmo 2, precisamos da seguinte hipótese adicional:

(H7) T é um operador localmente limitado e $G(T)$ é fechado.

Observação 3.2 *Quando se trabalha com um operador monótono maximal a hipótese (H7) sempre é verificada. Mas como estamos trabalhando com um relaxamento da monotonicidade, destacamos que operadores ponto-ponto contínuo e o subdiferencial de Clarke de uma função localmente Lipschitz também satisfazem a hipótese (H7), ver [29, Proposição 7.1.4].*

Teorema 3.3 *Seja T um operador pseudo-monótono, tal que (3.1) tenha solução. Supondo (H3),(H5)-(H7), então a sequência $\{x^k\}$ converge para um elemento do $SOL(T;C)$.*

Prova: Dado $\overline{x} \in SOL(T, C)$, existe $\overline{u} \in T(\overline{x})$ tal que $\langle \overline{u}, x - \overline{x} \rangle \geq 0$ para todo $x \in C$, em particular $\langle \overline{u}, x^k - \overline{x} \rangle \geq 0$. Da pseudo-monotonicidade do operador T, obtemos $\langle u^k, x^k - \overline{x} \rangle \geq 0$, onde $u^k \in T(x^k)$.
Adicionando-se isto com (3.1) e Lema 2.1 com $x = x^k, z = x^{k-1}$ e $w = \overline{x}$, segue que

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

onde θ é dado em (2.3). Desde que $\beta_k > 0$ e $\theta > 0$, a sequência $\{x^k\}$ é Fejér convergente para $SOL(T;C)$, com relação à função $\|\cdot\|_A$. Portanto, segue da Proposição 1.3 que $\{x^k\}$ é limitada.
Sejam x^* um ponto de acumulação de $\{x^k\}$ e $\{x^{k_j}\}$ uma subsequência convergente para x^* com $u^{k_j} \in T(x^{k_j})$. Como o operador T é localmente limitado, (H7), a sequência $\{u^{k_j}\}$ também é limitada. Assim, existe uma subsequência $\{u^{k_{j_l}}\}$ de $\{u^{k_j}\}$ convergente para algum u^*. Sendo $G(T)$ fechado, (H7), temos que $u^* \in T(x^*)$. Sem perda de generalidade, podemos assumir que $u^{k_j} \to u^* \in T(x^*)$. De forma semelhante ao que foi usado para provar (2.12), obtemos

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

logo,

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

o que implica $x^* \in SOL(T, C)$.
Aplicando novamente a Proposição 1.3 concluimos que a sequência $\{x^k\}$ converge para $x^* \in SOL(T;C)$. $\square$

Como casos particulares de operadores pseudo-monótonos satisfazendo o último teorema, destacamos: T um operador fracamente monótono com módulo $L > 0$, como no Teorema 3.1, ou T pseudo-monótono satisfazendo o Teorema 3.2.

Consideramos agora o $PDV(T;C)$ associado a um operador quase-monótono T. Para obtermos convergência global da sequência gerada pelo Algoritmo 2 a um ponto do problema proposto, consideramos também o seguinte subconjunto do $SOL(T;C)$:

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

Além disso, relembramos a definição do cone normal de C em x^*

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

Observação 3.3 *Se a solução do $PDV(T;C)$ for um ponto do interior do conjunto C, então o $PDV(T;C)$ reduz-se a encontrar $0 \in T(x)$, ou seja, estaremos trabalhando com o problema irrestrito. Neste trabalho estamos interessados no caso $SOL^*(T;C) \neq \emptyset$ o que implica $SOL(T,C) \cap \partial C \neq \emptyset$, onde ∂C denota a fronteira do conjunto C.*

Motivados pela Observação 3.3, assumimos a seguinte hipótese:

(H8) $SOL^*(T;C) \neq \emptyset$.

Proposição 3.1 *Suponha (H8). Se* $x^* \in SOL^*(T,C)$ *e* $w \in \operatorname{int} C$, *então* $\langle u^*, w - x^* \rangle > 0$.

Prova: Dado $x^* \in SOL^*(T,C) \subset SOL(T;C)$, existe $u^* \neq 0$ pertencente a $T(x^*)$ tal que $\langle u^*, x - x^* \rangle \geq 0$ para todo $x \in C$, assim,
(i) $-u^* \in N_C(x^*)$,
(ii) $\langle u^*, w - x^* \rangle \geq 0$, para todo $w \in \operatorname{int} C$.
Suponha, por absurdo, que $\langle u^*, w - x^* \rangle = 0$ para algum $w \in \operatorname{int} C$. Como $w \in \operatorname{int} C$, existe $B(w,r) \subset \operatorname{int} C$, onde B é uma bola de centro w e raio $r > 0$.
Como $u^* \neq 0$, existe $\epsilon > 0$ ($\epsilon < r\|u^*\|$) tal que $\overline{x} = w - \frac{\epsilon}{\|u^*\|^2} u^* \in B(w,r)$. Assim, encontramos um elemento $\overline{x} \in C$ tal que,

$$\langle -u^*, \overline{x} - x^* \rangle = \langle -u^*, w - \frac{\epsilon}{\|u^*\|^2} u^* - x^* \rangle = \epsilon + \langle -u^*, w - x^* \rangle = \epsilon > 0.$$

O que é uma contradição pois $-u^* \in N_C(x^*)$. □

Mostramos abaixo que a sequência $\{x^k\}$ converge para uma solução do $PDV(T;C)$. Denotamos por $Acc(x^k)$ o conjunto dos pontos de acumulação da sequência $\{x^k\}$.

Teorema 3.4 *Seja* T *um operador quase-monótono, tal que (3.1) tenha solução. Assumindo (H3),(H5)-(H8), temos:*

(i) $Acc(x^k) \neq \emptyset$ *e todo elemento do* $Acc(x^k)$ *é uma solução do* $PDV(T;C)$,

(ii) *Se* $Acc(x^k) \cap SOL^*(T;C) \neq \emptyset$, *então* $\{x^k\}$ *converge para um elemento do* $SOL(T;C)$.

Prova: **(i)** Dado $x \in SOL^*(T;C)$, da Proposição 3.1 segue que existe $u \in T(x)$ tal que $\langle u, x^k - x \rangle > 0$ para todo $k \in \mathbb{N}$. Como T é um operador quase-monótono, temos

$$\langle u^k, x^k - x \rangle \geq 0 \text{ para todo } k \in \mathbb{N} \text{ e } u^k \in T(x^k).$$

De forma análoga à prova do Teorema 3.3, obtemos que a sequência $\{x^k\}$ é Fejér convergente para $SOL^*(T;C)$, com relação à função $\|\cdot\|_A$ portanto, ela é limitada, logo $Acc(x^k) \neq \emptyset$ e todo ponto de $Acc(x^k)$ é uma solução do $PDV(T;C)$.
(ii) Seja $x^* \in Acc(x^k) \cap SOL^*(T;C)$, da Proposição 1.3, $\{x^k\}$ converge para x^*. □

Capítulo 4

Método Proximal Inviável para o Problema da Desigualdade Variacional

Neste capítulo usando a distância proximal apresentada por Auslender e Teboulle em [3], apresentamos os algoritmos 3 e 4 com o objetivo de resolver o problema da desigualdade variacional $PDV(T; C)$. Diferente do que foi requerido no capítulo anterior, não será exigido que o interior topológico do conjunto C seja não vazio. Entretanto, será exigido que o operador T seja monótono maximal. Mostramos que os algoritmos estão bem definidos. Além disso, obtemos convergência a um ponto solução do $PDV(T; C)$.

Sejam $T : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ um operador ponto-conjunto, monótono maximal e C o subconjunto do $\mathbb{R}^n$ definido em (2.1). O nosso objetivo neste capítulo é: obter um $x^* \in C$, tal que exista $u^* \in T(x^*)$, com $\langle u^*, x - x^* \rangle \geq 0$ para todo $x \in C$.

4.1 Algoritmo 3

A fim de resolvermos o $PDV(T; C)$, apresentamos as seguintes notações:

(N1) $C^k := \{x \in \mathbb{R}^n : Ax \leq b + \delta^k\}$, onde $\delta^k \in \mathbb{R}^m_{++}$;

(N2) Seja a_i a i-ésima linha da matriz A e $d \in \mathcal{D}(\mathbb{R}^n_{++})$ uma distância proximal. Para cada $x^k \in C^k$ defina:

$$y_i(x) := b_i + \delta_i^k - \langle a_i, x \rangle,$$

$$y^k(x^k) := (y_1^k(x^k), y_2^k(x^k), ..., y_m^k(x^k))^T,$$

$$D(x^k, z^k) := d(y^k(x^k), y^k(z^k)).$$

Observação 4.1 *De* (N1) *e* (N2), *decorrem as seguintes afirmações:*

(i) *Para cada* $x^k \in \operatorname{int} C^k$ *e* $x^{k-1} \in \operatorname{int} C^{k-1}$, *temos*

$$\nabla_1 D(x^k, x^{k-1}) = -A^T \nabla_1 d(y^k(x^k), y^{k-1}(x^{k-1})).$$

(ii) *Como* $C \subset \operatorname{int} C^k$, *se* $C \neq \emptyset$ *então* $\operatorname{int} C^k \neq \emptyset$ *para todo* k.

Agora encontramo-nos em condição de apresentar o nosso primeiro algoritmo deste capítulo.

Algoritmo 3

Inicialização. Escolha $x^0 \in \mathbb{R}^n$ *e* $y^0 \in \mathbb{R}^m_{++}$ *tal que* $\delta^0 := y^0 - (b - Ax^0) \in \mathbb{R}^m_{++}$.
Passo 1. Escolha $\lambda_k > 0, \rho_k \in (0,1)$ *e faça* $\delta^k := \rho_k \delta^{k-1}$.
Encontre $(x^k, y^k) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$ *e* $u^k \in \mathbb{R}^n$ *tal que:*

$$\begin{cases} u^k \in T(x^k), \\ u^k + \lambda_k \nabla_1 D(x^k, x^{k-1}) = 0, \\ y^k - (b - Ax^k) = \delta^k. \end{cases} \tag{4.1}$$

Passo 2. Se $x^{k+1} = x^k$ *e* $x^k \in C$, *então pare. Caso contrário,*
Passo 3. Faça $k := k+1$ *e retorne ao passo 1.*

Observação 4.2 *No Algoritmo 3, o ponto inicial pode ser escolhido arbitrariamente, desde que* $\delta^0 \in \mathbb{R}^m_{++}$, *ou seja, basta escolher* y^0 *suficientemente "grande" para garantir que* $\delta^0 \in \mathbb{R}^m_{++}$. *Além disso, decorre do Passo 1 que* $\sum_{k=0}^{+\infty} \|\delta^k\| < +\infty$.

Observação 4.3 *A sequência de números reais positivos* $\{\lambda_k\}$ *deve ser escolhida satisfazendo as seguintes condições:*

(i) $\lambda_k \to 0$;

(ii) $\sum_{k=0}^{+\infty} \lambda_k^{-1} \|\delta^k\| < +\infty$.

Relembramos que neste capítulo estamos trabalhando com um operador T monótono maximal. Além disso, vamos assumir as seguintes hipóteses:

(A1) T é coercivo ou $\nabla_1 D$ é coercivo;

(A2) $\text{ir}\, C \cap \text{ir}\, D(T) \neq \emptyset$;

(A3) A tem posto máximo.

A condição (A1) é naturalmente verificada quando se está trabalhando com operadores fortemente monótono ou com a distância proximal homogênea de segunda ordem, em particular a distância proximal log-quadrática. Pois esta distância proximal é uma função fortemente convexa, isto implica que o operador $\nabla_1 D$ é fortemente monótono e, portanto, coercivo. A condição (A2) é um relaxamento da hipótese $D(T) \cap \text{int}\, C \neq \emptyset$, que é assumida em diversos trabalhos, veja por exemplo [3, 5, 53]. A hipótese (A3) é padrão.

4.1.1 Boa Definição

O resultado a seguir garante a boa definição dos iterados gerados pelo nosso algoritmo.

Teorema 4.1 *Assuma (A1)-(A3). Para cada $\lambda_k > 0, \delta^k \succ 0$ e $(x^{k-1}, y^{k-1}) \in \text{int}\, C^k \times \mathbb{R}^m_{++}$, existe um par $(x^k, y^k) \in \text{int}\, C^k \times \mathbb{R}^m_{++}$, satisfazendo (4.1).*

Prova: Assuma T coercivo. Defina os operadores $T_1 := T$ e $T_2 := \lambda_k \nabla_1 D + N_{C^k}$. Para todo k é verdadeiro que:

(i) T_1 é regular e sobrejetivo, ver Teorema 1.4.

(ii) $\text{ir}\, D(T) \cap \text{ir}\, D(\nabla_1 D) \cap \text{ir}\, D(N_{C^k}) = \text{ir}\, D(T) \cap \text{ir}\, C^k \supset \text{ir}\, D(T) \cap \text{ir}\, C \neq \emptyset$, por (A2).

(ii) $T_1 + T_2$ é monótono maximal (por (ii) e Proposição 1.5).

Então pela Proposição 1.6 o operador $T + \lambda_k \nabla_1 D + N_{C^k}$ é sobrejetivo. Portanto, existe uma solução x^k da equação

$$0 \in T(x^k) + \lambda_k \nabla_1 D(x^k, x^{k-1}) + N_{C^k}(x^k).$$

Deste modo, existem $u^k \in T(x^k)$ e $w^k \in N_{C^k}(x^k)$, tais que

$$0 = u^k + \lambda_k \nabla_1 D(x^k, x^{k-1}) + w^k. \tag{4.2}$$

Fazendo $b + \delta^k - Ax^k = y^k$, temos que $y^k \in \mathbb{R}^m_{++}$, o que implica $x^k \in \operatorname{int} C^k$, daí $w^k = 0$. Assim, por (4.2) existe um par $(x^k, y^k) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$ satisfazendo

$$\begin{cases} u^k \in T(x^k), \\ u^k + \lambda_k \nabla_1 D(x^k, x^{k-1}) = 0, \\ y^k - (b - Ax^k) = \delta^k. \end{cases}$$

Agora, assuma que $\nabla_1 D$ é coercivo. Defina os operadores $T_1 := \nabla_1 D$ e $T_2 := T + N_{C^k}$ e a prova segue de forma semelhante ao caso anterior. $\square$

4.1.2 Análise de Convergência

Nesta seção, provamos a convergência do Algoritmo 3 para uma solução do problema proposto. Para mostrarmos o próximo resultado, precisamos da seguinte hipótese adicional:

(A4) O conjunto solução $SOL(T; C)$ é não vazio.

Proposição 4.1 *Assuma (A1)-(A4). Dados $(d, H) \in \mathcal{F}(\mathbb{R}^n_+), \overline{x} \in SOL(T; C)$ e $\overline{u} \in T(\overline{x})$. Defina $\overline{y} := b - A\overline{x}$, então para todo $k = 1, 2, ...$, temos:*

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle + \alpha \lambda_k^{-1} \|\delta^k\| \|\overline{u}\|, \tag{4.3}$$

onde α é uma constante positiva.

Prova: Tome $k > 0$. Seja $u^k \in T(x^k)$. Para todo $(x, u) \in G(T)$ temos,

$$\langle x - x^k, u - u^k \rangle \geq 0.$$

Como $u^k = -\lambda_k \nabla_1 D(x^k, x^{k-1})$, ver (4.1), segue-se que

$$\langle x - x^k, u \rangle \geq \langle x - x^k, u^k \rangle = \langle x - x^k, -\lambda_k \nabla_1 D(x^k, x^{k-1}) \rangle. \tag{4.4}$$

Combinando a desigualdade acima com a Observação 4.1 e (N2), temos

$$\begin{aligned} \langle x - x^k, u \rangle &\geq \lambda_k \langle x - x^k, A^T \nabla_1 d(y^k, y^{k-1}) \rangle \\ &= \lambda_k \langle A(x - x^k), \nabla_1 d(y^k, y^{k-1}) \rangle \\ &= \lambda_k \langle y^k - y - \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle \\ &= \lambda_k \langle y^k - y, \nabla_1 d(y^k, y^{k-1}) \rangle - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle. \end{aligned} \tag{4.5}$$

Usando agora a Definição 1.7 com $v = y^k, u = y^{k-1}$ e $w = y$ em (4.5), obtemos

$$\langle x - x^k, u \rangle \geq \lambda_k \left[H(y, y^k) - H(y, y^{k-1}) \right] - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle. \tag{4.6}$$

A desigualdade acima vale para todo $(x,u) \in G(T)$, em particular para $(\overline{x},\overline{u})$, onde $\overline{x} \in SOL(T;C)$, $\overline{u} \in T(\overline{x})$ e $\overline{y} := b - A\overline{x}$. Portanto,

$$\langle \overline{x} - x^k, \overline{u} \rangle \geq \lambda_k \left[H(\overline{y}, y^k) - H(\overline{y}, y^{k-1}) \right] - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle. \tag{4.7}$$

Por outro lado, como $\overline{x} \in SOL(T;C)$ e $\overline{u} \in T(\overline{x})$, temos

$$\langle \overline{x} - x, \overline{u} \rangle \leq 0 \;\; \forall x \in C.$$

Seja p^k a projeção de x^k sobre C. Desde que $p^k \in C$, tem-se

$$\langle \overline{x} - p^k, \overline{u} \rangle \leq 0.$$

Assim,

$$\langle \overline{x} - x^k + x^k - p^k, \overline{u} \rangle \leq 0,$$

e logo

$$\langle \overline{x} - x^k, \overline{u} \rangle \leq \langle p^k - x^k, \overline{u} \rangle.$$

Usando a desigualdade de Cauchy-Schwarz, obtemos

$$\langle \overline{x} - x^k, \overline{u} \rangle \leq \|p^k - x^k\| \|\overline{u}\|.$$

Portanto, segue do Lema 1.2 que

$$\langle \overline{x} - x^k, \overline{u} \rangle \leq \alpha \|\overline{u}\| \|\delta^k\|, \tag{4.8}$$

para algum $\alpha > 0$. Combinando (4.7) e (4.8) obtemos que

$$\alpha \|\overline{u}\| \|\delta^k\| \geq \lambda_k \left[H(\overline{y}, y^k) - H(\overline{y}, y^{k-1}) \right] - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle,$$

portanto,

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle + \alpha \lambda_k^{-1} \|\delta^k\| \|\overline{u}\|.$$

□

A fim de mostrarmos a convergência do nosso algoritmo, apresentamos uma observação que fundamenta a condição adicional requerida no próximo teorema.

Observação 4.4 *Seja $d \in \mathcal{D}(\mathbb{R}^n_{++})$.*

(i) *Distância proximal log-quadrático.*

Considere φ dada em (1.2). De (2.6) temos que

$$\nabla_1 d_\varphi(y^k, y^{k-1}) = \left(y_1^{k-1}\varphi'\left(\frac{y_1^k}{y_1^{k-1}}\right), ..., y_m^{k-1}\varphi'\left(\frac{y_m^k}{y_m^{k-1}}\right)\right)^T. \tag{4.9}$$

Tomando $t = \frac{y_j^k}{y_j^{k-1}} > 0$ para $j \in \{1, ..., m\}$ em (1.4), temos

$$\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right) \leq \varphi''(1)\left(\frac{y_j^k}{y_j^{k-1}} - 1\right).$$

Multiplicando a desigualdade acima por y_j^{k-1}, obtemos

$$y_j^{k-1}\varphi'\left(\frac{y_j^k}{y_j^{k-1}}\right) \leq \varphi''(1)\left(y_j^k - y_j^{k-1}\right). \tag{4.10}$$

Combinando (4.9), (4.10) e a desigualdade de Cauchy-Schwarz, obtemos

$$\langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle \leq \varphi''(1)\langle \delta^k, y^k - y^{k-1}\rangle \leq \varphi''(1)\|\delta^k\|\|y^k - y^{k-1}\|.$$

Portanto existe uma constante $L > 0$, $L = \varphi''(1)$, tal que $\langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle < L\|\delta^k\|\|y^k - y^{k-1}\|$.

(ii) *Distância proximal de Bregman*

Seja h uma função de Bregman com zona S. Da Proposição 1.2, temos

$$\nabla_1 d(y^k, y^{k-1}) = \nabla h(y^k) - \nabla h(y^{k-1}).$$

Suponha também que o gradiente da função h seja Lipschitz. Então existe uma constante $L > 0$ tal que

$$\begin{aligned}\langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle &\leq \|\delta^k\|\|\nabla_1 d(y^k, y^{k-1})\| = \|\delta^k\|\|\nabla h(y^k) - \nabla h(y^{k-1})\| \\ &\leq L\|\delta^k\|\|y^k - y^{k-1}\|.\end{aligned}$$

A observação acima nos permite considerar a seguinte hipótese:

(A5) Existe uma constante $L > 0$, tal que $\langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle \leq L\|\delta^k\|\|y^k - y^{k-1}\|$.

Esta condição será usada no próximo resultado, para garantir a convergência da sequência gerada pelo Algoritmo 3 a uma solução do $PDV(T;C)$. Com a finalidade de facilitar a notação, definimos:

$$\widehat{SOL}(T;C) := \{\overline{y} \in \mathbb{R}^m : \overline{y} = b - A\overline{x},\ \overline{x} \in SOL(T;C)\}.$$

Desde que $SOL(T;C)$ é diferente do vazio $(A4)$, o mesmo ocorre com $\widehat{SOL}(T;C)$.

Teorema 4.2 *Assuma (A1)-(A5). Dados* $(d,H) \in \mathcal{F}(\mathbb{R}^n_+), \overline{x} \in SOL(T;C)$ *e* $\overline{u} \in T(\overline{x})$. *Se* $\{x^k\}$ *for limitada, então:*

(i) *A sequência* $\{y^k\}$ *é quase-Fejér convergente para* $\widehat{SOL}(T;C)$, *com respeito à distância proximal induzida* H;

(ii) *A sequência* $\{H(\overline{y},y^k)\}$ *é convergente, onde* $\overline{y} \in \widehat{SOL}(T;C)$;

(iii) *A sequência* $\{x^k\}$ *converge para um elemento do conjunto* $SOL(T;C)$.

Prova: **(i)** De (4.3) e (A5) tem-se que

$$H(\overline{y},y^k) \leq H(\overline{y},y^{k-1}) + L\|\delta^k\|\|y^k - y^{k-1}\| + \alpha\lambda_k^{-1}\|\delta^k\|\|\overline{u}\|, \tag{4.11}$$

defina

$$\beta_k := L\|\delta^k\|\|y^k - y^{k-1}\| + \alpha\lambda_k^{-1}\|\delta^k\|\|\overline{u}\|.$$

Como $\{y^k\}$ é limitada, $\sum_{k=0}^{\infty}\|\delta^k\| < +\infty$ e $\sum_{k=0}^{\infty}\lambda_k^{-1}\|\delta^k\| < +\infty$, concluimos que $\sum\limits_{k=0}^{+\infty}\beta_k < +\infty$. Portanto

$$H(\overline{y},y^k) \leq H(\overline{y},y^{k-1}) + \beta_k \text{ com } \sum_{k=0}^{+\infty}\beta_k < +\infty,$$

ou seja, a sequência $\{y^k\}$ é quase-Fejér convergente para $\widehat{SOL}(T;C)$.

(ii) Defina $\sigma_{k+1} := H(\overline{y},y^k)$. Como as sequências $\{\beta_k\}$ e $\{\sigma_k\}$ satisfazem às condições exigidas pelo Lema 1.1, concluimos que $\{H(\overline{y},y^k)\}$ é convergente.

(iii) Dado $(x,u) \in G(T)$. De (4.6), Definição 1.7 e (A5), temos

$$\begin{aligned} \langle x - x^k, u\rangle &\geq \lambda_k\left[H(y,y^k) - H(y,y^{k-1}) - \langle\delta^k, \nabla_1 d(y^k,y^{k-1})\rangle\right] \\ &\geq \lambda_k\left[-H(y,y^{k-1}) - L\|\delta^k\|\|y^k - y^{k-1}\|\right] \end{aligned} \tag{4.12}$$

Sabemos que $\lambda_k \to 0$, $\{y^k\}$ é limitado, $H(y,\cdot)$ é nível limitado e $\sum_{k=0}^{\infty}\|\delta^k\| < +\infty$, então tomando o limite em (4.12) com $k \to +\infty$, obtemos

$$\lim_{k\to+\infty}\inf\langle x - x^k, u\rangle \geq 0. \tag{4.13}$$

Por outro lado, como $\{x^k\}$ é limitada, seja $\{x^{k_j}\} \subset \{x^k\}$ uma subsequência convergente para algum $\widehat{x}$. Então, por (4.13), obtemos

$$\langle x - \widehat{x}, u\rangle = \lim_{j\to+\infty}\langle x - x^{k_j}, u\rangle \geq \lim_{k\to+\infty}\inf\langle x - x^k, u\rangle \geq 0.$$

Assim,

$$\langle x - \widehat{x}, u \rangle \geq 0 \ \ \forall (x, u) \in G(T). \tag{4.14}$$

Por definição $y^{k_j} = b - Ax^{k_j} + \delta^{k_j}$, com $y^{k_j} > 0$. Passando ao limite quando $j \to \infty$ temos que $\widehat{y} = b - A\widehat{x}$ com $\widehat{y} \geq 0$. Portanto, $A\widehat{x} \leq b$, o que implica $\widehat{x} \in C$. Usando agora a definição de N_C, temos

$$\langle x - \widehat{x}, w \rangle \geq 0 \ \ \forall (x, w) \in G(N_C). \tag{4.15}$$

Combinando (4.14) e (4.15)

$$\langle x - \widehat{x}, u + w \rangle \geq 0 \ \ \forall (x, u + w) \in G(T + N_C).$$

Por (A2) e Proposição 1.5 o operador $T + N_C$ é monótono maximal. Portanto $0 \in (T + N_C)(\widehat{x})$, isto é, $\widehat{x} \in SOL(T; C)$. Por outro lado, isto implica que $\widehat{y} \in \widehat{SOL}(T; C)$. Como a sequência $\{y^k\}$ é quase-Fejér convergente para $\widehat{SOL}$, e além disso, possui um ponto de acumulação pertencente a este conjunto, concluimos pela Proposição 1.3 que $\{y^k\}$ converge para $\widehat{y}$, o que implica na convergência da sequência $\{x^k\}$ para $\widehat{x} \in SOL(T; C)$. □

4.2 Algoritmo 4

Para obter a convergência da sequência gerada pelo Algoritmo 3, foi necessário assumir a coercividade do operador T (ou $\nabla_1 D$). No método que apresentamos agora, substituimos esta exigência por uma regularização do tipo Tikhonov do operador $\nabla_1 D$. Tais regularizações exigem que a sequência dos parâmetros de regularização convirja a zero, a fim de estabelecer a convergência do método.

Seja $d \in \mathcal{D}(\mathbb{R}^n_{++})$ uma distância proximal. Para todo $x, z \in \operatorname{int} C^k$, defina:

$$D_k : \operatorname{int} C^k \times \operatorname{int} C^k \to \mathbb{R}.$$

$$D_k(x, z) := d(y(x), y(z)) + \frac{\nu_k}{2} \|x - z\|^2.$$

Relembramos que $y^k(x^k) = b - Ax^k + \delta^k$, onde $x^k \in C^k$.

Observação 4.5 *Segue diretamente da definição de D_k as seguintes afirmações:*

(i) *Para cada $x^k \in \operatorname{int} C^k$ e $x^{k-1} \in \operatorname{int} C^{k-1}$, temos*

$$\nabla_1 D_k(x^k, x^{k-1}) = -A^T \nabla_1 d(y^k(x^k), y^{k-1}(x^{k-1})) + \nu_k (x^k - x^{k-1}).$$

(ii) *Como $D_k(\cdot, z)$ é fortemente convexa, então $\nabla_1 D_k(\cdot, z)$ é fortemente monótono, o que implica coercividade do operador $\nabla_1 D_k(\cdot, z)$, ver Lema 1.3.*

Observação 4.6 *A sequência de números reais positivos* $\{\nu_k\}$ *deve ser escolhida de tal forma que* $\sum_{k=0}^{+\infty} \nu_k < +\infty$.

Exceto pela distância proximal a ser considerada, o algoritmo abaixo tem a mesma estrutura do Algoritmo 3.

Algoritmo 4

Inicialização. Escolha $x^0 \in \mathbb{R}^n$ *e* $y^0 \in \mathbb{R}^m_{++}$ *tal que* $\delta^0 := y^0 - (b - Ax^0) \in \mathbb{R}^m_{++}$.
Passo 1. Escolha $\lambda_k > 0, \rho_k \in (0,1)$ *e faça* $\delta^k := \rho_k \delta^{k-1}$.
Encontre $(x^k, y^k) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$ *e* $u^k \in \mathbb{R}^n$ *tal que:*

$$\begin{cases} u^k \in T(x^k), \\ u^k + \lambda_k \nabla_1 D_k(x^k, x^{k-1}) = 0, \\ y^k - (b - Ax^k) = \delta^k. \end{cases} \tag{4.16}$$

Passo 2. Se $x^{k+1} = x^k$ *e* $x^k \in C$, *então pare. Caso contrário,*
Passo 3. Faça $k := k + 1$ *e retorne ao passo 1.*

4.2.1 Boa definição

Mostramos abaixo que o Algoritmo 4 está bem definido e, que a sequência gerada pelo referido algoritmo é limitada.

Teorema 4.3 *Assuma (A2)-(A3). Para todo* $\lambda_k > 0, \delta^k \succ 0$ *e* $(x^{k-1}, y^{k-1}) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$, *existe um par* $(x^k, y^k) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$, *satisfazendo (4.16).*

Prova: Defina o operador,

$$J_{\lambda_k}(x) := T(x) + \lambda_k \nabla_1 D_k(x, x^{k-1}) + N_{C^k}(x).$$

Para todo k é verdadeiro que:

(i) $D(J_{\lambda_k}) = D(T) \cap D(\nabla_1 D_k) \cap D(N_{C^k}) = D(T) \cap C^k \neq \emptyset$, por (A2).

(ii) Como $\nabla_1 D_k$ é fortemente monótono, ver Observação 4.5, temos que J_{λ_k} tambem é fortemente monótono, e pelo Lema 1.3 J_{λ_k} é coercivo.

Usando (i),(ii) e o Teorema 1.4 concluimos que J_{λ_k} é sobrejetivo. Assim, existe uma solução x^k da equação

$$0 \in T(x^k) + \lambda_k \nabla_1 D_k(x^k, x^{k-1}) + N_{C^k}(x^k).$$

E de forma análoga à prova do Teorema 4.3, obtemos a existência de um par $(x^k, y^k) \in \operatorname{int} C^k \times \mathbb{R}^m_{++}$ satisfazendo (4.16). $\square$

4.2.2 Análise de Convergência

Nesta subseção provamos a convergência do Algoritmo 4 para uma solução do $PDV(T; C)$.

Proposição 4.2 *Assuma (A2)-(A3). Dados* $(d, H) \in \mathcal{F}(\mathbb{R}^n_+), \overline{x} \in SOL(T; C)$ *e* $\overline{u} \in T(\overline{x})$. *Defina* $\overline{y} := b - A\overline{x}$, *então para todo* $k = 1, 2, ...,$ *temos:*

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + \langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle + \nu_k \|\overline{x} - x^k\| \|x^k - x^{k-1}\| + \alpha\lambda_k^{-1} \|\delta^k\| \|\overline{u}\|, \tag{4.17}$$

onde α *é uma constante positiva.*

Prova: Tome $k > 0$. Seja $u^k \in T(x^k)$. Para todo $(x, u) \in G(T)$ temos,

$$\langle x - x^k, u - u^k\rangle \geq 0.$$

Desde que $u^k = -\lambda_k \nabla_1 D_k(x^k, x^{k-1})$, ver (4.16), segue-se que

$$\langle x - x^k, u\rangle \geq \langle x - x^k, u^k\rangle = \langle x - x^k, -\lambda_k \nabla_1 D_k(x^k, x^{k-1})\rangle. \tag{4.18}$$

Da Observação 4.5, (4.18) e de (N2), temos

$$\begin{aligned}
\langle x - x^k, u\rangle &\geq \lambda_k \langle x - x^k, A^T \nabla_1 d(y^k, y^{k-1}) - \nu_k (x^k - x^{k-1})\rangle \\
&= \lambda_k \langle A(x - x^k), \nabla_1 d(y^k, y^{k-1})\rangle - \lambda_k \nu_k \langle x - x^k, x^k - x^{k-1}\rangle \\
&= \lambda_k \langle y^k - y - \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle - \lambda_k \nu_k \langle x - x^k, x^k - x^{k-1}\rangle \\
&= \lambda_k \langle y^k - y, \nabla_1 d(y^k, y^{k-1})\rangle - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle \\
&- \lambda_k \nu_k \langle x - x^k, x^k - x^{k-1}\rangle.
\end{aligned}$$

Usando a última expressão com a Definição 1.7 com $v = y^k, u = y^{k-1}$ e $w = y$, obtemos

$$\langle x - x^k, u\rangle \geq \lambda_k \left[H(y, y^k) - H(y, y^{k-1})\right] - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle - \lambda_k \nu_k \langle x - x^k, x^k - x^{k-1}\rangle. \tag{4.19}$$

De forma análoga ao que foi feito na prova da Proposição 4.1, obtemos que

$$\alpha \|\overline{u}\| \|\delta^k\| \geq \lambda_k \left[H(\overline{y}, y^k) - H(\overline{y}, y^{k-1})\right] - \lambda_k \langle \delta^k, \nabla_1 d(y^k, y^{k-1})\rangle - \lambda_k \nu_k \langle \overline{x} - x^k, x^k - x^{k-1}\rangle,$$

onde $\overline{x} \in SOL(T;C)$, $\overline{u} \in T(\overline{x})$ e $\overline{y} = b - A\overline{x}$. Portanto, aplicando a desigualdade de Cauchy-Schwarz concluimos que

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle + \nu_k \|\overline{x} - x^k\| \|x^k - x^{k-1}\| + \alpha \lambda_k^{-1} \|\delta^k\| \|\overline{u}\|.$$

$\square$

Mostraremos agora que a sequência gerada pelo Algoritmo 4 converge a uma solução do problema variacional.

Teorema 4.4 *Assuma (A2)-(A5). Dados* $(d, H) \in \mathcal{F}(\mathbb{R}^n_+), \overline{x} \in SOL(T;C)$ *e* $\overline{u} \in T(\overline{x})$. *Se* $\{x^k\}$ *for limitada, então:*

(i) *A sequência* $\{y^k\}$ *é quase-Fejér convergente para* $\widehat{SOL}(T;C)$*, com respeito à distância proximal induzida* H*;*

(ii) *A sequência* $\{H(\overline{y}, y^k)\}$ *é convergente, onde* $\overline{y} \in \widehat{SOL}(T;C)$*;*

(iii) *A sequência* $\{x^k\}$ *converge para um elemento do conjunto* $SOL(T;C)$*.*

Prova: **(i)** De (4.17) e (A5), temos

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + L\|\delta^k\| \|y^k - y^{k-1}\| + \nu_k \|\overline{x} - x^k\| \|x^k - x^{k-1}\| + \alpha \lambda_k^{-1} \|\delta^k\| \|\overline{u}\|, \tag{4.20}$$

defina

$$\overline{\beta}_k := L\|\delta^k\| \|y^k - y^{k-1}\| + \nu_k \|\overline{x} - x^k\| \|x^k - x^{k-1}\| + \alpha \lambda_k^{-1} \|\delta^k\| \|\overline{u}\|.$$

Como $\sum_{k=0}^{\infty} \|\delta^k\| < +\infty$, $\{x^k\}$ e $\{y^k\}$ são limitadas, $\sum_{k=0}^{\infty} \|\nu^k\| < +\infty$ e $\sum_{k=0}^{\infty} \lambda_k^{-1} \|\delta^k\| < +\infty$, obtemos $\sum_{k=0}^{\infty} \overline{\beta}_k < +\infty$. Portanto,

$$H(\overline{y}, y^k) \leq H(\overline{y}, y^{k-1}) + \overline{\beta}_k \;\text{ com } \sum_{k=0}^{\infty} \overline{\beta}_k < +\infty.$$

ou seja, a sequência $\{y^k\}$ é quase-Fejér convergente para $\widehat{SOL}(T;C)$.

(ii) Defina $\overline{\sigma}_{k+1} := H(\overline{y}, y^k)$. Como as sequências $\{\overline{\beta}_k\}$ e $\{\overline{\sigma}_k\}$ satisfazem as condições exigidas pelo Lema 1.1, $\{H(\overline{y}, y^k)\}$ é convergente.

(iii) Dado $(x, u) \in G(T)$. De (4.19), (A5) e da desigualdade de Cauchy-Schwarz, obtemos

$$\begin{aligned} \langle x - x^k, u \rangle &\geq \lambda_k [H(y, y^k) - H(y, y^{k-1}) - \langle \delta^k, \nabla_1 d(y^k, y^{k-1}) \rangle - \nu_k \langle x - x^k, x^k - x^{k-1} \rangle] \\ &\geq \lambda_k \left[-H(y, y^{k-1}) - L\|\delta^k\| \|y^k - y^{k-1}\| - \nu_k \|x - x^k\| \|x^k - x^{k-1}\| \right]. \end{aligned} \tag{4.21}$$

Como $\lambda_k \to 0$, $H(y,\cdot)$ é nível limitada, $\sum_{k=0}^{\infty}\delta^k < +\infty$, $\{x^k\}$ e $\{y^k\}$ são limitadas e $\sum_{k=0}^{\infty}\nu_k < +\infty$, tomando o limite em (4.21) com $k \to +\infty$, obtemos

$$\lim_{k\to+\infty} \inf\langle x - x^k, u\rangle \geq 0. \tag{4.22}$$

Por outro lado, como $\{x^k\}$ é limitada, seja $\{x^{k_j}\} \subset \{x^k\}$ uma subsequência convergente para algum $\widehat{x}$. Então, por (4.22), obtemos

$$\langle x - \widehat{x}, u\rangle = \lim_{j\to+\infty} \langle x - x^{k_j}, u\rangle \geq \lim_{k\to+\infty} \inf\langle x - x^k, u\rangle \geq 0.$$

Assim,

$$\langle x - \widehat{x}, u\rangle \geq 0 \ \forall (x,u) \in G(T). \tag{4.23}$$

Por definição $y^{k_j} = b - Ax^{k_j} + \delta^{k_j}$, com $y^{k_j} > 0$. Passando o limite quando $j \to \infty$ temos que $\widehat{y} = b - A\widehat{x}$ com $\widehat{y} \geq 0$. Portanto, $A\widehat{x} \leq b$, o que implica $\widehat{x} \in C$. Usando agora a definição de N_C, temos

$$\langle x - \widehat{x}, w\rangle \geq 0 \ \ \forall (x,w) \in G(N_C). \tag{4.24}$$

Combinando (4.23) e (4.24)

$$\langle x - \widehat{x}, u + w\rangle \geq 0 \ \ \forall (x, u + w) \in G(T + N_C).$$

Por (A2) e Proposição 1.5 obtemos que o operador $T + N_C$ é monótono maximal. Portanto $0 \in (T + N_C)(\widehat{x})$, isto é, $\widehat{x} \in SOL(T;C)$. Por outro lado, isto implica que $\widehat{y} \in \widehat{SOL}(T;C)$. Como a sequência $\{y^k\}$ é quase-Fejér convergente para $\widehat{SOL}$ e, além disso, possui um ponto de acumulação pertencente a este conjunto, concluimos pela Proposição 1.3 que $\{y^k\}$ converge para $\widehat{y}$, o que implica convergência da sequência $\{x^k\}$ para $\widehat{x} \in SOL(T;C)$. $\square$

Capítulo 5

Considerações Finais

Neste trabalho apresentamos quatro algoritmos. O primeiro, denotado Algoritmo 1, tem como finalidade resolver o problema de otimização quase-convexa com restrições lineares (P). Mostramos convergência global a um ponto do conjunto solução do problema proposto, quando a sequência dos parâmetros de regularização converge a zero. Esta condição pode ser substituida pela limitação da sequência dos parâmetros de regularização, quando substituimos a quase-convexidade da função objetivo no problema (P) por pseudo-convexidade.

Os demais algoritmos, denotados por Algoritmo 2, 3 e 4, respectivamente, têm como objetivo resolver o problema da desigualdade variacional com restrições lineares $PDV(T;C)$.

Mostramos que a sequência gerada pelo Algoitmo 2 converge para uma solução do $PDV(T;C)$, quando T é um operador pseudo-monótono. Além disso, no Teorema 3.4, sob a hipótese adicional (H8), obtevemos convergência a um ponto solução do $PDV(T;C)$, quando T é quase-monótono, que é menos restritivo do que a monotonicidade.

Enquanto no Algoritmo 2 trabalhamos com um relaxamento da monotonicidade do operador T, nos Algoritmos 3 e 4 exigimos T monótono maximal. Estes algoritmos podem ser inicializados a partir de um ponto arbitrário do $\mathbb{R}^n$, e são aplicáveis, mesmo se o interior topológico do conjunto viável for vazio. No Algoritmo 3, para provarmos a boa definição, e obtermos convergência global a um ponto solução do problema variacional, foi necessário assumir a coercividade do operador T (ou $\nabla_1 D$), ver (A1). No Algoritmo 4, usando uma regularização do tipo Tikhonov, mostramos a boa definição da sequência gerada por este algoritmo. Ademais, obtevemos também convergência global a uma solução do $PDV(T;C)$.

A seguir, mencionamos alguns pontos que, de um modo natural, darão continuidade ao nosso trabalho:

1. Para o problema de otimização quase-convexa:

(a) Método proximal inviável quase-convexa, com distância proximal generalizada;

(b) Taxa de convergência;

(c) Implementação numérica.

2. Problema da desigualdade variacional

(a) Generalização do conjunto viável;

(b) Método proximal inviável quase-monótono;

(c) Taxa de convergência;

(d) Implementação numérica.

Referências Bibliográficas

[1] ABDELLAH B. "An LQP method for pseudomonotone variational inequalities". J. Glob. Optim. **36** (2006); 351-363.

[2] AUSLENDER, A., 1976. *Optimisation. Méthodes Numériques*. Masson Inc. Paris.

[3] AUSLENDER, A. and TEBOULLE, M. Interior gradient and proximal methods for convex and conic optimization. SIAM J. Optim. **16** (2006); 693-725.

[4] AUSLENDER, A., TEBOULLE, M. and BEN-TIBA, S. "Interior proximal and multiplier methods based on second order homogeneous kernels". Math. Oper. Res. **24** (1999), 645-668.

[5] AUSLENDER, A., TEBOULLE, M. and BEN-TIBA, S. "A logarithmic-quadratic proximal method for variational inequalities". Comput. Optim. Appl. **12** (1999), 31-40.

[6] AUSLENDER, A. and TEBOULLE, M., 2003, *Asymptotic cones and functions in optimization and variational inequalities*, New York, Springer.

[7] AUSLENDER, A., TEBOULLE, M. "Interior gradient and epsilon-subgradient descent methods for constrained convex minimization". Math. Oper. Res. **29** (2004), 1-26.

[8] BAJONA-XANDREI C., MARTINEZ-LEGAZ, J. E., "Lower subdifferentiability in minimax fractional programming", Optimization. **45** (1999), 1-22.

[9] BARRON, E. N. and LIU, W.,"Calculus of Variation in L^{∞}", Appl. Math. Optim. **35** (1997), 237-263.

[10] BNOUHACHEM, A., "An LQP method for pseudomonotone variational inequalities". J. Glob. Optim. **36** (2006), 351-363.

[11] BREGMAN, L.M.,“The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. Comput. Math. Math. Phys. **7** (1967), 200-217.

[12] BREZIS, H., Equations et Inequations Nonlineaires dans les Espaces Vectoriels en Dualité . Ann. Inst. Fourier, **18** (1968), 115-175.

[13] BREZIS, H., “A image d'une somm d'opérateurs monotones et applications”. Isr. J. Math. **7** (1976), 165-186.

[14] BRITO, A.S., CRUZ NETO, J.X., LOPES, J.O and OLIVEIRA, P.R. “Interior proximal algorithm for quasiconvex programming problems and variational inequalities with linear constraints”. J. Optim. Theory Appl., published online: February 2012.

[15] BURACHIK, R.S., LOPES, J. O. and SILVA, G.J.P. “An inexact interior proximal method for the variational inequality problem”. Comp. Appl. Math. **28** (2009), 1-22.

[16] BURACHIK, R.S. e IUSEM, A.N., “A generalized proximal point algorithm for the variational inequality problem in a Hilbert Space‘”, SIAM J. Optim. **8** (1998), 197-216.

[17] BURACHIK, R.S. and SVAITER. B.F., “A relative error tolerance for a family of generalized proximal point methods”. Math. Oper. Res. **26** (2001), 816-831.

[18] BURACHIK, R.S. and SCHEIMBERG, S., “A proximal point method for the variational inequality problem in banach spaces”. SIAM J. Cont. Optim. **39** (2001), 1633-1649.

[19] CENSOR, Y. and ZENIOS, S.A. “The proximal minimization algorithm with D-functions”. J. Optim. Theory Appl. **73** (1992), 451-464.

[20] CHEN, J-S. and PAN, S. “A proximal-like algorithm for a class of nonconvex programming”. Pacific J. Optim. **4** (2008), 319-333.

[21] CHEN,G., TEBOULLE, M., “Convergence analysis of a proximal-like minimization algorithm using bregman function”. SIAM J. Optim. **3** (1993), 538-543.

[22] CLARKE, F.H., 1983, *Optimization and Nonsmooth Analysis*. New York, Wiley Intercience.

[23] CROUZEIX, J-P., Generalized convexity and generalized monotonicity, Monografias del IMCA 17, Peru.

[24] CROUZEIX, J.-P., MARTINEZ-LEGAZ J.-E. and VOLLE, M. Generalized convexity, generalized monotonicity. Non Convex optimization and its applications. Kluwer Academic Publishers; 1998.

[25] CUNHA, F.G.M., CRUZ NETO, J.X. da, and OLIVEIRA, P.R. "A proximal point algorithm with φ- divergence to quasiconvex programming". Optimization.**5** (2010), 777-792.

[26] DANILLIDIS, A., HADJISAVVAS, N., "Characterization of nonsmooth semistrictly quasiconvex and strictly quasiconvex functions". J. Optim. Theory Appl. **102** (1999), 525-536.

[27] ECKSTEIN, J., "Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming". Math. Oper. Res. **18** (1993), 202-226.

[28] EL FAROUQ, N., "Pseudomonotone variational inequalities: Convergence of the auxiliar problem method". J. Optim. Theory Appl. **2** (2001), 305-326.

[29] FACCHINE, F., PANG, J.-S., 2003, *Finite-dimensional variational inequality and complementarity problems*. New York, Springer.

[30] FRANCU, J., "Monotone operators. A survey directed to applications to differential equations". Appl. Mathem. **35** (1990), 257-301.

[31] HOFFMAN, A. J., "On approximate solutions of system of linear inequalities. J. National Bureau of Standards, **49** (1952), 263-265.

[32] GROMICHO, J. Quasiconvex optimization and location theory. Kluwer Academic Publishers: Dordrecht, The Netherlands; 1998.

[33] GROETSCH, C.W., The Theory of Tikhonov regularization for Fredholm Equations of the First Kind, Pitman, Boston; 1984.

[34] HADJISAVVAS, N., "Continuity and maximality properties of pseudomonote operators". J. Convex Anal. 10 (2003), 459-469.

[35] HARKER, P. T., PANG, J. S., "Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms, and applications". Math. Prog. **48** (1990), 161-220.

[36] KIWIEL, K. C., “Proximal minimization methods with generalized bregman function”. SIAM J. Cont. Optim. **35** (1997), 1142-1168.

[37] IUSEM, A.N., SVAITER, B. and TEBOULLE, M. “Entropy-like proximal methods in convex programming”. Math. Oper. Res. **19** (1994), 790-814.

[38] MANGASARIAN, O.L., 1969. *Nonlinear Programming*, McGraw-Hill Series in Systems Science.

[39] MARTINET, B. *Pertubation des méthodes d'opimisation*, R.A.I.R.O, Analyse de Mathematique de France, 93 (1978), 153-171.

[40] LANGENBERG, N. “Pseudomonotone operators and the bregman proximal point algorithm”. J. Glob. Optim. **47** (2010), 537-555.

[41] LOPES, J. O. 2002, Métodos de aproximação exterior para o problema da desigualdade variacional. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

[42] KARAMARDIAN, S., “Complementarity problems over cones with monotone and pseudomonotone maps”. J. Optim. Theory Appl. **18** (1976), 445-455.

[43] ORTEGA, J.M. and RHEINBOLDT, W.C., 1970. *Interative Solution on Nonlinear Equation in Several Variables*. Academic Press, New York.

[44] PAPA, Q. A. and OLIVEIRA, P.R., “Proximal point methods for quasiconvex and convex function with Bregman distances on Hadamard manifolds”. J. Convex Anal. **355** (2009), 469-478.

[45] PAPA, Q. A. and OLIVEIRA, P.R., “An extension of proximal methods for quasiconvex minimization on the nonnegative orthant”. Eur. J. Oper. Res. **216** (2012) 26-32.

[46] POLYAK, B. T., “Introduction to optimization”, Translations Series in Mathematics and Engineering, Optimiization Software Inc. Publications Division, New York, 1987, Translated from the Russian, With a foreword by Dimitri P. Bertsekas.

[47] ROCKAFELLAR, R.T.,1970. *Convex Analysis*. NJ, Princeton University Press, Princeton.

[48] ROCKAFELLAR, R.T. “On the maximality of sums of nonlinear monotone operators”. Trans. Amer. Math. Soc. **149** (1970), 75-88.

[49] ROCKAFELLAR, R.T. "Monotone operators and the proximal point algorithm". SIAM J. Cont. Optim. **14** (1976), 877-898.

[50] ROCKAFELLAR, R.T., WETS, R., 1998. *Variational Analysis*. Berlin, Springer.

[51] SOUZA, S.S., OLIVEIRA. P.R., CRUZ NETO, J. X. da. and SOUBEYRAN. A. "A proximal method with separable distances Bregman for quasiconvex minimization over the nonnegative orthant". Eur. J. Oper. Res. **2** (2010), 365-376.

[52] TEBOULLE, M. "Entropic proximal mapping with applications to nonlinear programming". Math. Oper. Res. **17** (1992), 670-690.

[53] TEBOULLE, M. "Convergence of proximal-like algorithms". SIAM J. Optim. **7** (1997), 1069-1083.

[54] TIKHONOV, A.N.and ARSENIN.V.Y. Solutions of Ill-Posed Problems, John Wiley and Sons, Washington, D.C. 1977, Translation editor Fritz John (English).

[55] YAMASHITA, N., KANZOW, C., MORIMOTO, T. and FUKUSHIMA, M. "An infeasible interior proximal method for convex programming problems with linear constraints". J. Non. Convex Anal. **2** (2001), 139-156.