

UMA CLASSE DE MÉTODOS INTERIOR-PROXIMAIS COM MÉTRICA
VARIÁVEL PARA PROBLEMAS CONVEXOS E UM ALGORITMO
INTERIOR-PROXIMAL COM DISTÂNCIA DE BREGMAN PARA
OTIMIZAÇÃO QUASE-CONVEXA.

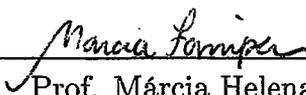
Sissy da Silva Souza

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



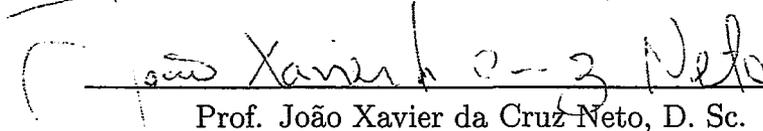
Prof. Paulo Roberto Oliveira, Dr. Ing.



Prof. Márcia Helena Costa Fampa, D. Sc.



Prof. José Mário Martinez, D. Sc.



Prof. João Xavier da Cruz Neto, D. Sc.



Prof. Carlile Campos Lavor, D. Sc.

RIO DE JANEIRO, RJ - BRASIL

JANEIRO DE 2007

SOUZA, SISSY DA SILVA

Uma classe de métodos interior-proximais com métrica variável para problemas convexos e um algoritmo interior-proximal com distância de Bregman para otimização quase-convexa [Rio de Janeiro] 2007

VIII, 53 p. 29,7 cm (COPPE/UFRJ, D.Sc., Engenharia de Sistemas e Computação, 2007)

Tese – Universidade Federal do Rio de Janeiro, COPPE

1 - Métodos Proximais

2 - Algoritmos de Pontos Interiores

3 - Programação Convexa

4 - Programação Quase-convexa

5 - Distâncias de Bregman

6 - Métrica Variável

I. COPPE/UFRJ II. Título (série)

"Tudo posso Naquele que me fortalece"
(Filipenses 4:13).

*À minha mãe Luzia,
às minhas irmãs Jaqueline e Anny,
aos meus sobrinhos Tazshida e Lucas.*

Agradecimentos

A Deus, por sua eterna proteção.

Ao Prof. Paulo Roberto Oliveira, pela orientação, incentivo e apoio durante todo o curso.

Ao prof. João Xavier, pela importante colaboração no desenvolvimento do trabalho.

Aos membros da banca examinadora, cujas observações, críticas e sugestões muito enriqueceram o trabalho.

A minha mãe Luzia, e minhas irmãs Anny e Jaquelline, pelo incentivo e apoio que me deram. Aos meus tios e primos, em especial Ana Zélia, Roberto e Eliana, que direta ou indiretamente, me ajudaram nesta caminhada.

À comunidade amazonense do PESC e colegas de Manaus: Roberto Cristóvão e Shirley, Flávia Morgana e Aldemir, Roberto Prata, Nilomar, Disney Douglas e Valtemir; e aos demais colegas do PESC: Paulo Sérgio, Gêvane, Felipe e Kely, Ronaldo, Sérgio, Valéria, Erik, Talita, Yuri, dentre tantos outros, com quem tive a satisfação de conviver, estudar, e trocar idéias importantes durante o curso. Aos demais amigos e familiares, pelo apoio direto e indireto durante toda a realização do doutorado.

Às secretárias, aos funcionários do suporte e aos demais servidores do PESC/COPPE/UFRJ por toda a atenção concedida.

A todas as pessoas cujos nomes não constam nesta lista, mas que de alguma forma contribuíram para a realização deste trabalho.

Ao CNPq, pelo apoio financeiro.

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

UMA CLASSE DE MÉTODOS INTERIOR-PROXIMAIS COM MÉTRICA
VARIÁVEL PARA PROBLEMAS CONVEXOS E UM ALGORITMO
INTERIOR-PROXIMAL COM DISTÂNCIA DE BREGMAN PARA
OTIMIZAÇÃO QUASE-CONVEXA

Sissy da Silva Souza

Janeiro/2007

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Neste trabalho, apresentamos métodos interior-proximais para resolver o problema de otimização com restrições de não-negatividade. Estudamos o caso convexo. Propomos um algoritmo com núcleo proximal definido por uma métrica variável e parâmetros de regularização convenientemente obtidos, procedemos à análise de convergência e apresentamos uma versão inexata para esta classe de métodos. Em seguida, analisamos o caso quase-convexo. Propomos um algoritmo cujo núcleo proximal é dado por uma distância de Bregman. Estabelecemos resultados de convergência da seqüência gerada pelo método.

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

A CLASS OF INTERIOR-PROXIMAL METHODS WITH VARIABLE METRIC
FOR CONVEX PROBLEMS AND AN INTERIOR-PROXIMAL ALGORITHM
WITH BREGMAN DISTANCE FOR QUASICONVEX OPTIMIZATION

Sissy da Silva Souza

January/2007

Advisor: Paulo Roberto Oliveira

Department: Systems and Computation Engineering

In this work, we present interior-proximal methods for solving the optimization problem with nonnegative constraints. We study the convex case. We propose an algorithm with proximal kernel defined by a variable metric and regularization parameters conveniently obtained, we proceed a convergence analysis and we present an inexact version for this class of methods. Next, we analyze the quasiconvex case. We propose an algorithm whose proximal kernel is given by a Bregman distance. We establish convergence results of the sequence generated by this method.

Sumário

Introdução	1
1 Preliminares	4
1.1 Operadores monótono maximais	4
1.2 Resultado de existência e definição de convergência fraca	7
1.3 Funções quase-convexas	8
1.4 Funções e distâncias de Bregman	9
2 Uma classe de métodos interior-proximais com métrica variável para problemas convexos	13
2.1 O Método IPMV	13
2.2 Existência e propriedades preliminares	15
2.3 Análise de convergência	17
3 Uma classe de métodos inexatos interior-proximais com métrica variável para problemas convexos	27
3.1 O Método IIPMV	27
3.2 Existência e propriedades preliminares	29
3.3 Análise de convergência	32
4 Um algoritmo interior-proximal com distância de Bregman para otimização quase-convexa	35
4.1 O Método IPDB	35
4.2 Existência e propriedades preliminares	37
4.3 Análise de convergência	40
Considerações finais	46
Referências Bibliográficas	49

Notações

\mathbb{R}_+^n	Pontos de \mathbb{R}^n cujas componentes são todas não-negativas.
\mathbb{R}_{++}^n	Interior de \mathbb{R}_+^n .
$\arg \min$	Conjunto dos pontos que realizam o mínimo.
$\mathcal{P}(X)$	Conjunto das partes de um conjunto X .
$\text{Im}(T)$	Imagem do operador T .
$\text{dom } T$	Domínio do operador T .
$Gr(T)$	Gráfico do operador T .
$\text{int } C$	Interior topológico do conjunto $C \subset \mathbb{R}^n$.
\overline{C}	Fecho topológico do conjunto $C \subset \mathbb{R}^n$.
$\mathcal{S}^*(P)$	Conjunto das soluções do problema (P).
$L_f(\alpha)$	Conjunto de nível de f em α .
$d(x, A)$	Distância de um ponto x a um conjunto A .
$B_X(K, r)$	Bola de raio r em volta de $K \subset X$.
$\nabla f(x)$	Gradiente de f em x .
$(\nabla f(x))_i$	i -ésima derivada parcial de f em x .
$\nabla_x f(x, y)$	Derivada parcial de f com respeito a x .
$\partial f(x)$	Subgradiente de f em x .
$\ \cdot\ $	Norma Euclideana.
$\ \cdot\ _X$	Norma definida por uma matriz X .
$x^r = (x_1^r, x_2^r, \dots, x_n^r)$	Vetor das r -potências das componentes de $x \in \mathbb{R}^n$.
X^r	Matriz diagonal cujos elementos são as componentes de x^r .
$\max(x^r)$	Maior componente de x^r .
$\min(x^r)$	Menor componente de x^r .
\mathcal{S}_{++}^n	Conjunto das matrizes simétricas definidas positivas.

Introdução

Neste trabalho apresentamos algoritmos para resolver o seguinte problema de otimização:

$$(P) \quad \min f(x) \\ \text{s.a } x \geq 0,$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria e contínua, e supomos que o conjunto solução de (P) é não-vazio. Estudamos os casos em que (P) é um problema convexo ou (P) é quase-convexo. Para isto, definimos métodos proximais.

Lembramos que o algoritmo de ponto proximal clássico, APP, para o problema de encontrar zeros de um operador T , veja [40, 49], é um método iterativo que inicia em um ponto $x^0 \in \mathbb{R}^n$ e gera sucessivamente uma seqüência de pontos x^{k+1} tais que $0 \in T_k(x)$, onde $T_k(x) = T(x) + \lambda_k(x - x^k)$ e $\{\lambda_k\}$ é uma seqüência de números reais positivos convenientemente escolhida. Observemos 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 distâncias de Bregman, ϕ -divergências e o método logarítmico-quadrático, veja por exemplo [2, 4, 5, 13, 25, 35, 37, 44, 46, 52, 55, 56] e suas referências.

O uso da métrica variável para o método proximal foi desenvolvido em [32] para o problema de minimização irrestrita, onde a função objetivo é suposta convexa, própria e fechada. Os autores mostraram que sob certas hipóteses a seqüência das iteradas é minimizante. Em [10] este método foi aprofundado. Os autores consideraram a métrica definida por uma matriz $M_k \in \mathcal{S}_{++}^n$ e $\sum_{k=1}^{+\infty} \frac{1}{\lambda_k} = +\infty$, e estabeleceram convergência global e taxa de convergência superlinear para a seqüência gerada pelo algoritmo proposto por eles.

A Geometria Riemanniana tem desempenhado um importante papel em programação matemática contínua, na análise da trajetória e convergência de algoritmos,

bem como na definição de novos métodos. A associação entre a métrica riemanniana e a direção de descida foi desenvolvida inicialmente em [39] e posteriormente em [29, 47]. Em [20] os autores exploraram esta relação e mostraram que a maioria dos métodos primais clássicos são métodos de gradiente em alguma métrica riemanniana apropriadamente escolhida. Por exemplo, com relação aos métodos multiplicativos de Eggermont, em [23], para resolver (P), os autores mostraram que a direção proposta representa o gradiente da função f relativo à métrica definida no interior do ortante positivo, dada pela hessiana da barreira entropia $h(x) = \sum_{i=1}^n x_i \log x_i$. Em [19] os autores desenvolveram, usando a métrica riemanniana, uma teoria de convergência semelhante aos clássicos da otimização irrestrita.

Algoritmos de pontos interiores têm sido desenvolvidos considerando \mathbb{R}_{++}^n com a métrica riemanniana dada por X^{-r} , com $r \geq 1$. Estas classes de métricas têm sido consideradas em algoritmos afim-escala primal e dual e métodos interior-proximais. Em [43] os autores generalizaram e analisaram o algoritmo proposto em [23]. Neste trabalho, os autores obtêm resultados de convergência fraca para o algoritmo apresentado em [23]. Em [42] os autores apresentaram uma nova classe de métodos interior-proximais, cujo núcleo é dado pela métrica riemanniana, para resolver (P) sob as condições de que a função objetivo é convexa, nível-limitada, com gradiente lipschitziano. Neste trabalho, os parâmetros de regularização são definidos explicitamente e convenientemente de forma a garantir a estrita viabilidade das iteradas; sua expressão depende da constante de Lipschitz do gradiente da função objetivo. Eles provam convergência fraca a um ponto KKT para funções convexas e convergência global no caso de funções estritamente convexas. Eles estabelecem também taxa de convergência linear quando a função é fortemente convexa no ótimo e sublinear caso contrário. Em [21] os autores desenvolveram uma nova classe de algoritmos proximais com a métrica riemanniana para o problema de complementaridade não-linear. Eles obtêm resultados de convergência para P_0 -funções.

Motivados por estes resultados obtidos com a métrica riemanniana, generalizamos o método apresentado em [42], propondo uma nova classe de algoritmos proximais de pontos interiores, com núcleo definido pela métrica variável riemanniana $(X^k)^{-r}$, com $r \geq 2$, para resolver o problema (P), sob hipóteses mais fracas que aquelas encontradas em [42], a saber, função objetivo convexa e nível-limitada, e conjunto solução não-vazio. Estabelecemos convergência fraca a um ponto KKT de (P), e

uma estimativa de convergência, de modo diferente à apresentada naquele trabalho. Apresentamos e analisamos uma versão inexata para o nosso algoritmo, adotando um critério sobre os erros similar ao apresentado em [42].

Problemas quase-convexos possuem um vasto campo de aplicações nas áreas de ciências e engenharias, tais como economia, teoria da locação, teoria do controle e teoria da aproximação, veja por exemplo [1, 6, 7, 30, 41]. Porém, ainda há poucos trabalhos voltados ao estudo de (P), quando a função objetivo é quase-convexa, veja [2, 17, 45, 46].

No recente trabalho apresentado em [17] os autores estudaram o problema (P) para o caso quase-convexo. Eles propuseram um método interior-proximal com ϕ -divergência e obtiveram propriedades de convergência. Motivados por estes resultados e pelas vastas aplicações deste problema, propomos um algoritmo proximal de pontos interiores com distância de Bregman para resolver (P) sob a quase-convexidade da função objetivo. Analisamos o problema sob as mesmas hipóteses consideradas em [17] e obtemos resultados de convergência.

Esta tese está organizada como segue. No primeiro capítulo revisamos alguns conceitos e resultados básicos para o desenvolvimento deste trabalho. Nos três capítulos seguintes apresentamos nossa contribuição ao estudo do problema (P). No Capítulo 2 definimos o método IPMV para resolver (P) no caso convexo, o qual denotamos por (PC), e procedemos à análise de convergência. Apresentamos uma estimativa de convergência e a propriedade Féjer na métrica riemanniana. No Capítulo 3 propomos e analisamos uma versão inexata de IPMV, obtemos as mesmas propriedades de convergência que o caso exato. Estudamos no Capítulo 4 o caso quase-convexo, isto é, o problema (P) quando f é uma função quase-convexa, o qual denotamos por (PQC). Apresentamos um algoritmo e analisamos sua convergência. Os resultados destes três últimos capítulos estão divulgados, como relatório técnico, em [53, 54], e submetidos a publicação. Finalizamos com uma breve discussão sobre os métodos propostos, os resultados obtidos e nossas perspectivas de continuidade das pesquisas.

Capítulo 1

Preliminares

Neste capítulo revisamos alguns conceitos e propriedades básicas sobre a teoria de operadores monótono maximais, e sobre o operador subdiferencial de funções convexas. Apresentamos a definição de convergência fraca e um conhecido resultado de existência, que usaremos nos Capítulos 2 e 3. Finalizamos com a teoria de funções quase-convexas e distâncias de Bregman.

1.1 Operadores monótono maximais

Nesta seção apresentamos conceitos e propriedades sobre aplicações ponto-a-conjunto e operadores monótono-maximais, para mais detalhes veja [3, 50]. Iniciamos revisando conceitos e propriedades sobre aplicações ponto-a-conjunto.

Definição 1.1.1. *Sejam X e Y dois subconjuntos de \mathbb{R}^n . Uma aplicação ponto-a-conjunto $T : X \rightarrow \mathcal{P}(Y)$ é uma aplicação que a cada $x \in X$ associa um conjunto $T(x) \subset \mathcal{P}(Y)$, onde $\mathcal{P}(Y)$ é o conjunto das partes de Y .*

Observação 1.1.1. *É natural caracterizar uma aplicação ponto-a-conjunto por seu gráfico. O gráfico de T é definido por:*

$$\text{Gr}(T) := \{(x, y) : y \in T(x)\}.$$

Reciprocamente, todo subconjunto G de $X \times \mathcal{P}(Y)$ é o gráfico da aplicação ponto-a-conjunto T , unicamente definida por:

$$T(x) = \{y \in \mathcal{P}(Y) : (x, y) \in G\}.$$

Exemplo 1.1.1. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e própria (isto é, $\text{dom}f := \{x \in \mathbb{R}^n : f(x) < +\infty\} \neq \emptyset$). O subdiferencial de f pode ser visto como*

uma aplicação ponto-a-conjunto $\partial f : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$, que a cada $x \in \mathbb{R}^n$, associa o conjunto $\partial f(x)$.

Observação 1.1.2. Lembramos que:

(i) Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é convexa quando, para quaisquer $x, y \in \mathbb{R}^n$, e $\lambda \in [0, 1]$, temos:

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

(ii) ξ é dito um subgradiente de uma função convexa f , em um ponto x , se, e somente se,

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

O conjunto de todos os subgradientes de f em x , $\partial f(x)$, é chamado o subdiferencial de f em x .

Revisamos em seguida caracterizações de semi-continuidade inferior e superior de aplicações ponto-a-conjunto.

Definição 1.1.2. Uma aplicação ponto-a-conjunto T é semi-contínua superiormente em $\bar{x} \in \text{dom} T$ se, e somente se, para qualquer vizinhança $U \subset T(\bar{x})$,

$$\exists \eta > 0 \text{ tal que } \forall \tilde{x} \in B_X(\bar{x}, \eta), \quad T(\tilde{x}) \subset U.$$

Definição 1.1.3. Uma aplicação ponto-a-conjunto $T : X \rightarrow \mathcal{P}(Y)$ é chamada semi-contínua inferiormente em $\bar{x} \in \text{dom} T$ se, e somente se, para qualquer $\bar{y} \in T(\bar{x})$ e para qualquer seqüência $\{x^k\} \subset \text{dom} T$ convergindo para \bar{x} , existe uma seqüência $\{y^k\} \subset T(x^k)$ convergindo a \bar{y} .

Definição 1.1.4. Dizemos que uma aplicação ponto-a-conjunto $T : X \rightarrow \mathcal{P}(Y)$ é contínua em \bar{x} se é semi-contínua superiormente e semi-contínua inferiormente em \bar{x} .

Apresentamos abaixo, caracterização para duas propriedades de aplicação ponto-a-conjunto: fecho e limitação local.

Definição 1.1.5. Uma aplicação ponto-a-conjunto $T : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^m)$ é fechada em \bar{x} se, para quaisquer seqüências $\{x^k\} \subset \mathbb{R}^n$ e $\{y^k\} \subset \mathbb{R}^m$, com $x^k \rightarrow \bar{x}$, $y^k \rightarrow \bar{y}$, $y^k \in T(x^k)$, então $\bar{y} \in T(\bar{x})$.

Proposição 1.1.1. *Uma aplicação $T : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^m)$ é localmente limitada se, e somente se, $T(B)$ é limitada para todo conjunto limitado B . Isto é equivalente à propriedade que quando $y^k \in T(x^k)$ e a seqüência $\{x^k\} \subset \mathbb{R}^n$ é limitada, então a seqüência $\{y^k\}$ é limitada.*

Revisamos agora a definição de operador monótono maximal, e uma propriedade geométrica, que será usada na justificativa do critério de parada dos algoritmos apresentados nos Capítulos 2 e 3.

Definição 1.1.6. *Uma aplicação $T : \mathbb{R}^m \rightarrow \mathcal{P}(\mathbb{R}^m)$ é dita monótona se, para quaisquer $x, y \in \text{dom } T$, vale a desigualdade*

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

para todo $u \in T(x)$, $v \in T(y)$.

Definição 1.1.7. *Uma aplicação $T : \mathbb{R}^m \rightarrow \mathcal{P}(\mathbb{R}^m)$ é dita monótona maximal se seu gráfico não está inteiramente contido no gráfico de qualquer outro operador monótono, o que é equivalente a*

$$\text{Se } \langle u - v, x - y \rangle \geq 0, \quad \forall v \in T(y), y \in \text{dom } T \text{ então } u \in T(x) .$$

Observação 1.1.3. *A propriedade de maximalidade assume para operadores monótonos o mesmo papel que a continuidade assume para funções.*

Proposição 1.1.2. *Seja $T : \mathbb{R}^m \rightarrow \mathcal{P}(\mathbb{R}^m)$ um operador monótono maximal. Então*
(i) $T(x)$ é convexo e fechado para todo $x \in \text{dom } T$;
(ii) o gráfico de T é fechado.

Exemplo 1.1.2. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e semi-contínua inferiormente. Então $\partial f : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$ é monótono maximal.*

Subdiferencial de funções convexas

A seguir listamos algumas propriedades satisfeitas pela aplicação subdiferencial de funções convexas. Para mais detalhes, veja [48, 50].

Primeiramente revisamos um critério que garante quando $\partial f(x)$ é não-vazio e limitado.

Proposição 1.1.3. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e própria. Então $\partial f(x)$ é não-vazio e limitado se, e somente se, $x \in \text{int dom } f$.*

Em seguida apresentamos um resultado bem conhecido sobre o subgradiente da soma de funções convexas.

Proposição 1.1.4. *Sejam $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funções convexas e próprias. Então*

(i) $\partial(f_1 + f_2)(x) \supset \partial f_1(x) + \partial f_2(x)$.

(ii) *Se existe um ponto em $\text{dom } f_1 \cap \text{dom } f_2$, então*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x),$$

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

Veremos a seguir condições para que a subdiferencial de funções convexas seja uma aplicação fechada.

Proposição 1.1.5. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e própria. Então ∂f é semi-contínua superiormente e localmente limitada em todo ponto $x \in \text{int dom } f$. Além disso, se f é semi-contínua inferiormente então ∂f é fechada.*

Finalizamos apresentando um importante resultado para subgradientes em otimização convexa.

Proposição 1.1.6. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa e própria. Um ponto $x \in \mathbb{R}^n$ é um minimizador de f se $0 \in \partial f(x)$.*

1.2 Resultado de existência e definição de convergência fraca

Nesta seção apresentamos um resultado de existência e a definição de convergência fraca a um conjunto, que serão usados no Capítulo 2.

Proposição 1.2.1. [13, Lema 5] *Sejam $T : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$ um operador monótono maximal e ∂g o subdiferencial de uma função convexa, própria e fechada em \mathbb{R}^n , tais que $\text{dom } T \cap \text{int dom } g \neq \emptyset$. Se ∂g é sobrejetor, então $T + \partial g$ é sobrejetor.*

Consideramos a seguinte noção sobre convergência fraca.

Definição 1.2.1. [34, Definição 1] *Uma seqüência $\{x^k\} \subset \mathbb{R}^n$ é dita fracamente convergente a um conjunto $U \subset \mathbb{R}^n$ se:*

- (i) *A seqüência $\{x^k\}$ é limitada;*
- (ii) $\lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0$;
- (iii) *todo ponto limite de $\{x^k\}$ pertence a U .*

1.3 Funções quase-convexas

Nesta seção apresentamos a definição de funções quase-convexas e duas de suas caracterizações, uma geométrica e outra para o caso de função diferenciável. Para mais detalhes, veja [9, 16, 28].

Definição 1.3.1. *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita quase-convexa se, para cada x e $y \in \mathbb{R}^n$,*

$$f(tx + (1-t)y) \leq \max\{f(x), f(y)\} \quad \forall t \in (0, 1).$$

Observação 1.3.1. *É consequência imediata da definição de funções quase-convexas, que toda função convexa é quase-convexa e o domínio de uma função quase-convexa é convexo.*

Exemplo 1.3.1. *São quase-convexas as funções:*

1. *não decrescentes, $f_i : \mathbb{R} \rightarrow \mathbb{R}$, $i=1,2,3$, dadas por*

$$f_1(x) = x, \quad f_2(x) = \log x, \quad f_3(x) = x^3.$$

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

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

Observação 1.3.2. *Muitas propriedades de funções convexas não valem para funções quase-convexas. Por exemplo, a soma de funções quase-convexas não é em geral uma função quase-convexa. Isto pode ser ilustrado através da soma das funções $f(x) = x$ e*

$$g(x) = \begin{cases} g(x) = x^2 & \text{para } |x| \leq 1 \\ g(x) = 1 & \text{caso contrário} \end{cases}.$$

Em seguida apresentamos uma interpretação geométrica para quase-convexidade.

Proposição 1.3.1. [9, Teorema 3.5.2] *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é quase-convexa se, e somente se,*

$$L_f(\alpha) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\} \text{ é convexo } \forall \alpha \in \mathbb{R}.$$

Finalizamos apresentando a seguinte caracterização para funções quase-convexas diferenciáveis.

Teorema 1.3.1. [9, Teorema 3.5.4] *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável. Então, f é quase-convexa se, e somente se, vale o seguinte:*

$$\text{Se } x, y \in \mathbb{R}^n \text{ e } f(x) \leq f(y), \text{ então } \langle \nabla f(y), y - x \rangle \geq 0. \quad (1.1)$$

Observação 1.3.3. *A condição (1.1) dá uma interpretação geométrica simples quando $\nabla f(y) \neq 0$. Neste caso, $\nabla f(y)$ define um hiperplano suporte para o conjunto de nível $\{x \in \mathbb{R}^n : f(x) \leq f(y)\}$ no ponto y .*

1.4 Funções e distâncias de Bregman

Nesta seção apresentamos a definição de funções e distâncias de Bregman e suas propriedades usuais. Para mais detalhes, veja [8, 11, 15, 25, 52].

Sejam S um subconjunto convexo e aberto de \mathbb{R}^n e \bar{S} seu fecho. Consideramos uma função convexa $h : \bar{S} \rightarrow \mathbb{R}$. Seja $D_h : \bar{S} \times S \rightarrow \mathbb{R}$ tal que

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle. \quad (1.2)$$

Definição 1.4.1. [52, Definição 2.1] *A função h é chamada uma função de Bregman com zona S se:*

- (B1) *h é estritamente convexa e contínua sobre \bar{S} ;*
- (B2) *h é continuamente diferenciável sobre S ;*
- (B3) *Dados quaisquer $x \in \bar{S}$ 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.*

Observação 1.4.1. A definição clássica de função de Bregman, apresentada pela primeira vez em [15] e baseada em [11], requer duas condições adicionais:

(B3') O conjunto parcial de nível $L_{D_h}(\delta, y) = \{x \in \bar{S} : D_h(x, y) \leq \delta\}$ é limitado para todo $y \in S$.

(B5) Se $\{z^k\} \subset \bar{S}$ é limitada, $\{y^k\} \subset S$ converge para y , e $\lim_{k \rightarrow +\infty} D_h(z^k, y^k) = 0$, então $\lim_{k \rightarrow +\infty} z^k = y$.

No entanto, estas condições são redundantes. Em [52] os autores mostram que a condição (B5) é uma consequência da Definição 1.4.1. A restrição (B3') também é obtida facilmente de (B1)-(B4), como provado por exemplo em [8, 52],.

Observação 1.4.2. A condição (B5) é também conhecida como Consistência de Convergência ([52]).

Observação 1.4.3. Se h é uma função de Bregman então D_h é chamada a distância de Bregman associada a ela.

Da definição de D_h dada em (1.2), e da condição (B1), temos que, para todo $x \in \bar{S}$ e $y \in S$,

$$\begin{aligned} D_h(x, y) &\geq 0 && \text{e} \\ D_h(x, y) &= 0 && \Leftrightarrow x = y. \end{aligned}$$

Em seguida, apresentamos alguns exemplos de funções e distâncias de Bregman.

Exemplo 1.4.1.

1. Sejam $S = \mathbb{R}^n$ e $h(x) = \frac{1}{2}\|x\|^2$. Neste caso, $D_h(x, y) = \frac{1}{2}\|x - y\|^2$.

2. Sejam $S = \mathbb{R}_{++}^n$ e $h(x) = \sum_{i=1}^n x_i \log x_i$. h pode ser estendida continuamente para o bordo de \mathbb{R}_{++}^n quando convencionamos $0 \log 0 = 0$. Neste caso, $D_h(x, y) = \sum_{i=1}^n \left(x_i \log \frac{x_i}{y_i} + y_i - x_i \right)$, conhecida na literatura por função divergência Kullback-Leibler.

3. Sejam $S = \mathbb{R}_{++}^n$ e $h(x) = \sum_{i=1}^n (x_i^\alpha - x_i^\beta)$, com $\alpha \geq 1$ e $0 < \beta_k < 1$.

Para $\alpha = 2$, $\beta = \frac{1}{2}$, obtemos $D_h(x, y) = \|x - y\|^2 + \sum_{i=1}^n \frac{1}{2\sqrt{y_i}} (\sqrt{x_i} - \sqrt{y_i})^2$, e para $\alpha = 1$, $\beta = \frac{1}{2}$, temos $D_h(x, y) = \sum_{i=1}^n \frac{1}{2\sqrt{y_i}} (\sqrt{x_i} - \sqrt{y_i})^2$.

Apresentamos a seguir uma conhecida propriedade sobre distâncias de Bregman, que segue diretamente de (1.2) e das condições (B1)-(B4).

Proposição 1.4.1. *Seja h uma função de Bregman com zona S . Então,*

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

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

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

Consideramos a subclasse de funções de Bregman que satisfazem:

(B6) Para todo $y \in \mathbb{R}^n$ existe $x \in S$ tal que $\nabla h(x) = y$.

Observação 1.4.4. *A condição (B6) é chamada de zona coercividade ([13]).*

Exemplo 1.4.2. *Todas as funções do Exemplo 1.4.1 são zona coercivas, excetuando o caso (3) quando $\alpha = 1$. Se consideramos $S = \mathbb{R}_{++}^n$, a função h como definida no Exemplo 1.4.1 (1) não é zona coerciva.*

Finalizamos definindo funções de Bregman separáveis, apresentando exemplos, e listando propriedades necessárias ao estudo apresentado no Capítulo 4.

Definição 1.4.2. *Quando uma função de Bregman h tem a forma*

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

dizemos que h , e a sua distância de Bregman associada, D_h , são separáveis.

Exemplo 1.4.3.

1. Quando $h_i(t) = t^2 \forall i = 1, \dots, n$, então $h(x) = \sum_{i=1}^n h_i(x_i)$ é uma função de Bregman separável e $D_h(x, y)$ é a distância Euclidiana entre x e y .

2. Quando $h_i(t) = t \log t \forall i = 1, \dots, n$, então $h(x) = \sum_{i=1}^n h_i(x_i)$ é uma função de Bregman separável e $D_h(x, y)$ é a função divergência Kullback-Leibler.

Observação 1.4.5. Quando h é uma função de Bregman separável com zona $S = \mathbb{R}_{++}^n$, a condição (B6) significa que $h'_i(0, +\infty) = (-\infty, +\infty)$ (isto é, h'_i é sobrejetora em \mathbb{R}_{++}^n), e como h'_i é crescente (por (B1)), temos que $\lim_{t \rightarrow 0^+} h'_i(t) = -\infty$.

Capítulo 2

Uma classe de métodos interior-proximais com métrica variável para problemas convexos

Neste capítulo apresentamos uma família de métodos proximais de pontos interiores com métrica variável, para resolver problemas convexos sob restrições de não-negatividade. Propomos um algoritmo, denotado por IPMV (interior proximal com métrica variável), cujo núcleo proximal é uma métrica determinada por matrizes diagonais que são atualizadas a cada iteração. Os parâmetros de regularização são convenientemente definidos de forma a garantir que os pontos sejam estritamente viáveis. Mostramos que o método está bem definido, que a seqüência das iteradas funcionais decresce e converge, e outras propriedades da seqüência gerada pelo Algoritmo IPMV. Finalizamos estabelecendo convergência fraca ao conjunto solução do problema, uma estimativa de convergência e a propriedade Fejér.

2.1 O Método IPMV

Nesta seção definimos o problema estudado neste capítulo e o Algoritmo IPMV que propomos para resolvê-lo.

Consideramos o problema

$$\begin{aligned} \text{(PC)} \quad & \min f(x) \\ & \text{s.a } x \geq 0, \end{aligned}$$

onde

H1. $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função convexa, própria, contínua e nível-limitada;

H2. $\mathcal{S}^*(\text{PC})$, o conjunto solução de (PC), é não-vazio.

Para propor o método IPMV, escolhemos um número real $r \geq 2$, e uma seqüência de números reais positivos $\{\alpha_k\}$, tal que $0 < a < \alpha_k \leq 1 \quad \forall k \in \mathbb{N}$, para algum $0 < a < 1$.

Definimos o **Algoritmo IPMV** como segue:

Algoritmo IPMV

Inicialização: Escolha

$$(x^0, g^0) \in \mathbb{R}_{++}^n \times \mathbb{R}^n, g^0 \in \partial f(x^0) \quad (2.1)$$

Passo Principal: Dado $(x^k, g^k) \in \mathbb{R}_{++}^n \times \mathbb{R}^n, g^k \in \partial f(x^k)$, calcule

$$\beta_k = \|g^k\|(\max(x^k))^{r-1} + \alpha_k, \quad (2.2)$$

e encontre $(x^{k+1}, g^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^n, g^{k+1} \in \partial f(x^{k+1})$, satisfazendo

$$g^{k+1} + \beta_k (X^k)^{-r} (x^{k+1} - x^k) = 0. \quad (2.3)$$

Se $g^k = 0$ ou $x^{k+1} = x^k$ pare, senão

Faça $k = k + 1$.

Observação 2.1.1. O critério de parada do Algoritmo IPMV está bem definido. De fato, quando $g^k = 0$, temos imediatamente que $x^k > 0$ é uma solução do problema (PC); e quando $x^{k+1} = x^k$, por (2.3) obtemos $g^{k+1} = 0$, o qual por sua vez implica que $0 \in \partial f(x^k)$, donde $x^k \in \mathcal{S}^*(PC)$.

Observação 2.1.2. Pela Proposição 1.1.6, temos que resolver a equação dada em (2.3) é equivalente a encontrar x^{k+1} tal que $x^{k+1} = \arg \min_{x \in \mathbb{R}^n} \{f_k(x)\}$, onde

$$f_k(x) = f(x) + \frac{\beta_k}{2} \|x - x^k\|_{(X^k)^{-r}}^2,$$

lembrando que $\|x - x^k\|_{(X^k)^{-r}}^2 = \langle (X^k)^{-r} x - x^k, x - x^k \rangle$.

2.2 Existência e propriedades preliminares

Nesta seção estabelecemos a boa-definição da seqüência das iteradas e as propriedades básicas necessárias à análise de convergência.

Iniciamos com a prova da boa-definição do Algoritmo IPMV.

Proposição 2.2.1. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPMV está bem definida.*

Prova: Seja $h(x) = \frac{\beta_k}{2} \|x - x^k\|_{(X^k)^{-r}}^2$ e seu gradiente $\nabla h(x) = \beta_k (X^k)^{-r} (x - x^k)$. Seja $\tilde{T}(x) = \partial f(x) + \nabla h(x)$. Sabemos que ∂f é um operador monótono maximal, h é uma função própria, contínua e estritamente convexa e, ∇h é sobrejetor. Então, pela Proposição 1.2.1, \tilde{T} é também sobrejetor. Assim, x^{k+1} existe. Da estrita convexidade de h , temos que $\tilde{T} = \partial f + \nabla h$ é estritamente monótono, e portanto x^{k+1} obtido é único. A positividade de x^{k+1} será mostrada na Proposição 2.3.2. \square

O resultado abaixo é usado para provar a limitação da seqüência $\{x^{k+1}\}$.

Proposição 2.2.2. *Seja $\{x^{k+1}\}$ a seqüência gerada pelo Algoritmo IPMV. Então $\{f(x^{k+1})\}$ é decrescente e convergente.*

Prova: Da desigualdade do subgradiente para funções convexas temos

$$\langle g^{k+1}, x^k - x^{k+1} \rangle \leq f(x^k) - f(x^{k+1}).$$

Assim, substituindo a expressão de g^{k+1} , dada por (2.3), obtemos

$$\langle -\beta_k (X^k)^{-r} (x^{k+1} - x^k), x^k - x^{k+1} \rangle \leq f(x^k) - f(x^{k+1}),$$

isto é,

$$\beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 \leq f(x^k) - f(x^{k+1}), \quad (2.4)$$

donde

$$f(x^{k+1}) \leq f(x^k). \quad (2.5)$$

Logo $\{f(x^{k+1})\}$ é decrescente. Agora, por (2.5), temos que $\{f(x^{k+1})\}$ é limitada superiormente por $f(x^0)$, e como f atinge seu mínimo em \mathbb{R}_+^n (por H2), ela também é limitada inferiormente, concluindo a prova. \square

O seguinte lema é aplicado na prova da convergência a zero da diferença entre as iteradas consecutivas.

Lema 2.2.1. *Sejam $\{(x^{k+1}, g^{k+1})\}$ e $\{\beta_k\}$ as seqüências geradas pelo Algoritmo IPMV. Então:*

$$(i) \sum_{k=0}^{+\infty} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 < +\infty;$$

$$(ii) \sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 < +\infty;$$

$$(iii) \sum_{k=0}^{+\infty} \frac{1}{\beta_k} \|g^{k+1}\|_{(X^k)^r}^2 < +\infty.$$

Prova: Parte (i). Somando a desigualdade (2.4), de $k = 0$ até l , obtemos

$$\begin{aligned} \sum_{k=0}^l \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 &\leq \sum_{k=0}^l (f(x^k) - f(x^{k+1})) \\ &= f(x^0) - f(x^{l+1}) \\ &\leq f(x^0) - f(x^*), \end{aligned}$$

onde $x^* \in \mathcal{S}^*(PC)$. Portanto

$$\sum_{k=0}^{+\infty} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 < +\infty.$$

Parte (ii). Da definição de β_k e da limitação superior de α_k , obtemos

$$\begin{aligned} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 &= (\|g^k\|(\max(x^k))^{r-1} + \alpha_k) \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 \\ &\geq a \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2, \end{aligned}$$

e, conseqüentemente,

$$\begin{aligned} \sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 &\leq \frac{1}{a} \sum_{k=0}^{+\infty} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 \\ &< +\infty. \end{aligned}$$

Parte (iii). Conseqüência direta da parte (i). De fato, de (2.3) temos

$$x^{k+1} - x^k = -\frac{1}{\beta_k} (X^k)^r g^{k+1}, \quad (2.6)$$

logo

$$\begin{aligned} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 &= \beta_k \left\| \frac{1}{\beta_k} (X^k)^r g^{k+1} \right\|_{(X^k)^{-r}}^2 \\ &= \left\langle g^{k+1}, \frac{1}{\beta_k} (X^k)^r g^{k+1} \right\rangle \\ &= \frac{1}{\beta_k} \|g^{k+1}\|_{(X^k)^r}^2. \end{aligned}$$

Assim

$$\begin{aligned} \sum_{k=0}^{+\infty} \frac{1}{\beta_k} \|g^{k+1}\|_{(X^k)^r}^2 &= \sum_{k=0}^{+\infty} \beta_k \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 \\ &< +\infty. \end{aligned}$$

□

2.3 Análise de convergência

Nesta subseção provamos que todos os elementos da seqüência $\{x^{k+1}\}$ são estritamente viáveis. Estabelecemos convergência fraca de $\{x^{k+1}\}$ ao conjunto solução do problema (PC), no sentido da Definição 1.2.1 (dada no Capítulo 1). Finalizamos apresentando uma estimativa de convergência e a propriedade Féjer.

Nos três próximos resultados mostramos a limitação das seqüências $\{x^{k+1}\}$, $\{g^{k+1}\}$ e $\{\beta_k\}$ geradas pelo Método IPMV.

Proposição 2.3.1. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPMV é limitada.*

Prova: Pela Proposição 2.2.2(i) obtemos $f(x^{k+1}) \leq f(x^0) \quad \forall k \in \mathbb{N}$. Logo $x^{k+1} \in L_f(f(x^0)) \quad \forall k \in \mathbb{N}$, onde $L_f(f(x^0)) = \{x \in \mathbb{R}^n : f(x) \leq f(x^0)\}$. Sabemos que f é nível limitada (por H1); portanto $L_f(f(x^0))$ e conseqüentemente $\{x^{k+1}\}$ são limitados. □

Corolário 2.3.1. *A seqüência $\{g^{k+1}\}$ gerada pelo Algoritmo IPMV é limitada.*

Prova: Da Proposição 2.3.1 (acima), temos que $\{x^{k+1}\}$ é limitada. Logo, como ∂f é localmente limitada (pela Proposição 1.1.5), e por definição, $g^{k+1} \in \partial f(x^{k+1})$, segue como conseqüência da Proposição 1.1.1 que $\{g^{k+1}\}$ também é limitada. □

Corolário 2.3.2. *A seqüência $\{\beta_k\}$ gerada pelo Algoritmo IPMV é limitada.*

Prova: Conseqüência direta da definição de β_k , da limitação de α_k (tomada por hipótese), e dos resultados da Proposição 2.3.1 e do último corolário. □

A seguir estabelecemos que os pontos da seqüência $\{x^{k+1}\}$ são estritamente viáveis.

Proposição 2.3.2. *Sejam $\{(x^{k+1}, g^{k+1})\}$ e $\{\beta_k\}$ as seqüências geradas pelo IPMV. Então $x^{k+1} \in \mathbb{R}_{++}^n$.*

Prova: Por indução. É verdadeiro para x^0 devido a (2.1). Suponha que $x^k > 0$. Da monotonicidade de ∂f temos que

$$\langle g^{k+1} - g^k, x^k - x^{k+1} \rangle \leq 0. \quad (2.7)$$

Da expressão de g^{k+1} , dada por (2.3), aplicada em (2.7), resulta

$$\langle -\beta_k (X^k)^{-r} (x^{k+1} - x^k) - g^k, x^k - x^{k+1} \rangle \leq 0,$$

isto é,

$$\langle \beta_k (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^k \rangle \leq \langle g^k, x^k - x^{k+1} \rangle.$$

Substituindo $(X^k)^{-r} = (X^k)^{-1} (X^k)^{-r+2} (X^k)^{-1}$ na última desigualdade, obtemos

$$\beta_k \langle (X^k)^{-r+2} ((X^k)^{-1} (x^{k+1} - x^k)), (X^k)^{-1} (x^{k+1} - x^k) \rangle \leq \langle g^k, x^k - x^{k+1} \rangle. \quad (2.8)$$

Lembrando que $r \geq 2$, de (2.8) resulta

$$\beta_k (\max(x^k))^{-r+2} \|(X^k)^{-1} (x^{k+1} - x^k)\|^2 \leq \langle g^k, x^k - x^{k+1} \rangle. \quad (2.9)$$

Agora, como

$$\begin{aligned} \|(X^k)^{-1} (x^{k+1} - x^k)\|^2 &= \sum_{i=1}^n ((x_i^k)^{-1} (x_i^{k+1} - x_i^k))^2 \\ &= \sum_{i=1}^n (x_i^k)^{-2} (x_i^{k+1} - x_i^k)^2 \\ &\geq (\max(x^k))^{-2} \|x^{k+1} - x^k\|^2, \end{aligned} \quad (2.10)$$

segue que, devido a positividade dos elementos em (2.10),

$$\|(X^k)^{-1} (x^{k+1} - x^k)\| \geq (\max(x^k))^{-1} \|x^{k+1} - x^k\|. \quad (2.11)$$

Assim, multiplicando ambos os membros de (2.11) pela expressão positiva

$$\beta_k (\max(x^k))^{-r+2} \|(X^k)^{-1} (x^{k+1} - x^k)\|,$$

obtemos:

$$\beta_k (\max(x^k))^{-r+2} \|(X^k)^{-1} (x^{k+1} - x^k)\|^2 \geq \beta_k (\max(x^k))^{-r+1} \|(X^k)^{-1} (x^{k+1} - x^k)\| \|x^{k+1} - x^k\|. \quad (2.12)$$

De (2.9) e (2.12),

$$\beta_k (\max(x^k))^{-r+1} \|(X^k)^{-1} (x^{k+1} - x^k)\| \|x^{k+1} - x^k\| \leq \langle g^k, x^k - x^{k+1} \rangle.$$

Aplicando a desigualdade de Cauchy-Schwarz no lado direito da desigualdade acima, segue que

$$\beta_k(\max(x^k))^{-r+1}\|(X^k)^{-1}(x^{k+1} - x^k)\|\|x^{k+1} - x^k\| \leq \|g^k\|\|x^k - x^{k+1}\|.$$

Logo

$$\|(X^k)^{-1}(x^{k+1} - x^k)\| \leq \frac{\|g^k\|(\max(x^k))^{r-1}}{\beta_k}. \quad (2.13)$$

A definição de β_k (em (2.2)) e (2.13) implicam que

$$\|(X^k)^{-1}(x^{k+1} - x^k)\| \leq \frac{\|g^k\|(\max(x^k))^{r-1}}{\|g^k\|(\max(x^k))^{r-1} + \alpha_k} < 1, \quad (2.14)$$

onde a estrita desigualdade é assegurada pelo fato que $\alpha_k > 0 \forall k$.

Da desigualdade (2.14), resulta

$$\left| \frac{x_i^{k+1}}{x_i^k} - 1 \right| < 1 \quad \forall i = 1, \dots, n,$$

logo

$$0 < x_i^{k+1} < 2x_i^k \quad \forall i = 1, \dots, n.$$

Portanto $x^{k+1} \in \mathbb{R}_{++}^n$. □

O resultado seguinte é uma conseqüência da Proposição 2.3.1, do Lema 2.2.1(ii), e da positividade obtida na propriedade anterior.

Proposição 2.3.3. *Seja $\{x^{k+1}\}$ a seqüência gerada pelo Algoritmo IPMV. Então*

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty \text{ e, conseqüentemente, } \lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0.$$

Prova: Observemos que

$$\begin{aligned} \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 &= \langle (X^k)^{-r}(x^{k+1} - x^k), x^{k+1} - x^k \rangle \\ &= \sum_{i=1}^n (x_i^k)^{-r} (x_i^{k+1} - x_i^k)^2 \\ &\geq \min((x^k)^{-r}) \sum_{i=1}^n (x_i^{k+1} - x_i^k)^2 \\ &= (\max(x^k))^{-r} \|x^{k+1} - x^k\|^2. \end{aligned}$$

Logo

$$\|x^{k+1} - x^k\|^2 \leq (\max(x^k))^r \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2,$$

e

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 \leq \sum_{k=0}^{+\infty} (\max(x^k))^r \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2. \quad (2.15)$$

Como $\{x^k\}$ é limitada (pela Proposição 2.3.1) e a série $\sum \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2$ converge (assegurado pelo Lema 2.2.1(ii)), obtemos de (2.15) a convergência da série $\sum \|x^{k+1} - x^k\|^2$. \square

No lema seguinte mostramos que a seqüência $\{(x^{k+1}, g^{k+1})\}$ verifica a condição de complementaridade.

Lema 2.3.1. *Sejam $\{(x^{k+1}, g^{k+1})\}$ e $\{\beta_k\}$ as seqüências geradas pelo Algoritmo IPMV. Então todo ponto de acumulação (\bar{x}, \bar{g}) de $\{(x^{k+1}, g^{k+1})\}$ satisfaz:*

$$\bar{x}_i \cdot \bar{g}_i = 0 \quad \forall i = 1, \dots, n.$$

Prova: De (2.3) temos

$$\frac{1}{\beta_k} (X^k)^r g^{k+1} = x^k - x^{k+1}. \quad (2.16)$$

Agora, pela Proposição 2.3.3, passando ao limite em (2.16), resulta

$$\lim_{k \rightarrow +\infty} \frac{1}{\beta_k} (X^k)^r g^{k+1} = 0.$$

Visto que $\{\beta_k\}$ é limitada (pelo Corolário 2.3.2), a última desigualdade implica:

$$\lim_{k \rightarrow +\infty} \beta_k \cdot \left(\frac{1}{\beta_k} (X^k)^r g^{k+1} \right) = 0,$$

isto é,

$$\lim_{k \rightarrow +\infty} (X^k)^r g^{k+1} = 0. \quad (2.17)$$

Seja $\{(x^{k_j+1}, g^{k_j+1})\}$ uma subseqüência de $\{(x^{k+1}, g^{k+1})\}$ tal que

$$(x^{k_j+1}, g^{k_j+1}) \rightarrow (\bar{x}, \bar{g}). \quad (2.18)$$

Como a seqüência $\{x^{k+1} - x^k\}$ converge a zero, então o mesmo ocorre com a subseqüência $\{x^{k_j+1} - x^{k_j}\}$, ou seja,

$$\lim_{j \rightarrow +\infty} (x^{k_j+1} - x^{k_j}) = 0. \quad (2.19)$$

De (2.18) e (2.19), resulta

$$\bar{x} = \lim_{j \rightarrow +\infty} x^{k_j+1} = \lim_{j \rightarrow +\infty} x^{k_j}. \quad (2.20)$$

Como a seqüência $\{(X^k)^r g^{k+1}\}$ vai a zero, isto também ocorre com a subsequência $\{(X^{k_j})^r g^{k_j+1}\}$, isto é,

$$\lim_{j \rightarrow +\infty} (X^{k_j})^r g^{k_j+1} = 0. \quad (2.21)$$

Logo, de (2.18), (2.20) e (2.21),

$$0 = \lim_{j \rightarrow +\infty} (X^{k_j})^r g^{k_j+1} = \overline{X}^r \overline{g}. \quad (2.22)$$

Concluimos a prova observando que como $r \geq 2$,

$$\overline{X}^r \overline{g} = 0 \iff \overline{X} \overline{g} = 0 \iff \overline{x}_i \cdot \overline{g}_i = 0 \quad \forall i = 1, \dots, n. \quad (2.23)$$

□

A seguir provamos que todo ponto de acumulação de $\{(x^{k+1}, g^{k+1})\}$ tem todas as suas coordenadas não-negativas.

Lema 2.3.2. *Seja $\{(x^{k+1}, g^{k+1})\}$ a seqüência gerada pelo Algoritmo IPV. Então todo ponto limite $(\overline{x}, \overline{g})$ de $\{(x^{k+1}, g^{k+1})\}$ satisfaz*

$$(\overline{x}, \overline{g}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n. \quad (2.24)$$

Prova: Pela Proposição 2.3.2 temos que $\overline{x} \in \mathbb{R}_+^n$. Assim é suficiente mostrar que $\overline{g} \in \mathbb{R}_+^n$.

Seja $\{(x^{k_j+1}, g^{k_j+1})\}$ uma subsequência de $\{(x^{k+1}, g^{k+1})\}$ tal que $\lim_{j \rightarrow +\infty} (x^{k_j+1}, g^{k_j+1}) = (\overline{x}, \overline{g})$. Suponha por absurdo que existe $i \in \{1, \dots, n\}$ tal que

$$\overline{g}_i < 0.$$

Para tal i , temos

$$0 > \overline{g}_i = \lim_{j \rightarrow +\infty} g_i^{k_j+1} = - \lim_{j \rightarrow +\infty} \beta_{k_j} (x_i^{k_j})^{-r} (x_i^{k_j+1} - x_i^{k_j}).$$

Logo existe uma subsequência, cujos índices denotamos também por k_j , e existe $k_{j_0} \in \mathbb{N}$, tal que para todo $k_j \geq k_{j_0}$,

$$\beta_{k_j} (x_i^{k_j})^{-r} (x_i^{k_j+1} - x_i^{k_j}) > 0.$$

Como β_{k_j} e $(x_i^{k_j})^{-r}$ são positivos (para cada $k_j \in \mathbb{N}$), obtemos

$$x_i^{k_j+1} > x_i^{k_j} \quad \forall k_j \geq k_{j_0},$$

e, conseqüentemente,

$$x_i^{k_j+1} > x_i^{k_{j_0}} \quad \forall k_j \geq k_{j_0}.$$

Assim, passando ao limite quando $j \rightarrow +\infty$, resulta

$$\bar{x}_i = \lim_{j \rightarrow +\infty} x_i^{k_j+1} \geq x_i^{k_{j_0}} > 0,$$

isto é, $\bar{x}_i > 0$.

Por outro lado, pela hipótese considerada, isto é, $\bar{g}_i < 0$, segue que $\bar{x}_i \bar{g}_i < 0$, contradizendo à condição de complementaridade provada no Lema 2.3.1. A prova está completa. \square

Os Lemas 2.3.1 e 2.3.2 mostram que todo ponto limite \bar{x} de $\{x^{k+1}\}$ satisfaz as condições de KKT do problema (PC), a saber,

$$\bar{x} \geq 0, \quad \bar{g} \geq 0, \quad \bar{x}_i \cdot \bar{g}_i = 0 \quad \forall i = 1, \dots, n, \quad (2.25)$$

$$\bar{g} \in \partial f(\bar{x}).$$

Portanto, temos o seguinte resultado.

Proposição 2.3.4. *Todo ponto limite da seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPMV pertence ao conjunto solução $\mathcal{S}^*(PC)$.* \square

Estabelecemos a convergência fraca no seguinte resultado.

Teorema 2.3.1. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPMV converge fracamente ao conjunto solução $\mathcal{S}^*(PC)$.*

Prova: É uma conseqüência direta das Proposições 2.3.1, 2.3.3 e 2.3.4. \square

A seguir apresentamos uma propriedade visando obter uma estimativa de convergência para o método IPMV.

Da desigualdade do subgradiente para funções convexas temos

$$f(x) - f(x^{k+1}) \geq \langle g^{k+1}, x - x^{k+1} \rangle.$$

Substituindo, na desigualdade acima, a expressão de g^{k+1} dada por (2.3), obtemos

$$f(x) - f(x^{k+1}) \geq \beta_k \langle (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x \rangle,$$

isto é,

$$\frac{2}{\beta_k} f(x^{k+1}) \leq \frac{2}{\beta_k} f(x) - 2\langle (X^k)^{-r}(x^{k+1} - x^k), x^{k+1} - x \rangle. \quad (2.26)$$

Por outro lado, sabendo que

$$\begin{aligned} \|x - x^k\|_{(X^k)^{-r}}^2 &= \|(x^{k+1} - x^k) - (x^{k+1} - x)\|_{(X^k)^{-r}}^2 \\ &= \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2 + \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \\ &\quad - 2\langle (X^k)^{-r}(x^{k+1} - x^k), x^{k+1} - x \rangle, \end{aligned}$$

temos

$$-2\langle (X^k)^{-r}(x^{k+1} - x^k), x^{k+1} - x \rangle = \|x - x^k\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2. \quad (2.27)$$

Assim, por (2.26) e (2.27),

$$\frac{2}{\beta_k} f(x^{k+1}) \leq \frac{2}{\beta_k} f(x) + \|x - x^k\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x^k\|_{(X^k)^{-r}}^2. \quad (2.28)$$

Observação 2.3.1. *A desigualdade (2.28) pode também ser obtida quando aplicamos o Lema 6.2 [26] à métrica variável considerada neste capítulo.*

A seguir apresentamos uma estimativa de convergência usando uma idéia bem conhecida na literatura, veja por exemplo [2, 5, 26, 31, 55]. Precisamos inicialmente definir a seqüência $\sigma_k = \sum_{j=0}^{k-1} \frac{1}{\beta_j}$, onde convencionamos $\sigma_0 = 0$.

Proposição 2.3.5. *Sejam $\{(x^{k+1}, g^{k+1})\}$ e $\{\beta_k\}$ as seqüências geradas pelo Algoritmo IPMV. Então, para qualquer $x \in \mathbb{R}_+^n$, temos a seguinte estimativa de convergência:*

$$f(x^m) - f(x) \leq \max_{k=0, \dots, m-1} \{|\langle x^k - x, g^{k+1} \rangle|\}.$$

Prova: Do Lema 2.2.2(i) sabemos que

$$f(x^{k+1}) \leq f(x^k),$$

do qual, multiplicando σ_k a ambos lados, obtemos

$$\sigma_k f(x^{k+1}) \leq \sigma_k f(x^k). \quad (2.29)$$

Usando a definição de σ_k , podemos escrever

$$\sigma_{k+1} = \frac{1}{\beta_k} + \sigma_k. \quad (2.30)$$

De (2.30) em (2.29), resulta

$$\sigma_{k+1}f(x^{k+1}) - \sigma_k f(x^k) \leq \frac{1}{\beta_k} f(x^{k+1}). \quad (2.31)$$

Somando (2.31), de $k = 0$ até $m - 1$, segue que

$$\sigma_m f(x^m) \leq \sum_{k=0}^{m-1} \frac{1}{\beta_k} f(x^{k+1}). \quad (2.32)$$

Agora, somando (2.28), de $k = 0$ até $m - 1$, e usando a definição de σ_m , obtemos

$$\begin{aligned} \sum_{k=0}^{m-1} \frac{2}{\beta_k} f(x^{k+1}) &\leq 2\sigma_m f(x) + \sum_{k=0}^{m-1} \left(\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \right) \\ &\quad - \sum_{k=0}^{m-1} \|x^k - x^{k+1}\|_{(X^k)^{-r}}^2. \end{aligned}$$

Logo

$$\sum_{k=0}^{m-1} \frac{2}{\beta_k} f(x^{k+1}) \leq 2\sigma_m f(x) + \sum_{k=0}^{m-1} \left(\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \right). \quad (2.33)$$

As desigualdades (2.32) e (2.33) implicam

$$2\sigma_m (f(x^m) - f(x)) \leq \sum_{k=0}^{m-1} \left(\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \right). \quad (2.34)$$

Por outro lado, considerando a definição de x^{k+1} , dada por (2.3), temos

$$\begin{aligned} \|x^{k+1} - x\|_{(X^k)^{-r}}^2 &= \left\| x^k - \frac{1}{\beta_k} (X^k)^r g^{k+1} - x \right\|_{(X^k)^{-r}}^2 \\ &= \|x^k - x\|_{(X^k)^{-r}}^2 + \left\| \frac{1}{\beta_k} (X^k)^r g^{k+1} \right\|_{(X^k)^{-r}}^2 \\ &\quad - 2 \left\langle x^k - x, (X^k)^{-r} \frac{1}{\beta_k} (X^k)^r g^{k+1} \right\rangle, \end{aligned}$$

isto é,

$$\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 = 2 \left\langle x^k - x, \frac{1}{\beta_k} g^{k+1} \right\rangle - \left\| \frac{1}{\beta_k} (X^k)^r g^{k+1} \right\|_{(X^k)^{-r}}^2,$$

portanto

$$\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \leq 2 \left\langle x^k - x, \frac{1}{\beta_k} g^{k+1} \right\rangle,$$

e, conseqüentemente,

$$\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \leq 2 \left| \left\langle x^k - x, \frac{1}{\beta_k} g^{k+1} \right\rangle \right|.$$

Somando a última desigualdade, de $k = 0$ até $m - 1$, obtemos

$$\sum_{k=0}^{m-1} \left(\|x^k - x\|_{(X^k)^{-r}}^2 - \|x^{k+1} - x\|_{(X^k)^{-r}}^2 \right) \leq \sum_{k=0}^{m-1} \frac{2}{\beta_k} |\langle x^k - x, g^{k+1} \rangle|. \quad (2.35)$$

Então, de (2.34) e (2.35),

$$2\sigma_m(f(x^m) - f(x)) \leq \sum_{k=0}^{m-1} \frac{2}{\beta_k} |\langle x^k - x, g^{k+1} \rangle|.$$

Logo

$$2\sigma_m(f(x^m) - f(x)) \leq \max_{k=0, \dots, m-1} \{|\langle x^k - x, g^{k+1} \rangle|\} \sum_{k=0}^{m-1} \frac{2}{\beta_k}.$$

Da definição de σ_m e desigualdade acima, segue que

$$2\sigma_m(f(x^m) - f(x)) \leq 2\sigma_m \max_{k=0, \dots, m-1} \{|\langle x^k - x, g^{k+1} \rangle|\}.$$

Portanto

$$f(x^m) - f(x) \leq \max_{k=0, \dots, m-1} \{|\langle x^k - x, g^{k+1} \rangle|\}. \quad (2.36)$$

A prova está completa. \square

Finalmente apresentamos a propriedade Fejér (que implica a convergência forte para soluções interiores).

Proposição 2.3.6. *Seja $\{x^{k+1}\}$ gerada pelo Algoritmo IPMV. Então*

$$\|x^{k+1} - x^*\|_{(X^k)^{-r}} \leq \|x^k - x^*\|_{(X^k)^{-r}},$$

para todo $x^* \in \mathcal{S}^*(PC)$.

Prova: Seja x^* uma solução de (PC). Temos

$$\langle g^{k+1}, x^* - x^{k+1} \rangle \leq f(x^*) - f(x^{k+1}) \leq 0 \quad \forall k \in \mathbb{N}, \quad (2.37)$$

onde a primeira desigualdade é consequência da convexidade de f e a última segue da otimalidade de x^* . Substituindo a expressão de g^{k+1} , dada por (2.3), em (2.37), obtemos

$$\langle -\beta_k (X^k)^{-r} (x^{k+1} - x^k), x^* - x^{k+1} \rangle \leq 0.$$

Como $\beta_k > 0 \quad \forall k \in \mathbb{N}$, resulta

$$\langle (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^* \rangle \leq 0,$$

isto é,

$$\langle (X^k)^{-\frac{r}{2}} (x^{k+1} - x^k), (X^k)^{-\frac{r}{2}} (x^{k+1} - x^*) \rangle \leq 0. \quad (2.38)$$

Sabendo que dados $a, b \in \mathbb{R}^n$, vale

$$2\langle a, b \rangle = \|a\|^2 + \|b\|^2 - \|a - b\|^2,$$

substituímos

$$a = (X^k)^{-\frac{r}{2}} (x^{k+1} - x^*) \quad \text{e} \quad b = (X^k)^{-\frac{r}{2}} (x^{k+1} - x^k),$$

de onde obtemos

$$\begin{aligned} \langle (X^k)^{-\frac{r}{2}} (x^{k+1} - x^*), (X^k)^{-\frac{r}{2}} (x^{k+1} - x^k) \rangle &= \frac{1}{2} \|(X^k)^{-\frac{r}{2}} (x^{k+1} - x^*)\|^2 + \\ &+ \frac{1}{2} \|(X^k)^{-\frac{r}{2}} (x^{k+1} - x^k)\|^2 \\ &- \frac{1}{2} \|(X^k)^{-\frac{r}{2}} (x^k - x^*)\|^2, \end{aligned}$$

a qual, junto com (2.38), resulta

$$\|(X^k)^{-r/2} (x^{k+1} - x^*)\|^2 \leq \|(X^k)^{-r/2} (x^k - x^*)\|^2,$$

isto é,

$$\|x^{k+1} - x^*\|_{(X^k)^{-r}}^2 \leq \|x^k - x^*\|_{(X^k)^{-r}}^2,$$

concluindo a prova. □

No próximo capítulo apresentamos e analisamos uma versão inexata para o Algoritmo IPMV.

Capítulo 3

Uma classe de métodos inexatos interior-proximais com métrica variável para problemas convexos

Neste capítulo apresentamos uma versão inexata do Método IPMV, denotada por IIPMV, para resolver o problema (PC). Como no caso exato, os parâmetros de regularização são definidos de forma a garantir a estrita viabilidade dos pontos gerados. Mostramos que o algoritmo está bem definido. Provamos que a seqüência dos erros converge a zero. Finalizamos estabelecendo convergência fraca ao conjunto solução de (PC).

3.1 O Método IIPMV

Nesta seção definimos o Algoritmo IIPMV para resolver o problema (PC). Para facilitar a leitura, lembramos o problema estudado:

$$\begin{aligned} \text{(PC)} \quad & \min f(x) \\ & \text{s.a } x \geq 0, \end{aligned}$$

onde:

H1. $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função convexa, própria, contínua e nível-limitada;

H2. $\mathcal{S}^*(\text{PC})$, o conjunto solução de (PC), é não-vazio.

Escolhemos um número real $r \geq 2$. Definimos o **Algoritmo IIPMV** como segue.

Algoritmo IIPMV

Inicialização: *Escolha*

$$(x^0, g^0) \in \mathbb{R}_{++}^n \times \mathbb{R}^n, \quad g^0 \in \partial f(x^0) \quad (3.1)$$

Passo Principal: *Dado*

$$(x^k, g^k) \in \mathbb{R}_{++}^n \times \mathbb{R}^n, \quad g^k \in \partial f(x^k), \quad (3.2)$$

calcule

$$\beta_k = 2\|g^k\|(\max(x^k))^{r-1} + 2(\max(x^k))^r, \quad (3.3)$$

e encontre $(x^{k+1}, g^{k+1}, e^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^n$, $g^{k+1} \in \partial f(x^{k+1})$, *satisfazendo*

$$g^{k+1} + \beta_k (X^k)^{-r} (x^{k+1} - x^k) = e^{k+1}, \quad (3.4)$$

com

$$\|e^{k+1}\| \leq \min\{\|g^k\|, \|x^{k+1} - x^k\|\} \quad (3.5)$$

Se $g^k = 0$ *ou* $x^{k+1} = x^k$ *pare, senão*

Faça $k = k + 1$.

Observação 3.1.1. *O critério de parada do algoritmo está bem definido. De fato, quando* $g^k = 0$, *temos imediatamente que* $x^k > 0$ *é uma solução do problema (PC). Quando* $x^{k+1} = x^k$, *de (3.4) obtemos* $g^{k+1} = e^{k+1}$; *visto que* $\|e^{k+1}\| \leq \|x^{k+1} - x^k\| = 0$, *resulta* $e^{k+1} = 0$ *e, conseqüentemente,* $g^{k+1} = 0$. *Portanto,* $0 \in \partial f(x^k)$ *e assim* $0 < x^k \in \mathcal{S}^*(PC)$.

Observação 3.1.2. *De (3.4) e (3.5) temos que em cada passo o erro está definido com respeito à iterada corrente, que é desconhecida "a priori". Porém, há vários trabalhos na literatura cujos critérios sobre os erros também dependem da iterada atual, citamos como exemplo, [4, 22, 24, 33, 51]. Uma vantagem da condição (3.5) é esta ser mais construtiva que aquelas encontradas, por exemplo, em [4, 24], a saber:*

$$\sum \|e^{k+1}\| < +\infty \quad e \quad \sum |\langle e^{k+1}, x^{k+1} \rangle| < +\infty.$$

Alguns resultados podem ser obtidos através de simples adaptações do Método IPMV. Por isso, simplificaremos as provas quando possível.

3.2 Existência e propriedades preliminares

Nesta seção estabelecemos a boa-definição do Algoritmo IIPMV, ao mostrarmos existência, unicidade e positividade de x^{k+1} , e que o elemento e^{k+1} é obtido em um número finito de passos.

Proposição 3.2.1. *A seqüência $\{x^{k+1}, e^{k+1}\}$ gerada pelo Algoritmo IIPMV está bem definida.*

Prova: Dividimos a prova em três passos.

Passo 1. *Existe único x^{k+1} satisfazendo (3.4) $\forall k \geq -1$*

Por indução. Sejam $h(x) = \frac{\beta_k}{2} \|x - x^k\|_{(X^k)^{-r}}^2$ e seu gradiente $\nabla h(x) = \beta_k (X^k)^{-r} (x - x^k)$. Considere $\tilde{T}(x) = \partial f(x) + \nabla h(x)$. Como ∂f é um operador monótono maximal, h é uma função própria, contínua e estritamente convexa, e ∇h é sobrejetor, então, pela Proposição 1.2.1, \tilde{T} é sobrejetor. Assim, dados x^k , β_k e e^{k+1} temos que x^{k+1} existe. Da estrita convexidade de h , obtemos que $\tilde{T} = \partial f + \nabla h$ é estritamente monótono, logo x^{k+1} obtido é único.

Passo 2. *x^{k+1} satisfazendo (3.4) é estritamente viável para todo $k \geq -1$.*

Por indução. Pela inicialização do algoritmo temos que $x^0 > 0$. Suponha que $x^k > 0$. Da monotonicidade de ∂f , temos que

$$\langle g^{k+1} - g^k, x^k - x^{k+1} \rangle \leq 0. \quad (3.6)$$

Da expressão de g^{k+1} , dada por (3.4), aplicada em (3.6), resulta

$$\langle e^{k+1} - \beta_k (X^k)^{-r} (x^{k+1} - x^k), x^k - x^{k+1} \rangle \leq \langle g^k, x^k - x^{k+1} \rangle.$$

Prosseguindo analogamente como na prova da versão exata (veja Proposição 2.3.2), obtemos:

$$\|(X^k)^{-1} (x^{k+1} - x^k)\| \leq \frac{(\|g^k\| + \|e^{k+1}\|)(\max(x^k))^{r-1}}{\beta_k}.$$

Da definição de β_k em (3.3), da condição sobre os erros em (3.5), e da última desigualdade, resulta

$$\|(X^k)^{-1}(x^{k+1} - x^k)\| \leq \frac{2\|g^k\|(\max(x^k))^{r-1}}{2\|g^k\|(\max(x^k))^{r-1} + 2(\max(x^k))^r} < 1, \quad (3.7)$$

onde a desigualdade estrita é válida visto que $x_k > 0$. Portanto, de (3.7), temos

$$0 < x_i^{k+1} < 2x_i^k \quad \forall i = 1, \dots, n,$$

concluindo que $x^{k+1} \in \mathbb{R}_{++}^n$.

Passo 3. e^{k+1} é *finitamente obtido satisfazendo (3.5)*.

Supomos por absurdo que na k -ésima iteração isto não ocorre. Sendo assim, existe um número infinito de $l \in \mathbb{N}$ tais que

$$\|e^{k+l}\| > \min\{\|g^k\|, \|x^{k+l} - x^k\|\}. \quad (3.8)$$

Sem perda de generalidade, consideramos uma seqüência $\{e^{k+l}\} \subset \mathbb{R}_+^n$ satisfazendo

$\lim_{k \rightarrow +\infty} e^{k+l} = 0$. Então, passando ao limite em ambos os lados de (3.8) quando $l \rightarrow +\infty$, obtemos

$$\lim_{l \rightarrow +\infty} \min\{\|g^k\|, \|x^{k+l} - x^k\|\} \leq \lim_{l \rightarrow +\infty} \|e^{k+l}\| = 0,$$

a qual implica que

$$\|g^k\| = 0 \quad \text{ou} \quad \lim_{l \rightarrow +\infty} \|x^{k+l} - x^k\| = 0.$$

Quando $\|g^k\| = 0$ o algoritmo pára e $x^k > 0$ é uma solução para (PC). Por outro lado, supondo que $\lim_{l \rightarrow +\infty} \|x^{k+l} - x^k\| = 0$, temos, usando (3.4), que

$$\lim_{l \rightarrow +\infty} g^{k+l} = \lim_{l \rightarrow +\infty} e^{k+l} = 0 \quad (3.9)$$

e

$$\lim_{l \rightarrow +\infty} x^{k+l} = x^k. \quad (3.10)$$

Como ∂f é fechada (pela Proposição 1.1.5), segue de (3.9) e (3.10) que $0 \in \partial f(x^k)$. Logo $x^k > 0$ é uma solução para o problema (PC). Portanto, ou e^{k+1} é obtido finitamente, ou $x^k \in \mathcal{S}^*(PC)$, concluindo a prova. \square

Na próxima proposição nosso objetivo é mostrar que a seqüência das iteradas funcionais decresce e converge.

Proposição 3.2.2. *Seja $\{x^{k+1}\}$ a seqüência gerada pelo Algoritmo IIPVM. Então:*

- (i) $f(x^{k+1}) \leq f(x^k) \quad \forall k \geq 0$;
- (ii) a seqüência $\{f(x^{k+1})\}$ é convergente.

Prova: Parte (i). Da desigualdade do subgradiente para funções convexas temos

$$\langle g^{k+1}, x^k - x^{k+1} \rangle \leq f(x^k) - f(x^{k+1}).$$

Substituindo a expressão de g^{k+1} definida em (3.4) obtemos

$$\langle e^{k+1} - \beta_k (X^k)^{-r} (x^{k+1} - x^k), x^k - x^{k+1} \rangle \leq f(x^k) - f(x^{k+1}),$$

logo

$$\langle e^{k+1}, x^k - x^{k+1} \rangle + \langle \beta_k (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^k \rangle \leq f(x^k) - f(x^{k+1}).$$

Aplicando a desigualdade de Cauchy-Schwarz no primeiro termo do lado esquerdo da expressão acima resulta

$$- \|e^{k+1}\| \|x^k - x^{k+1}\| + \beta_k \langle (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^k \rangle \leq f(x^k) - f(x^{k+1}).$$

Considerando a condição sobre os erros (3.5), da última desigualdade segue que

$$- \|x^{k+1} - x^k\|^2 + \beta_k \langle (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^k \rangle \leq f(x^k) - f(x^{k+1}). \quad (3.11)$$

Lembrando que, por (2.10),

$$\langle (X^k)^{-r} (x^{k+1} - x^k), x^{k+1} - x^k \rangle \geq (\max(x^k))^{-r} \|x^{k+1} - x^k\|^2,$$

a desigualdade (3.11) implica

$$- \|x^{k+1} - x^k\|^2 + \beta_k (\max(x^k))^{-r} \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}).$$

Logo

$$(\beta_k (\max(x^k))^{-r} - 1) \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}),$$

isto é,

$$\frac{\beta_k - (\max(x^k))^r}{(\max(x^k))^r} \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}).$$

Da definição de β_k (veja (3.3)), e da desigualdade acima, obtemos

$$\frac{2\|g^k\|(\max(x^k))^{r-1} + 2(\max(x^k))^r - (\max(x^k))^r}{(\max(x^k))^r} \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}),$$

de onde

$$\left(\frac{2\|g^k\|}{\max(x^k)} + 1 \right) \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}).$$

Assim

$$\|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1}). \quad (3.12)$$

Portanto

$$f(x^{k+1}) \leq f(x^k).$$

Parte (ii). Conseqüência direta da parte (i). De fato, da primeira parte obtemos que $\{f(x^k)\}$ é decrescente, logo limitada superiormente por $f(x^0)$. Além disso, como f atinge mínimo em \mathbb{R}_+^n (por H2), ela é também limitada inferiormente, concluindo a prova. \square

3.3 Análise de convergência

Nesta seção estabelecemos convergência fraca de $\{x^{k+1}\}$ a um ponto KKT de (PC). Os três próximos resultados são dedicados à prova da limitação das seqüências $\{x^{k+1}\}$, $\{g^{k+1}\}$ e $\{\beta_k\}$ geradas pelo Algoritmo IIPVM.

Proposição 3.3.1. *A seqüência $\{x^{k+1}\}$ gerada por IIPMV é limitada.*

Prova: Conseqüência direta da nível-limitação de f (dada por H1) e do decréscimo de $f(x^k)$ (provado na Proposição 3.2.2(i)). \square

Corolário 3.3.1. *A seqüência $\{g^{k+1}\}$ gerada por IIPMV é limitada.*

Prova: Analogamente à prova do caso exato (veja Corolário 2.3.1), este resultado é uma conseqüência imediata da limitação de $\{x^{k+1}\}$, mostrada na proposição anterior, e do fato que ∂f é localmente limitada, pela Proposição 1.1.1. \square

Corolário 3.3.2. *A seqüência $\{\beta_k\}$ gerada pelo Algoritmo IIPMV é limitada.*

Prova: Imediato da definição de β_k e da limitação das seqüências $\{x^{k+1}\}$ e $\{g^{k+1}\}$, respectivamente provadas na Proposição 3.3.1 e no último corolário. \square

Em seguida mostramos que a seqüência da diferença das iteradas sucessivas vai a zero.

Proposição 3.3.2. *Seja $\{x^{k+1}\}$ a seqüência gerada pelo Algoritmo IIPMV. Então, $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$; e, conseqüentemente, $\lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0$.*

Prova: Somando (3.12), de $k = 0$ até l , obtemos

$$\begin{aligned} \sum_{k=0}^l \|x^{k+1} - x^k\|^2 &\leq \sum_{k=0}^l (f(x^k) - f(x^{k+1})) \\ &\leq f(x^0) - f(x^{l+1}) \\ &\leq f(x^0) - f(x^*), \end{aligned}$$

onde $x^* \in \mathcal{S}^*(PC)$.

Logo, a série $\sum \|x^{k+1} - x^k\|^2$ converge. \square

Estabelecemos a seguir uma importante propriedade satisfeita pela seqüência dos erros.

Corolário 3.3.3. *Seja $\{e^{k+1}\}$ gerada pelo Algoritmo IIPMV. Então*

$$\lim_{k \rightarrow +\infty} e^{k+1} = 0.$$

Prova: Imediato da condição sobre os erros dada em (3.5), e pela Proposição 3.3.2. \square

Nos três próximos resultados mostramos que todo ponto limite da seqüência $\{(x^{k+1}, g^{k+1})\}$ gerada pelo Algoritmo IIPMV verifica as condições de KKT.

Lema 3.3.1. *Sejam $\{x^{k+1}, g^{k+1}, e^{k+1}\}$ e $\{\beta_k\}$ as seqüências geradas pelo Algoritmo IIPMV. Então todo ponto limite (\bar{x}, \bar{g}) de $\{(x^{k+1}, g^{k+1})\}$ satisfaz:*

$$\bar{x}_i \cdot \bar{g}_i = 0 \quad \forall i = 1, \dots, n.$$

Prova: Da equação (3.4) segue que

$$\beta_k(x^{k+1} - x^k) = (X^k)^r(e^{k+1} - g^{k+1}),$$

isto é,

$$(X^k)^r g^{k+1} = (X^k)^r e^{k+1} - \beta_k(x^{k+1} - x^k). \quad (3.13)$$

Agora, como as seqüências $\{e^{k+1}\}$ e $\{x^{k+1} - x^k\}$ convergem para zero (respectivamente provadas em Corolário 3.3.3, e Proposição 3.3.2), e as seqüências $\{x^k\}$ e

$\{\beta_k\}$ são limitadas (respectivamente obtidas em Proposição 3.3.1 e Corolário 3.3.2), passando ao limite em (3.13), quando $k \rightarrow +\infty$, resulta

$$\lim_{k \rightarrow +\infty} (X^k)^r g^{k+1} = 0. \quad (3.14)$$

Seja $\{(x^{k_j+1}, g^{k_j+1})\}$ uma subsequência de $\{(x^{k+1}, g^{k+1})\}$ tal que

$$(x^{k_j+1}, g^{k_j+1}) \rightarrow (\bar{x}, \bar{g}). \quad (3.15)$$

Procedendo analogamente como na análise do caso exato (veja Lema 2.3.1), obtemos

$$0 = \lim_{j \rightarrow +\infty} (X^{k_j})^r g^{k_j+1} = \bar{X}^r \bar{g},$$

donde,

$$\bar{x}_i \cdot \bar{g}_i = 0 \quad \forall i = 1, \dots, n.$$

A prova está completa. \square

Lema 3.3.2. *Seja $\{(x^{k+1}, g^{k+1})\}$ a seqüência gerada pelo Algoritmo IIPMV. Então, todo ponto limite (\bar{x}, \bar{g}) de $\{(x^{k+1}, g^{k+1})\}$ satisfaz*

$$(\bar{x}, \bar{g}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n. \quad (3.16)$$

Prova: Semelhante aos passos da prova do Lema 2.3.2, apenas observando que a seqüência de erros vai a zero. \square

Veremos a seguir, como consequência dos últimos lemas, que todo ponto de acumulação da seqüência das iteradas é um ponto KKT de (PC).

Proposição 3.3.3. *Todo ponto limite da seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IIPMV pertence a $\mathcal{S}^*(PC)$.*

Prova: É suficiente mostrar que qualquer ponto limite \bar{x} de $\{x^{k+1}\}$ satisfaz as condições de KKT do problema (PC), isto é:

$$\bar{x} \geq 0, \quad \bar{g} \geq 0, \quad \bar{x}_i \cdot \bar{g}_i = 0 \quad \forall i = 1, \dots, n, \quad \bar{g} \in \partial f(\bar{x}), \quad (3.17)$$

o qual é garantido pelos Lemas 3.3.1 e 3.3.2, finalizando a prova. \square

Concluimos esta seção enunciando o principal resultado, consequência imediata das Proposições 3.3.1, 3.3.2 e 3.3.3.

Teorema 3.3.1. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IIPMV converge fracamente ao conjunto solução $\mathcal{S}^*(PC)$.* \square

Capítulo 4

Um algoritmo interior-proximal com distância de Bregman para otimização quase-convexa

Neste capítulo apresentamos um método proximal de pontos interiores com distância de Bregman, denotado por IPDB, para resolver o problema de minimizar uma função quase-convexa sob restrições de não-negatividade. Consideramos funções de Bregman separáveis, com a zona sendo o interior do ortante positivo. Estabelecemos a boa-definição da seqüência gerada por nosso método, e mostramos que o algoritmo não cicla. Finalmente provamos convergência a um ponto KKT do problema quando os parâmetros de regularização são apenas limitados, e, quando a seqüência dos parâmetros tende a zero, obtemos convergência a um ponto solução do problema estudado.

4.1 O Método IPDB

Nesta seção definimos o problema estudado e o Algoritmo IPDB que propomos para resolvê-lo.

Consideramos o problema

$$\begin{aligned} \text{(PQC)} \quad & \min f(x) \\ & \text{s.a } x \geq 0, \end{aligned}$$

A1. $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função quase-convexa, própria, e continuamente diferenciável.

A2. $\mathcal{S}^*(PQC)$ é não-vazio.

Para definir o método IPDB, escolhamos uma seqüência de números reais positivos $\{\beta_k\}$, satisfazendo:

R1. $0 < \beta_k \leq \bar{\beta}$, para algum $\bar{\beta} > 0$,

e consideramos $D_h(x, y)$ uma distância de Bregman associada a uma função de Bregman h , com zona $S = \mathbb{R}_{++}^n$, e tal que

R2. h é separável e zona coerciva.

Algoritmo IPDB

Inicialização: Seja $x^0 \in \mathbb{R}_{++}^n$

Passo principal: Dado $x^k \in \mathbb{R}_{++}^n$,

Se $\nabla f(x^k) = 0$ pare, senão

Encontre

$$x^{k+1} \in \{z \in \mathbb{R}_{++}^n : z \in \arg \min_{x \geq 0} \{f_k(x)\}\} \quad (4.1)$$

onde

$$f_k(x) = f(x) + \beta_k D_h(x, x^k). \quad (4.2)$$

Faça $k = k + 1$.

4.2 Existência e propriedades preliminares

Nesta seção estabelecemos a boa-definição do Métodoo IPDB. Provamos que o algoritmo não cicla. Finalizamos com algumas propriedades básicas necessárias à análise de convergência.

Proposição 4.2.1. *Existe $x^{k+1} > 0$ satisfazendo (4.1) $\forall k \geq -1$.*

Prova: Por indução. $x^0 > 0$ existe pela inicialização do algoritmo. Seja $x^k > 0$. Então, como f é limitada inferiormente (devido a hipótese A2) e $D_h(x, x^k)$ é nível limitada (veja Observação 1.4.1(1)), obtemos que f_k é também nível limitada. Agora, pela continuidade de f_k , segue que os conjuntos de nível de f_k são compactos. Logo existe $z \in \mathbb{R}_+^n$ que resolve o problema em (4.1), o qual pode não ser único devido à não-convexidade de f .

Agora vamos mostrar que $z \in \mathbb{R}_{++}^n$. Primeiro note que z satisfaz as condições de otimalidade, isto é:

$$z_i \geq 0, \quad (\nabla f_k(z))_i \geq 0, \quad z_i(\nabla f_k(z))_i = 0 \quad \forall i = 1, \dots, n, \quad (4.3)$$

onde

$$\nabla f_k(z) = \nabla f(z) + \beta_k \nabla_x D_h(z, x^k).$$

Pela Proposição 1.4.1(ii) sabemos que $\nabla_x D_h(x, x^k) = \nabla h(x) - \nabla h(x^k)$, e como h é separável, então

$$(\nabla_x D_h(x, x^k))_i = h'_i(x_i) - h'_i(x_i^k) \quad \forall i = 1, \dots, n.$$

Pela hipótese A1 temos que ∇f é contínua, e visto que $L_{f_k}(\alpha)$ é limitada e $z \in L_{f_k}(\alpha)$, segue que existe $c > 0$ tal que $(\nabla f(z))_i \leq c \quad \forall i = 1, \dots, n$. Assim

$$\begin{aligned} (\nabla f_k(z))_i &= (\nabla f(z))_i + \beta_k (\nabla_x D_h(x, x^k))_i \\ &= (\nabla f(z))_i + \beta_k (h'_i(x_i) - h'_i(x_i^k)) \\ &\leq c + \beta_k (h'_i(x_i) - h'_i(x_i^k)). \end{aligned} \quad (4.4)$$

Pelo fato de que h satisfaz R2, isto é, h é separável e zona coerciva, então, da Observação 1.4.5, resulta

$$\lim_{z_i \rightarrow 0^+} h'_i(z_i) = -\infty.$$

Logo, passando ao limite quando $z_i \rightarrow 0^+$ em (4.4), resulta

$$\lim_{z_i \rightarrow 0^+} (\nabla f_k(z))_i = -\infty,$$

o que contradiz a otimalidade de z (na segunda desigualdade em (4.3)). Assim não pode ocorrer $z_i = 0$. Então $z > 0$. Portanto existe $x^{k+1} > 0$ satisfazendo (4.1) $\forall k \geq -1$. \square

Observação 4.2.1. *Da última proposição, $z \in \operatorname{argmin}_{x \geq 0} \{f_k(x), z \geq 0\}$ é tal que $z > 0$. Logo $\nabla f_k(z) = 0$, e, conseqüentemente, $\nabla f(z) = -\beta_k \nabla_x D_h(z, x^k) \forall k \in \mathbb{N}$. Em particular,*

$$\nabla f(x^{k+1}) = -\beta_k \nabla_x D_h(x^{k+1}, x^k) \quad \forall k \in \mathbb{N}.$$

Em seguida provamos que o Algoritmo IPBD não gera pontos iguais entre si.

Proposição 4.2.2. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPBD não cicla.*

Prova: Dividimos a prova em dois passos.

Passo 1. *A seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD satisfaz $x^{k+1} \neq x^k \forall k \geq 0$. Seja $k \geq 0$. Suponha que $\nabla f(x^k) \neq 0$. Seja x^{k+1} tal que $x^{k+1} = x^k$. Da Observação 4.2.1, obtemos*

$$\nabla f(x^{k+1}) = -\beta_k \nabla_x D_h(x^{k+1}, x^k) = -\beta_k \nabla_x D_h(x^k, x^k) = 0.$$

Logo $\nabla f(x^k) = \nabla f(x^{k+1}) = 0$, o que contradiz a hipótese considerada. Portanto $x^{k+1} \neq x^k \forall k \geq 0$.

Passo 2. *A seqüência $\{x^{k+1}\}$ gerada pelo Algoritmo IPBD satisfaz $x^{k+1} \neq x^l$ para todo $k \geq 1$ e $l < k$.*

Da definição de x^{k+1} , dada em (4.1), temos

$$f(x^{k+1}) + \beta_k D_h(x^{k+1}, x^k) \leq f(x^k) + \beta_k D_h(x^k, x^k) = f(x^k). \quad (4.5)$$

Por outro lado, como $D_h(x^{k+1}, x^k) > 0 \forall k \geq 0$ segue que

$$f(x^{k+1}) + \beta_k D_h(x^{k+1}, x^k) > f(x^{k+1}). \quad (4.6)$$

Logo, de (4.5) e (4.6),

$$f(x^{k+1}) < f(x^k). \quad (4.7)$$

Então $\{f(x^k)\}$ é uma seqüência estritamente decrescente.

Suponha que $x^{k+1} = x^l$ para algum $k \geq 1$ e $l < k$. De (4.5) resulta

$$f(x^l) + \beta_k D_h(x^l, x^k) \leq f(x^k). \quad (4.8)$$

Visto que $x^l = x^{k+1} \neq x^k$, temos $D_h(x^l, x^k) > 0$. Assim, pela positividade de β_k , a desigualdade (4.8) implica

$$f(x^l) < f(x^k),$$

o que contradiz o fato de que $\{f(x^k)\}$ é estritamente decrescente. Portanto $x^{k+1} \neq x^l$ para todo $k \geq 1$ e $l < k$.

Então, dos resultados obtidos nos passos 1 e 2, a prova está completa. \square

Em seguida mostramos um importante resultado para nossa análise de convergência.

Proposição 4.2.3. *Seja $\{x^k\}$ a seqüência gerada pelo Algoritmo IPBD. Então,*

- (i) $0 < \beta_k D_h(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) \quad \forall k \in \mathbb{N}$;
- (ii) $\sum_{k=0}^{+\infty} \beta_k D_h(x^{k+1}, x^k) < +\infty$ e, conseqüentemente, $\lim_{k \rightarrow +\infty} \beta_k D_h(x^{k+1}, x^k) = 0$;
- (iii) $\{f(x^k)\}$ é uma seqüência estritamente decrescente e convergente.

Prova: Parte (i). Pela definição de x^{k+1} dada por IPBD em (4.1), temos

$$f(x^{k+1}) + \beta_k D_h(x^{k+1}, x^k) \leq f(x^k) + \beta_k D_h(x^k, x^k) = f(x^k) \quad \forall k \in \mathbb{N}.$$

Logo

$$\beta_k D_h(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}) \quad \forall k \in \mathbb{N}. \quad (4.9)$$

Por outro lado, como $x^{k+1} \neq x^k$ e $\beta_k > 0$, segue que

$$\beta_k D_h(x^{k+1}, x^k) > 0.$$

Assim, da última desigualdade e de (4.9), resulta a parte (i).

Parte (ii). Somando (4.9), de $k = 0$ até l , obtemos

$$\begin{aligned} \sum_{k=0}^l \beta_k D_h(x^{k+1}, x^k) &\leq \sum_{k=0}^l [f(x^k) - f(x^{k+1})] \\ &= f(x^0) - f(x^{l+1}) \\ &\leq f(x^0) - f(x^*), \end{aligned}$$

onde $x^* \in \mathcal{S}^*(PQC)$.

Pela parte (i) sabemos que $\beta_k D_h(x^{k+1}, x^k) > 0$. Logo $\sum_{k=0}^{\infty} \beta_k D_h(x^{k+1}, x^k) < \infty$.

Part (iii). Na prova do passo 2 da Proposição 4.2.2, vimos que $\{f(x^k)\}$ é estritamente decrescente. Logo $\{f(x^k)\}$ é limitada por cima por $f(x^0)$. Por outro lado, pela hipótese A2, temos que f atinge seu mínimo em \mathbb{R}_+^n e, portanto, é limitada superiormente, concluindo a prova. \square

4.3 Análise de convergência

Nesta seção mostramos a limitação da seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD. Estabelecemos convergência para um ponto KKT do problema (PQC) quando $\{\beta_k\}$ atende a R1, isto é, assumimos apenas limitação; e, provamos convergência para uma solução de (PQC) quando admitimos a seguinte hipótese adicional:

R3. $\lim_{k \rightarrow +\infty} \beta_k = 0$.

Usamos o seguinte resultado para mostrar a limitação da seqüência das iterações geradas por IPDB.

Lema 4.3.1. $D_h(x^*, x^{k+1}) \leq D_h(x^*, x^k) - D_h(x^{k+1}, x^k) \forall k \in \mathbb{N}$ e qualquer $x^* \in \mathcal{S}^*(PQC)$.

Prova: Seja $x^* \in \mathcal{S}^*(PQC)$. Fazendo $x = x^*$, $y = x^k$ e $z = x^{k+1}$ na Proposição 1.4.1(i), obtemos

$$D_h(x^*, x^k) - D_h(x^*, x^{k+1}) - D_h(x^{k+1}, x^k) = \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^{k+1} - x^* \rangle. \quad (4.10)$$

Da Observação 4.2.1 e da Proposição 1.4.1(ii), segue que

$$\nabla h(x^k) - \nabla h(x^{k+1}) = \frac{1}{\beta_k} \nabla f(x^{k+1}),$$

a qual, substituída em (4.10), resulta

$$D_h(x^*, x^k) - D_h(x^*, x^{k+1}) - D_h(x^{k+1}, x^k) = \frac{1}{\beta_k} \langle \nabla f(x^{k+1}), x^{k+1} - x^* \rangle. \quad (4.11)$$

Visto que $f(x^*) \leq f(x^{k+1}) \quad \forall k \in \mathbb{N}$, a quase-convexidade de f , pelo Teorema 1.3.1, implica que

$$\langle \nabla f(x^{k+1}), x^{k+1} - x^* \rangle \geq 0.$$

Logo, da última desigualdade e de (4.11), obtemos

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

A prova está completa. □

A seguir mostraremos que $\{x^k\}$ admite subsequência convergente.

Proposição 4.3.1. *A seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD é limitada.*

Prova: Pelo Lema 4.3.1 (acima) $D_h(x^*, x^{k+1}) \leq D_h(x^*, x^k)$. Logo

$$D_h(x^*, x^{k+1}) \leq D_h(x^*, x^0). \quad (4.12)$$

Agora, da condição (B3), temos que os conjuntos de nível de $D_h(x^*, \cdot)$ são limitados, implicando, portanto, a limitação de $\{x^{k+1}\}$. □

Nos três próximos resultados estabelecemos convergência da seqüência $\{x^k\}$.

Proposição 4.3.2. *A seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD é convergente.*

Prova: Seja \bar{x} um ponto de acumulação de $\{x^k\}$, cuja existência é assegurada pela Proposição 4.3.1 (acima). Da Proposição 4.2.3(iii), temos

$$f(x^{k+1}) < f(x^k) \quad \forall k \in \mathbb{N}.$$

Logo, pela continuidade de f , resulta

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

Analogamente ao desenvolvimento apresentado na prova do Lema 4.3.1, obtemos que $\{D_h(\bar{x}, x^k)\}$ é decrescente.

Agora seja $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim_{j \rightarrow +\infty} x^{k_j} = \bar{x}$. Então, da condição (B4), segue que

$$\lim_{j \rightarrow +\infty} D_h(\bar{x}, x^{k_j}) = 0.$$

Assim $\{D_h(\bar{x}, x^k)\}$ é uma seqüência decrescente, não-negativa, que possui uma subsequência convergindo a 0. Logo $\{D_h(\bar{x}, x^k)\}$ converge a 0, isto é,

$$\lim_{k \rightarrow +\infty} D_h(\bar{x}, x^k) = 0. \quad (4.13)$$

Seja $\{x^{k_l}\}$ uma outra subsequência de $\{x^k\}$ tal que $\lim_{l \rightarrow +\infty} x^{k_l} = \tilde{x}$. Então, por (4.13), temos

$$\lim_{l \rightarrow +\infty} D_h(\bar{x}, x^{k_l}) = 0.$$

Fazendo $y^k = x^{k_l}$, $y = \tilde{x}$ e $z^k = \bar{x}$ na Observação 1.4.1 (B5), resulta

$$\lim_{k \rightarrow +\infty} z^k = \tilde{x},$$

isto é, $\tilde{x} = \bar{x}$.

Portanto, $\{x^k\}$ é convergente, concluindo a prova. □

Teorema 4.3.1. *A seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD converge para um ponto KKT de (PQC).*

Prova: Da última proposição sabemos que $\{x^k\}$ é convergente. Seja $\bar{x} = \lim_{k \rightarrow +\infty} x^k$. Mostraremos que \bar{x} é um ponto KKT de (PQC), isto é,

$$\bar{x}_i \geq 0, \quad (\nabla f(\bar{x}))_i \geq 0 \quad \text{e} \quad \bar{x}_i (\nabla f(\bar{x}))_i = 0 \quad \forall i = 1, \dots, n. \quad (4.14)$$

Pela Proposição 4.2.1 a primeira condição em (4.14) é verdadeira. Para provar as outras duas condições de (4.14), consideramos os conjuntos:

$$I(\bar{x}) = \{i \in \{1, \dots, n\} : \bar{x}_i = 0\},$$

$$J(\bar{x}) = \{i \in \{1, \dots, n\} : \bar{x}_i > 0\},$$

e analisamos os casos quando $i \in I(\bar{x})$, e quando $i \in J(\bar{x})$.

[Se $i \in I(\bar{x})$.] Neste caso supomos por contradição que $(\nabla f(\bar{x}))_i < 0$. Visto que f é continuamente diferenciável (pela hipótese A1), temos que

$$\lim_{k \rightarrow +\infty} (\nabla f(x^{k+1}))_i = (\nabla f(\bar{x}))_i < 0,$$

o que, por sua vez, implica na existência de uma subsequência, também denotada por $\{x^k\}$, e um índice $k_0 \in \mathbb{N}$, tais que

$$(\nabla f(x^{k+1}))_i < 0 \quad \forall k \geq k_0. \quad (4.15)$$

Da Observação 4.2.1, da Proposição 1.4.1(ii) e do fato de que h é separável, resulta:

$$\begin{aligned} (\nabla f(x^{k+1}))_i &= -\beta_k (\nabla_x D_h(x^{k+1}, x^k))_i \\ &= \beta_k (h'_i(x_i^k) - h'_i(x_i^{k+1})) \quad \forall k \geq k_0. \end{aligned} \quad (4.16)$$

Visto que $\beta_k > 0$, então, de (4.15) e (4.16),

$$h'_i(x_i^k) - h'_i(x_i^{k+1}) < 0.$$

Pela condição (B1) temos que $h'_i(x_i)$ é crescente para todo $i = 1, \dots, n$. Logo, da última desigualdade, segue que

$$x_i^{k+1} > x_i^k \quad \forall k \geq k_0,$$

de onde

$$x_i^{k+1} > x_i^k > \dots > x_i^{k_0} \quad \forall k \geq k_0.$$

No entanto,

$$0 = \bar{x}_i = \lim_{k \rightarrow +\infty} x_i^k \geq x_i^{k_0} > 0,$$

o que é uma contradição.

Portanto $(\nabla f(\bar{x}))_i \geq 0 \quad \forall i \in I(\bar{x})$.

[Se $i \in J(\bar{x})$.] Neste caso, como $\lim_{k \rightarrow +\infty} x_i^k = \bar{x}_i > 0$, obtemos

$$\lim_{k \rightarrow +\infty} \frac{x_i^{k+1}}{x_i^k} = 1,$$

isto é,

$$\lim_{k \rightarrow +\infty} x_i^{k+1} = \lim_{k \rightarrow +\infty} x_i^k = \bar{x}_i.$$

Esta última igualdade e a continuidade de h'_i (assegurada pela condição (B2)) implicam que

$$\lim_{k \rightarrow +\infty} h'_i(x^k) - h'_i(x^{k+1}) = 0.$$

Logo, como β_k é limitada (pela hipótese R1),

$$(\nabla f(\bar{x}))_i = \lim_{k \rightarrow +\infty} (\nabla f(x^{k+1}))_i = \lim_{k \rightarrow +\infty} \beta_k \left(h'_i(x_i^k) - h'_i(x_i^{k+1}) \right) = 0,$$

onde a segunda desigualdade ocorre devido a (4.16).

Portanto $(\nabla f(\bar{x}))_i = 0 \quad \forall i \in J(\bar{x})$.

A prova está completa. □

Teorema 4.3.2. *Se $\{\beta_k\}$ satisfaz R3 então a seqüência $\{x^k\}$ gerada pelo Algoritmo IPBD converge para uma solução de (PQC).*

Prova: Pela definição de x^{k+1} , dada por IPBD em (4.1), temos

$$f(x^{k+1}) + \beta_k D_h(x^{k+1}, x^k) \leq f(x) + \beta_k D_h(x, x^k) \quad \forall x \in R_+^n.$$

Logo, podemos tomar $x = x^*$, donde

$$f(x^{k+1}) + \beta_k D_h(x^{k+1}, x^k) \leq f(x^*) + \beta_k D_h(x^*, x^k)$$

Da última proposição sabemos que $\{x^k\}$ é convergente. Seja $\bar{x} = \lim_{k \rightarrow +\infty} x^k$. Visto que $\lim_{k \rightarrow +\infty} \beta_k D_h(x^{k+1}, x^k) = 0$ pela Proposição 4.2.3(ii), f é contínua pela hipótese A1, a seqüência $\{D_h(x^*, x^k)\}$ é limitada (veja (4.12)), e $\lim_{k \rightarrow +\infty} \beta_k = 0$, passando ao limite quando $k \rightarrow +\infty$, obtemos

$$f(\bar{x}) \leq f(x^*), \tag{4.17}$$

donde temos imediatamente que $\bar{x} \in \mathcal{S}^*(PQC)$, concluindo a prova. □

Finalizamos o capítulo apresentando uma estimativa de convergência. Como no Capítulo 2 definimos $\sigma_m = \sum_{k=0}^{m-1} \frac{1}{\beta_k}$ e convencionamos $\sigma_0 = 0$.

Proposição 4.3.3. *Seja $\{x^k\}$ a seqüência gerada pelo Algoritmo IPDB. Então, para todo $x \in \mathbb{R}_+^n$, temos a seguinte estimativa de convergência:*

$$f(x^m) - f(x) < \frac{1}{\sigma_m} (D_h(x, x^0) - D_h(x, x^m)) - \min_{j=0, \dots, m-1} r^{j+1} (|x - x^{j+1}|)$$

onde $\lim_{v \rightarrow 0} \frac{r^{j+1}(v)}{|v|} = 0$, $\forall j = 0, \dots, m-1$.

Prova: Da Proposição 4.2.3(iii) temos que

$$f(x^{k+1}) < f(x^k) \quad \forall k \in \mathbb{N}.$$

Então, multiplicando ambos os membros da desigualdade acima por σ_k , e observando que $\sigma_{k+1} = \frac{1}{\beta_k} + \sigma_k$, obtemos

$$\sigma_{k+1} f(x^{k+1}) - \sigma_k f(x^k) < \frac{1}{\beta_k} f(x^{k+1}).$$

Logo, somando esta última desigualdade de $k = 0$ até $m-1$, resulta

$$\sigma_m f(x^m) < \sum_{k=0}^{m-1} \frac{1}{\beta_k} f(x^{k+1}). \quad (4.18)$$

Escrevendo a fórmula de Taylor para f segue que

$$f(x) = f(x^{k+1}) + \langle \nabla f(x^{k+1}), x - x^{k+1} \rangle + r^{k+1} (|x - x^{k+1}|), \quad (4.19)$$

onde $r^{k+1}(v)$ é tal que $\lim_{v \rightarrow 0} \frac{r^{k+1}(v)}{|v|} = 0$.

Considerando $x \in \mathbb{R}_+^n$ qualquer e fazendo $y = x^k$ e $z = x^{k+1}$ na Proposição 1.5.1(i), procedendo como na prova do Lema 4.3.1, obtemos:

$$\langle \nabla f(x^{k+1}), x^{k+1} - x \rangle = \beta_k [D_h(x, x^k) - D_h(x, x^{k+1}) - D_h(x^{k+1}, x^k)],$$

donde

$$\langle \nabla f(x^{k+1}), x^{k+1} - x \rangle \leq \beta_k [D_h(x, x^k) - D_h(x, x^{k+1})].$$

De (4.19) e da última desigualdade, resulta

$$f(x^{k+1}) - f(x) \leq \beta_k [D_h(x, x^k) - D_h(x, x^{k+1})] - r^{k+1} (|x - x^{k+1}|),$$

isto é,

$$\frac{1}{\beta_k} (f(x^{k+1}) - f(x)) \leq D_h(x, x^k) - D_h(x, x^{k+1}) - \frac{1}{\beta_k} r^{k+1} (|x - x^{k+1}|).$$

Somando a desigualdade acima, de $k = 0$ até $m - 1$, segue que

$$\sum_{k=0}^{m-1} \frac{1}{\beta_k} f(x^{k+1}) - \sigma_m f(x) \leq D_h(x, x^0) - D_h(x, x^m) - \sum_{k=0}^{m-1} \frac{1}{\beta_k} r^{k+1} (|x - x^{k+1}|). \quad (4.20)$$

De (4.18) e (4.20) obtemos

$$\sigma_m (f(x^m) - f(x)) < D_h(x, x^0) - D_h(x, x^m) - \sum_{k=0}^{m-1} \frac{1}{\beta_k} r^{k+1} (|x - x^{k+1}|). \quad (4.21)$$

Como

$$\min_{j=0, \dots, m-1} r^{j+1} (|x - x^{j+1}|) \leq r^{k+1} (|x - x^{k+1}|) \quad \forall k = 0, \dots, m-1, \quad (4.22)$$

e $\beta_k > 0$, então, somando (4.22), de $k = 0$ até $m - 1$, resulta

$$\sum_{k=0}^{m-1} \frac{1}{\beta_k} \min_{j=0, \dots, m-1} r^{j+1} (|x - x^{j+1}|) \leq \sum_{k=0}^{m-1} \frac{1}{\beta_k} r^{k+1} (|x - x^{k+1}|),$$

donde

$$\sigma_m \min_{k=0, \dots, m-1} r^{j+1} (|x - x^{k+1}|) \leq \sum_{k=0}^{m-1} \frac{1}{\beta_k} r^{k+1} (|x - x^{k+1}|). \quad (4.23)$$

Agora, das desigualdades (4.21) e (4.23),

$$\sigma_m (f(x^m) - f(x)) < D_h(x, x^0) - D_h(x, x^m) - \sigma_m \min_{j=0, \dots, m-1} r^{j+1} (|x - x^{j+1}|),$$

isto é,

$$f(x^m) - f(x) < \frac{1}{\sigma_m} (D_h(x, x^0) - D_h(x, x^m)) - \min_{j=0, \dots, m-1} r^{j+1} (|x - x^{j+1}|).$$

□

Considerações finais

Neste trabalho desenvolvemos métodos para resolver problemas de otimização sob restrição de positividade, denotado por (P).

Definimos dois algoritmos do tipo interior-proximal com métrica variável para o problema (P) com a função objetivo suposta convexa, própria, contínua e nível-limitada. O primeiro, denotado por IPMV, considera o passo proximal com uma métrica induzida pela matriz diagonal $(X^k)^{-r}$, com $r \geq 2$ dado, e os parâmetros de regularização obtidos explicitamente de forma a implicar a positividade dos iterados. Obtemos resultados clássicos como o decrescimento e a convergência da seqüência das iteradas funcionais. Na análise de convergência, mostramos que todo ponto limite de seqüência das iteradas é um ponto KKT de (P). O segundo algoritmo corresponde a uma versão inexata para o método IPMV, e é denotado por IIPMV. Na análise de convergência, assim como no caso exato, obtivemos convergência fraca ao conjunto solução do problema estudado.

Finalmente estudamos o problema (P) para o caso em que a função objetivo é quase-convexa, própria e continuamente diferenciável. O passo proximal é definido por uma distância de Bregman separável, associada a uma função de Bregman zona coerciva, cuja zona é o interior do ortante positivo. Obtivemos convergência a um ponto KKT de (P), supondo limitação dos parâmetros de regularização. Desde que a seqüência dos parâmetros convirja a zero, mostramos convergência a um ponto solução do problema estudado.

Como pesquisas futuras, mencionamos alguns pontos em que estamos particularmente interessados:

1. Sobre os Métodos IPMV e IIPMV:

- (i) Generalização de ambos os métodos para problemas quase-convexos e de desigualdade variacional;
- (ii) obter condições sobre o problema de forma a garantir convergência global;
- (iii) estudar a dependência do parâmetro r .

2. Sobre o Método IPDB:

- (i) Obter convergência a um ponto solução do problema sob hipótese mais fraca que A5;
- (ii) estender o método para problemas quase-convexos não diferenciáveis;
- (iii) estabelecer uma localização do ponto limite da seqüência das iteradas;
- (iv) estudar uma versão contínua para IPDB.

Referências Bibliográficas

- [1] Arrow, K. J. e Enthoven, A. C., Quasi-concave programming, *Econometrica* 29, 4, 779-800, 1961.
- [2] Attouch, H. e Teboulle, M., Regularized Lotka-Volterra dynamical system as continuous proximal-like method in optimization, *J. Optim. Theory Appl.* 121, 541-570, 2004.
- [3] Aubin, J-P., *Set Valued Analysis*, Birkhäuser, Berlim, 1990.
- [4] Auslender, A., Teboulle, M. e Ben-Tiba, S., A logarithmic-quadratic proximal method for variational inequalities, *Computational Optimization and Applications* 12, 31-40, 1999.
- [5] Auslender, A., Teboulle, M. e Ben-Tiba, S., Interior proximal and multiplier methods based on second order homogeneous kernels, *Math. Oper. Res.* 24, 645-668, 1999.
- [6] Bajona-Xandri, C. e Martinez-Legaz, J. E., Lower subdifferentiability in minimax fractional programming, *Optimization* 45, 1-12, 1999.
- [7] Barron, E.N. e Liu, W., Calculus of variation in L^∞ , *Applied Mathematics and Optimization* 35, 237-263, 1997.
- [8] Bauschke, H.H. e Borwein, J.M., Legendre functions and the method of random Bregman projections, *Journal of Convex Analysis* 4, 27-67, 1997.
- [9] Bazaraa, M.S., Sherali, H.D. e Shetty, C.M., *Nonlinear programming: theory and algorithms*, 2nd Edition, John Wiley and Sons Inc., New York, 1993.
- [10] Bonnans, J.F., Gilbert, J.C., Lemaréchal, C. e Sagastizábal, C.A., A family of variable metric proximal methods, *Math. Program.* 68, 15-47, 1995.

- [11] Bregman, L.M., The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Computational Mathematics and Mathematical Physics* 7, 200-217, 1967.
- [12] Burachik, R.S. e Iusem, A.N., A generalized proximal point algorithm for the nonlinear complementarity problem, *RAIRO Oper. Res.* 33, 447-479, 1999.
- [13] 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, 1, 197-216, 1998.
- [14] Byrne, C. e Censor, Y., Proximity functions minimization using multiple Bregman projections, with applications to split feasibility and Kullback-Leibler distance minimization, *Annal of Operations Research* 105, 77-98, 2001.
- [15] Censor, Y. e Lent, A., An iterative row-action method for interval convex programming, *JOTA* 34, 321-353, 1981.
- [16] Crouzeix, J-P., *Generalized convexity and generalized monotonicity*, Monografias del IMCA 17, Peru.
- [17] Cruz Neto, J.X. da, Cunha, F.G.M., e Oliveira, P.R., A proximal point algorithm with φ -divergence to quasiconvex programming, Preprint, 2006.
- [18] Cruz Neto, J.X. da, Cunha, F.G.M., Oliveira, P.R. e Pinto, A.W.M, Generalization of the primal and dual affine scaling algorithms, ES 675/05, 2005, PESC/COPPE, Federal University of Rio de Janeiro.
- [19] Cruz Neto, J.X. da, Lima, L.L. e Oliveira, P.R., Geodesic algorithms in Riemannian manifolds, *Balkan Journal of Geometry and Its Applications* 3, 89-100, 1998.
- [20] Cruz Neto, J.X. da e Oliveira, P.R, Métodos geodésicos na programação matemática. Tese de doutorado, COPPE/UFRJ, Rio de Janeiro, 1995.
- [21] Da Silva, G.J.P. e Oliveira, P.R., A new class of proximal algorithms for the nonlinear complementarity problem, *Optimization and Control With Applications*, L. Qi, K. Teo e X. Yang (Eds.), Applied Optimization Series 96, 549-561, Springer, 2005.

- [22] Dong, Y., A new relative error criterion for the proximal point algorithm, *Non-linear Analysis* 64, 2143-2148, 2006.
- [23] Eggermont, P.P.B., Multiplicative iterative algorithms for convex programming, *Linear Algebra Appl.* 130, 25-42, 1990.
- [24] Eckstein, J., Approximate iterations in Bregman-function-based proximal algorithms, *Math. Program.* 83, 113-123, 1998.
- [25] Eckstein, J., Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, *Math. Oper. Res.* 18, 1, 202-226, 1993.
- [26] Ferreira, O.P. e Oliveira, P.R., Proximal point algorithm on Riemannian manifolds, *Optimization* 51, 257-270, 2002.
- [27] Ferreira, O.P. e Oliveira, P.R., Subgradient algorithm on Riemannian manifolds, *J. Optim. Theory Appl.* 97, 93-104, 1998.
- [28] Frenk, J.B.G. e Kassay, G., Introduction to Convex and Quasiconvex Analysis, ERIM Report Series Reference No. ERS-2004-075-LIS, August 2004.
- [29] Gabay, D., Minimizing a differentiable function over a differentiable manifold, *J. Optim. Theory Appl.* 37, 2, 1982.
- [30] Gromicho, J., *Quasiconvex optimization and location theory*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
- [31] Güler, O., On the convergence of the proximal point algorithm for convex minimization, *SIAM J. Control Optim.* 29, 403-919, 1991.
- [32] Hiriart-Urruty, J.B. e Lemaréchal, C., *Convex analysis in minimization algorithms*, vols. 1-2, Springer, New York, 1993.
- [33] Humes Júnior, C. e Silva, P.J.S., Inexact proximal point algorithms and descent methods in optimization, *Optim. Eng.* 6, 257-271, 2005.
- [34] Iusem, A.N., An interior multiplicative method for optimization under positivity constraints, *Acta Appl. Math.* 38, 163-184, 1995.

- [35] Iusem, A. N., Svaiter, B. F. e Teboulle, M., Entropy like proximal methods in convex programming, *Mathematics of Operations Research* 19, 790-814, 1994.
- [36] Kaplan, A. e Tichatschke, R., Convergence analysis of non-quadratic proximal methods for variational inequalities in Hilbert spaces, *J. Global Optim.* 22, 119-136, 2002.
- [37] Kiwiel, K.C., Proximal minimization methods with generalized Bregman functions, *SIAM Journal on Control and Optimization* 35, 1142-1168, 1997.
- [38] Lemaire, B., The proximal point, in *International Series of Numerical Mathematics*, J.P. Penot, Ed. Birkhauser, Basel 87, 73-87, 1989.
- [39] Luenberger, D.G., The gradient projection methods along geodesics, *Management Sci.* 18, 620-631, 1972.
- [40] Martinet, B., Regularization d'inequations variationnelles par approximations successives, *Rev. Franc. Inform. Rech. Oper.* 4, 154-159, 1970.
- [41] Mas-Colell, A., Whinston, M. D. e Green, J.R. , Microeconomic theory, Oxford University Press, New York, NY, USA, 1995.
- [42] Oliveira, G.L. e Oliveira, P.R., A new class of proximal point methods with variable metric for optimization under positivity constraints, ES-570/02, 2002, PESC/COPPE, Federal University of Rio de Janeiro.
- [43] Pinto, A., Oliveira, P.R. e Cruz Neto, J.X. da, A new class of potential affine algorithms for linear convex programming, ES-576/02, 2002, PESC/COPPE, Federal University of Rio de Janeiro.
- [44] Polyak, R. e Teboulle, M., Nonlinear rescaling and proximal-like methods in convex programming, *Math. Program.* 76, 265-284, 1997.
- [45] Quiroz, E.P. e Oliveira, P.R., Convergence of the Proximal Point Method for Quasiconvex Minimization, preprint 2006.
- [46] Quiroz, E.P. e Oliveira, P.R., Proximal Point Methods for Quasiconvex and Convex Functions with Bregman Distances on Hadamard Manifolds, to appear in *J. Convex Anal.*

- [47] Rapcsák, T. e Thang, T.T., A class of polynomial variable-metric algorithms for linear optimization, *Math. Program.* 74, 319-331, 1996.
- [48] Rockafellar, R.T., *Convex analysis*, Princeton University Press: New Jersey, 1970.
- [49] Rockafellar, R.T., Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14, 877-898, 1976.
- [50] Rockafellar, R.T. e Wets, R.J-B., *Variational analysis*, Springer-Verlag: New York, 1998.
- [51] Solodov, M.V. e Svaiter, B.F., A hybrid projection-proximal point algorithm, *J. Convex Anal.* 6, 59-70, 1999.
- [52] Solodov, M.V. e Svaiter, B.F., An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, *Math. Oper. Res.* 25, 2,214-230, 2000.
- [53] Souza, S. e Oliveira, P.R., A New Class of Interior Proximal Methods for Optimization over the Positive Orthant, *Optimization Online eprint*, 2006.
- [54] Souza, S., Oliveira, P.R., Cruz Neto, J.X. da, e Soubeyran, A., A Proximal Point Algorithm with Bregman Distances for Quasiconvex Optimization over the Positive Orthant, *Optimization Online eprint*, 2006.
- [55] Teboulle, M., Convergence of proximal-like algorithms, *SIAM J. Optim.* 7, 1069-1083, 1997.
- [56] Xu, Y., He, B. e Yuan, X., A hybrid inexact Logarithmic-Quadratic proximal method for nonlinear complementarity problems, *J. Math. Anal. Appl.* 322, 276-287, 2006.