



MÉTODO DE PONTO PROXIMAL PARA OTIMIZAÇÃO VETORIAL E  
MÉTODO DE REGIÃO DE CONFIANÇA PARA OTIMIZAÇÃO  
MULTIOBJETIVO

Kely Diana Villacorta Villacorta

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

Orientador: Paulo Roberto Oliveira

Rio de Janeiro  
Junho de 2011

MÉTODO DE PONTO PROXIMAL PARA OTIMIZAÇÃO VETORIAL E  
MÉTODO DE REGIÃO DE CONFIANÇA PARA OTIMIZAÇÃO  
MULTIOBJETIVO

Kely Diana Villacorta Villacorta

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

Examinada por:



Prof. Paulo Roberto Oliveira, D.Ing.



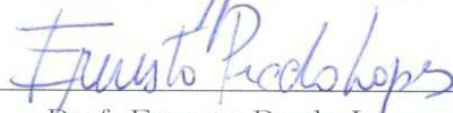
Prof. João Xavier da Cruz Neto, D. Sc.



Prof. Orizon Pereira Ferreira, D.Sc.



Prof. Rolci de Almeida Cipolatti, Dr.



Prof. Ernesto Prado Lopes, Ph.D.

RIO DE JANEIRO, RJ – BRASIL  
JUNHO DE 2011

Villacorta, Kely Diana Villacorta

Método de ponto proximal para otimização vetorial e método de região de confiança para otimização multiobjetivo/Kely Diana Villacorta Villacorta. – Rio de Janeiro: UFRJ/COPPE, 2011.

XI, 69 p. 29, 7cm.

Orientador: Paulo Roberto Oliveira

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2011.

Referências Bibliográficas: p. 64 – 69.

1. método de ponto proximal. 2. método de região de confiança. 3. funções  $C$ -convexas. 4. eficiência. 5. eficiência fraca. 6. Pareto otimalidade. 7. Pareto crítico. 8. otimização vetorial. 9. otimização multiobjetivo. I. Oliveira, Paulo Roberto. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*Ao meu Felipe;  
Aos meus pais Silveria e Alberto;  
Aos meus sogros Isabel e Felipe.*

# Agradecimentos

Agradeço a Deus, em primeiro lugar, pois com Ele tudo é possível.

Ao Prof. Paulo Roberto Oliveira, por ter acreditado em mim num momento em que muito precisei, por estes anos de orientação, pelo entusiasmo ao escutar a minhas ideias, pela amizade e confiança.

Aos Professores João Xavier da Cruz Neto, Orizon Pereira Ferreira, Rolci de Almeida Cipolatti e Ernesto Prado Lopes, por terem participado da banca examinadora e pelas sugestões dadas que contribuíram para a melhoria desta tese.

Ao Prof. Angel Guillermo Coca Balta através do qual tive meu primeiro contato com a otimização (durante a orientação de minha iniciação científica).

Ao Prof. Maculam, pela confiança no novo projeto de Pós-Doutorado.

Ao meu amado Felipe, por ter acreditado que tudo daria certo há 8 anos, pelo seu carinho e força em todos os momentos. Além das preciosas contribuições até na minha pesquisa, pois também foi um colega de sala excepcional.

Aos meus pais, irmãs, sogros e cunhados, que me deram o suporte necessário para a realização deste sonho.

Às minhas queridas amigas: Claudia Susie Camargo Rodrigues, Flávia Morgana Jacinto, Maria de Fátima Cruz Marques, Mariela Morveli Espinoza, Valéria Saldanha Motta, por terem entrado na minha vida no momento certo e aceitar o meu jeitinho peculiar de ser.

A todos os professores do PESC/COPE-UFRJ.

A todos os colegas que tive a oportunidade de conhecer durante estes sete anos no PESC e na UFRJ, em especial ao Aldo Bazán, Erika & Hebert, Glaydston, Guilherme, Hellena, Jesus, Nilomar, Paulinho & Sissy, Roberto Cristóvão & Shirley e a os demais amigos que de alguma forma contribuíram para a realização deste trabalho.

A todos os Funcionários do PESC/COPPE-UFRJ: Adilson, Ana Paula, Ari, Cláudia, Deda, Eliah, Gutierrez, Itamar, Lourdes, Mercedes, Nathalia, Patrícia, Rosa, Sônia, Solange e Taísa.

Às pessoas que nós apoiaram muito nessa caminhada como Cristiane, Margarete & Jorge, Dona Magnólia, Paulo Fernando e ao povo Carioca em geral.

Ao pessoal da biblioteca do CT e do IM, por ter-me brindado muitas vezes a bibliografia que precisava, em especial à Daniele, Jorge e Zoraide.

À CAPES/CNPq-IEL Nacional - Brasil, 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.)

MÉTODO DE PONTO PROXIMAL PARA OTIMIZAÇÃO VETORIAL E  
MÉTODO DE REGIÃO DE CONFIANÇA PARA OTIMIZAÇÃO  
MULTIOBJETIVO

Kely Diana Villacorta Villacorta

Junho/2011

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Neste trabalho considera-se o problema de encontrar soluções fracamente eficientes do problema de otimização vetorial convexo, com respeito à ordem parcial induzida por um cone convexo, fechado e pontudo, com interior não vazio. Para este problema desenvolve-se uma extensão do método de ponto proximal escalar convexo e considera-se também um enfraquecimento nas condições de convergência usuais, que permite construir um método interior proximal para “resolver” este problema sobre conjuntos não poliedrais.

Sob uma outra perspectiva, apresenta-se também um método de região de confiança para otimização multiobjetivo irrestrita. Sob hipóteses apropriadas, que podem ser vistas como uma extensão natural das exigidas para o caso escalar, prova-se que a sequência gerada pelo método converge para um ponto Pareto crítico.

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

PROXIMAL METHOD FOR VECTOR OPTIMIZATION AND  
TRUST-REGION METHOD FOR MULTIOBJECTIVE OPTIMIZATION

Kely Diana Villacorta Villacorta

June/2011

Advisor: Paulo Roberto Oliveira

Department: Systems Engineering and Computer Science

In this work, we consider the problem of finding weakly efficient solutions for convex vector optimization problem with respect to the partial order induced by a closed, convex and pointed cone with a nonempty interior. For this problem, we develop an extension of the convex scalar proximal point method and also it is considered a weakness in the usual conditions of convergence, which allows us to build an interior proximal method to “solve” this problem on nonpolyhedral sets.

In another perspective, we present a trust region method for unconstrained multiobjective optimization. Under appropriate assumptions, which can be viewed as natural extensions of those required for the scalar case, we prove that the sequence generated by this method converges to a Pareto critical point.



# Sumário

<b>1</b>	<b>Preliminares</b>	<b>5</b>
1.1	Noções Básicas . . . . .	5
1.2	Otimização vetorial . . . . .	10
1.2.1	Ordem parcial e espaços objetivos estendidos. . . . .	10
1.2.2	Tópicos de otimização vetorial . . . . .	13
1.2.3	Escalarização . . . . .	23
<b>2</b>	<b>Distância Proximal</b>	<b>29</b>
<b>3</b>	<b>Método Interior Proximal para o Problema de Otimização Vetorial</b>	
	<i>K</i> -convexo	<b>31</b>
3.1	O problema . . . . .	31
3.2	Um método interior proximal exato . . . . .	35
3.3	Um método interior proximal inexato . . . . .	42
<b>4</b>	<b>Método de Região de Confiança para o Problema de Otimização</b>	
	<b>Multiobjetivo Irrestrito</b>	<b>47</b>
4.1	O problema . . . . .	47
4.2	Um método de Região de Confiança . . . . .	49
4.3	Análise de Convergência . . . . .	50
<b>5</b>	<b>Conclusão</b>	<b>62</b>
5.1	Considerações finais . . . . .	62
5.2	Trabalhos futuros . . . . .	63
	<b>Referências Bibliográficas</b>	<b>64</b>

# Notações

$\mathbb{R}^n$	Espaço Euclidiano $n$ -dimensional.
$\mathbb{R}_+^n$	Ortante não-negativo de $\mathbb{R}^n$ .
$\mathbb{R}_{++}^n$	Interior de $\mathbb{R}_+^n$ .
$\mathbb{R}^n \setminus U$	Complementar de $U$ relativo a $\mathbb{R}^n$ .
$\mathbb{S}^n$	Conjunto das matrizes simétricas, $n \times n$ .
$\text{int}(U)$	Interior do conjunto $U$ .
$\text{dim}(U)$	Dimensão do conjunto $U$ .
$\text{cl}(U)$	Fecho do conjunto $U$ .
$\text{conv}(U)$	Envoltória convexa do conjunto $U$ .
$\text{cone}(U)$	Envoltória cônica do conjunto $U$ .
$O^+(U)$	Cone de recessão do conjunto $U$ .
$\text{ir}(U)$	Interior relativo do conjunto $U$ .
$\mathcal{P}(U)$	Conjunto das partes de $U$ .
$I_U$	Função indicadora do conjunto $U$ .
$N_U(x)$	Cone normal de $U$ em $x$ .
$U^*$	Cone polar positivo do cone $U$ .
$\text{dist}(x, U)$	Distância de $x$ ao conjunto $U$ .
$\langle x, y \rangle$	Produto interno Euclidiano de $x$ e $y$ em $\mathbb{R}^n$ .
$x \preceq_C y$	$y - x \in C$ .
$x \prec_C y$	$y - x \in \text{int}(C)$ .
$\text{dom}(f)$	Domínio efetivo de $f$ .
$\text{Im}(f)$	Imagem de $f$ .
$\nabla f(x)$	Gradiente da função $f$ em $x$ .
$\nabla^2 f(x)$	Matriz hessiana da função $f$ em $x$ .
$\partial f(x)$	Subdiferencial de $f$ em $x$ .
$\partial_\varepsilon f(x)$	$\varepsilon$ -subdiferencial de $f$ em $x$ .
$\partial_\circ f(x)$	Subdiferencial de Clarke de $f$ em $x$ .

$G(U)$	Imagem da aplicação $G$ sobre $U$ .
$\mathcal{J}G(x)$	Matriz jacobiana da aplicação $G$ em $x$ .
$\arg \min$	Conjunto dos pontos que realizam o mínimo.
$C - \arg \min$	Conjunto das soluções eficientes.
$C - \arg \min_w$	Conjunto das soluções eficientes fracas.
POVC	Problema de Otimização Vetorial Convexo.
POM	Problema de Otimização Multiobjetivo.
$\kappa_{shF_i}$	Limite superior da matriz <u>h</u> essiana de $F_i$ .
$\kappa_{sahm}$	Limite superior da <u>a</u> proximação da matriz <u>h</u> essiana de $m_k$ .
$\kappa_{lsi}$	<u>L</u> imite <u>s</u> uperior de <u>i</u> nclinação.
$\kappa_{bck}$	Constante de <u>b</u> acktracking.
$\kappa_{dpa}$	constante de <u>d</u> ecréscimo no <u>p</u> onto <u>a</u> proximado.
$\kappa_{dasm}$	constante de <u>d</u> ecréscimo na <u>a</u> proximação da <u>s</u> olução de $m_k$ .
$\kappa_{li\omega}$	<u>L</u> imite <u>i</u> nferior para $\omega$ .
$\kappa_{li\Delta}$	<u>L</u> imite <u>i</u> nferior para $\Delta$ .
BFGS	Broyden-Fletcher-Goldfarb-Shanno.

# Introdução

Considerando que situações da vida cotidiana envolvem inevitavelmente a tomada de decisões, é natural querer que as escolhas a serem feitas sejam as melhores possíveis. A dificuldade surge no conflito entre os vários objetivos e/ou metas que cada um a tem. Da mesma forma, existem áreas de estudo onde os problemas de otimização multiobjetivo e/ou vetorial têm encontrado muitas aplicações, como por exemplo, na teoria econômica, gestão de negócios, entre outras. Embora a grande maioria dos problemas da vida real sejam formulados como problemas de otimização multiobjetivo, ou seja, aqueles onde a ordem parcial é induzida pelos cones Paretianos,  $\mathbb{R}_+^m$ , existem muitos outros que requerem ordens de preferência induzidas por cones convexos fechados, não necessariamente Paretianos, como por exemplo, os cones que foram analisados em [3] (com base nos resultados teóricos de [2]).

Assim, o problema a ser estudado na primeira parte deste trabalho é o problema de otimização vetorial convexo. Antes de apresentá-lo formalmente, lembre que dado  $K$  um cone convexo, fechado e pontudo em  $\mathbb{R}^m$ , com interior não vazio, diz-se que  $K$  induz uma ordem parcial  $\preceq_K$  em  $\mathbb{R}^m$  se, dados  $y, y' \in \mathbb{R}^m$ ,  $y \preceq_K y'$ , se, e somente se,  $y' - y \in K$ . De modo similar, define-se a relação associada  $\prec_K$  como segue: dados  $y, y' \in \mathbb{R}^m$ ,  $y \prec_K y'$ , se, e somente se,  $y' - y \in \text{int}(K)$ .

Então o problema de otimização vetorial convexo é dado por:

$$(POVC) \quad K - \min_{x \in \Omega} F(x), \quad (1)$$

onde  $\Omega$  é um conjunto convexo e fechado, com interior não vazio,  $K$  é o cone que define a ordem parcial  $\preceq_K$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_K\}$  é uma aplicação própria,  $K$ -convexa e positivamente semi-contínua inferior.

Portanto, o interesse relacionado a este problema é desenvolver métodos para encontrar soluções eficientes fracas para (POVC), isto é, métodos para encontrar  $x^* \in \Omega$  tal que

$$\nexists x \in \Omega : F(x) \prec_K F(x^*).$$

Recentemente, Bonnel *et al.* em [11], Ceng e Yao em [13], Fliege *et al.* em [19–23], Graña *et al.* em [26–29], Gregório e Oliveira em [31], Gutiérrez e Jiménez em [32], Huang e Yang em [34], entre outros, têm trabalhado em otimização multiob-

jetivo e/ou vetorial, conseguindo estender para estes problemas resultados teóricos e métodos numéricos utilizados na literatura para otimização escalar. Por exemplo, tem-se a extensão do conceito de ponto crítico para otimização vetorial, e com isto, a extensão das condições necessárias de otimalidade de primeira ordem para problemas deste tipo. E entre os métodos estendidos cita-se, por exemplo, o método do gradiente, o método do gradiente projetado e os métodos do tipo proximal.

Nos últimos anos foram obtidos progressos na teoria do método de ponto proximal para problemas escalares convexos, baseados em distâncias generalizadas [5, 38–40]. Uma breve descrição é dada a seguir. Considere o seguinte problema de minimização convexa:

$$\inf\{f(x) : x \in \Omega\}, \quad (2)$$

onde  $\Omega$  é um conjunto convexo e fechado, com interior não vazio,  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  é uma função própria, convexa e semi-contínua inferior. Logo, o método de ponto proximal gera uma sequência  $\{x^k\} \subset \mathbb{R}^n$  do modo seguinte:

$$\nu^{k+1} + \beta_k \nabla_1 d(x^{k+1}, x^k) = 0, \quad (3)$$

com  $\nu^{k+1} \in \partial_{\varepsilon_k} f(x^k)$ ,  $\{\beta_k\} \subset \mathbb{R}_{++}$ ,  $\nabla_1 d(\cdot, y)$  denotando o gradiente do termo proximal  $d(\cdot, y)$  com respeito à primeira variável e  $x^k$  sendo o iterando atual. Com a escolha  $d(x, y) = 2^{-1}\|x - y\|^2$  e  $\varepsilon_k = 0$ , recupera-se o algoritmo de ponto proximal clássico, o qual foi amplamente analisado, por exemplo, por Moreau em [46], Martinet em [43, 44] e Rockafellar em [49]. No entanto, nesse caso, a sequência  $\{x^k\}$  produzida não está necessariamente contida em  $\text{int}(\Omega)$ . Por conseguinte, o termo proximal  $d(x, y)$  desempenhará o papel de uma função tipo penalização, forçando ao iterando a permanecer no interior do conjunto restrição.

É importante também analisar a versão inexata deste tipo de método, essencial para a obtenção da convergência nas implementações. Uma versão inexata do método clássico foi proposta por Rockafellar em [49]. Nessa versão,  $x^{k+1}$  não necessariamente é uma solução exata do subproblema (3), senão apenas uma solução aproximada deste. Em outras palavras, o  $k$ -ésimo subproblema pode ser resolvido dentro de uma tolerância prescrita  $\tau_k > 0$ , com  $\sum_{k=0}^{\infty} \tau_k < \infty$ . Trabalhos posteriores que propuseram a versão inexata, considerado o termo proximal geral  $d(\cdot, \cdot)$ , adotam também semelhantes critérios sobre o erro  $\tau_k$ , veja por exemplo [5, 14, 38, 41].

Foi provado em [5] que a sequência  $\{x^k\}$  gerada por (3) está contida em  $\text{int}(\Omega)$  e converge para a solução de (2), sob certas condições.

Por outro lado, Kaplan e Tichatschke em [38, 39] apresentaram uma função tipo Bregman com uma modificação na “condição de convergência”, a qual lhes permitiu construir um método proximal generalizado para resolver (2), restrito a conjuntos não necessariamente poliedrais. Para isso, eles admitiram aproximações sucessivas

do operador  $\partial f$  e um cálculo inexato dos iterandos proximais. Esse método gera sequências  $\{x^k\} \subset \mathbb{R}^n$  e  $\{\theta^k\} \subset \mathbb{R}^n$  do modo seguinte:

$$\theta^{k+1} \in Q^k(x^{k+1}) + \beta_k \nabla_1 D_{\hat{h}}(x^{k+1}, x^k), \quad (4)$$

onde  $\{\beta_k\} \subset \mathbb{R}_{++}$  é uma sequência limitada,  $D_{\hat{h}}$  é a distância de Bregman induzida por  $\hat{h} = h + \eta$  e  $\partial f \subset Q^k \subset \partial_{\varepsilon_k} f$ . Se  $h$ ,  $\eta$ ,  $\theta^k$  e  $\varepsilon_k$  satisfazem certas condições, os autores provaram que a sequência  $\{x^k\}$  gerada por (4) está contida em  $\text{int}(\Omega)$  e converge para a solução do problema (2).

Portanto, a essência da primeira parte deste trabalho consiste na extensão da versão exata do método (3) e do método (4) ao contexto vetorial. Nestas extensões, são combinadas as teorias apresentadas por Auslender e Teboulle em [5] e por Kaplan e Tichatschke em [38, 39]. Ao mesmo tempo, o termo proximal usado é mais geral que aquele apresentado em [5], pois a “condição de convergência” é enfraquecida. Vale ressaltar que os resultados obtidos nesta parte da tese foram aceitos pela *European Journal of Operational Research*, [55].

Outro tema de interesse nesta tese foi considerar os métodos de região de confiança para resolver problemas escalares não-convexos, bem conhecidos e utilizados em otimização não-convexa com ou sem restrições. Nestes métodos, é considerado um modelo  $m_k$  da função objetivo, suposto adequado em uma “região de confiança”, a qual é uma vizinhança da iteração atual. Esta vizinhança é frequentemente representada por uma bola em alguma norma, cujo raio  $\Delta_k$  é “atualizado” de iteração em iteração, em relação a quão bem  $m_k$  aproximou à função objetivo no “ponto de prova”.

Vale a pena mencionar que ao longo dos últimos anos estes métodos desfrutam de uma boa reputação com base na sua notável confiança numérica, além de uma teoria sólida e completa de convergência, veja por exemplo [1, 7, 17, 18, 24, 36, 59] e suas respectivas referências. Embora regiões de confiança sejam uma abordagem eficiente para resolver problemas de otimização não-convexa escalar, poucos trabalhos são encontrados na literatura que a apliquem para otimização multiobjetivo e/ou vetorial, veja por exemplo [47, 56, 57].

Assim, no caso particular em que  $K = \mathbb{R}_+^m$  e  $\Omega = \mathbb{R}^n$ , é obtido o segundo problema de estudo, que se descreve a seguir. Considere o Problema de Otimização Multiobjetivo Irrestrito, denotado por:

$$(POMI) \quad \mathbb{R}_+^m - \min_{x \in \mathbb{R}^n} F(x),$$

onde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tem cada uma das suas componentes duas vezes diferenciável.

Logo, o interesse relacionado a este problema é desenvolver métodos para encontrar pontos Paretos críticos para o (POMI), isto é, métodos para encontrar  $x^* \in \Omega$

tal que, para todo  $d \in \mathbb{R}^n$ , existe  $i_0 = i_0(d) \in \{1, \dots, m\}$  com

$$\langle \nabla F_{i_0}(x^*), d \rangle \geq 0.$$

Portanto, na segunda parte desta tese, estabelece-se um algoritmo de região de confiança para encontrar pontos Pareto críticos. Vale a pena mencionar que o algoritmo apresentado está baseado no algoritmo básico apresentado por Conn *et al.* em [17], e embora pareça simples, a análise de convergência aqui apresentada, segue o roteiro dos métodos no caso escalar. Não conhecemos trabalhos similares na literatura.

Este trabalho é organizado como segue. No Capítulo 1 recordam-se alguns conceitos e resultados preliminares. No Capítulo 2 é apresentada a família de distâncias proximais e distâncias proximais induzidas a serem usadas. No capítulo 3 define-se formalmente o primeiro problema de estudo, e é apresentado um método interior proximal para a obtenção de soluções eficientes fracas. Além disso, uma completa análise de convergência é apresentada para as versões exata e inexata, respectivamente. No Capítulo 4, o problema de otimização multiobjetivo é considerado como o segundo problema a ser estudado e é apresentado um algoritmo de região de confiança para a obtenção de pontos Pareto críticos. Mais ainda, assim como no Capítulo 3, a análise de convergência relativa a este método também é apresentada. Finalmente, no Capítulo 5, são feitas considerações finais e são apresentadas algumas possibilidades de pesquisas futuras.

# Capítulo 1

## Preliminares

A finalidade deste capítulo é facilitar a leitura desta tese fornecendo as definições e propriedades que são necessárias para o desenvolvimento dos demais capítulos. Esses fatos sobre conjuntos e funções com imagens em  $(-\infty, +\infty]$  podem ser encontrados em diversos livros de análise convexa, tais como em [4, 48, 50, 51]. Os demais conceitos e/ou propriedades foram recompilados de diversos livros e trabalhos tais como por exemplo [21, 22, 25, 29, 33, 35, 42, 58].

### 1.1 Noções Básicas

Um resultado importante que envolve sequências de números reais não-negativas é o seguinte:

**Lema 1.1.1.** [48, Lema 2.2.2] *Sejam  $\{\xi_k\}$ ,  $\{v_k\}$  e  $\{\zeta_k\}$  sequências de números reais não-negativas. Se*

$$\xi_{k+1} \leq (1 + v_k)\xi_k + \zeta_k, \quad \sum_{k=1}^{\infty} \zeta_k < \infty \quad e \quad \sum_{k=1}^{\infty} v_k < \infty,$$

*então a sequência  $\{\xi_k\}$  é convergente.*

Abaixo são listados alguns tipos de conjuntos que serão utilizados ao longo deste trabalho.

**Definição 1.1.1.** Seja  $U \subseteq \mathbb{R}^n$  um conjunto e  $x^* \in \mathbb{R}^n$ . Diz-se que:

*i)  $U$  é **convexo** se, para qualquer  $x, y \in U$  e  $\lambda \in [0, 1]$ , tem-se que*

$$\lambda x + (1 - \lambda)y \in U;$$

*ii)  $U$  é um **cone** se, para qualquer  $\lambda \geq 0$ ,*

$$\lambda U \subset U;$$



iii)  $U$  é um **cone pontudo** se,

$$U \cap (-U) = \{0\};$$

iv) O **cone polar positivo** de  $U$ ,  $U^*$ , é definido por

$$U^* := \{z \in \mathbb{R}^n : \langle z, y \rangle \geq 0, \quad y \in U\};$$

v) O **cone normal** de  $U$  em  $x^*$ ,  $N_U(x^*)$ , é definido por

$$N_U(x^*) := \{z \in \mathbb{R}^n : \langle z, y - x^* \rangle \leq 0, \quad y \in U\}.$$

**Definição 1.1.2.** Seja  $U \subseteq \mathbb{R}^n$  um conjunto. A **função indicadora** de  $U$ ,  $I_U$ , é definida por

$$I_U(x) = \begin{cases} 0, & x \in U \\ +\infty, & \text{caso contrário.} \end{cases}$$

*Observação 1.1.1.*  $U$  é um conjunto convexo se, e somente se,  $I_U$  é uma função convexa em  $\mathbb{R}^n$ .

**Lema 1.1.2.** [50, Corolário 6.5.2] *Sejam  $U_1$  e  $U_2$  conjuntos convexos. Se  $U_2 \subseteq \text{cl}(U_1)$ , com  $U_2$  não inteiramente contido na fronteira relativa de  $U_1$ , então  $\text{ir}(U_2) \subset \text{ir}(U_1)$ .*

**Definição 1.1.3.** Seja  $f : U \subseteq \mathbb{R}^n \rightarrow (-\infty, +\infty]$  uma função e  $x^* \in \mathbb{R}^m$ .

i) O **domínio efetivo** de  $f$ ,  $\text{dom}(f)$ , é definido por

$$\text{dom}(f) := \{x \in U : f(x) < +\infty\};$$

ii)  $f$  é **própria** se, e somente se,  $\text{dom}(f) \neq \emptyset$ ;

iii) Se  $U$  é um conjunto convexo, então  $f$  é **convexa** em  $U$  se, e somente se, para qualquer  $x, y \in U$  e  $\lambda \in [0, 1]$  tem-se que

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

iv)  $f$  é **Lipschitz** em  $U \subseteq \text{dom}(f)$  se, e somente se, existe  $L \geq 0$  tal que

$$|f(x) - f(y)| \leq L\|x - y\|, \quad \forall x, y \in U;$$

v)  $f$  é **localmente Lipschitz** em  $U \subseteq \text{dom}(f)$  se, e somente se, para cada  $x \in U$  existe  $\varepsilon > 0$  tal que

$f$  é Lipschitz em  $B(x, \varepsilon)$ ;

v)  $f$  é **coerciva** se, e somente se,

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty;$$

vi)  $f$  é **1-coerciva** se, e somente se,

$$\lim_{\|x\| \rightarrow \infty} \frac{f(x)}{\|x\|} = +\infty;$$

vii)  $f$  é **semi-contínua inferior**, *sci*, em  $x^* \in U$  se, e somente se, para qualquer sequência  $\{x^k\} \subset U$  que converge para  $x^*$ ,

$$f(x^*) \leq \liminf_{k \rightarrow +\infty} f(x_k);$$

viii)  $f$  é **nível limitada** se, e somente se, para qualquer  $\delta \in \mathbb{R}$  o conjunto de nível

$$\mathcal{L}(f, \delta) := \{x \in U : f(x) \leq \delta\}$$

é um conjunto limitado (possivelmente vazio).

*Observação 1.1.2.*

i) Se  $f$  é própria e convexa, então  $f$  é localmente Lipschitz em  $\text{ir}(\text{dom}(f))$ .

ii) Se  $f$  é Lipschitz, então  $f$  é uniformemente contínua.

Para funções não-diferenciáveis, a convexidade desempenha um papel fundamental que permite definir o conceito de subgradientes e subdiferencial.

**Definição 1.1.4.** [4] Seja  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  uma função convexa, e seja  $x \in \text{dom}(f)$ . O vetor  $\xi \in \mathbb{R}^n$  é dito ser um **subgradiente** de  $f$  em  $x$  se

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

O conjunto de todos os subgradientes de  $f$  em  $x \in \text{dom}(f)$  é chamado de **subdiferencial** de  $f$  em  $x$  e denotado por  $\partial f(x)$ ; isto é,

$$\partial f(x) := \{g \in \mathbb{R}^n : f(z) \geq f(x) + \langle g, z - x \rangle, \quad z \in \mathbb{R}^n\}.$$

*Observação 1.1.3.*

i)  $\partial f(x)$  é um conjunto convexo e fechado, para todo  $x \in \mathbb{R}^n$ .

- ii) Se  $\partial f(x) \neq \emptyset$ , então  $f$  é subdiferenciável em  $x$ .
- iii) Uma função própria, convexa  $f$  é diferenciável em  $x \in \text{dom}(f)$  se, e somente se, o subdiferencial reduz-se a um conjunto unitário, o qual é o gradiente de  $f$  em  $x$ ,  $\nabla f(x)$ .
- iv) Se  $U$  é um conjunto não-vazio e convexo, então

$$\partial I_U(x) = \begin{cases} N_U(x), & x \in U \\ \emptyset, & \text{caso contrário.} \end{cases}$$

**Lema 1.1.3.** [50, Teorema 27.4] *Seja  $f : U \rightarrow (-\infty, +\infty]$  uma função própria convexa, e seja  $U$  um conjunto não vazio e convexo. A fim de que  $x$  seja um ponto onde o ínfimo de  $f$  relativo a  $U$  é atingido, é suficiente que exista um vetor  $\nu \in \partial f(x)$  tal que  $-\nu$  é normal a  $U$  em  $x$ . Esta condição também é necessária se  $\text{ir}(\text{dom}(f))$  intersecta  $\text{ir}(U)$ , ou se  $U$  é um conjunto poliedral, e  $\text{ir}(\text{dom}(f))$ , simplesmente, intersecta  $U$ .*

A seguir são enunciados alguns conceitos de generalizações do subdiferencial.

**Definição 1.1.5.** [4] *Seja  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  uma função semi-contínua inferior. Para qualquer  $x \in \text{dom}(f)$  e qualquer  $\varepsilon > 0$ , o  $\varepsilon$ -**subdiferencial** de  $f$  em  $x$ , denotado por  $\partial_\varepsilon f(x)$ , é definido por*

$$\partial_\varepsilon f(x) := \{g \in \mathbb{R}^n : f(z) + \varepsilon \geq f(x) + \langle g, z - x \rangle, \quad z \in \mathbb{R}^n\};$$

Dentre as propriedades que o  $\varepsilon$ -subdiferencial goza, tem-se em particular o seguinte lema.

**Lema 1.1.4.** [4, Theorem 4.1.3] *Seja  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  uma função própria, convexa e semi-contínua inferior, e seja  $\varepsilon \geq 0$ . Então, para cada  $x \in \text{dom}(f)$ ,  $\partial_\varepsilon f(x)$  é um conjunto não vazio, convexo e fechado.*

Note que se  $0 \leq \varepsilon_1 \leq \varepsilon_2 < +\infty$ , então

$$\partial f(x) = \partial_0 f(x) \subseteq \partial_{\varepsilon_1} f(x) \subseteq \partial_{\varepsilon_2} f(x).$$

**Definição 1.1.6.** [12] *Seja  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  uma função localmente Lipschitz. A **derivada direcional de Clarke** de  $f$  em  $x \in \text{dom}(f)$  na direção  $d$ , denotada por  $f^\circ(x; d)$ , é definida por*

$$\begin{aligned} f^\circ(x; d) &= \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y+td) - f(y)}{t} \\ &= \inf_{\delta > 0} \sup_{\|y-x\| \leq \delta, 0 < t < \delta} \frac{f(y+td) - f(y)}{t}. \end{aligned}$$

Assim, o **subdiferencial de Clarke** de  $f$  em  $x$ , denotado por  $\partial_{\circ}f(x)$ , é definido por

$$\partial_{\circ}f(x) = \{\xi \in \mathbb{R}^n : f^{\circ}(x; d) \geq \langle \xi, d \rangle, \quad d \in \mathbb{R}^n\}.$$

*Observação 1.1.4.*

- i) Se  $f$  é diferenciável, então  $f^{\circ}(x; d) = \langle \nabla f(x), d \rangle$ ;
- ii) Se  $f$  é convexa, então  $\partial_{\circ}f(x) = \partial f(x)$ .

**Lema 1.1.5.** [58, Lemma 2.1] *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função localmente Lipschitz. Então*

- i)  $\partial_{\circ}f(x)$  é um subconjunto não vazio, compacto e convexo de  $\mathbb{R}^n$ .
- ii)  $f^{\circ}(x; d) = \max\{\langle \xi, d \rangle : \xi \in \partial_{\circ}f(x)\}$ .
- iii) A aplicação ponto conjunto  $\partial_{\circ}f(\cdot)$  é localmente limitada e semi-contínua superior.
- iv) Seja  $f(x) = \max\{f_l(x) : l \in \Lambda\}$ , onde o conjunto de índices  $\Lambda \subset \mathbb{R}$  é compacto. Se cada  $f_l$  é continuamente diferenciável em  $x$ , então

$$\partial_{\circ}f(x) = \text{conv}\{\nabla f_l(x) : l \in \Lambda(x)\},$$

$$\text{onde } \Lambda(x) = \{l \in \Lambda : f_l(x) = f(x)\}.$$

Vários resultados importantes em otimização vetorial são estabelecidos analisando a intersecção das direções assintóticas do conjunto formado pelas imagens da função objetivo e o cone que define a ordem parcial. Abaixo, é enunciada a definição relacionada a este conceito e são apresentadas suas principais propriedades.

**Definição 1.1.7.** [42, Definição 2.2, Observação 2.3] *Seja  $U \subseteq \mathbb{R}^n$  um conjunto não vazio. O **cone de recessão** de  $U$ ,  $O^+(U)$ , é dado por*

$$O^+(U) := \cap \{\text{cl}((0, \alpha]U) : \alpha > 0\},$$

onde  $(0, \alpha]U = \{tU : 0 < t \leq \alpha\}$ .

*Observação 1.1.5.* O conceito de cone de recessão foi introduzido, inicialmente, para conjuntos convexos, veja por exemplo [50]. Assim, se  $U$  é convexo, então a definição acima se reduz a:

$$O^+(U) := \{v \in \mathbb{R}^n : U + tv \subset U, \quad t \geq 0\}.$$

Luc em [42] apresenta os seguintes resultados envolvendo cones de recessão.

**Lema 1.1.6.** [42, Proposição 2.5, Lema 2.11, Observação 2.6] *Sejam  $U, V \subseteq \mathbb{R}^n$  conjuntos não vazios. As seguintes propriedades são satisfeitas:*

i)  $U$  é limitado se, e somente se,

$$O^+(U) = \{0\}, \quad e \quad O^+(U + v) = O^+(v);$$

ii) Se  $U \subseteq V$ , então

$$O^+(U) \subseteq O^+(v);$$

iii) Para qualquer  $t \in \mathbb{R}$ ,

$$O^+(tU) = \text{ sinal}(t)O^+(U),$$

onde  $\text{ sinal}(t)$  é 1, 0 ou  $-1$  se  $t$  é positivo, zero ou negativo, respectivamente;

iv)  $O^+(U) = O^+(\text{cl}(U))$ ;

v) Se  $U$  é um cone, então

$$O^+(U) = \text{cl}(U);$$

vi)  $O^+(U \cup v) = O^+(U) \cup O^+(v)$ ;

vii)  $O^+(U \cap v) \subseteq O^+(U) \cap O^+(v)$ .

## 1.2 Otimização vetorial

Nesta seção são abordados os conceitos básicos de “otimalidade” ou eficiência no espaço Euclidiano  $\mathbb{R}^m$ . Antes será estabelecida a relação de ordem entre os elementos deste espaço para compará-los. Na literatura, estes conceitos estão relacionados com uma ordem parcial. Neste trabalho, somente é considerada a ordem parcial induzida por cones convexos, fechados e pontudos, com interior não vazio. Para mais detalhes veja por exemplo [25, 42, 45].

### 1.2.1 Ordem parcial e espaços objetivos estendidos.

**Definição 1.2.1.** [25] Uma relação binária é chamada *ordem parcial* se ela é reflexiva, transitiva e anti-simétrica.

A seguir, define-se a classe de ordens parciais determinadas por cones em  $\mathbb{R}^m$ .

**Definição 1.2.2.** Dado  $C \subset \mathbb{R}^m$  um cone não vazio, convexo, fechado e pontudo. A relação binária associada a  $C$ , denotada por  $\preceq_C$ , é definida como:

$$y_1 \preceq_C y_2 \quad \text{para} \quad y_1, y_2 \in \mathbb{R}^m \quad \iff \quad y_2 - y_1 \in C;$$

e se  $C$  tem interior não vazio, a relação binária estrita associada a  $C$ , denotada por  $\prec_C$ , é definida como:

$$y_1 \prec_C y_2 \quad \text{para} \quad y_1, y_2 \in \mathbb{R}^m \quad \iff \quad y_2 - y_1 \in \text{int}(C).$$

Pelo Teorema 2.1.13 em [25] tem-se que  $\preceq_C$  é uma relação binária reflexiva, transitiva e anti-simétrica. Portanto,  $\preceq_C$  é uma ordem parcial.

Além disso, esta ordem parcial é compatível com a soma entre vetores e o produto com um escalar positivo, ou seja, dados  $y, y_1, y_2 \in \mathbb{R}^m$  e  $\lambda \in \mathbb{R}$  tem-se que

$$y_1 \preceq_C y_2, \quad \lambda \geq 0 \quad \Rightarrow \quad \lambda y_1 \preceq_C \lambda y_2,$$

e

$$y_1 \preceq_C y_2 \quad \Rightarrow \quad y_1 + y \preceq_C y_2 + y.$$

**Definição 1.2.3.** Sejam  $C \subset \mathbb{R}^m$  um cone não vazio, convexo, fechado e pontudo, com interior não vazio, e dois pontos  $y_1, y_2 \in \mathbb{R}^m$ . Diz-se que:

$y_2$  **domina**  $y_1$  se, e somente se,  $y_1 \preceq_C y_2$ , e  
 $y_2$  **domina estritamente**  $y_1$  se, e somente se,  $y_1 \prec_C y_2$ .

As seguintes figuras ilustram estes conceitos.

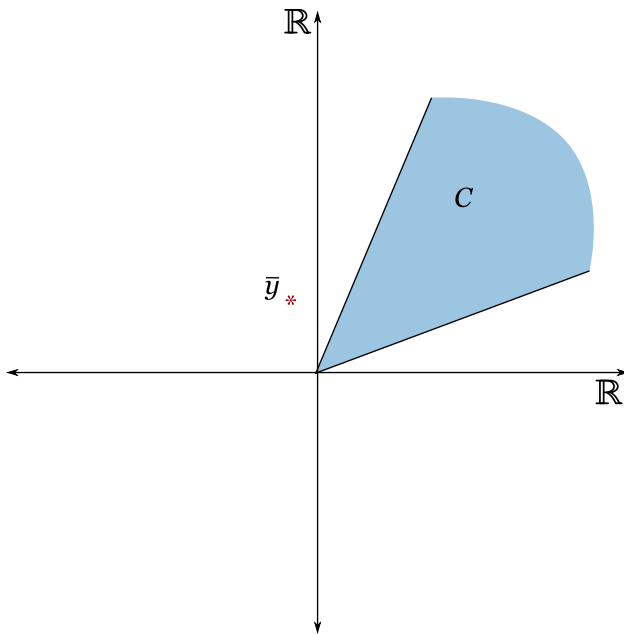


Figura 1: Sejam  $\bar{y} \in \mathbb{R}^2$  e  $C \subset \mathbb{R}^2$  um cone não vazio, fechado, convexo e pontudo, com interior não vazio.

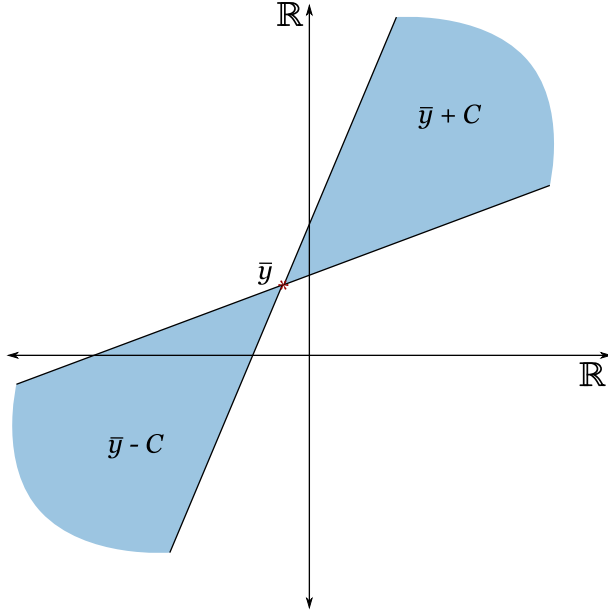


Figura 2: Todos os vetores no cone trasladado  $\bar{y} + C$  dominam  $\bar{y}$ , isto é,  $\bar{y} \preceq_C w$  para todo  $w \in \bar{y} + C$ , enquanto que todos os vetores no cone trasladado  $\bar{y} - C$  são dominados por  $\bar{y}$ , isto é,  $u \preceq_C \bar{y}$  para todo  $u \in \bar{y} - C$ .

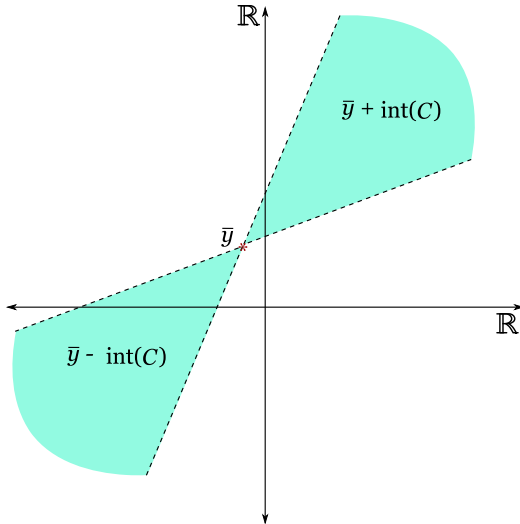


Figura 3: Todos os vetores no cone trasladado  $\bar{y} + \text{int}(C)$  dominam estritamente  $\bar{y}$ , isto é,  $\bar{y} \prec_C w$  para todo  $w \in \bar{y} + \text{int}(C)$ , enquanto que todos os vetores no cone trasladado  $\bar{y} - \text{int}(C)$  são dominados estritamente por  $\bar{y}$ , isto é,  $u \prec_C \bar{y}$  para todo  $u \in \bar{y} - \text{int}(C)$ .

Observa-se que, quando  $C = \mathbb{R}_+^m$ , dados  $y^1, y^2 \in \mathbb{R}^m$ , tem-se que  $y^2$  domina  $y^1$  se, e somente se,

$$y_i^1 \leq y_i^2, \quad \text{para todo } i \in \{1, \dots, m\}, \quad \text{e}$$

$y^2$  domina estritamente  $y^1$  se, e somente se,

$$y_i^1 < y_i^2, \quad \text{para todo } i \in \{1, \dots, m\}.$$

Observa-se que a teoria em funções definidas sobre  $\mathbb{R}^n$  a valores estendidos,  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ , é, frequentemente, usada em análise convexa real. Uma das vantagens é, por exemplo, a unificação das funções definidas sobre todo o espaço  $\mathbb{R}^n$  a valores reais,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , e as funções definidas sobre subconjuntos de  $\mathbb{R}^n$ ,  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ; isto é, podem-se identificar  $f$  pela função  $\tilde{f} : \mathbb{R}^n \rightarrow (-\infty, +\infty]$

dada por

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

Por esta e outras razões práticas, neste trabalho será considerada uma extensão de  $\mathbb{R}^m$  em relação de  $\preceq_C$ , isto é, a partir de agora será considerado o espaço  $\mathbb{R}^m \cup \{+\infty_C\}$  onde  $+\infty_C$  é um elemento que não pertence a  $\mathbb{R}^m$  (alguns outros resultados a respeito de espaços estendidos podem ser encontrados em [9, 10] e suas referências). Assim, as relações  $\preceq_C$  e  $\prec_C$  são estendidas para  $\mathbb{R}^m \cup \{+\infty_C\}$  da seguinte forma:

$$y \preceq +\infty_C \quad \text{e} \quad y \prec +\infty_C, \quad y \in \mathbb{R}^m,$$

e serão consideradas aplicações do tipo  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ .

**Definição 1.2.4.** Seja uma aplicação  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ .

i) O **domínio efetivo** de  $G$ ,  $\text{dom}(G)$ , é dado por

$$\text{dom}(G) = \{x \in \mathbb{R}^n : G(x) \neq +\infty_C\}.$$

ii)  $G$  é **própria** se  $\text{dom}(G) \neq \emptyset$ .

## 1.2.2 Tópicos de otimização vetorial

Nesta seção são apresentadas definições e propriedades relacionadas à otimização vetorial, úteis para o desenvolvimento deste trabalho. Considerando  $\mathbb{R}^m \cup \{+\infty_C\}$  parcialmente ordenado por  $C \subseteq \mathbb{R}^m$ , um cone não vazio, convexo e fechado, com  $\text{int}(C) \neq \emptyset$ , seja o problema de otimização vetorial correspondente a  $U$  e  $G$  dado por:

$$C - \min_{x \in U} G(x) \tag{1.1}$$

onde  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ , com  $G = [G_1, \dots, G_m]^T$ , e  $U \subset \text{dom}(G)$ .

Observa-se que este problema faz sentido, exatamente, porque em geral não existe algum  $\bar{x} \in U$  que minimize, simultaneamente, todas as diferentes funções objetivo. Além disso, este tipo de problema quase sempre possui mais de um “valor ótimo”. Assim, estabelecer o conceito de “otimalidade” e/ou eficiência vetorial faz-se necessário.

**Definição 1.2.5.** Sejam  $x^* \in U$  e  $G(U)$  a imagem pela aplicação  $G$  do conjunto  $U$ . Diz-se que  $x^* \in U$  é:

i) **Eficiente (ou Pareto)** para  $G(U)$  se não existe  $x \in U$  tal que

$$G(x) \preceq_C G(x^*) \quad \text{e} \quad G(x) \neq G(x^*);$$



o conjunto das soluções eficientes é denotado por

$$C - \arg \min \{G(x) : x \in U\}.$$

ii) **Eficiente fraco** para  $G(U)$  se não existe  $x \in U$  tal que

$$G(x) \prec_C G(x^*);$$

o conjunto das soluções eficientes fracas é denotado por

$$C - \operatorname{argmin}_w \{G(x) : x \in U\}.$$

Se as condições acima são verificadas em  $\mathcal{V} \cap U$ , onde  $\mathcal{V}$  é uma vizinhança de  $x^*$ , então diz-se que  $x^*$  é eficiente local ou eficiente fraco local para  $G(U)$ , respectivamente.

No caso escalar ( $m = 1$ ) o conceito de ponto eficiente (local) ou eficiente fraco (local) se reduz à definição usual de ponto mínimo (local).

É imediato ver que

$$C - \arg \min \{G(x) : x \in U\} \subseteq C - \operatorname{argmin}_w \{G(x) : x \in U\}.$$

*Observação 1.2.1.* Quando  $C = \mathbb{R}_+^m$ , o problema de otimização vetorial é conhecido, na literatura, como problema de otimização multiobjetivo, e as noções de otimalidade para este problema reduzem-se a:

$x^* \in U$  é **Pareto** para  $G(U)$  se não existe  $x \in U$  tal que

$$G_i(x) \leq G_i(x^*), \quad \text{para todo } i \in \{1, \dots, m\}, \quad \text{e}$$

$$G_j(x) < G_j(x^*), \quad \text{para algum } j \in \{1, \dots, m\};$$

$x^* \in U$  é **Pareto fraco** para  $G(U)$  se não existe  $x \in U$  tal que

$$G_i(x) < G_i(x^*), \quad \text{para todo } i = 1, \dots, m.$$

**Proposição 1.2.1.** [42, Proposição 2.4] *Sejam  $C_1$  e  $C_2$  dois cones em  $\mathbb{R}^m$ . Se  $C_1 \subseteq C_2$  e  $C_2$  é pontudo então*

$$C_2 - \arg \min \{G(x) : x \in U\} \subseteq C_1 - \arg \min \{G(x) : x \in U\}$$

e

$$C_2 - \operatorname{argmin}_w \{G(x) : x \in U\} \subseteq C_1 - \operatorname{argmin}_w \{G(x) : x \in U\}.$$

O seguintes resultados dizem respeito à extensão da ideia de limitação inferior para funções vetoriais.

**Definição 1.2.6.** [25] Sejam  $A \subseteq \mathbb{R}^m$  e um cone convexo  $K \subset \mathbb{R}^m$ . O conjunto  $A$  é dito  $K$ -**limitado inferior** se existe  $\bar{y} \in \mathbb{R}^m$  tal que

$$A \subseteq \bar{y} + K.$$

*Observação 1.2.2.*

- i) Em geral, dado um conjunto limitado  $A$  em  $\mathbb{R}^m$  e  $K$  um cone convexo em  $\mathbb{R}^m$ , se  $\dim(A) \leq \dim(K)$ , então  $A$  é  $K$ -limitado inferior.
- ii) Ressalta-se que a última definição será usada, somente, para o espaço objetivo da função vetorial  $G$ ,  $\mathbb{R}^m$ . Assim, a definição acima reduz-se a:

Sejam um conjunto  $U \subseteq \mathbb{R}^n$  e uma aplicação  $G : U \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ . Então  $G(U)$  é  $K$ -limitado inferior se existe  $\bar{y} \in \mathbb{R}^m$  tal que

$$G(U) \subseteq \bar{y} + K.$$

No caso  $m = 1$  e  $K = \mathbb{R}_+$  recupera-se a definição usual de limitação inferior para funções escalares.

A seguir dá-se um exemplo que ilustra o item (ii) da observação acima.

**Exemplo 1.2.1.** [29] Seja  $U = (-1, +\infty) \times [0, +\infty) \subseteq \mathbb{R}^2$  e  $G : U \rightarrow \mathbb{R}^2$  dada por

$$G(x) := \begin{bmatrix} x_1 \\ x_2 - \ln(x_1 + 1) \end{bmatrix}.$$

Logo,

$$G(U) = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 : \begin{array}{l} -\ln(y_1 + 1) \leq y_2, \\ -1 < y_1 \end{array} \right\}.$$

Agora, consideram-se os seguintes cones:

$$K_1 = \mathbb{R}_+^2 \quad \text{e}$$

$$K_2 = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 : y_1 + y_2 \geq 0, \quad y_1 \geq 0 \right\}.$$

Veja figuras 5 e 6, respectivamente.

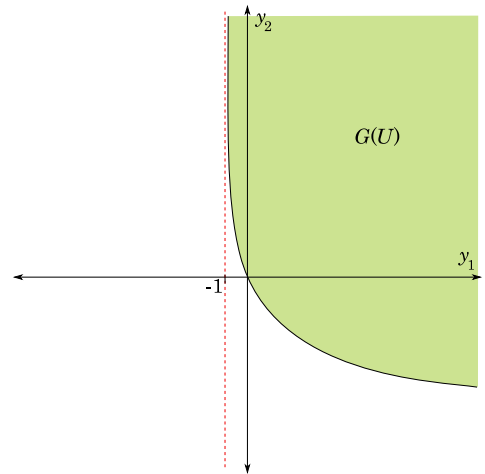


Figura 4: Imagem pela aplicação  $G$  do conjunto  $U$ .

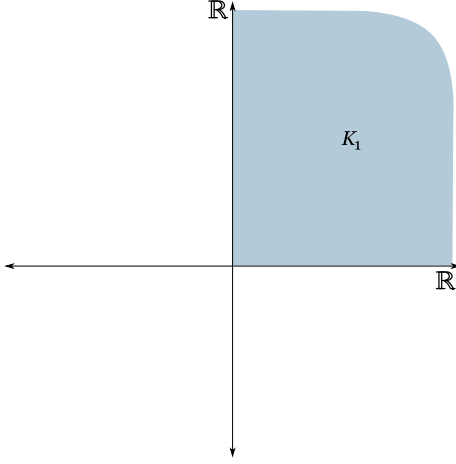


Figura 5.

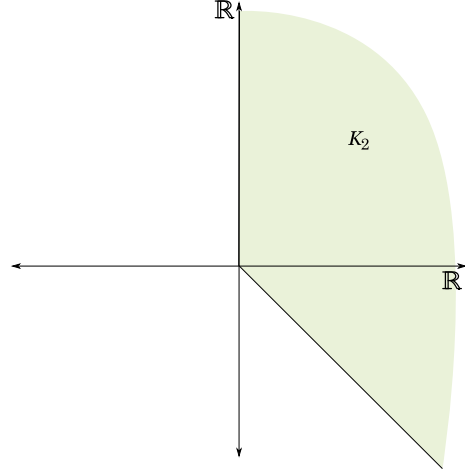


Figura 6.

Desde que para qualquer  $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in G(U)$  tem-se que

$$-\ln(y_1 + 1) \leq y_2 \quad \text{e} \quad y_2 \in \mathbb{R},$$

resulta que  $G(U)$  não é  $K_1$ -limitado inferior. Por outro lado,

$$-1 - y_1 < -\ln(y_1 + 1), \quad y_1 \geq 0.$$

Assim,

$$\exists \bar{y} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{tal que} \quad G(U) \subset \bar{y} + K_2.$$

Portanto,  $G(U)$  é  $K_2$ -limitado inferior.

*Observação 1.2.3.* Sejam dados dois cones convexos  $K_1$  e  $K_2 \subset \mathbb{R}^m$ . Se  $A$  é  $K_1$ -limitado inferior e  $K_1 \subseteq K_2$ , então  $A$  é  $K_2$ -limitado inferior.

No exemplo acima tem-se que  $G(U)$  é  $K$ -limitado inferior para todo cone convexo fechado  $K$  que contém  $K_2$ .

**Lema 1.2.1.** [42, Proposition 3.6] *Se  $A$  é  $K$ -limitado inferior, então*

$$O^+(A) \subseteq \text{cl}(K).$$

**Afirmção 1.2.1.** *Se  $G_i : U \rightarrow (-\infty, +\infty]$  é limitada inferiormente para cada  $i = 1, \dots, m$ , então  $G(U)$  é  $\mathbb{R}_+^m$ -limitada inferior.*

*Demonstração.* Visto que  $G_i$  é limitada inferiormente para cada  $i = 1, \dots, m$ , existe  $\alpha_i \in \mathbb{R}$ ,  $i = 1, \dots, m$ , tal que

$$\alpha_i \leq \inf_{x \in U} G_i(x), \quad i = 1, \dots, m.$$

Fazendo

$$\bar{y} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \quad \text{obtem-se que } G(U) \subset \bar{y} + \mathbb{R}_+^m. \quad \square$$

**Afirmação 1.2.2.** [34] *Sejam os problemas de otimização multiobjetivo dados por:*

$$\mathbb{R}_+^m - \min_{x \in U} G(x), \quad (1.2)$$

$$\mathbb{R}_+^m - \min_{x \in U} \exp(G(x)), \quad (1.3)$$

onde  $G_i : U \rightarrow (-\infty, +\infty]$ ,  $i = 1, \dots, m$  e  $\exp(G) = [\exp(G_1), \dots, \exp(G_m)]^T$ .  
Então, o problema (1.2) é equivalente ao problema (1.3) “no sentido fraco”, isto é,

$$\mathbb{R}_+^m - \operatorname{argmin}_w \{G(x) : x \in U\} = \mathbb{R}_+^m - \operatorname{argmin}_w \{\exp(G(x)) : x \in U\}.$$

*Demonstração.* Prova-se o resultado verificando as inclusões entre esses conjuntos de soluções eficientes fracas.

[ $\subseteq$ ] Dado  $\bar{x} \in \mathbb{R}_+^m - \operatorname{argmin}_w \{G(x) : x \in U\}$  suponha-se que

$$\bar{x} \notin \mathbb{R}_+^m - \operatorname{argmin}_w \{\exp(G(x)) : x \in U\}.$$

Logo, existe  $\tilde{x} \in U$  tal que

$$\exp(G_i(\tilde{x})) < \exp(G_i(\bar{x})), \quad \text{para todo } i = 1, \dots, m.$$

Assim, como a função exponencial é monótona crescente, tem-se que

$$G_i(\tilde{x}) < G_i(\bar{x}), \quad \text{para todo } i = 1, \dots, m,$$

implicando que

$$G(\tilde{x}) < G(\bar{x}),$$

o que contradiz a eficiência fraca de  $\bar{x}$  para o problema (1.2).

[ $\supseteq$ ] Esta inclusão é obtida de forma análoga. □

*Observação 1.2.4.* Seja o problema multiobjetivo,  $C = \mathbb{R}_+^m$ ,

$$\mathbb{R}_+^m - \min_{x \in U} G(x). \quad (1.4)$$

Das Afirmações 1.2.1 e 1.2.2, pode-se supor sem perda de generalidade que

$G(U)$  é  $\mathbb{R}_+^m$ -limitado inferior.

**Definição 1.2.7.** Sejam  $A \subseteq \mathbb{R}^m$  e  $y \in A$ . O conjunto  $(y - C) \cap A$  é chamado de *seção* de  $A$  em  $y$ .

**Afirmção 1.2.3.** Sejam  $A \subseteq \mathbb{R}^m$ ,  $\bar{y} \in A$ ,  $C \subset \mathbb{R}^m$  um cone convexo, fechado e pontudo, e um semi-espaço fechado  $\mathcal{H}$ . Se  $(-C) \cap \mathcal{H} = \{0\}$  e  $A$  é  $\mathcal{H}$ -limitado inferior, então a seção de  $A$  em  $\bar{y}$  é um conjunto limitado.

*Demonstração.* Visto que  $\bar{y} \in A$ , obtém-se

$$(\bar{y} - C) \cap A \neq \emptyset.$$

Por outro lado, dos Lemas 1.1.6 e 1.2.1, as seguintes inclusões são verificadas:

$$O^+((\bar{y} - C) \cap A) \subseteq O^+(\bar{y} - C) \cap O^+(A),$$

$$O^+(A) \subseteq \mathcal{H} \quad \text{e} \quad O^+(\bar{y} - C) = -C.$$

Assim,

$$O^+((\bar{y} - C) \cap A) \subseteq (-C) \cap \mathcal{H} = \{0\}.$$

Portanto,  $(\bar{y} - C) \cap A$  é um conjunto limitado. □

Observe-se que este resultado continua sendo válido se  $A$  é  $C$ -limitado inferior. Além disso, se  $A$  é um conjunto fechado e as hipóteses da afirmação acima são verificadas, então

$$(\bar{y} - C) \cap A \text{ é um conjunto compacto.}$$

A seguir, apresenta-se um resultado importante, o qual implica na existência de pontos eficientes para o problema vetorial (1.1), quando  $C$  é um cone convexo em  $\mathbb{R}^m$ , por exemplo, veja [25, 42], mas antes são necessárias as seguintes definições.

**Definição 1.2.8.** [25] Seja  $I$  um conjunto de índices e uma sequência  $\{y_\alpha\}_{\alpha \in I}$  em  $\mathbb{R}^m$ . Diz-se que a sequência  $\{y_\alpha\}_{\alpha \in I}$  é  $C$ -**decrecente**, com respeito a  $C$ , se

$$y_\alpha \preceq_C y_\beta, \text{ para cada } \alpha, \beta \in I, \alpha > \beta.$$

**Definição 1.2.9.** [25] Seja um conjunto  $A \subseteq \mathbb{R}^m$ . Diz-se que o conjunto  $A$  é  $C$ -**completo**, se cada sequência  $C$ -**decrecente** de  $A$  é  $C$ -limitada inferior por um elemento de  $A$ .

**Teorema 1.2.1.** [42, Teorema 3.4] Sejam  $C \subset \mathbb{R}^m$  um cone não vazio, convexo, fechado e o problema vetorial

$$C - \min_{x \in U} G(x).$$

Então  $C - \operatorname{argmin}\{G(x) : x \in U\}$  é não vazio se, e somente se,  $G(U)$  tem uma seção  $C$ -completa.

Para finalizar esta parte, recordam-se resultados que estendem o conceito de ponto crítico (ou estacionário), de convexidade e semi-continuidade inferior para funções vetoriais; define-se também o cone polar positivo de  $C$ , pois ferramentas e resultados essenciais em otimização vetorial estão associadas a ele. Para mais detalhes veja [27, 28, 30, 42].

**Definição 1.2.10.** Seja  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma função continuamente diferenciável. Um ponto  $\bar{x} \in \mathbb{R}^n$  é  $C$ -crítico se

$$\operatorname{Im}(\mathcal{J}G(\bar{x})) \cap (-\operatorname{int}(C)) = \emptyset, \quad (1.5)$$

onde  $\mathcal{J}G(\bar{x})$  denota a matriz Jacobiana da aplicação  $G$  no ponto  $\bar{x}$ .

Assim, se  $\bar{x}$  não é  $C$ -crítico, então existe uma direção  $\tilde{d} \in \mathbb{R}^n$  tal que

$$\mathcal{J}G(\bar{x})\tilde{d} \in -\operatorname{int}(C) \quad (1.6)$$

ou, equivalentemente,

$$\mathcal{J}G(\bar{x})\tilde{d} \prec_C 0. \quad (1.7)$$

*Observação 1.2.5.* A partir da definição acima, tem-se que:

- i) Para otimização escalar ( $m = 1$  e  $C = \mathbb{R}_+$ ), esta definição reduz-se à condição clássica de primeira ordem “gradiente igual a zero”.
- ii) Para o caso multiobjetivo,  $C = \mathbb{R}_+^m$ ,  $\bar{x}$  é conhecido na literatura como Pareto crítico e (1.5) reduz-se a:

Para todo  $d \in \mathbb{R}^n$ , existe  $i_0 = i_0(d) \in \{1, \dots, m\}$  tal que

$$\langle \nabla G_{i_0}(\bar{x}), d \rangle \geq 0.$$

e se  $\bar{x}$  não é Pareto crítico, então existe  $\tilde{d} \in \mathbb{R}^n$  tal que

$$\langle \nabla G_i(\bar{x}), \tilde{d} \rangle < 0, \quad \text{para todo } i = 1, \dots, m. \quad (1.8)$$

Além disso,  $G$  diferenciável implica que

$$\lim_{t \rightarrow 0} \frac{G_i(\bar{x} + t\tilde{d}) - G_i(\bar{x})}{t} = \langle \nabla G_i(\bar{x}), \tilde{d} \rangle < 0, \quad \text{para todo } i = 1, \dots, m. \quad (1.9)$$

Assim,  $\tilde{d}$  é uma direção descendente para  $G_i$  em  $\bar{x}$ ; isto é, existe  $t_0 > 0$  tal que

$$G_i(\bar{x} + t\tilde{d}) \leq G_i(\bar{x}) \quad \text{para todo } t \in (0, t_0]. \quad (1.10)$$

iii) Seguindo um raciocínio semelhante ao do item anterior, tem-se que:

Se  $\bar{x}$  não é  $C$ -crítico, então existe  $t_0 > 0$  tal que

$$G(\bar{x} + t\tilde{d}) \prec G(\bar{x}) \quad \text{para todo } t \in (0, t_0]. \quad (1.11)$$

Assim,  $\tilde{d}$  é uma direção  $C$ -descendente para  $G$  em  $\bar{x}$ .

**Definição 1.2.11.** Seja  $U \subseteq \mathbb{R}^n$  um conjunto convexo e  $G : U \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ .  $G$  é dita  $C$ -**convexa** se

$$G(\lambda x^1 + (1 - \lambda)x^2) \preceq_C \lambda G(x^1) + (1 - \lambda)G(x^2), \quad x^1, x^2 \in U, \quad \lambda \in [0, 1].$$

Em particular, se  $m = 1$  e  $C = \mathbb{R}_+$ , então obtém-se a definição usual de função convexa.

**Lema 1.2.2.** [42, Corolário 6.6] *Se  $C = \mathbb{R}_+^m$ , então  $G$  é  $C$ -convexa se, e somente se,  $G_i$  é convexa, para cada  $i = 1, \dots, m$ .*

O exemplo seguinte mostra que se  $C \neq \mathbb{R}_+^m$  e  $G_i$  é convexa para cada  $i = 1, \dots, m$ , então  $G$  não, necessariamente, é  $C$ -convexa.

**Exemplo 1.2.2.** [11] Seja  $U = \mathbb{R}^2$  e  $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \cup \{+\infty_C\}$  dada por

$$G(x) := \begin{bmatrix} x_1^2 - x_2 \\ x_2 \end{bmatrix}.$$

Então,

$$G(U) = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 : y_1 + y_2 \geq 0 \right\}.$$

Considerando os pontos  $x^1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  e  $x^2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ , e  $\lambda = \frac{1}{2}$  tem-se que

$$x_\lambda = \lambda x^1 + (1 - \lambda)x^2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

$$G(x^1) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad G(x^2) = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \quad G(x_\lambda) = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

$$\lambda G(x^1) + (1 - \lambda)G(x^2) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Ilustram-se todos estes pontos na Figura 8.

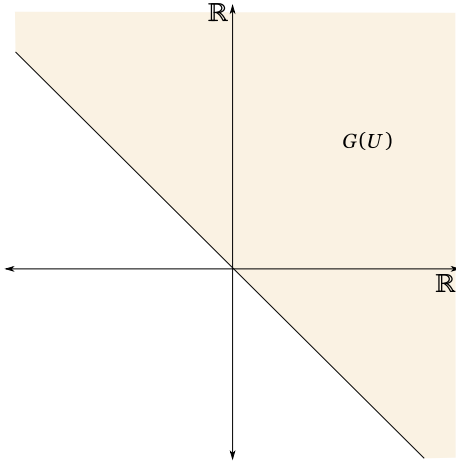


Figura 7: Imagem pela aplicação  $G$  do conjunto  $U$ .

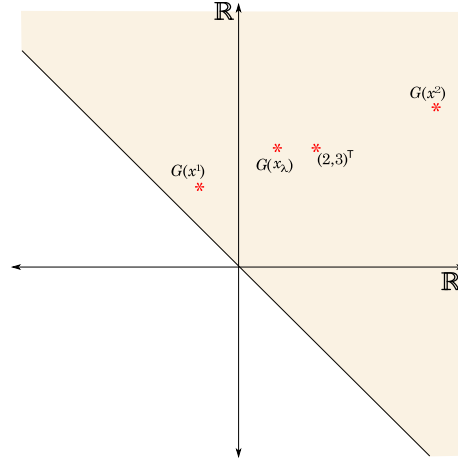


Figura 8.

Visto que  $G_1(x) = x_1^2 - x_2$  e  $G_2(x) = x_2$  são funções convexas, se  $C = C_1 = \mathbb{R}_+^2$ , veja Figura 9, então  $G$  é  $C_1$ -convexa, ou seja, os vetores  $\lambda G(x^1) + (1 - \lambda)G(x^2)$  pertencem ao cone trasladado  $G(x_\lambda) + C_1$ , para cada  $\lambda \in [0, 1]$ . Este fato é ilustrado, para  $\lambda = \frac{1}{2}$ , na Figura 10.

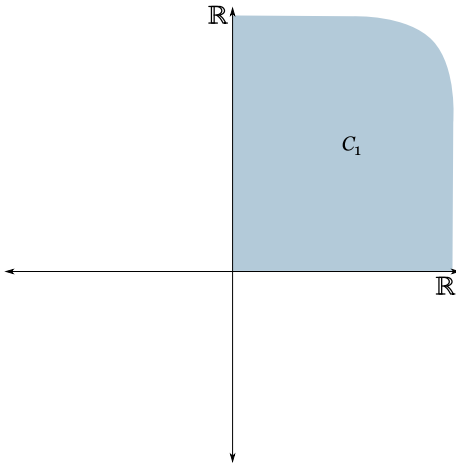


Figura 9.

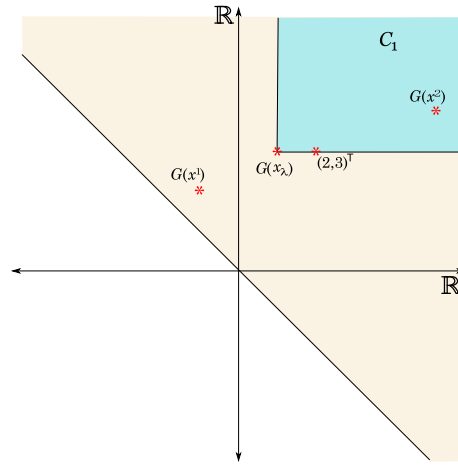


Figura 10.

Por outro lado, se  $C = C_2 = \text{cone} \left( \text{conv} \left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\} \right)$ , veja Figura 11, tem-se que

$$G(x_\lambda) \notin_{C_2} \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Portanto,  $G$  não é  $C_2$ -convexa. Este fato é ilustrado na Figura 12.



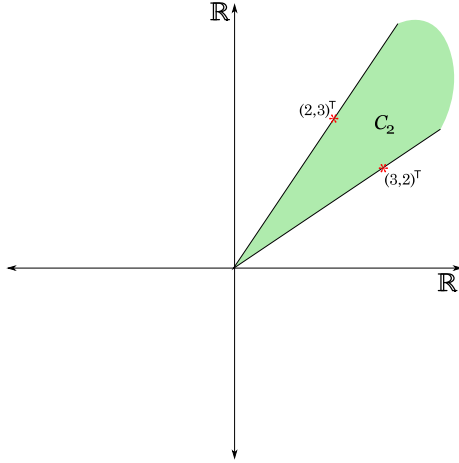


Figura 11.

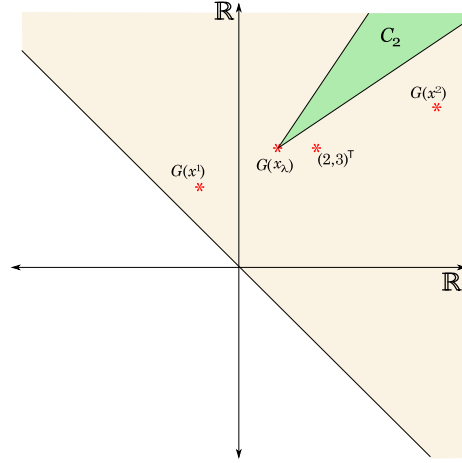


Figura 12.

**Lema 1.2.3.** [42, Proposição 6.3] *Seja  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  uma aplicação. Se  $G$  é  $C$ -convexa, então seus conjunto de nível são convexos, isto é,*

$$\{x \in \mathbb{R}^n : G(x) \preceq_C \bar{y}\},$$

*é convexo para qualquer  $\bar{y} \in \mathbb{R}^m$ .*

Assim como em otimização escalar, um problema vetorial é dito Problema Vetorial  $C$ -convexo, se seu conjunto viável é convexo e sua aplicação é  $C$ -convexa.

O resultado seguinte estabelece as relações existentes entre pontos eficientes e pontos estacionários.

**Teorema 1.2.2.** [21] *Sejam  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação continuamente diferenciável e  $C$  o cone que define a ordem parcial  $\preceq_C$ .*

- i) Se  $\bar{x}$  é um ponto eficiente fraco local, então  $\bar{x}$  é um ponto  $C$ -crítico de  $G$ .*
- ii) Se  $G$  é  $C$ -convexa e  $\bar{x} \in \mathbb{R}^n$  é um ponto  $C$ -crítico de  $G$ , então  $\bar{x}$  é um ponto eficiente fraco.*
- iii) Se  $C = \mathbb{R}_+^m$ ,  $G$  é duas vezes continuamente diferenciável,  $\nabla^2 G_i(x)$  é definida positiva, para todo  $i \in \{1, \dots, m\}$  e todo  $x \in \mathbb{R}^n$ , e se  $\bar{x} \in \mathbb{R}^n$  é um ponto Pareto crítico de  $G$ , então  $\bar{x}$  é um ponto eficiente.*

Nesta parte, considera-se  $C$  sendo um cone não vazio, convexo fechado e pontudo em  $\mathbb{R}^m$ , com interior não vazio, e serão revistas algumas definições e propriedades relacionadas a seu cone polar positivo.

**Definição 1.2.12.** *Seja  $C \subseteq \mathbb{R}^m$  um cone convexo. O cone polar positivo de  $C$ ,  $C^*$ , é dado por*

$$C^* := \{z \in \mathbb{R}^m : \langle y, z \rangle \geq 0, \quad y \in C\}.$$

*Observação 1.2.6.* Se  $C$  é um cone convexo, fechado e pontudo, com  $\text{int}(C) \neq \emptyset$ , então tem-se que:

- i)  $C^*$  é um cone convexo, fechado e pontudo;
- ii)  $(C^*)^* = C$ ;
- iii)  $C^* \cap (C \setminus \{0\}) \neq \emptyset$ ;
- iv)  $\text{int}(C^*) \neq \emptyset$ ;
- v)  $C^* = \mathbb{R}_+^m$ , quando  $C = \mathbb{R}_+^m$ .

Estende-se por continuidade, para cada  $z \in C^* \setminus \{0\}$ , a  $\mathbb{R}^m \cup \{+\infty_C\}$  através de  $\langle +\infty_C, z \rangle = +\infty$ .

O lema abaixo relaciona o cone  $C$  com  $C^*$  e  $\text{int}(C)$ .

**Lema 1.2.4.** [9, Lemma 2.2] *Seja  $C \subset \mathbb{R}^m$  um cone não vazio, convexo fechado e pontudo, com  $\text{int}(C) \neq \emptyset$ . Então, para todo  $e \in \text{int}(C)$  tem-se*

$$\inf\{\langle e, z \rangle : z \in C^*, \|z\| = 1\} = \text{dist}(e, \mathbb{R}^m \setminus C).$$

A propriedade seguinte relaciona aplicações  $C$ -convexas com o cone  $C^*$ .

**Lema 1.2.5.** [11] *Seja  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  uma aplicação.  $G$  é  $C$ -convexa se, e somente se,  $\langle G(\cdot), z \rangle$  é convexa para cada  $z \in C^*$ .*

**Definição 1.2.13.** [11] *Seja  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  uma aplicação. Diz-se que  $G$  é **positivamente semi-contínua inferior** se  $\langle G(\cdot), z \rangle$  é semi-contínua inferior, para cada  $z \in C^*$ .*

**Lema 1.2.6.** [42, Teorema 5.9] *Seja  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  uma aplicação. Se  $G$  é positivamente semi-contínua inferior, então seus conjuntos de nível são fechados, isto é,*

$$\{x \in \mathbb{R}^n : G(x) \preceq_C \bar{y}\},$$

*é fechado para qualquer  $\bar{y} \in \mathbb{R}^m$ .*

### 1.2.3 Escalarização

Nesta parte, são apresentadas algumas técnicas úteis para superar as dificuldades causadas pela não completude da ordem estabelecida no espaço objetivo  $\mathbb{R}^m$ . Estas técnicas, quase sempre, substituem o problema vetorial original por um outro problema, ou família de problemas, do tipo escalar. A principal vantagem desta substituição é poder usufruir da teoria e métodos já existentes em otimização escalar.

Considere o problema de otimização vetorial

$$(Pv) \quad C - \min_{x \in U} G(x),$$

onde  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$ ,  $C$  é o cone que define a ordem parcial em  $\mathbb{R}^m$  e  $U \subseteq \mathbb{R}^n$ , e o problema de otimização escalar

$$(PE) \quad \min_{x \in U} g(x),$$

onde  $g : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ .

**Definição 1.2.14.** [42, Definição 2.1] Diz-se que

i)  $(PE)$  é uma **representação escalar** de  $(Pv)$ , com respeito a  $C$ , se para cada  $x$  e  $\tilde{x} \in U$ ,

$$G(x) \preceq_C G(\tilde{x}) \quad \text{implica} \quad g(x) \leq g(\tilde{x}), \quad \text{e}$$

$$G(x) \preceq_C G(\tilde{x}) \quad \text{e} \quad G(x) \neq G(\tilde{x}) \quad \text{implicam} \quad g(x) < g(\tilde{x});$$

ii)  $(PE)$  é uma **representação escalar estrita** de  $(Pv)$ , com respeito a  $C$ , se para cada  $x$  e  $\tilde{x} \in U$ ,

$$G(x) \preceq_C G(\tilde{x}) \quad \text{implica} \quad g(x) \leq g(\tilde{x}), \quad \text{e}$$

$$G(x) \prec_C G(\tilde{x}) \quad \text{implica} \quad g(x) < g(\tilde{x});$$

iii)  $(PE)$  é uma **representação escalar fraca** de  $(Pv)$ , com respeito a  $C$ , se para cada  $x$  e  $\tilde{x} \in U$ ,

$$G(x) \prec_C G(\tilde{x}) \quad \text{implica} \quad g(x) < g(\tilde{x}).$$

*Observação 1.2.7.* Note da definição acima que

$$i) \Rightarrow ii) \Rightarrow iii)$$

**Proposição 1.2.2.** [42, Proposição 2.2] *Sejam os problemas  $(Pv)$  e  $(PE)$ . As seguintes propriedades são satisfeitas:*

i) *Se  $(PE)$  é uma representação escalar de  $(Pv)$ , então*

$$\arg \min\{g(x) : x \in U\} \subseteq C - \operatorname{argmin}\{G(x) : x \in U\};$$

ii) Se  $(PE)$  é uma representação escalar fraca de  $(Pv)$ , então

$$\arg \min\{g(x) : x \in U\} \subseteq C - \operatorname{argmin}_w\{G(x) : x \in U\}.$$

*Observação 1.2.8.* [42] Sejam  $C$  um cone convexo, fechado e pontudo,  $C^*$  o cone polar positivo associado a  $C$  e  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  uma aplicação. Dados  $x, \tilde{x} \in \operatorname{dom}(G)$ :

Se  $G(x) \preceq_C G(\tilde{x})$ , então, da definição de  $\preceq_C$  tem-se que  $G(\tilde{x}) - G(x) \in C$ .

Logo, da definição de  $C^*$  segue que, para todo  $z \in C^*$

$$\langle G(\tilde{x}) - G(x), z \rangle \geq 0,$$

ou, equivalentemente,

$$\langle G(\tilde{x}), z \rangle \geq \langle G(x), z \rangle.$$

Além disso, se  $G(x) \prec_C G(\tilde{x})$ , então  $G(\tilde{x}) - G(x) \in \operatorname{int}(C)$ . Assim, para todo  $z \in C^* \setminus \{0\}$

$$\langle G(\tilde{x}) - G(x), z \rangle > 0 \quad \Rightarrow \quad \langle G(\tilde{x}), z \rangle > \langle G(x), z \rangle.$$

Portanto, o problema

$$\min_{x \in U} \langle G(x), z \rangle$$

é uma representação escalar estrita do problema vetorial

$$C - \min_{x \in U} G(x),$$

para cada  $z \in C^* \setminus \{0\}$ . Esta escalarização é conhecida na literatura como escalarização linear.

O resultado seguinte relaciona o conjunto solução do problema vetorial  $C$ -convexo e o conjunto solução da sua escalarização linear.

**Teorema 1.2.3.** [11] *Suponha que  $U \subseteq \mathbb{R}^m$  é um conjunto convexo, a aplicação  $G : U \rightarrow \mathbb{R}^m \cup \{+\infty_C\}$  é própria e  $C$ -convexa. Então, dado o problema vetorial*

$$C - \min_{x \in U} G(x),$$

tem-se que

$$C - \operatorname{argmin}_w\{G(x) : x \in U\} = \bigcup_{z \in C^* \setminus \{0\}} \operatorname{argmin}\{\langle G(x), z \rangle : x \in U\}.$$

Note que o conjunto  $\arg \min\{\langle G(x), z \rangle : x \in U\}$  pode ser vazio para algum  $z \in C^*$ , o exemplo seguinte ilustra este fato.

**Exemplo 1.2.3.** [11] Seja  $U = \mathbb{R}^2$ ,  $C = \mathbb{R}_+^2$  e  $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \cup \{+\infty_C\}$  dada por

$$G(x) := \begin{bmatrix} x_1^2 - x_2 \\ x_2 \end{bmatrix}.$$

No Exemplo 1.2.2 vê-se que  $G$  é  $\mathbb{R}_+^2$ -convexa, e desde que  $(\mathbb{R}_+^2)^* = \mathbb{R}_+^2$  obtém-se que

$$\mathbb{R}_+^2 - \arg \min_w \{G(x) : x \in U\} = \bigcup_{z \in \mathbb{R}_+^2 \setminus \{0\}} \arg \min\{\langle G(x), z \rangle : x \in U\}.$$

Porém, desde que

$$G(U) = \left\{ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 : y_1 + y_2 \geq 0 \right\},$$

tem-se que para cada  $z = (z_1, z_2) \in \mathbb{R}_+^2$ , com  $z_1 \neq z_2$ , a função

$$z_1 x_1^2 + (z_2 - z_1)x_2,$$

é ilimitada inferiormente. Implica que  $\arg \min\{\langle G(x), z \rangle : x \in U\}$  é vazio, para cada  $z = (z_1, z_2) \in \mathbb{R}_+^2$ , com  $z_1 \neq z_2$ . Portanto,

$$\mathbb{R}_+^2 - \arg \min_w \{G(x) : x \in U\} = \arg \min\{x_1^2 : x \in U\} = \{0\} \times \mathbb{R}.$$

Por outro lado, Graña *et al.* em [29], supondo que  $G(U)$  é fechado e  $C$  um cone pontudo, estabeleceram o seguinte resultado.

**Proposição 1.2.3.** [29, Corolário 1] *Seja  $G : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Suponha que  $G(U)$  é fechado. Se*

$$C^* \cap \text{int}([0^+(\text{cl}(\text{conv}(G(U))))]^*) \neq \emptyset, \quad (1.12)$$

então

$$(Pv) \quad C - \min_{x \in U} G(x)$$

tem soluções eficientes fracas. Além disso, para qualquer  $z \neq 0$  na interseção acima, o problema escalar

$$\min_{x \in U} \langle G(x), z \rangle$$

tem solução ótima, a qual é uma solução eficiente fraca para  $(Pv)$ .

*Observação 1.2.9.* No exemplo acima,  $G(U)$  é um semi-espaco fechado, logo

$$\text{int}([0^+(\text{cl}(\text{conv}(G(U))))]^*) = \emptyset.$$

Assim, a intersecção em (1.12) é sempre vazia, para qualquer cone. Porém, nesse exemplo obteve-se

$$\mathbb{R}_+^2 - \arg \min_w \{G(x) : x \in U\} \neq \emptyset.$$

**Afirmção 1.2.4.** *Sejam  $A \subseteq \mathbb{R}^m$  e  $C$  um cone convexo. Se  $A$  é  $C$ -limitado inferior, então para cada  $z \in C^*$*

$$-\infty < \inf \{ \langle w, z \rangle : w \in A \}.$$

*Demonstração.* Visto que  $A$  é  $C$ -limitado inferior, existe  $\bar{y} \in \mathbb{R}^m$  tal que

$$A - \bar{y} \subset C. \quad (1.13)$$

Por outro lado, para cada  $z \in C^*$ , tem-se que

$$0 \leq \langle y, z \rangle, \quad \text{para todo } y \in C. \quad (1.14)$$

Portanto, combinando (1.13) e (1.14) segue o resultado.  $\square$

A seguir apresenta-se um outro exemplo de escalarização, para o problema de otimização multiobjetivo.

**Proposição 1.2.4.** *Seja  $G : U \rightarrow \mathbb{R}^m$  uma aplicação e considere o problema de otimização multiobjetivo*

$$\mathbb{R}_+^m - \min_{x \in U} G(x). \quad (1.15)$$

*Então*

$$\arg \min \{ \phi(x) : x \in U \} \subseteq \mathbb{R}_+^m - \operatorname{argmin}_w \{ G(x) : x \in U \}, \quad (1.16)$$

*onde*

$$\phi(x) := \max_{i=1, \dots, m} G_i(x).$$

*Demonstração.* Considere  $\tilde{x} \in \arg \min \{ \phi(x) : x \in U \}$ , e suponha, por contradição, que  $\tilde{x} \notin \mathbb{R}_+^m - \operatorname{argmin}_w \{ G(x) : x \in U \}$ .

Desde que  $\tilde{x} \in U$  não é Pareto fraco de (1.15),  $\exists \hat{x} \in \mathbb{R}^n$  tal que

$$G_i(\hat{x}) < G_i(\tilde{x}), \quad \text{para cada } i = 1 \dots, m.$$

Logo, considerando o máximo sobre cada lado da desigualdade, obtém-se

$$\max_{i=1, \dots, m} \{ G_i(\hat{x}) \} < \max_{i=1, \dots, m} \{ G_i(\tilde{x}) \},$$

ou, equivalentemente,

$$\phi(\hat{x}) < \phi(\tilde{x}).$$

Portanto, chega-se a uma contradição. Logo,

$$\tilde{x} \in \mathbb{R}_+^m - \operatorname{argmin}_w \{G(x) : x \in U\}.$$

□

# Capítulo 2

## Distância Proximal

Neste capítulo são recordadas definições de distâncias proximais e distâncias proximais induzidas, apresentadas por Auslender e Teboulle em [5]. Além disso, como primeira contribuição, será introduzida uma subclasse da distância proximal induzida, a qual permite enfraquecer a condição de convergência dada em [5].

**Definição 2.0.1.** Sejam  $S$  um conjunto não vazio, aberto e convexo de  $\mathbb{R}^n$  e  $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$  uma função. Diz-se que  $d$  é uma **distância proximal**, com respeito a  $S$ , se, para cada  $y \in S$ ,  $d$  verifica as seguintes condições:

- (P<sub>1</sub>)  $d(\cdot, y)$  é própria, semi-contínua inferior, convexa, e continuamente diferenciável sobre  $S$ ;
- (P<sub>2</sub>)  $\text{dom } d(\cdot, y) \subset \text{cl}(S)$  e  $\text{dom } \nabla_1 d(\cdot, y) = S$ , onde  $\nabla_1 d(\cdot, y)$  denota o gradiente da função  $d(\cdot, y)$  com respeito à primeira variável;
- (P<sub>3</sub>)  $d(\cdot, y)$  é coerciva em  $\mathbb{R}^n$ , isto é,  $\lim_{\|x\| \rightarrow \infty} d(x, y) = +\infty$ ;
- (P<sub>4</sub>)  $d(y, y) = 0$ .

Como em [5], neste trabalho também denota-se por  $\mathcal{D}(S)$  à família de funções  $d$  que satisfazem a Definição 2.0.1.

A próxima definição associa a cada distância proximal  $d \in \mathcal{D}(S)$  uma outra função satisfazendo algumas propriedades desejáveis.

**Definição 2.0.2.** Sejam  $S$  um conjunto não vazio, aberto e convexo de  $\mathbb{R}^n$ ,  $d \in \mathcal{D}(S)$  e  $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$  uma função. Diz-se que  $H$  é uma **distância proximal induzida** por  $d$ , se  $H$  verifica as seguintes condições:

- (P<sub>5</sub>)  $H$  tem valores finitos sobre  $\text{cl}(S) \times S$ ;
- (P<sub>6</sub>)  $H(y, y) = 0, \quad y \in S$ ;
- (P<sub>7</sub>)  $\langle x - z, \nabla_1 d(z, y) \rangle \leq H(x, y) - H(x, z), \quad y, z \in S, x \in \text{cl}(S)$ ;



(P<sub>8</sub>)  $H(x, \cdot)$  é nível limitado sobre  $S$ , para todo  $x \in \text{cl}(S)$ .

Denota-se por  $(d, H) \in \mathcal{F}(\text{cl}(S))$  para especificar que o par  $(d, H)$  verifica as premissas da Definição 2.0.2 com respeito a  $\text{cl}(S)$ .

A definição seguinte diz respeito a uma subclasse de distâncias proximais induzidas, a qual é útil para obter a convergência das sequências geradas pelo método proximal apresentado no próximo capítulo. Note que esta definição é uma contribuição à teoria de distâncias proximais, pois ela enfraquece a condição de convergência ( $a_2$ ) dada em [5], isto é,

$$\forall \bar{y} \in \text{cl}(S) \text{ e } \forall \{y^k\} \subset S \text{ convergindo para } \bar{y}, \text{ tem-se que } \lim_{k \rightarrow +\infty} H(\bar{y}, y^k) = 0.$$

**Definição 2.0.3.** Sejam  $S$  um conjunto não vazio, aberto e convexo de  $\mathbb{R}^n$  e  $(d, H) \in \mathcal{F}(\text{cl}(S))$ . Denota-se por  $\mathcal{F}^*(\text{cl}(S))$  o conjunto dos pares  $(d, H)$  tais que  $H$  satisfaz às seguintes condições:

(P<sub>9</sub>) Se  $\{y^k\} \subset S$  e  $\bar{y} \in \text{cl}(S)$  tais que  $\lim_{k \rightarrow +\infty} H(\bar{y}, y^k) = 0$ , então  $\lim_{k \rightarrow +\infty} y^k = \bar{y}$ ;

(P<sub>10</sub>) Se  $\{y^k\} \subset S$  converge para  $\bar{y}$ , então pelo menos uma das seguintes possibilidades é verificada:

$$\lim_{k \rightarrow +\infty} H(\bar{y}, y^k) = 0, \tag{2.1a}$$

$$\lim_{k \rightarrow +\infty} H(y, y^k) = +\infty, \quad y \neq \bar{y}, \quad y \in \text{cl}(S) \tag{2.1b}$$

Vale ressaltar que esta definição foi introduzida por Kaplan e Tichatschke em [38] para funções tipo Bregman.

Note que no contexto de distâncias proximais apresentado por Auslender e Teboulle em [5], duas escolhas conhecidas para  $d$  incluem distâncias de Bregman ou  $\varphi$ -divergências, porém nesse trabalho, bem como em outros existentes na literatura, veja por exemplo [6, 14, 41, 53, 54], o conjunto  $\text{cl}(S)$  considerado é um conjunto poliedral e em particular  $\text{cl}(S) = \mathbb{R}_+^n$ . Desta maneira, a Definição 2.0.3 é uma alternativa interessante para a teoria de distâncias proximais, pois possibilita o trabalho sobre conjuntos não necessariamente poliedrais, como exemplo pode-se considerar as distâncias tipo Bregman, apresentadas em [38–40].

# Capítulo 3

## Método Interior Proximal para o Problema de Otimização Vetorial $K$ -convexo

Neste capítulo estabelece-se o primeiro problema a ser abordado e propõe-se um método interior proximal exato e sua versão inexata para determinar soluções eficientes fracas. Além disso, seguindo as ideias de Bonnel *et al.* em [11], uma análise completa de convergência é apresentada. Vale ressaltar que este texto corresponde a um artigo aceito para publicação [55].

### 3.1 O problema

O problema sob estudo é o *Problema de Otimização Vetorial  $K$ -convexo*:

$$(POVC) \quad K - \min_{x \in \Omega} F(x),$$

onde

- i)  $\Omega \subseteq \text{dom}(F)$  é um conjunto convexo e fechado, com interior não vazio;
- ii)  $K$  é um cone convexo, fechado e pontudo, com  $\text{int}(K) \neq \emptyset$ , o qual define a ordem parcial  $\preceq_K$ ;
- iii) A aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_K\}$  é própria,  $K$ -convexa e positivamente semi-contínua inferior.

O principal interesse neste capítulo é encontrar *soluções eficientes fracas* para (POVC), isto é:

Encontrar  $x^* \in \Omega$  tal que

$$\nexists x \in \Omega : F(x) \prec_K F(x^*).$$

Assim, para estabelecer o método interior proximal exato, denotado por MIPE, e a sua respectiva versão inexata, denotado por MIPI, precisa-se revisar algumas propriedades referentes ao conjunto de nível da aplicação objetivo em um ponto viável.

*Observação 3.1.1.* [11] Considere  $u \in \Omega$  e o conjunto de nível  $\Omega_u$ , definido por

$$\Omega_u := \{x \in \Omega : F(x) \preceq_K F(u)\}.$$

Então:

- i)  $\emptyset \neq \Omega_u \subseteq \Omega$ ;
- ii)  $u \in \Omega_u \cap \Omega$ ;
- iii)  $F$   $K$ -convexa implica que o conjunto  $\Omega_u$  é convexo;
- iv)  $F$  positivamente semi-contínua inferior implica que o conjunto  $\Omega_u$  é fechado;
- v) Se  $u \in \text{int}(\Omega)$ , então os itens (i)-(iv) e Lema 1.1.2 implicam que  $\text{ir}(\Omega_u) \subseteq \text{int}(\Omega)$ ;
- vi)  $F(\Omega_u) = (F(u) - K) \cap F(\Omega)$ .

**Proposição 3.1.1.** *Seja  $d \in \mathcal{D}(\text{int}(\Omega))$  e, para cada  $u \in \text{int}(\Omega)$  considere o seguinte problema de otimização vetorial*

$$K - \min\{F(x) + d(x, u)e : x \in \Omega_u\}, \quad (3.1)$$

onde  $e \in \text{int}(K)$ . Se uma das condições seguintes é verificada

H1)  $F(\Omega_u)$  é um conjunto limitado;

H2)  $d(\cdot, u)$  é 1-coerciva em  $\mathbb{R}^n$ .

Então

$$K - \text{argmin}_w\{F(x) + d(x, u)e : x \in \Omega_u\}$$

é não vazio e um subconjunto de  $\text{int}(\Omega)$ .

*Demonstração.*

• Se ocorre a condição H1, então de  $\dim(F(\Omega_u)) \leq \dim(K)$  e do item *i*) da Observação 1.2.2, tem-se que

$$F(\Omega_u) \text{ é } K\text{-limitado inferior.}$$

Assim, pela Afirmação 1.2.4, para cada  $z \in K^*$ , existe  $M_z \in \mathbb{R}$  tal que

$$M_z < \langle F(x), z \rangle, \quad x \in \Omega_u. \quad (3.2)$$

Agora considere o seguinte problema escalar:

$$\min\{F_z(x) + \langle e, z \rangle d(x, u) : x \in \Omega_u\}, \quad (3.3)$$

onde  $F_z = \langle F(\cdot), z \rangle : \Omega \rightarrow (-\infty, +\infty]$ . Então, desde que

$$M_z + \langle e, z \rangle d(x, u) \leq F_z(x) + \langle e, z \rangle d(x, u), \quad x \in \Omega_u, \quad (3.4)$$

e os conjuntos de nível de  $d(\cdot, u)$  são limitados, tem-se que os conjuntos de nível de  $F_z + \langle e, z \rangle d(\cdot, u)$  são também limitados.

Por outro lado,  $(P_1)$  da Definição 2.0.1 implica que  $d(\cdot, u)$  é própria, semi-contínua inferior, convexa, e continuamente diferenciável sobre  $\text{int}(\Omega)$ . Mais ainda,  $e \in \text{int}(K)$  e  $z \in K^* \setminus \{0\}$  implicam que  $\langle e, z \rangle > 0$ . Logo,  $\langle e, z \rangle d(\cdot, u)$  é semi-contínua inferior e convexa.

Além disso, a definição de  $(POVC)$  implica que  $F_z$  é semi-contínua inferior e convexa, para cada  $z \in K^* \setminus \{0\}$ . Logo,  $F_z + \langle e, z \rangle d(\cdot, u)$  é semi-contínua inferior e convexa.

Em consequência, a minimização em (3.3) se reduz à minimização sobre um conjunto compacto e o mínimo é atingido. Além disso, como  $\Omega_u$  é um conjunto convexo e a aplicação  $F + d(\cdot, u)e$  é própria e  $C$ -convexa, então as hipóteses do Teorema 1.2.3 são verificadas por  $F + d(\cdot, u)e$ , portanto

$$K\text{-argmin}_w\{F(x) + d(x, u)e : x \in \Omega_u\} = \bigcup_{z \in K^* \setminus \{0\}} \arg \min\{F_z(x) + \langle e, z \rangle d(x, u) : x \in \Omega_u\}.$$

Logo

$$K\text{-argmin}_w\{F(x) + d(x, u)e : x \in \Omega_u\} \neq \emptyset. \quad (3.5)$$

Agora mostra-se que todas as soluções eficientes fracas de  $F + d(\cdot, u)e$  sobre  $\Omega_u$  pertencem a  $\text{int}(\Omega)$ . Com efeito, suponha que  $\exists \hat{z} \in K^* \setminus \{0\}$  tal que

$$\emptyset \neq \arg \min\{F_{\hat{z}}(x) + \langle e, \hat{z} \rangle d(x, u) : x \in \Omega_u\} \not\subseteq \text{int}(\Omega). \quad (3.6)$$

Assim, a condição de otimalidade de primeira ordem do problema escalar acima, para cada  $x \in \arg \min\{F_{\hat{z}}(x) + \langle e, \hat{z} \rangle d(x, u) : x \in \Omega_u\}$ , é

$$0 \in \partial(F_{\hat{z}} + \langle e, \hat{z} \rangle d(\cdot, u) + I_{\Omega_u})(x). \quad (3.7)$$

Além disso, as definições de  $F$  e  $d(\cdot, u)$ , e o item v) da Observação 3.1.1 implicam que

$$\text{ir}(\Omega_u) \cap \text{int}(\text{dom } F_{\hat{z}}) \cap \text{int}(\text{dom } d(\cdot, u)) \neq \emptyset.$$

Nestas condições, o subdiferencial da soma de funções é igual a soma dos subdiferenciais, veja por exemplo [50, Teorema 23.8], ou seja,

$$\partial(F_{\hat{z}} + \langle e, \hat{z} \rangle d(\cdot, u) + I_{\Omega_u})(x) = \partial F_{\hat{z}}(x) + \langle e, \hat{z} \rangle \partial_1 d(x, u) + N_{\Omega_u}(x), \quad (3.8)$$

para cada  $x \in \arg \min\{F_{\hat{z}}(x) + \langle e, \hat{z} \rangle d(x, u) : x \in \Omega_u\}$ . Assim, de (3.7) e (3.8) obtém-se que

$$0 \in \partial F_{\hat{z}}(x) + \langle e, \hat{z} \rangle \partial_1 d(x, u) + N_{\Omega_u}(x), \quad x \in \Omega. \quad (3.9)$$

por quanto que  $\text{dom}(\partial_1 d(\cdot, u)) = \text{int}(\Omega)$  conclui-se que  $x \in \text{int}(\Omega)$  e  $\partial_1 d(x, u) = \nabla_1 d(x, u)$ . Por conseguinte, todos os zeros de  $\partial F_{\hat{z}} + \langle e, \hat{z} \rangle \nabla_1 d(\cdot, u) + N_{\Omega_u}$  pertencem ao  $\text{int}(\Omega)$ , o que contradiz (3.6). Portanto segue o resultado.

- Se ocorre a condição H2, é suficiente mostrar que, para qualquer  $z \in K^* \setminus \{0\}$ , todos os conjuntos de nível de  $F_z + \langle e, z \rangle d(\cdot, u)$  são limitados. Os outros resultados são provados analogamente.

De fato, suponha por contradição que existe  $\tilde{\delta}$  tal que

$$\mathcal{L}(F_z + \langle e, z \rangle d(\cdot, u), \tilde{\delta}) := \{x \in \Omega : F_z(x) + \langle e, z \rangle d(x, u) \leq \tilde{\delta}\} \neq \emptyset \text{ é ilimitado.}$$

Em outras palavras, existe uma sequência  $\{y^k\} \subset \mathcal{L}(F_z + \langle e, z \rangle d(\cdot, u), \tilde{\delta})$  tal que

$$\lim_{k \rightarrow \infty} \|y^k\| = +\infty,$$

e, sem perda de generalidade, pode-se supor que  $\|y^k\| \neq 0$  para todo  $k \in \mathbb{N}$ .

Uma vez que  $\{y^k\} \subset \mathcal{L}(F_z + \langle e, z \rangle d(\cdot, u), \tilde{\delta})$  e  $F_z$  é convexa, segue que

$$\tilde{\delta} \geq F_z(y^k) + \langle e, z \rangle d(y^k, u) \quad (3.10)$$

$$\geq F_z(u) + \langle \zeta, y^k - u \rangle + \langle e, z \rangle d(y^k, u) \quad (3.11)$$

$$\geq F_z(u) - \langle \zeta, u \rangle - \|\zeta\| \|y^k\| + \langle e, z \rangle d(y^k, u), \quad (3.12)$$

$$\forall k \in \mathbb{N}, \quad \forall \zeta \in \partial F_z(u).$$

Dividindo cada extremo da desigualdade acima por  $\|y^k\|$ , obtém-se

$$\frac{F_z(u) - \langle \zeta, u \rangle}{\|y^k\|} - \|\zeta\| + \langle e, z \rangle \frac{d(y^k, u)}{\|y^k\|} \leq \frac{\tilde{\delta}}{\|y^k\|}, \quad (3.13)$$

$$\forall k \in \mathbb{N}, \quad \forall \zeta \in \partial F_z(u).$$

Assim, o lado direito vai para zero, enquanto o lado esquerdo vai para  $+\infty$ , quando  $k$  vai para  $+\infty$ , pois  $d(\cdot, u)$  é 1-coerciva e  $\langle e, z \rangle > 0$ . Chega-se a uma contradição, o que prova a afirmação.  $\square$

*Observação 3.1.2.* Pela Afirmação 1.2.3 nota-se que a condição H1 é verificada, para todo  $x \in \Omega$ , se existe um semi-espço fechado  $\mathcal{H}$  tal que

$$(-K) \cap \mathcal{H} = \{0\} \text{ e } F(\Omega) \text{ é } \mathcal{H}\text{-limitado inferior.}$$

Além disso, para problemas de otimização multiobjetivo,  $K = \mathbb{R}_+^m$  esta condição pode ser considerada verdadeira sem perda de generalidade. Caso contrário, pela Afirmação 1.2.3, substitua  $F$  por  $(\exp(F_1), \dots, \exp(F_m))^T$ . Então considere o seguinte problema de otimização multiobjetivo:

$$\mathbb{R}_+^m - \min\{(\exp(F_1(x)), \dots, \exp(F_m(x)))^T : x \in \Omega\}. \quad (3.14)$$

Por outro lado, no caso geral, se a condição H1 não é verificada, então pode-se ser um pouco mais restrito com a distância proximal, isto é, pode-se considerar uma distância proximal sendo 1-coerciva, o que em geral é uma hipótese menos forte que exigir que ela seja fortemente convexa, como acontece nas distâncias proximais usadas em outros trabalhos, veja por exemplo [11, 13, 16].

*Observação 3.1.3.* Sejam  $u^*$  e  $\hat{u} \in \Omega$ . Se  $u^*$  verifica a condição H1 e  $F(\hat{u}) \preceq_K F(u^*)$ , então  $\hat{u}$  também verifica a condição H1.

## 3.2 Um método interior proximal exato

O método seguinte, denotado por “MIPE”, gera uma sequência  $\{x^k\} \subset \text{int}(\Omega)$  correspondente à seguinte recursão:

Dada a iterada atual  $x^k$  ( $x^0 \in \text{int}(\Omega)$ ), encontrar  $x^{k+1} \in \text{int}(\Omega)$  tal que

$$x^{k+1} \in K - \operatorname{argmin}_w \{F(x) + \beta_k d(x, x^k) e_k : x \in \Omega_k\}, \quad (3.15)$$

onde  $d$  é uma distância proximal com respeito a  $\text{int}(\Omega)$ ,  $0 < \beta_k < \tilde{\beta}$  para algum  $\tilde{\beta} > 0$ ,  $e_k \in \text{int}(K)$  com  $\|e_k\| = 1$ , e  $\Omega_k := \{x \in \Omega : F(x) \preceq_K F(x^k)\}$ .

O corolário seguinte é uma consequência imediata da Proposição 3.1.1.

**Corolário 3.2.1.** *Suponha que  $F(\Omega_0)$  satisfaz à condição H1 ou  $d(\cdot, x)$  satisfaz à condição H2, para cada  $x \in \text{int}(\Omega)$ . Então a recursão gerada pelo MIPE está bem definida.*

Agora, para estabelecer a limitação da sequência gerada pelo MIPE, suponha-se a seguinte condição utilizada em [11, 13, 15, 16].

**Hipótese 3.2.1.** O conjunto  $F(\Omega_0)$  é  $K$ -completo, onde  $F(\Omega_0)$  é a seção inicial de  $F(\Omega)$  em  $x^0$ .

A partir desta hipótese tem-se que o conjunto de limites inferiores da seção inicial, denotado por  $E$ , é não vazio, isto é

$$E := \{x \in \Omega : F(x) \preceq_K F(x^k), \quad k \in \mathbb{N}\} \neq \emptyset.$$

**Proposição 3.2.1.** *Sejam  $H$  uma distância proximal induzida por  $d$  e  $\{x^k\}$  uma sequência gerada pelo MIPE. Se  $(d, H) \in \mathcal{F}(\Omega)$  e a Hipótese 3.2.1 acima é satisfeita, então*

- i) para cada  $\bar{x} \in E$ ,  $\{H(\bar{x}, x^k)\}$  é uma sequência convergente;*
- ii)  $\{x^k\}$  é uma sequência limitada e todos seus pontos de acumulação são soluções eficientes fracas de (POVC).*

*Demonstração.* *i)* Visto que  $x^{k+1} \in K - \text{argmin}_w \{F(x) + \beta_k d(x, x^k) e_k : x \in \Omega_k\}$ , do Teorema 1.2.3 existe  $z_k \in K^* \setminus \{0\}$  tal que  $x^{k+1}$  é solução do seguinte problema:

$$\min \{F_k(x) + \beta_k \langle e_k, z_k \rangle d(x, x^k) : x \in \Omega_k\}, \quad (3.16)$$

onde  $F_k := \langle F(\cdot), z_k \rangle$ . Mais ainda, sem perda de generalidade, suponha que  $\|z_k\| = 1$  para todo  $k \in \mathbb{N}$ , já que a solução deste problema não é alterada se sua função objetivo é multiplicada por um escalar positivo.

Então, como  $\text{int}(\text{dom}(F_k)) \cap \text{int}(\text{dom}(d(\cdot, x^k))) \cap \text{ir}(\Omega_k) \neq \emptyset$ , segue do Lema 1.1.3 que  $x^{k+1}$  satisfaz à condição de primeira ordem para o problema (3.16), isto é, existe  $u_k \in \mathbb{R}^n$  tal que

$$u_k \in \partial(F_k + \beta_k \langle e_k, z_k \rangle d(\cdot, x^k))(x^{k+1}), \quad (3.17)$$

e

$$0 \leq \langle x - x^{k+1}, u_k \rangle, \quad (3.18)$$

para todo  $x \in \Omega_k$ . De (3.17),  $x^{k+1} \in \text{int}(\Omega)$ , e do fato que  $\text{dom} \partial_1 d(\cdot, x^k) = \text{int}(\Omega)$ , resulta que

$$u_k = v_k + \beta_k \langle e_k, z_k \rangle \nabla_1 d(x^{k+1}, x^k), \quad (3.19)$$

para algum

$$v_k \in \partial F_k(x^{k+1}). \quad (3.20)$$

Além do mais, da definição de  $E$  tem-se, para todo  $k \in \mathbb{N}$ , que  $E \subseteq \Omega_k$ . Considerando  $\bar{x} \in E$  e fazendo  $x = \bar{x}$  em (3.18), junto com (3.19), implicam em

$$0 \leq \langle \bar{x} - x^{k+1}, v_k \rangle + \beta_k \langle e_k, z_k \rangle \langle \bar{x} - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle. \quad (3.21)$$

Além disso, da convexidade de  $F_k$  e (3.20) tem-se, para  $\bar{x} \in E$ , que

$$\langle \bar{x} - x^{k+1}, v_k \rangle \leq F_k(\bar{x}) - F_k(x^{k+1}) \stackrel{\text{def}}{=} \langle F(\bar{x}) - F(x^{k+1}), z_k \rangle. \quad (3.22)$$

Dado que  $z_k \in K^* \setminus \{0\}$  e  $F(\bar{x}) - F(x^{k+1}) \preceq_K 0$  implicam em

$$\langle F(\bar{x}) - F(x^{k+1}), z_k \rangle \leq 0.$$

Juntando com o fato de  $\beta_k \langle e_k, z_k \rangle > 0$ , (3.21) e (3.22), fornecem

$$0 \leq \langle \bar{x} - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle. \quad (3.23)$$

Por outro lado, fazendo  $y = x^k$ ,  $z = x^{k+1}$  e  $x = \bar{x}$ , no item  $(P_7)$  da Definição 2.0.2, pois  $(d, H) \in \mathcal{F}(\Omega)$ , obtém-se

$$\langle \bar{x} - x^{k+1}, \nabla_1 d(x^k, x^{k+1}) \rangle \leq H(\bar{x}, x^k) - H(\bar{x}, x^{k+1}),$$

que, junto com (3.23), fornece

$$H(\bar{x}, x^{k+1}) \leq H(\bar{x}, x^k).$$

Logo, para todo  $\bar{x} \in E$ ,  $\{H(\bar{x}, x^k)\}$  é uma sequência decrescente e não negativa. Portanto,  $\{H(\bar{x}, x^k)\}$  é convergente, além disso

$$\lim_{k \rightarrow +\infty} \langle \bar{x} - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle = 0. \quad (3.24)$$

*ii)* Inicialmente, observa-se que o item  $(P_8)$  da Definição 2.0.2 e a convergência de  $\{H(\bar{x}, x^k)\}$  implicam que  $\{x^k\}$  é uma sequência limitada. Assim, a existência de pontos de acumulação de  $\{x^k\}$  é garantida. Agora mostra-se que todos são soluções eficiente fracas de  $(POVC)$ .



Seja  $\hat{x}$  um ponto de acumulação de  $\{x^k\}$  e seja  $\{x^{j_k}\}$  a subsequência de  $\{x^k\}$  convergindo para  $\hat{x}$ . Afirma-se que  $\hat{x} \in E$  ou, equivalentemente,

$$F(\hat{x}) \preceq_K F(x^k), \quad k \in \mathbb{N}. \quad (3.25)$$

De fato, defina a função  $F_z : \mathbb{R}^n \rightarrow (-\infty, +\infty]$  por

$$F_z(x) := \langle F(x), z \rangle, \quad \text{com } z \in K^*.$$

Uma vez que  $F$  é positivamente semi-contínua inferior e  $K$ -convexa implicam que  $F_z$  é semi-contínua inferior e convexa, tem-se

$$F_z(\hat{x}) \leq \lim_{k \rightarrow \infty} F_z(x^{j_k}).$$

Além disso  $F(x^{k+1}) \preceq_K F(x^k)$  para todo  $k \in \mathbb{N}$  implica que

$$F_z(x^{k+1}) \leq F_z(x^k), \quad k \in \mathbb{N},$$

logo,  $\lim_{k \rightarrow \infty} F_z(x^{j_k}) = \inf\{F_z(x^k)\}$ . Combinando isto com as duas desigualdades acima tem-se que

$$F_z(\hat{x}) \leq F_z(x^k), \quad (3.26)$$

para todo  $z \in K^*$  e todo  $k \in \mathbb{N}$ . Assim, a afirmação é verificada desde que  $F(x^k) - F(\hat{x}) \in K$  e  $(K^*)^* = K$ .

Para concluir esta prova, suponha, por contradição, que  $\hat{x}$  não é uma solução eficiente fraca para (POVC), isto é,

$$\exists \check{x} \in \Omega : \quad F(\check{x}) \prec_K F(\hat{x}).$$

Considere agora  $z_k$  como em (3.16). Desde que  $\|z_k\| = 1$  para todo  $k \in \mathbb{N}$ , existe um ponto de acumulação de  $\{z_k\}$ , seja  $\bar{z}$ , o qual é o limite da subsequência  $\{z_{j_k}\}$  de  $\{z_k\}$ . Visto que  $K^*$  é fechado,  $\bar{z}$  pertence a  $K^*$ . Por conseguinte, de (3.25) e da relação acima se segue

$$\langle F(\check{x}) - F(\hat{x}), z_{j_k} \rangle \geq \langle F(\check{x}) - F(x^{j_k+1}), z_{j_k} \rangle \stackrel{def}{=} F_{j_k}(\check{x}) - F_{j_k}(x^{j_k+1}). \quad (3.27)$$

Por outro lado, a convexidade de  $F_{j_k}$  e  $x^{j_k+1} \in \text{int}(\Omega)$  implicam que existe  $v_{j_k} \in \partial F_{j_k}(x^{j_k+1})$  tal que

$$\langle F(\check{x}) - F(x^{j_k+1}), z_{j_k} \rangle \geq \langle \check{x} - x^{j_k+1}, v_{j_k} \rangle. \quad (3.28)$$

Acrescenta-se que de (3.17)-(3.19) existe  $u_{j_k} \in \mathbb{R}^n$  tal que

$$v_{j_k} = u_{j_k} - \beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle \nabla_1 d(x^{j_k+1}, x^{j_k}) \quad (3.29)$$

e

$$\langle x - x^{j_k+1}, u_{j_k} \rangle \geq 0. \quad (3.30)$$

para cada  $x \in \Omega_{j_k}$ . Além disso,  $\check{x}$  pertence a  $E$ , pois  $F(\check{x}) \prec_K F(\hat{x})$ .

Então, fazendo  $x = \check{x}$  em (3.30), obtém-se de (3.27)-(3.29)

$$\langle F(\check{x}) - F(\hat{x}), z_{j_k} \rangle \geq -\beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle \langle \check{x} - x^{j_k+1}, \nabla_1 d(x^{j_k+1}, x^{j_k}) \rangle. \quad (3.31)$$

Como (3.24) implica que

$$\lim_{k \rightarrow \infty} \langle \check{x} - x^{j_k+1}, \nabla_1 d(x^{j_k+1}, x^{j_k}) \rangle = 0$$

e  $\{-\beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle\}$  é uma sequência limitada, pode-se considerar o limite, quando  $k$  vai a  $+\infty$ , na desigualdade (3.31), para concluir que

$$\langle F(\check{x}) - F(\hat{x}), \bar{z} \rangle \geq 0. \quad (3.32)$$

Mais ainda, afirma-se que  $\bar{z} \neq 0$ . De fato, fazendo  $e = F(\check{x}) - F(\hat{x})$  e  $C = K$  no Lema 1.2.4, obtém-se  $\inf_{k \in \mathbb{N}} \{\langle F(\check{x}) - F(\hat{x}), z_k \rangle\} = d(F(\check{x}) - F(\hat{x}), \mathbb{R}^n \setminus K) > 0$ . Assim,  $z_{j_k} \rightarrow \bar{z}$  implica que

$$\langle F(\check{x}) - F(\hat{x}), \bar{z} \rangle > 0. \quad (3.33)$$

e à afirmação esta provada. Logo  $\bar{z} \neq 0$  e (3.33) contradizem o fato de que  $\bar{z} \in K^*$  e  $F(\check{x}) \prec_K F(\hat{x})$  implicam em

$$\langle F(\check{x}) - F(\hat{x}), \bar{z} \rangle < 0. \quad (3.34)$$

Portanto  $\hat{x}$  é uma solução eficiente fraca de (POVC).  $\square$

O teorema seguinte estabelece a convergência da sequência gerada pelo MIPE.

**Teorema 3.2.1.** *Seja  $(d, H) \in \mathcal{F}^*(\Omega)$  e seja  $\{x^k\}$  a sequência gerada pelo MIPE. Suponha que a Hipótese 3.2.1 é verificada. Então  $\{x^k\}$  converge para uma solução eficiente fraca de (POVC).*

*Demonstração.* Visto que  $\mathcal{F}^*(\Omega) \subseteq \mathcal{F}(\Omega)$ , a Proposição 3.2.1 implica que  $\{x^k\}$  é uma sequência limitada, que por sua vez implica na existência de pontos de acumulação. Assim, considere  $\hat{x}$  e  $x^*$  dois pontos de acumulação e,  $\{x^{j_k}\}$  e  $\{x^{l_k}\}$  duas

subsequências de  $\{x^k\}$  tais que

$$\lim_{k \rightarrow \infty} x^{j_k} = \hat{x}, \quad \lim_{k \rightarrow \infty} x^{l_k} = x^*.$$

Destas definições e de (3.25) segue que  $\hat{x}$  e  $x^* \in E$ . Logo o item *i*) da Proposição 3.2.1 implica que  $\{H(\hat{x}, x^k)\}$  e  $\{H(x^*, x^k)\}$  são sequências convergentes.

Agora, fazendo  $\bar{y} := x^*$ ,  $y^k := x^{l_k}$  e  $y := \hat{x}$  em  $(P_{10})$  da Definição 2.0.3, há dois casos a analisar:

- Se (2.1a) é válida, então  $\lim_{k \rightarrow \infty} H(x^*, x^{l_k}) = 0$ , o que por sua vez implica que

$$\lim_{k \rightarrow \infty} H(x^*, x^k) = 0.$$

Assim, como

$$\{x^{j_k}\} \subset \text{int}(\Omega), \quad x^* \in \Omega \quad \text{e} \quad \lim_{k \rightarrow \infty} H(x^*, x^{j_k}) = 0,$$

fazendo  $\bar{y} := x^*$  e  $y^k := x^{j_k}$  em  $(P_9)$ , tem-se que

$$\lim_{k \rightarrow \infty} x^{j_k} = x^*.$$

Logo,  $x^* = \hat{x}$ .

• Se, por outro lado, (2.1b) é válida e  $\hat{x} \neq x^*$ , então  $\lim_{k \rightarrow \infty} H(\hat{x}, x^{l_k}) = +\infty$ , o que contradiz a convergência de  $\{H(\hat{x}, x^k)\}$ .

Portanto,  $\{x^k\}$  converge para uma solução eficiente fraca de  $(POVC)$ .  $\square$

*Observação 3.2.1.* Neste trabalho separa-se cada resultado obtido, pois acha-se importante ressaltar o papel de cada uma destas classes:

$$\mathcal{D}(\text{int}(\Omega)), \quad \mathcal{F}(\Omega) \quad \text{e} \quad \mathcal{F}^*(\Omega).$$

Mais ainda, a classe  $\mathcal{F}(\Omega)$ , com hipóteses mais fracas, garante que a sequência gerada pelo MIPE é limitada e seus pontos de acumulação são soluções eficientes fracas do  $(POVC)$ , o que poderia ser útil, dependendo do problema.

Embora a classe  $\mathcal{F}^*(\Omega)$  seja mais restritiva, ela somente é necessária para garantir a convergência da sequência.

Agora analisa-se a Hipótese 3.2.1, isto é, a  $K$ -completude da seção inicial. Note que para o problema de minimização escalar,  $m = 1$  e  $K = \mathbb{R}_+$ , sempre que o conjunto de minimizadores é não vazio, resulta que cada seção  $F(\Omega_0)$  é  $K$ -completa. Além disso, Luc em [42, Lemma 3.5], apresentou o seguinte resultado:

*Observação 3.2.2.* Cada uma das condições seguintes é suficiente para a Hipótese 3.2.1:

i) O conjunto  $F(\Omega_0)$  é compacto.

ii)  $F(\Omega_0)$  é fechado,  $K$ -limitado inferior e o cone  $K$  tem a propriedade de Daniell (isto é, qualquer sequência  $K$ -decrecente e  $K$ -limitada inferior, converge para seu ínfimo).

Em situações práticas a Hipótese 3.2.1 pode ser difícil de ser verificada. Assim, é interessante estabelecer alguns resultados de convergência, embora mais fracos que aquele estabelecido na Proposição 3.2.1, para situações em que este pressuposto não seja verificado.

**Proposição 3.2.2.** *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_K\}$  uma aplicação própria,  $K$ -convexa, positivamente semi-contínua inferior. Se  $(d, H) \in \mathcal{F}^*(\Omega)$  e a sequência  $\{x^k\}$  gerada pelo MIPE (sem a hipótese 3.2.1) tem um ponto de acumulação, então esta sequência converge a uma solução eficiente fraca de (POVC).*

*Demonstração.* Por hipótese, existe uma subsequência  $\{x^{j_k}\}$  de  $\{x^k\}$  cujo limite é algum  $\hat{x} \in \Omega$ . Desde que a sequência  $\{F(x^k)\}$  é  $K$ -decrecente, tem-se que, para todo  $z \in K^*$ , a sequência  $\{\langle F(x^{j_k}), z \rangle\}$  é decrescente. Por outro lado, pela definição de  $F$ , tem-se que  $\langle F(\cdot), z \rangle$  é convexa e semi-contínua inferior, assim  $\langle F(\hat{x}), z \rangle$  é um limite inferior de  $\{\langle F(x^{j_k}), z \rangle\}$ . Visto que isto é verificado para todo  $z \in K^*$ , obtém-se que  $F(\hat{x}) \preceq_K F(x^{j_k})$  para todo  $k \in \mathbb{N}$ , implicando que

$$F(\hat{x}) \preceq_K F(x^k).$$

Deste ponto em diante, o argumento continua sendo o mesmo que na prova da Proposição 3.2.1 após de (3.25).  $\square$

Antes de enunciar o próximo resultado, precisa-se definir o conjunto de pontos ínfimos fracos de  $F(\Omega)$ :

$$K - \text{INF}_w(F(\Omega)) := \{\bar{y} \in \text{cl}(F(\Omega)) : \nexists x \in \Omega \text{ tal que } F(x) \prec_K \bar{y}\}.$$

**Proposição 3.2.3.** *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_K\}$  uma aplicação própria,  $K$ -convexa, positivamente semi-contínua inferior. Seja  $\{x^k\}$  uma sequência gerada pelo MIPE. Suponha que  $(d, H) \in \mathcal{F}^*(\Omega)$ ,  $\{x^k\}$  é ilimitada e que o conjunto de pontos de acumulação de  $\{F(x^k)\}$ , diferentes de  $\infty_K$ , é não vazio. Então a sequência  $\{F(x^k)\}$  converge para um elemento de  $K - \text{INF}_w(F(\Omega))$ .*

*Demonstração.* Sejam  $\bar{y} \in \mathbb{R}^m$  um ponto de acumulação e  $\{F(x^{j_k})\}$  uma subsequência de  $\{F(x^k)\}$  tal que  $F(x^{j_k}) \rightarrow \bar{y}$ .

Provar-se a seguir que a sequência toda converge para  $\bar{y}$ . De fato,  $\bar{y} \in \text{cl}(F(\Omega))$ , já que  $\text{cl}(F(\Omega))$  é um conjunto fechado.

Mais ainda, como  $F(x^{k+1}) \preceq_K F(x^k)$ , resulta que  $\{\langle F(x^k), z \rangle\}$  é uma sequência decrescente, e

$$\lim_{k \rightarrow +\infty} \langle F(x^{j_k}), z \rangle = \langle \bar{y}, z \rangle, \quad \forall z \in K^*.$$

Logo

$$\lim_{k \rightarrow +\infty} \langle F(x^k), z \rangle = \langle \bar{y}, z \rangle, \quad \forall z \in K^*. \quad (3.35)$$

Por outro lado,  $F(x^k)$   $K$ -decrescente,  $K$  fechado e (3.35) implicam em

$$\bar{y} \preceq_K F(x^k), \quad \forall k \in \mathbb{N}.$$

Considerando novamente o Lema 1.2.4, fazendo  $C = K^*$ ,  $z = \frac{F(x^k) - \bar{y}}{\|F(x^k) - \bar{y}\|}$  e considerando algum  $\tilde{z} \in \text{int}(K^*)$  obtém-se que

$$\langle F(x^k) - \bar{y}, \tilde{z} \rangle = \text{dist}(\tilde{z}, \mathbb{R}^m \setminus K^*) \|F(x^k) - \bar{y}\|, \quad (3.36)$$

para todo  $k \in \mathbb{N}$ . Uma vez que  $\langle F(x^k) - \bar{y}, \tilde{z} \rangle \rightarrow 0$  quando  $k \rightarrow \infty$ , e  $\text{dist}(\tilde{z}, \mathbb{R}^m \setminus K^*) > 0$ , obtém-se que  $\{F(x^k)\}$  converge para  $\bar{y}$ .

Se  $\bar{y} \notin \text{INF}_w(F(\Omega))$ , então existe  $\bar{x} \in \Omega$  com  $F(\bar{x}) \prec_K \bar{y}$ . Usando o mesmo argumento que na prova da Proposição 3.2.2, obtém-se que  $F(\bar{x}) \preceq_K F(x^k)$  para todo  $k \in \mathbb{N}$ . Então se estabelece a convergência de  $\{x^k\}$  para uma solução eficiente fraca de (POVC), seguindo o mesmo argumento que na prova da Proposição 3.2.1, de (3.25) em diante. Como consequência, contradiz-se a hipótese de  $\{x^k\}$  ser ilimitada, estabelecendo o resultado.  $\square$

### 3.3 Um método interior proximal inexato

Nesta parte apresenta-se a versão inexata do método apresentado na seção anterior, o qual é denotado por MIPI. O método abaixo gera sequências  $\{x^k\} \subset \text{int}(\Omega)$  e  $\{w^k\} \subset \mathbb{R}^n$  correspondentes à seguinte recursão:

Dada a iterada atual  $x^k$  ( $x^0 \in \text{int}(\Omega)$ ), encontrar  $(x^{k+1}, w^{k+1}) \in \text{int}(\Omega) \times \mathbb{R}^n$  tal que

$$w^{k+1} \in Q^k(x^{k+1}) + \beta_k \langle e_k, z_k \rangle \nabla_1 \tilde{d}(x^{k+1}, x^k), \quad (3.37)$$

onde  $\partial(\langle F(\cdot), z_k \rangle + I_{\Omega_k}) \subset Q^k \subset \partial_{\varepsilon_k}(\langle F(\cdot), z_k \rangle + I_{\Omega_k})$ ,  $\Omega_k = \{x \in \Omega : F(x) \preceq_K F(x^k)\}$ ,  $\tilde{d}$  é uma regularização dupla, isto é,  $\tilde{d} = d_1 + d_2$ , com  $d_i \in \mathcal{D}(\text{int}(\Omega))$   $i = 1, 2$ ,  $0 < \beta_k \leq \tilde{\beta}$ , para algum  $\tilde{\beta} > 0$ ,  $e_k \in \text{int}(C)$ , com  $\|e_k\| = 1$ , e  $z_k \in C^*$ , com  $\|z_k\| = 1$ .

*Observação 3.3.1.* Note que neste método admite-se tanto aproximações sucessivas do subdiferencial da representação escalar  $\partial(\langle F(\cdot), z \rangle + I_{\Omega_k})$  quanto um cálculo inexato das iteradas.

**Corolário 3.3.1.** *Suponha que  $F(\Omega_0)$  satisfaz à condição H1 ou  $d(\cdot, x)$  satisfaz à condição H2, para cada  $x \in \text{int}(\Omega)$ . Então a recursão gerada pelo MIPI está bem definida.*

*Demonstração.* Se  $F(\Omega_0)$  satisfaz à condição H1 ou  $d(\cdot, x)$  satisfaz à condição H2, para cada  $x \in \text{int}(\Omega)$ , então, do Corolário 3.2.1, garante-se a existência de  $\tilde{x} \in \text{int}(\Omega)$  tal que

$$0 \in \partial(\langle F(\cdot), z_k \rangle + I_{\Omega_k})(\tilde{x}) + \beta_k \langle e_k, z_k \rangle \nabla_1 \tilde{d}(\tilde{x}, x^k).$$

Por outro lado, considerando que

$$\partial(\langle F(\cdot), z_k \rangle + I_{\Omega_k})(\tilde{x}) \subset Q^k(\tilde{x}),$$

existem  $w^{k+1} \in \mathbb{R}^n$  e  $x^{k+1} \in \text{int}(\Omega)$  tais que o par  $(x^{k+1}, w^{k+1})$  verifica (3.37).  $\square$

Para estabelecer a limitação da sequência  $\{x^k\}$  gerada por MIPI, são necessárias hipóteses adicionais sobre as sequências  $\{\varepsilon_k\}$ ,  $\{w^k\}$  e a distância proximal induzida por  $d_2, H_2$ .

**Hipótese 3.3.1.**  $\sum_{k=1}^{\infty} \delta_k < \infty$ , onde  $\delta_k := \max\left\{\frac{\|w^{k+1}\|}{\beta_k \langle e_k, z_k \rangle}, \frac{\varepsilon_k}{\beta_k \langle e_k, z_k \rangle}\right\}$ .

**Hipótese 3.3.2.** Para cada  $x \in \Omega$  existem constantes  $\alpha(x) > 0$  e  $c(x)$ :

$$H_2(x, \nu) + c(x) \geq \alpha(x) \|x - \nu\|, \quad \nu \in \text{int}(\Omega)$$

onde  $H_2$  é uma distância proximal induzida de  $d_2$ .

*Observação 3.3.2.*

- i) Kaplan e Tichatschke em [37, section 3] mostraram que a Hipótese 3.3.2 é válida, em particular, para

$$\eta : x \mapsto \sum_{i=1}^n |x_i|^\rho, \quad \rho > 1.$$

- ii) A Hipótese 3.3.2 não é necessária se se considera  $w^{k+1} \equiv 0$ , para todo  $k \in \mathbb{N}$ .

**Proposição 3.3.1.** *Seja  $H_i$  uma distância proximal induzida por  $d_i$ ,  $i = 1, 2$ , e seja  $\{x^k\}$  uma sequência gerada pelo MIPI. Se  $(d_i, H_i) \in \mathcal{F}(\Omega)$ ,  $i = 1, 2$ , e as Hipóteses 3.2.1, 3.3.1 e 3.3.2 acima são verificadas, então*

i) para cada  $\bar{x} \in E$ ,  $\{\tilde{H}(\bar{x}, x^k)\}$  é uma sequência convergente;

ii)  $\{x^k\}$  é uma sequência limitada e todos seus pontos de acumulação são soluções eficientes fracas de (POVC);

onde  $\tilde{H} = H_1 + H_2$ ;

*Demonstração.* i) Dado que  $(d_i, H_i) \in \mathcal{F}(\Omega)$ , para  $i = 1, 2$ , de  $(P_7)$  da Definição 2.0.2, tem-se

$$\langle \nabla_1 \tilde{d}(x^{k+1}, x^k), x - x^{k+1} \rangle \leq \tilde{H}(x, x^k) - \tilde{H}(x, x^{k+1}), \quad x \in \Omega. \quad (3.38)$$

Mais ainda, da recursão deste método, (3.37), existe  $q^{k+1} \in Q^k(x^{k+1})$  tal que

$$\langle q^{k+1} + \beta_k \langle e_k, z_k \rangle \nabla_1 \tilde{d}(x^{k+1}, x^k), x - x^{k+1} \rangle \geq -\|w^{k+1}\| \|x - x^{k+1}\|, \quad x \in \mathbb{R}^n. \quad (3.39)$$

Logo, combinados (3.39) e (3.38), obtém-se

$$\tilde{H}(x, x^{k+1}) - \tilde{H}(x, x^k) \leq \frac{1}{\beta_k \langle e_k, z_k \rangle} \langle q^{k+1}, x - x^{k+1} \rangle + \frac{\|w^{k+1}\|}{\beta_k \langle e_k, z_k \rangle} \|x - x^{k+1}\|, \quad x \in \Omega. \quad (3.40)$$

Considerando  $\bar{x} \in E$  e o fato de que  $\partial(\langle F(\cdot), z_k \rangle + I_{\Omega_k})(x^{k+1}) \subset Q^k(x^{k+1}) \subset \partial_{\varepsilon_k}(\langle F(\cdot), z_k \rangle + I_{\Omega_k})(x^{k+1})$  tem-se

$$\langle q^{k+1}, \bar{x} - x^{k+1} \rangle - \varepsilon_k \leq \langle F(\bar{x}) - F(x^{k+1}), z_k \rangle \leq 0.$$

Assim

$$\langle q^{k+1}, \bar{x} - x^{k+1} \rangle \leq \varepsilon_k. \quad (3.41)$$

Por outro lado, fazendo  $z = \bar{x}$  e  $\nu = x^{k+1}$  na Hipótese 3.3.2, segue que

$$\|\bar{x} - x^{k+1}\| \leq \frac{1}{\alpha(\bar{x})} [H_2(\bar{x}, x^{k+1}) + c(\bar{x})] \leq \frac{1}{\alpha(\bar{x})} [\tilde{H}(\bar{x}, x^{k+1}) + c(\bar{x})]. \quad (3.42)$$

Agora, substitui-se  $\bar{x}$  por  $x$  em (3.40), que, junto com (3.41) e (3.42) implicam em

$$\tilde{H}(\bar{x}, x^{k+1}) - \tilde{H}(\bar{x}, x^k) \leq \frac{\delta_k}{\alpha(\bar{x})} \tilde{H}(\bar{x}, x^{k+1}) + \left(1 + \frac{c(\bar{x})}{\alpha(\bar{x})}\right) \delta_k. \quad (3.43)$$

Além disso, a Hipótese 3.3.1 garante que

$$\frac{\delta_k}{\alpha(\bar{x})} < \frac{1}{2}, \quad k > k_0,$$

para  $k_0$  suficientemente grande. Assim,

$$1 \leq \left(1 - \frac{\delta_k}{\alpha(\bar{x})}\right)^{-1} \leq \left(1 + \frac{2\delta_k}{\alpha(\bar{x})}\right) < 2, \quad \text{para } k \geq k_0,$$

o que junto com (3.43) resulta em

$$\tilde{H}(\bar{x}, x^{k+1}) \leq \left(1 + \frac{2\delta_k}{\alpha(\bar{x})}\right) \tilde{H}(\bar{x}, x^k) + 2 \left(1 + \frac{c(\bar{x})}{\alpha(\bar{x})}\right) \delta_k. \quad (3.44)$$

Desde que  $\sum_{k=1}^{\infty} \delta_k < \infty$ , aplicando o Lema 1.1.1 à desigualdade (3.44), obtém-se a convergência de  $\{\tilde{H}(\bar{x}, x^k)\}$ , para cada  $\bar{x} \in E$ .

*ii)* Note que a convergência de  $\{\tilde{H}(\bar{x}, x^k)\}$ , obtida no item acima, implica que

$$\tilde{H}(\bar{x}, x^k) \leq \tilde{M}, \quad k \in \mathbb{N},$$

onde  $\tilde{M} := \sup_{k \geq 0} \tilde{H}(\bar{x}, x^k) < +\infty$ .

Nestas condições,  $\{x^k\} \subset \text{int}(\Omega)$  e  $(P_8)$  da Definição 2.0.2 implicam na limitação de  $\{x^k\}$ , que, por sua vez, garante que o conjunto de pontos de acumulação desta sequência é não vazio. A seguir prova-se que todos estes são soluções eficientes fracas de  $(POVC)$ .

Sejam  $\hat{x}$  um ponto de acumulação e  $\{x^{j_k}\}$  uma subsequência de  $\{x^k\}$  tal que  $x^{j_k} \rightarrow \hat{x}$ . De forma análoga à prova da Proposição 3.2.1, obtém-se que

$$F(\hat{x}) \preceq_K F(x^k), \quad k \in \mathbb{N}. \quad (3.45)$$

Agora, suponha que  $\hat{x}$  não é uma solução eficiente fraca de  $(POVC)$ , isto é,

$$\exists \bar{x} \in \Omega : \quad F(\bar{x}) \prec_K F(\hat{x}).$$

Note que  $\bar{x} \in E$ , e assim,  $\{\tilde{H}(\bar{x}, x^k)\}$  é uma sequência convergente.

Por outro lado, desde que  $\|z_k\| = 1$  para todo  $k \in \mathbb{N}$ , existe um ponto de acumulação de  $\{z_k\}$ , seja  $\bar{z}$ , o qual é o limite da sequência  $\{z_{j_k}\}$ . Visto que  $C^*$  é fechado, tem-se que  $\bar{z} \in C^*$ . Além disso de (3.45) resulta que

$$\langle F(\bar{x}) - F(\hat{x}), z_{j_k} \rangle \geq \langle F(\bar{x}) - F(x^{j_k+1}), z_{j_k} \rangle. \quad (3.46)$$

Devido ao fato que  $\langle F(\cdot), z_k \rangle + I_{\Omega_k}$  é convexa,  $q^{k+1} \in Q^k(x^{k+1})$  e  $Q^k(x^{k+1}) \subset \partial_{\varepsilon_k}(\langle F(\cdot), z_k \rangle + I_{\Omega_k})(x^{k+1})$ , a relação (3.39), com  $k = j_k$ , implica que

$$\begin{aligned} \langle F(\bar{x}) - F(x^{j_k+1}), z_{j_k} \rangle &\geq -\beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle \langle \nabla_1 \tilde{d}(x^{j_k+1}, x^{j_k}), x - x^{j_k+1} \rangle \\ &\quad - \|w^{j_k+1}\| \|\bar{x} - x^{j_k+1}\| - \varepsilon_{j_k}. \end{aligned} \quad (3.47)$$



Assim, combinando (3.38), (3.46) e (3.47), segue que

$$\begin{aligned} \langle F(\bar{x}) - F(\hat{x}), z_{j_k} \rangle &\geq \beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle (\tilde{H}(\bar{x}, x^{j_k+1}) - \tilde{H}(\bar{x}, x^{j_k})) \\ &\quad - \|w^{j_k+1}\| \|\bar{x} - x^{j_k+1}\| - \varepsilon_{j_k}. \end{aligned} \quad (3.48)$$

Logo, da Hipótese 3.3.1 e a última desigualdade, tem-se que

$$\langle F(\bar{x}) - F(\hat{x}), z_{j_k} \rangle \geq \beta_{j_k} \langle e_{j_k}, z_{j_k} \rangle \left( \tilde{H}(\bar{x}, x^{j_k+1}) - \tilde{H}(\bar{x}, x^{j_k}) - \delta_{j_k+1} (\|\bar{x} - x^{j_k+1}\| + 1) \right). \quad (3.49)$$

Visto que  $\{\beta_k \langle e_k, z_k \rangle\}$  e  $\{x^k\}$  são sequências limitadas e  $z_{j_k} \rightarrow \bar{z}$  e  $\delta_k \rightarrow 0$ , quando  $k \rightarrow +\infty$ , obtém-se que

$$\langle F(\bar{x}) - F(\hat{x}), \bar{z} \rangle \geq 0. \quad (3.50)$$

Seguindo os mesmos argumentos usados no item *ii*) da Proposição 3.2.1 tem-se  $\bar{z} \neq 0$ , o que por sua vez implica em uma contradição. Portanto tal pressuposto é falso e  $\hat{x}$  é uma solução eficiente fraca de (POVC).  $\square$

**Teorema 3.3.1.** *Sejam  $(d_i, H_i) \in \mathcal{F}^*(\Omega)$ ,  $i = 1, 2$ , e seja  $\{x^k\}$  a sequência gerada pelo MIPI. Então a sequência  $\{x^k\}$  converge para uma solução eficiente fraca de (POVC).*

*Demonstração.* A prova deste teorema é análoga à prova do Teorema 3.2.1.  $\square$

# Capítulo 4

## Método de Região de Confiança para o Problema de Otimização Multiobjetivo Irrestrito

Neste capítulo considera-se o caso particular em que  $K = \mathbb{R}_+^m$  e  $\Omega = \mathbb{R}^n$  para estabelecer o segundo problema a ser tratado e propõe-se um método de região de confiança para determinar pontos Pareto críticos. Além disso, uma completa análise de convergência é apresentada seguindo as ideias de Conn *et al.* em [17].

### 4.1 O problema

O problema sob consideração é o *Problema de Otimização Multiobjetivo Irrestrito*:

$$(POMI) \quad \mathbb{R}_+^m - \min_{x \in \mathbb{R}^n} F(x),$$

onde

- i)  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $F = (F_1, \dots, F_m)^T$  e, para cada  $i = 1, \dots, m$ ,  $F_i$  é duas vezes continuamente diferenciável;
- ii) Existe  $j \in \{1, \dots, m\}$  tal que

$$-\infty < \inf\{F_j(x) : x \in \mathbb{R}^n\}.$$

O principal interesse neste capítulo é encontrar *pontos Pareto críticos* para (POMI), ou seja:

Encontrar  $x^* \in \mathbb{R}^n$  tal que

$$\forall d \in \mathbb{R}^n, \quad \exists i_0 = i_0(d) \in \{1, \dots, m\} : \quad \langle \nabla F_{i_0}(x^*), d \rangle \geq 0.$$

Note que para o caso escalar,  $m = 1$ , esta condição se reduz à condição de primeira ordem “gradiente igual a zero”

Desta maneira, para estabelecer o método de região de confiança, denotado por MRCPM, é necessário definir algumas funções auxiliares e revisar algumas propriedades referentes a elas.

Começa-se definindo a função marginal  $\omega : \mathbb{R}^n \rightarrow \mathbb{R}$  dada abaixo

$$\omega(x) := - \min_{\|d\| \leq 1} \left( \max_{i=1, \dots, m} \left\{ \langle \nabla F_i(x), d \rangle \right\} \right), \quad (4.1)$$

Repare que quando  $m = 1$  então  $d^*(x) = -\frac{\nabla F(x)}{\|\nabla F(x)\|}$  e  $\omega(x) = \|\nabla F(x)\|$ . Por outro lado, Fliege e Svaiter em [22] mostram algumas propriedades de  $\omega(x)$  relacionadas com o conceito de pontos Pareto críticos. A fim de lembrar estas propriedades denota-se o conjunto solução de (4.1) por  $\mathcal{D}(x)$ .

**Lema 4.1.1.** [22, Lemma 3]

- i)  $\omega(x) \geq 0$ , para todo  $x \in \mathbb{R}^n$ ;*
- ii) Se  $x$  é Pareto crítico de (POMI), então  $0 \in \mathcal{D}(x)$  e  $\omega(x) = 0$ ;*
- iii) Se  $x$  não é Pareto crítico de (POMI), então  $\omega(x) > 0$  e para qualquer  $d \in \mathcal{D}(x)$  tem-se que*

$$\langle \nabla F_j(x), d \rangle \leq \max_{i=1, \dots, m} \{ \langle \nabla F_i(x), d \rangle \} < 0, \quad j = 1, \dots, m,$$

*isto é,  $d$  é uma direção descendente de (POMI);*

- iv) A aplicação  $x \mapsto \omega(x)$  é contínua;*
- v) Se  $x^k$  converge para  $\bar{x}$ ,  $d_k \in \mathcal{D}(x^k)$  e  $d_k$  converge para  $\bar{d}$ , então  $\bar{d} \in \mathcal{D}(\bar{x})$ .*

Agora considere outra função escalar  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , dada por

$$\phi(x) := \max_{i=1, \dots, m} F_i(x).$$

Observe que a principal justificativa para considerar  $\phi$  é que o problema seguinte

$$\min_{x \in \mathbb{R}^n} \phi(x),$$

é uma representação escalar estrita de (*POMI*). Assim a Observação 1.2.7 e a Proposição 1.2.2 implicam que

$$\arg \min\{\phi(x) : x \in \mathbb{R}^n\} \subseteq \mathbb{R}_+^m - \arg \min_w\{F(x) : x \in \mathbb{R}^n\}. \quad (4.2)$$

*Observação 4.1.1.* Da definição de (*POMI*) obtém-se que

$$-\infty < \inf\{\phi(x) : x \in \mathbb{R}^n\}.$$

## 4.2 Um método de Região de Confiança

O método seguinte, denotado por “MRCPM”, está baseado no algoritmo básico de região de confiança para o problema de otimização escalar irrestrito, apresentado por Conn *et al.* em [17] e gera uma sequência  $\{x^k\} \subset \mathbb{R}^n$  correspondente ao método a seguir descrito. Antes é necessário definir a seguinte aproximação local de  $\phi$ :

$$m(x, H, d) = \max_{i=1, \dots, m} \{F_i(x) + \langle \nabla F_i(x), d \rangle\} + \frac{1}{2} \langle d, Hd \rangle, \quad \text{com } H \in \mathbb{S}^{n \times n}.$$

*Observação 4.2.1.* Embora as funções  $\phi$  e  $m$  não sejam necessariamente diferenciáveis nem convexas, de suas respectivas definições e do Lema 1.1.5, tem-se que, dado  $x \in \mathbb{R}^n$ .

$$\partial_o \phi(x) = \text{conv}\{\nabla F_i(x) : i \in I(x)\}$$

onde  $I(x) = \{j \in \{1, \dots, m\} : \phi(x) = F_j(x)\}$ ; e dado  $d \in \mathbb{R}^n$ , com  $x \in \mathbb{R}^n$  e  $H \in \mathbb{S}^{n \times n}$  fixos, tem-se que

$$\partial_o^d m(x, H, d) = \text{conv}\{\nabla F_i(x) : i \in J(d)\} + Hd$$

onde  $J(d) = \{j \in \{1, \dots, m\} : m(x, H, d) = F_j(x) + \langle \nabla F_j(x), d \rangle\}$  e  $\partial_o^d m(x, H, d)$  denota o subdiferencial de Clarke de  $m$  em relação de  $d$ .

Assim, a função  $m$  em  $(x, H, 0)$  é uma “boa” aproximação de primeira ordem da função  $\phi$  em  $x$ , por que se tem

$$m(x, H, 0) = \phi(x) \quad \text{e} \quad \partial_o^d m(x, H, 0) = \partial_o \phi(x).$$

Já que na  $k$ -ésima iteração,  $x^k$ ,  $\Delta_k$  e  $H_k$  são dados, estabelece-se que

$$m_k(\cdot) := m(x^k, H_k, \cdot)$$

e a região de confiança relativa à  $k$ -ésima iteração é denotada por

$$\mathcal{B}_k := \{d \in \mathbb{R}^n : \|d\| \leq \Delta_k\}$$

## Método RCPM

**Passo 0.** DADOS:

$$0 < \eta_1 \leq \eta_2 < 1 \quad \text{e} \quad 0 < \gamma_1 \leq \gamma_2 < 1. \quad (4.3)$$

**Passo 1.** INICIALIZAÇÃO: Escolha  $x^0$ ,  $H_0$  e  $\Delta_0$ . Faça  $k := 0$ .

**Passo 2.** CÁLCULO DA DIREÇÃO: Encontre uma direção  $d_k \in \mathcal{B}_k$  que “reduz suficientemente” à função  $m_k$ .

**Passo 3.** ACEITAÇÃO DO PONTO: Calcule  $\phi(x^k + d_k)$  e defina

$$\rho_k = \frac{\phi(x^k) - \phi(x^k + d_k)}{m_k(0) - m_k(d_k)}. \quad (4.4)$$

Se  $\rho_k \geq \eta_1$ , então defina  $x^{k+1} = x^k + d_k$ ; caso contrário defina  $x^{k+1} = x^k$ .

**Passo 4.** ATUALIZAÇÃO DA REGIÃO DE CONFIANÇA: Faça

$$\Delta_{k+1} \in \begin{cases} [\Delta_k, \infty), & \text{se } \rho_k \geq \eta_2; \\ [\gamma_2 \Delta_k, \Delta_k], & \text{se } \rho_k \in [\eta_1, \eta_2); \\ [\gamma_1 \Delta_k, \gamma_2 \Delta_k], & \text{se } \rho_k < \eta_1; \end{cases} \quad (4.5)$$

Atualize  $H_k$  para  $H_{k+1}$ . Faça  $k = k + 1$ , e volte ao **Passo 1**.

## 4.3 Análise de Convergência

As hipóteses necessárias para estabelecer a convergência de MRCPM serão dadas a seguir.

**Hipótese 4.3.1.** Existe uma constante positiva  $\kappa_F$  tal que, para todo  $x \in \mathbb{R}^n$ ,

$$\|\nabla^2 F_i(x)\| \leq \kappa_F, \quad \text{para todo } i = 1, \dots, m.$$

**Hipótese 4.3.2.** A matriz  $H_k$  é uniformemente limitada, ou seja, existe uma constante  $\kappa_{\text{sahm}} \geq 1$  tal que:

$$\|H_k\| \leq \kappa_{\text{sahm}} - 1, \quad \text{para todo } k.$$

Observe-se que as hipóteses acima são uma extensão natural daquelas usadas para o método de região de confiança escalar estabelecido em [17]. Embora existam trabalhos com hipóteses mais fracas, para o caso escalar, veja por exemplo Shi e Guo em [52] e Ahookhosh e Amini em [1] e suas respectivas referências, existem também trabalhos recentes considerando as hipóteses que utilizamos, veja [7]. Além disso, observa-se que a análise de convergência apresentada neste trabalho estende naturalmente ao apresentado para  $m = 1$  em [17].

*Observação 4.3.1.* Note que a Hipótese 4.3.1 implica que  $\nabla F_i$  é Lipschitz, para cada  $i \in \{1, \dots, m\}$ , que, por sua vez implica na continuidade uniforme de  $\omega$ .

Por outro lado, um ponto crucial neste algoritmo é a determinação, no **Passo 2**, da direção  $d_k$  a qual “reduz suficientemente” a função  $m_k$  dentro da região de confiança  $\mathcal{B}_k$ , porque esta redução pode garantir ou não a convergência global. Antes de definir o que significa “reduz suficientemente” estabelecemos dois importantes resultados.

**Lema 4.3.1.** *Seja o seguinte problema escalar*

$$\min_{d \in \mathcal{B}_k} m_k(d). \quad (4.6)$$

Se  $d_k$  é uma solução do problema (4.6), então

$$m_k(0) - m_k(d_k) \geq \frac{1}{2} \omega(x^k) \min \left\{ \frac{\omega(x^k)}{\beta_k}, \Delta_k \right\}, \quad (4.7)$$

onde  $\omega(x)$  é definido por (4.1) e

$$\beta_k = 1 + \|H_k\|. \quad (4.8)$$

*Demonstração.* Visto que  $d_k$  é uma solução do problema (4.6), segue que

$$m_k(0) - m_k(d_k) \geq m_k(0) - m_k(d), \quad \forall d \in \mathcal{B}_k. \quad (4.9)$$

Por outro lado, da definição de  $\omega(x)$  em  $x^k$  existe  $d_k^* \in \mathbb{R}^n$  tal que  $\|d_k^*\| \leq 1$  e

$$\omega(x^k) = - \max_{i=1, \dots, m} \{ \langle \nabla F_i(x^k), d_k^* \rangle \}. \quad (4.10)$$

Logo, para todo  $\alpha \in [0, 1]$ , vale  $\alpha \Delta_k d_k^* \in \mathcal{B}_k$ . Como consequência, segue de (4.9),

da desigualdade de Cauchy-Schwarz, (4.8) e (4.10) que

$$m_k(0) - m_k(d_k) \geq m_k(0) - m_k(\alpha\Delta_k d_k^*) \quad (4.11)$$

$$\geq -\alpha\Delta_k \max_{i=1,\dots,m} \{\langle \nabla F_i(x^k), d_k^* \rangle\} - \frac{1}{2}(\alpha\Delta_k)^2 \|H_k\| \quad (4.12)$$

$$\geq \alpha\Delta_k\omega(x^k) - \frac{1}{2}(\alpha\Delta_k)^2\beta_k, \quad (4.13)$$

para todo  $\alpha \in [0, 1]$ , o que implica que

$$m_k(0) - m_k(d_k) \geq \max_{0 \leq \alpha \leq 1} \left\{ \alpha\Delta_k\omega(x^k) - \frac{1}{2}\alpha^2\Delta_k^2\beta_k \right\}. \quad (4.14)$$

Agora, desde que a função  $\alpha \mapsto \alpha\Delta_k\omega(x^k) - \frac{1}{2}\alpha^2\Delta_k^2\beta_k$  é côncava, precisa-se analisar as duas possibilidades para o máximo acima. Antes denota-se o ponto crítico desta função por  $\alpha^*$ , assim

$$\alpha^* = \frac{\omega(x^k)}{\Delta_k\beta_k} \geq 0 \text{ é o maximizador global desta função e} \quad (4.15)$$

$$\frac{1}{2} \frac{\omega(x^k)^2}{\beta_k} \geq 0 \text{ é o valor máximo global desta função.} \quad (4.16)$$

a) Se  $\alpha^* \in [0, 1)$ , então

$$\max_{0 \leq \alpha \leq 1} \left\{ \alpha\Delta_k\omega(x^k) - \frac{1}{2}\alpha^2\Delta_k^2\beta_k \right\} = \frac{1}{2} \frac{\omega(x^k)^2}{\beta_k}. \quad (4.17)$$

a) Se  $\alpha^* \geq 1$ , então

$$\max_{0 \leq \alpha \leq 1} \left\{ \alpha\Delta_k\omega(x^k) - \frac{1}{2}\alpha^2\Delta_k^2\beta_k \right\} = \Delta_k\omega(x^k) - \frac{1}{2}\Delta_k^2\beta_k. \quad (4.18)$$

Porém,  $\alpha^* \geq 1$  implica também que

$$\Delta_k\omega(x^k) - \frac{1}{2}\Delta_k^2\beta_k \geq \frac{1}{2}\Delta_k\omega(x^k). \quad (4.19)$$

Por conseguinte

$$m_k(0) - m_k(d_k) \geq \min \left\{ \frac{1}{2} \frac{\omega(x^k)^2}{\beta_k}, \frac{1}{2} \Delta_k\omega(x^k) \right\} \quad (4.20)$$

$$= \frac{1}{2}\omega(x^k) \min \left\{ \frac{\omega(x^k)}{\beta_k}, \Delta_k \right\}. \quad \square \quad (4.21)$$

*Observação 4.3.2.* O resultado acima é estabelecido inicialmente por Xi e Shi em [56], porém o mínimo em (4.1) é considerando sobre  $\mathcal{B}_k$ .

Embora o Lema 4.3.1 seja suficiente “teoricamente” para estabelecer a convergência da sequência gerada por MRCPM, obter um minimizador exato do problema (4.6) poderia ser uma tarefa difícil, devido a que  $m_k$  não é necessariamente diferenciável nem convexa em geral. Como na prática, não é desejável se ter um alto custo computacional na obtenção do iterando, algoritmos inexatos devem ser considerados.

Nestas condições, é interessante observar que seguindo a ideia da busca linear de Armijo, veja por exemplo [8], pode-se considerar uma técnica de *backtracking* em alguma direção  $d \in \mathcal{D}(x^k)$  (visto que todas estas direções são direções descendentes, se  $x^k$  não é Pareto crítico), para assim obter um tamanho de passo que proverá uma “boa” redução da função  $m_k$ . Mais precisamente, determina-se o menor inteiro não negativo  $j = j^*$  tal que a direção

$$d_k^*(j) = (\kappa_{\text{bck}})^j \Delta_k d_k^* \quad (4.22)$$

satisfaz a condição

$$m_k(d_k^*(j)) \leq m_k(0) - \kappa_{\text{lsi}} (\kappa_{\text{bck}})^j \Delta_k \omega(x^k), \quad (4.23)$$

onde  $d_k^*$  verifica (4.10),  $\kappa_{\text{lsi}} \in (0, \frac{1}{2})$  e  $\kappa_{\text{bck}} \in (0, 1)$  são constantes dadas. Como é esperável, no caso escalar,  $m = 1$ , (4.23) se reduz à busca linear de Armijo, na direção  $-\nabla F(x^k)$ . Logo, definindo esta direção aproximada como

$$d_k^{\text{apx}} := d_k^*(j^*),$$

estabelece-se o seguinte resultado.

**Lema 4.3.2.** *A direção aproximada  $d_k^{\text{apx}}$  está bem definida, isto é,  $j^*$  é finito. Além disso,*

$$m_k(0) - m_k(d_k^{\text{apx}}) \geq \kappa_{\text{dpa}} \omega(x^k) \min \left\{ \frac{\omega(x^k)}{\beta_k}, \Delta_k \right\}, \quad (4.24)$$

onde  $\kappa_{\text{dpa}} \in (0, 1)$  é uma constante que independe de  $k$ .

*Demonstração.* Primeiro considere o caso em que a condição (4.23) não é verificada para algum  $j$ , ou seja,

$$m_k(d_k^*(j)) > m_k(0) - \kappa_{\text{lsi}} (\kappa_{\text{bck}})^j \Delta_k \omega(x^k). \quad (4.25)$$



Agora, da definição de  $m_k$  no lado esquerdo da desigualdade (4.25) obtém-se

$$m_k(d_k^*(j)) = \max_{1 \leq i \leq m} \{F_i(x^k) + \langle \nabla F_i(x^k), d_k^*(j) \rangle\} + \frac{1}{2} \langle d_k^*(j), H_k d_k^*(j) \rangle \quad (4.26)$$

$$\begin{aligned} &\leq m_k(0) + (\kappa_{\text{bck}})^j \Delta_k \max_{1 \leq i \leq m} \{ \langle \nabla F_i(x^k), d_k^* \rangle \} \\ &\quad + \frac{1}{2} ((\kappa_{\text{bck}})^j \Delta_k)^2 \langle d_k^*, H_k d_k^* \rangle \end{aligned} \quad (4.27)$$

$$= m_k(0) - (\kappa_{\text{bck}})^j \Delta_k \omega(x^k) + \frac{1}{2} ((\kappa_{\text{bck}})^j \Delta_k)^2 \langle d_k^*, H_k d_k^* \rangle. \quad (4.28)$$

Como  $\|H_k\| \leq \beta_k$  e  $\|d_k^*\| \leq 1$ , obtém-se da desigualdade de Cauchy-Schwarz que

$$\langle d_k^*, H_k d_k^* \rangle \leq \beta_k. \quad (4.29)$$

Agora, combinando (4.29) com (4.28) e (4.25), é fácil ver que

$$(\kappa_{\text{bck}})^j > \frac{2(1 - \kappa_{\text{lsi}})\omega(x^k)}{\Delta_k \beta_k}. \quad (4.30)$$

Porém,  $\kappa_{\text{bck}} < 1$ , deve portanto existir um primeiro  $j^*$  finito tal que

$$(\kappa_{\text{bck}})^{j^*} < \frac{2(1 - \kappa_{\text{lsi}})\omega(x^k)}{\Delta_k \beta_k}, \quad (4.31)$$

para o qual (4.23) seja verificado. Então, a direção aproximada  $d_k^{\text{apx}}$  está bem definida, e obtém-se de (4.22) e (4.23) que

$$m_k(0) - m_k(d_k^{\text{apx}}) \geq \kappa_{\text{lsi}} (\kappa_{\text{bck}})^{j^*} \Delta_k \omega(x^k). \quad (4.32)$$

Quando  $j^* \geq 1$ , já que este é o menor  $j$  que garante (4.31), pode-se deduzir que

$$(\kappa_{\text{bck}})^{j^*} = \kappa_{\text{bck}} (\kappa_{\text{bck}})^{j^*-1} \geq 2\kappa_{\text{bck}} (1 - \kappa_{\text{lsi}}) \frac{\omega(x^k)}{\Delta_k \beta_k},$$

o que junto com (4.32) implicam que

$$m_k(0) - m_k(d_k^{\text{apx}}) \geq 2\kappa_{\text{bck}} \kappa_{\text{lsi}} (1 - \kappa_{\text{lsi}}) \frac{\omega(x^k)^2}{\beta_k}. \quad (4.33)$$

Por outro lado, se  $j^* = 0$ , então (4.32) reduz-se a

$$m_k(0) - m_k(d_k^{\text{apx}}) \geq \kappa_{\text{lsi}} \Delta_k \omega(x^k). \quad (4.34)$$

Combinando (4.32) e (4.34), conclui-se que (4.24) é verificada com

$$\kappa_{\text{dpa}} = \min\{\kappa_{\text{lsi}}, 2\kappa_{\text{lsi}}\kappa_{\text{bck}}(1 - \kappa_{\text{lsi}})\} < 1.$$

□

Então, em vista aos dois lemas acima, é aceitável requer-se que o “decrécimo suficiente” da função  $m_k$  seja estabelecido pela seguinte hipótese.

**Hipótese 4.3.3.** Para todo  $k$ ,

$$m_k(0) - m_k(d_k) \geq \kappa_{\text{dasm}} \omega(x^k) \min \left\{ \frac{\omega(x^k)}{\beta_k}, \Delta_k \right\}, \quad (4.35)$$

para alguma constante  $\kappa_{\text{dasm}} \in (0, 1)$ , onde  $\omega(x^k)$  e  $\beta_k$  são definidos por (4.1) e (4.8), respectivamente.

*Observação 4.3.3.* Como notado anteriormente, no caso escalar tem-se que  $\omega(x^k) = \|\nabla F(x^k)\|$ , e com isto recupera-se a Hipótese A.A.1 requerida por Conn *et al.* em [17].

O resultado de convergência é apresentado a seguir, ou seja, prova-se que toda sequência gerada por MRCPM converge para um ponto Pareto crítico de (POMI).

Primeiro prova-se que o erro existente entre o valor da função  $\phi$  e sua aproximação  $m_k$  na iteração  $x^k$  é limitado.

**Proposição 4.3.1.** *Suponha que as Hipóteses 4.3.1 e 4.3.2 são verificadas. Então*

$$|\phi(x^k + d_k) - m_k(d_k)| \leq \kappa_{\text{lsh}} \Delta_k^2. \quad (4.36)$$

onde

$$\kappa_{\text{lsh}} := \max\{\kappa_F, \kappa_{\text{sahm}}\} \quad (4.37)$$

*Demonstração.*  $F_i$  é duas vezes continuamente diferenciável, logo do Teorema do valor médio, para cada  $i \in \{1, \dots, m\}$ , existe  $\alpha_i \in [0, 1]$ :

$$F_i(x^k + d_k) = F_i(x^k) + \langle \nabla F_i(x^k), d_k \rangle + \frac{1}{2} \langle d_k, \nabla^2 F_i(\xi_k^i) d_k \rangle, \quad (4.38)$$

onde  $\xi_k^i = x^k + \alpha_i d_k$ . Consequentemente,

$$\begin{aligned} \max_{i=1, \dots, m} \{F_i(x^k + d_k)\} &\leq \max_{i=1, \dots, m} \{F_i(x^k) + \langle \nabla F_i(x^k), d_k \rangle\} \\ &\quad + \max_{i=1, \dots, m} \left\{ \frac{1}{2} \langle d_k, \nabla^2 F_i(\xi_k^i) d_k \rangle \right\}. \end{aligned} \quad (4.39)$$

Agora, somando e subtraindo  $\frac{1}{2} \langle d_k, H_k d_k \rangle$  no lado direito da desigualdade acima, obtém-se

$$\phi(x^k + d_k) - m_k(d_k) \leq \max_{i=1, \dots, m} \left\{ \frac{1}{2} \langle d_k, (\nabla^2 F_i(\xi_k^i) - H_k) d_k \rangle \right\}. \quad (4.40)$$

Por outro lado, (4.38) também pode ser escrita como

$$F_i(x^k + d_k) - \frac{1}{2}\langle d_k, \nabla^2 F_i(\xi_k^i) d_k \rangle = F_i(x^k) + \langle \nabla F_i(x^k), d_k \rangle \quad (4.41)$$

para cada  $i \in \{1, \dots, m\}$ . Então

$$\phi(x^k + d_k) \geq \max_{i=1, \dots, m} \left\{ F_i(x^k) + \langle \nabla F_i(x^k), d_k \rangle \right\} - \max_{i=1, \dots, m} \left\{ -\frac{1}{2}\langle d_k, \nabla^2 F_i(\xi_k^i) d_k \rangle \right\},$$

ou equivalentemente,

$$\phi(x^k + d_k) \geq \max_{i=1, \dots, m} \left\{ F_i(x^k) + \langle \nabla F_i(x^k), d_k \rangle \right\} + \min_{i=1, \dots, m} \left\{ \frac{1}{2}\langle d_k, \nabla^2 F_i(\xi_k^i) d_k \rangle \right\}.$$

Novamente, somando e subtraindo  $\frac{1}{2}\langle d_k, H_k d_k \rangle$  no lado direito da desigualdade acima resulta que

$$\phi(x^k + d_k) - m_k(d_k) \geq \min_{i=1, \dots, m} \left\{ \frac{1}{2}\langle d_k, (\nabla^2 F_i(\xi_k^i) - H_k) d_k \rangle \right\}. \quad (4.42)$$

Assim, combinando (4.40) com (4.42) e o fato de  $\|d_k\| \leq \Delta_k$  obtém-se que

$$\begin{aligned} \left| \phi(x^k + d_k) - m_k(d_k) \right| &\leq \max \left\{ \left| \min_{i=1, \dots, m} \left\{ \frac{1}{2}\langle d_k, (\nabla^2 F_i(\xi_k^i) - H_k) d_k \rangle \right\} \right|, \right. \\ &\quad \left. \left| \max_{i=1, \dots, m} \left\{ \frac{1}{2}\langle d_k, (\nabla^2 F_i(\xi_k^i) - H_k) d_k \rangle \right\} \right| \right\} \\ &\leq \max_{i=1, \dots, m} \left| \frac{1}{2}\langle d_k, (\nabla^2 F_i(\xi_k^i) - H_k) d_k \rangle \right| \\ &\leq \frac{1}{2} \left( \max_{i=1, \dots, m} \|\nabla^2 F_i(\xi_k^i)\| + \|H_k\| \right) \Delta_k^2. \end{aligned}$$

Por conseguinte as Hipóteses 4.3.1 e 4.3.2 implicam que

$$|\phi(x^k + d_k) - m_k(d_k)| \leq \kappa_{\text{lsh}} \Delta_k^2.$$

onde

$$\kappa_{\text{lsh}} := \max \{ \kappa_F, \kappa_{\text{sahm}} \}. \quad \square$$

A seguir define-se o conjunto das iterações *bem sucedidas*  $\mathcal{S}$ , por

$$\mathcal{S} = \{k \in \mathbb{N} : \rho_k \geq \eta_1\},$$

e similarmente, o conjunto das iterações  *muito bem sucedidas*   $\mathcal{V}$ , é dado por

$$\mathcal{V} = \{k \in \mathbb{N} : \rho_k \geq \eta_2\}.$$

**Proposição 4.3.2.** *Suponha que as Hipóteses 4.3.1–4.3.3 são verificadas. Se  $x^k$  não é Pareto crítico e*

$$\Delta_k \leq \frac{\kappa_{\text{dasm}}(1 - \eta_2)\omega(x^k)}{\kappa_{\text{lsh}}}, \quad (4.43)$$

*então a iteração  $k$  pertence a  $\mathcal{V}$  e*

$$\Delta_{k+1} \geq \Delta_k. \quad (4.44)$$

*Demonstração.* Primeiro note que as constantes  $\eta_2, \kappa_{\text{dpa}} \in (0, 1)$ , por conseguinte

$$\kappa_{\text{dasm}}(1 - \eta_2) < 1.$$

Assim, a condição (4.43) e  $\beta_k = 1 + \|H_k\|$  implicam que

$$\Delta_k < \frac{\omega(x^k)}{\beta_k}. \quad (4.45)$$

Logo, a Hipótese 4.3.3 implica diretamente que

$$m_k(0) - m_k(d_k) \geq \kappa_{\text{dasm}}\omega(x^k) \min \left\{ \frac{\omega(x^k)}{\beta_k}, \Delta_k \right\} = \kappa_{\text{dasm}}\omega(x^k)\Delta_k. \quad (4.46)$$

Nestas condições, de (4.36), a desigualdade acima e (4.45) deduz-se que

$$|\rho_k - 1| = \left| \frac{\phi(x^k + d_k) - m_k(d_k)}{m_k(0) - m_k(d_k)} \right| \leq \frac{\kappa_{\text{lsh}}}{\kappa_{\text{dasm}}\omega(x^k)}\Delta_k \leq 1 - \eta_2. \quad (4.47)$$

Portanto,  $\rho_k \geq \eta_2$ , e a iteração é  *muito bem sucedida* . Além disso, a atualização de  $\Delta_k$ , (4.5), garante que (4.44) é verificado.  $\square$

**Proposição 4.3.3.** *Suponha que as Hipóteses 4.3.1–4.3.3 são verificadas. Além disso, suponha que existe uma constante  $\kappa_{\text{li}\omega} > 0$  tal que  $\omega(x^k) \geq \kappa_{\text{li}\omega}$  para todo  $k$ . Então existe uma constante  $\kappa_{\text{li}\Delta} > 0$  tal que*

$$\Delta_k \geq \kappa_{\text{li}\Delta}, \quad \text{para todo } k. \quad (4.48)$$

*Demonstração.* Suponha por contradição que para cada  $\kappa > 0$  existe  $\hat{k}$  tal que

$$\Delta_{\hat{k}} < \kappa. \quad (4.49)$$

Em particular, considere que

$$\kappa = \frac{\gamma_1 \kappa_{\text{dasm}} \kappa_{\text{li}\omega} (1 - \eta_2)}{\kappa_{\text{lsh}}},$$

e seja  $\hat{k}$  a primeira iteração tal que

$$\Delta_{\hat{k}} < \frac{\gamma_1 \kappa_{\text{dasm}} \kappa_{\text{li}\omega} (1 - \eta_2)}{\kappa_{\text{lsh}}}. \quad (4.50)$$

Por outro lado, da atualização da região de confiança, (4.5), obtém-se que

$$\gamma_1 \Delta_{\hat{k}-1} \leq \Delta_{\hat{k}}.$$

Assim, estas duas últimas desigualdades implicam que

$$\Delta_{\hat{k}-1} < \frac{\kappa_{\text{dasm}} \kappa_{\text{li}\omega} (1 - \eta_2)}{\kappa_{\text{lsh}}}. \quad (4.51)$$

Desta maneira a hipótese sobre  $\omega(x^k)$  e (4.51) implicam que (4.43) é verificado. Então a iteração  $\hat{k} - 1$  é *muito bem sucedida* e

$$\Delta_{\hat{k}-1} \leq \Delta_{\hat{k}}.$$

No entanto isto é uma uma contradição com o fato de que  $\hat{k}$  é a primeira iteração que verifica (4.50). Portanto a assunção inicial é impossível.  $\square$

**Proposição 4.3.4.** *Suponha que as Hipóteses 4.3.1–4.3.3 são verificadas. Além disso, suponha que dado  $x^0$  existe somente um número finito de iterações bem sucedidas. Então  $x^k = x^*$  para todo  $k$  suficientemente grande e  $x^*$  é um Pareto crítico.*

*Demonstração.* Da definição de MRCPM garante-se que  $x^* = x^{k_0+1} = x^{k_0+j}$  para todo  $j \in \mathbb{N}$ , onde  $k_0$  é o índice da última iteração bem sucedida. Além disso, para toda iteração *sem sucesso*, (4.3) e (4.5) implicam que a sequência  $\Delta_k$  vai para zero, quando  $k$  vai para  $+\infty$ . Por conseguinte, se  $x^{k_0+1}$  não é Pareto crítico, a Proposição 4.3.2 implica que deve existir uma iteração bem sucedida de índice maior que  $k_0$ , o que é uma contradição. Portanto,  $x^{k_0+1} = x^*$  é Pareto crítico.  $\square$

**Proposição 4.3.5.** *Suponha que as Hipóteses 4.3.1–4.3.3 são verificadas. Então*

$$\liminf_{k \rightarrow \infty} \omega(x^k) = 0 \quad (4.52)$$

*Demonstração.* Suponha, com o propósito de chegar a uma contradição, que para todo  $k$ ,

$$\omega(x^k) \geq \epsilon \quad (4.53)$$

para algum  $\epsilon > 0$ . Pela Proposição 4.3.3 tem-se que existe  $\kappa_{\text{li}\Delta} > 0$  tal que

$$\Delta_k \geq \kappa_{\text{li}\Delta}, \quad \forall k. \quad (4.54)$$

Agora considere uma iteração *bem sucedida* com índice  $k$ . O fato de que  $k \in \mathcal{S}$ , junto com a Hipótese 4.3.3,

$$\beta_k \leq \kappa_{\text{sahm}}$$

e a desigualdade acima implicam que

$$\phi(x^k) - \phi(x^{k+1}) \geq \eta_1(m_k(0) - m_k(d_k)) \geq \kappa_{\text{dasm}}\epsilon\eta_1 \min \left\{ \frac{\epsilon}{\kappa_{\text{sahm}}}, \kappa_{\text{li}\Delta} \right\}. \quad (4.55)$$

Somando agora as iterações bem sucedidas de 0 até  $k$  tem-se que

$$\phi(x^0) - \phi(x^{k+1}) = \sum_{\substack{j=0 \\ k \in \mathcal{S}}}^k (\phi(x^j) - \phi(x^{j+1})) \geq \sigma_k \kappa_{\text{dasm}} \epsilon \eta_1 \min \left\{ \frac{\epsilon}{\kappa_{\text{sahm}}}, \kappa_{\text{li}\Delta} \right\},$$

onde  $\sigma_k$  é o número de iterações bem sucedidas até a iteração  $k$ . Porém, desde que estas são infinitas obtém-se que

$$\lim_{k \rightarrow \infty} \sigma_k = +\infty,$$

e a diferença entre  $\phi(x^0)$  e  $\phi(x^{k+1})$  é ilimitada, o que contradiz a limitação inferior de  $\phi(x)$ , comentada na Observação 4.1.1. Portanto o resultado segue.  $\square$

**Teorema 4.3.1.** *Suponha que as Hipóteses 4.3.1–4.3.3 são verificadas. Então*

$$\lim_{k \rightarrow \infty} \omega(x^k) = 0 \quad (4.56)$$

*Demonstração.* Suponha por contradição que existe uma subsequência de iterações *bem sucedidas*, com índices  $\{t_j\} \subseteq \mathcal{S}$  tal que

$$\omega(x^{t_j}) \geq 2\epsilon > 0 \quad (4.57)$$

para algum  $\epsilon > 0$  e para todo  $j$ . A Proposição 4.3.5 garante a existência, para cada  $t_j$ , de uma primeira iteração bem sucedida  $l(t_j) > t_j$  tal que  $\omega(x^{l(t_j)}) < \epsilon$ . Denotando  $l_j = l(t_j)$  tem-se que existe uma outra subsequência de  $\mathcal{S}$  com índices  $\{l_j\}$  tal que

$$\omega(x^k) \geq \epsilon \quad \text{para } t_j \leq k < l_j \quad \text{e} \quad \omega(x^{l_j}) < \epsilon. \quad (4.58)$$

Agora analisa-se a subsequência de iterações cujos índices se encontram em

$$\mathcal{K} := \{k \in \mathcal{S} : t_j \leq k < l_j\}$$

com  $t_j$  e  $l_j$  pertencendo às duas subsequências definidas acima.

Observa-se que a Hipótese 4.3.3, o fato de  $\mathcal{K} \subseteq \mathcal{S}$  e (4.58) implicam que, para  $k \in \mathcal{K}$ ,

$$\phi(x^k) - \phi(x^{k+1}) \geq \eta_1(m_k(0) - m_k(d_k)) \geq \kappa_{\text{dasm}}\epsilon\eta_1 \min \left\{ \frac{\epsilon}{\kappa_{\text{sahm}}}, \Delta_k \right\}. \quad (4.59)$$

Porém, a sequência  $\{\phi(x^k)\}$  é monótona decrescente e limitada inferiormente, pela Observação 4.1.1, portanto convergente. Por conseguinte, o lado esquerdo de (4.59) vai para zero quando  $k$  vai para  $+\infty$ . Assim, obtém-se que

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \Delta_k = 0.$$

Como uma consequência, do mínimo de (4.59) obtém-se que, para  $k \in \mathcal{K}$  suficientemente grande,

$$\Delta_k \leq \frac{1}{\kappa_{\text{dasm}}\epsilon\eta_1} (\phi(x^k) - \phi(x^{k+1})).$$

Logo para  $j$  suficientemente grande

$$\|x^{t_j} - x^{l_j}\| \leq \sum_{\substack{i=t_j \\ i \in \mathcal{K}}}^{l_j-1} \|x^i - x^{i+1}\| \leq \sum_{\substack{i=t_j \\ i \in \mathcal{K}}}^{l_j-1} \Delta_i \leq \frac{1}{\kappa_{\text{dasm}}\epsilon\eta_1} (\phi(x^{t_j}) - \phi(x^{l_j})). \quad (4.60)$$

Novamente, da limitação inferior de  $\phi$  e da monotonicidade da sequência  $\{\phi(x^k)\}$  tem-se que o lado direito da desigualdade vai para zero, e portanto obtém-se que

$$\lim_{j \rightarrow \infty} \|x^{t_j} - x^{l_j}\| = 0.$$

Assim, da continuidade uniforme de  $\omega$ , Observação 4.3.1, deduz-se que

$$\lim_{j \rightarrow \infty} |\omega(x^{t_j}) - \omega(x^{l_j})| = 0.$$

No entanto, da definição das sequências  $\{t_j\}$  e  $\{l_j\}$  segue que

$$|\omega(x^{t_j}) - \omega(x^{l_j})| \geq \epsilon.$$

Portanto não existem subsequências satisfazendo (4.57), e assim o teorema é provado.  $\square$

*Observação 4.3.4.* Não é considerado neste trabalho a atualização da matriz  $H_k$ . No caso escalar ela é realizada mediante uma aproximação da matriz hessiana  $\nabla^2 F(x^k)$ , por exemplo usando a técnica de atualização Quase-Newton BFGS, veja para mais detalhes [17].

Yigui e Qian em [58] também apresentaram uma forma de obter uma atualização de  $H_k$  simétrica e definida positiva. No entanto, ela depende tanto da solução exata quanto dos multiplicadores de Lagrange a ela associados relativo ao problema 4.6, que não são hipóteses realísticas do ponto de vista de implementação.



# Capítulo 5

## Conclusão

Este capítulo apresenta as considerações finais bem como os planos para a próxima etapa de pesquisa.

### 5.1 Considerações finais

Este texto é uma exposição unificada do trabalho aceito [55] e do trabalho ainda “básico” no assunto de regiões de confiança. Para ser mais explícito, na primeira parte desta tese enfraquece-se ainda mais o conceito de distância proximal induzida e estuda-se o problema de otimização vetorial:

$$(POVC) \quad K - \min_{x \in \Omega} F(x),$$

onde  $\Omega$  é um conjunto convexo e fechado, com interior não vazio,  $K \subset \mathbb{R}^m$  é o cone convexo, fechado e pontudo que define a ordem parcial  $\preceq_K$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_K\}$  é uma aplicação própria,  $K$ -convexa e positivamente semi-contínua inferior.

Propõe-se um método interior proximal para este problema e analisa-se suas versões exata e inexata.

Mostra-se a boa definição da sequência  $\{x^k\}$  gerada por cada versão, **MIPE** e **MIPI**, nos Corolários 3.2.1 e 3.3.1, respectivamente; nas Proposições 3.2.1 e 3.3.1 mostra-se que toda sequência gerada por alguma versão do método é limitada e que todos os pontos de acumulação são soluções eficientes fracas de  $(POVC)$ ; além disso, nos Teoremas 3.2.1 e 3.3.1, supondo hipóteses um pouco mais restritivas sobre as regularizações, estabelece-se a convergência de cada sequência gerada.

Na segunda parte, particulariza-se  $K = \mathbb{R}_+^m$  e  $\Omega = \mathbb{R}^n$ , para assim considerar o problema de otimização multiobjetivo:

$$(POMi) \quad \mathbb{R}_+^m - \min_{x \in \mathbb{R}^n} F(x),$$

onde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $F = (F_1, \dots, F_m)$  e, para cada  $i = 1, \dots, m$ ,  $F_i$  é duas vezes continuamente diferenciável e existe  $j \in \{1, \dots, m\}$  tal que  $-\infty < \inf\{F_j(x) : x \in \mathbb{R}^n\}$ .

Propõe-se um método de região de confiança para este problema, denotado por **MRCPM**, e mostram-se propriedades importantes da sequência  $\{x^k\}$  gerada pelo método, para que, finalmente, no Teorema 4.3.1, se estabeleça que a sequência  $\{x^k\}$  converge para um ponto Pareto crítico de  $(POMi)$ . No entanto, vê-se estes resultados como um passo inicial no âmbito de REGIÕES DE CONFIANÇA.

Embora, o que sabe-se ainda é muito pouco. Estabelecer novas propriedades sobre o problema de otimização multiobjetivo e/ou vetorial, assim como desenvolver outros métodos, são desafios para a próxima etapa do trabalho.

## 5.2 Trabalhos futuros

A seguir, descrevem-se linhas de futuras pesquisas que dão continuidade natural a este trabalho:

1. Sobre os Métodos **MIPE** e **MIPI**:
  - i*) Generalizar ambos métodos para problemas  $K$ -quaseconvexos e de desigualdade variacional vetorial;
2. Sobre o Método **MRCPM**:
  - i*) Analisar extensões eficientes para a atualização da matriz  $H_k$ .
  - ii*) Obter convergência a um ponto Pareto crítico do problema sob hipótese mais fraca que 4.3.1 e 4.3.2.
  - iii*) Obter convergência a um ponto Pareto.
  - iv*) Aprofundarmos no estudo de MÉTODOS DE REGIÃO DE CONFIANÇA, com o intuito de obter um método que demonstre boa performance computacional. Em particular, na próxima etapa será dedicado um grande esforço na implementação de **MRCPM** e ilustrar o comportamento numérico deste esquema através de exemplos.

# Referências Bibliográficas

- [1] AHOOKHOSH, M. AND AMINI, K., *A nonmonotone trust region method with adaptive radius for unconstrained optimization problems*, Comput. Math. Appl. **60** (2010), no. 3, 411–422. MR 2665646
- [2] ALIPRANTIS, C. D. AND FLORENZANO, M. AND MARTINS-DA-ROCHA, V. F. AND TOURKY, R., *Equilibrium analysis in financial markets with countably many securities*, J. Math. Econom. **40** (2004), no. 6, 683–699. MR 2070962 (2006g:91061)
- [3] ALIPRANTIS, C. D. AND FLORENZANO, M. AND TOURKY, R., *General equilibrium analysis in ordered topological vector spaces*, J. Math. Econom. **40** (2004), no. 3-4, 247–269. MR 2070697 (2005e:91112)
- [4] AUSLENDER, A. AND TEBoulLE, M., *Asymptotic cones and functions in optimization and variational inequalities*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2003. MR 1931309 (2003i:49002)
- [5] ———, *Interior gradient and proximal methods for convex and conic optimization*, SIAM J. Optim. **16** (2006), no. 3, 697–725 (electronic). MR 2197553 (2006i:90048)
- [6] AUSLENDER, A. AND TEBoulLE, M. AND BEN-TIBA, S., *Interior proximal and multiplier methods based on second order homogeneous kernels*, Math. Oper. Res. **24** (1999), no. 3, 645–668. MR 1854247 (2002f:90159)
- [7] BASTIN, F., MALMEDY, V., MOUFFE, M. TOINT, P. L. AND TOMANOS, D., *A retrospective trust-region method for unconstrained optimization*, Math. Program. **123** (2010), no. 2, Ser. A, 395–418. MR 2595919
- [8] BERTSEKAS, D. P., *Nonlinear programming*, Springer Monographs in Mathematics, Athena Scientific, Belmont, Mass., 1995.
- [9] BOLINTINÉANU, S., *Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems*, J. Optim. Theory Appl. **106** (2000), no. 2, 265–296. MR 1788925 (2001i:90071)

- [10] ———, *Vector variational principles;  $\epsilon$ -efficiency and scalar stationarity*, J. Convex Anal. **8** (2001), no. 1, 71–85. MR 1829056 (2002b:90113)
- [11] BONNEL, H. AND IUSEM, A. N. AND SVAITER, B. F., *Proximal methods in vector optimization*, SIAM J. Optim. **15** (2005), no. 4, 953–970 (electronic). MR 2178482 (2006f:90058)
- [12] BORWEIN, J. M. AND LEWIS, A. S., *Convex analysis and nonlinear optimization*, second ed., CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3, Springer, New York, 2006, Theory and examples. MR 2184742 (2006f:49001)
- [13] CENG, L.-C. AND YAO, J.-C., *Approximate proximal methods in vector optimization*, European J. Oper. Res. **183** (2007), no. 1, 1–19. MR 2329436 (2008e:90076)
- [14] CHEN, G. AND TEBoulLE, M., *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, SIAM J. Optim. **3** (1993), no. 3, 538–543. MR 1230155 (94e:90093)
- [15] CHEN, Z. AND HUANG, H. AND ZHAO, K., *Approximate generalized proximal-type method for convex vector optimization problem in Banach spaces*, Comput. Math. Appl. **57** (2009), no. 7, 1196–1203. MR 2508550
- [16] CHEN, Z. AND ZHAO, K., *A proximal-type method for convex vector optimization problem in Banach spaces*, Numer. Funct. Anal. Optim. **30** (2009), no. 1-2, 70–81. MR 2492074
- [17] CONN, A. R. AND GOULD, N. I. M. AND TOINT, P. L., *Trust-region methods*, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. MR 1774899 (2003e:90002)
- [18] ERWAY, J. B. AND GILL, P. E., *A subspace minimization method for the trust-region step*, SIAM J. Optim. **20** (2009), no. 3, 1439–1461. MR 2587735
- [19] FLIEGE, J., *An efficient interior-point method for convex multicriteria optimization problems*, Math. Oper. Res. **31** (2006), no. 4, 825–845. MR 2281232 (2008e:90077)
- [20] ———, *The effects of adding objectives to an optimisation problem on the solution set*, Oper. Res. Lett. **35** (2007), no. 6, 782–790. MR 2361048 (2008g:90102)

- [21] FLIEGE, J. AND GRAÑA DRUMMOND, L. M. AND SVAITER, B. F., *Newton's method for multiobjective optimization*, SIAM J. Optim. **20** (2009), no. 2, 602–626. MR 2515788 (2010b:90117)
- [22] FLIEGE, J. AND SVAITER, B. F., *Steepest descent methods for multicriteria optimization*, Math. Methods Oper. Res. **51** (2000), no. 3, 479–494. MR 1778656 (2001e:90078)
- [23] FLIEGE, J. AND VICENTE, L. N., *Multicriteria approach to bilevel optimization*, J. Optim. Theory Appl. **131** (2006), no. 2, 209–225. MR 2279595 (2007h:90073)
- [24] GARDAŠEVIĆ FILIPOVIĆ, M., *A trust region method using subgradient for minimizing a nondifferentiable function*, Yugosl. J. Oper. Res. **19** (2009), no. 2, 249–262. MR 2578615
- [25] GÖPFERT, A. AND RIAHI, H., TAMMER, C. AND ZĂLINESCU, C., *Variational methods in partially ordered spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 17, Springer-Verlag, New York, 2003. MR 1994718 (2004h:90001)
- [26] GRAÑA DRUMMOND, L. M. AND IUSEM, A. N., *First order conditions for ideal minimization of matrix-valued problems*, J. Convex Anal. **10** (2003), no. 1, 129–147. MR 1999918 (2004e:90106)
- [27] ———, *A projected gradient method for vector optimization problems*, Comput. Optim. Appl. **28** (2004), no. 1, 5–29. MR 2049673 (2005a:90139)
- [28] GRAÑA DRUMMOND, L. M. AND IUSEM, A. N. AND SVAITER, B. F., *On first order optimality conditions for vector optimization*, Acta Math. Appl. Sin. Engl. Ser. **19** (2003), no. 3, 371–386. MR 2012502 (2004g:90099)
- [29] GRAÑA DRUMMOND, L. M. AND MACULAN, N. AND SVAITER, B. F., *On the choice of parameters for the weighting method in vector optimization*, Math. Program. **111** (2008), no. 1-2, Ser. B, 201–216. MR 2322889 (2008j:90115)
- [30] GRAÑA DRUMMOND, L. M. AND SVAITER, B. F., *A steepest descent method for vector optimization*, J. Comput. Appl. Math. **175** (2005), no. 2, 395–414. MR 2108583 (2005i:90057)
- [31] GREGÓRIO, R. AND OLIVEIRA, P. R., *A logarithmic-quadratic proximal point scalarization method for multiobjective programming*, J. Global Optim. **49** (2011), no. 2, 281–291. MR 2754381

- [32] GUTIÉRREZ, C., JIMÉNEZ, B. AND NOVO, V., *Optimality conditions via scalarization for a new  $\epsilon$ -efficiency concept in vector optimization problems*, European J. Oper. Res. **201** (2010), no. 1, 11–22. MR 2573326 (2010k:90160)
- [33] HADJISAVVAS, N. AND KOMLÓSI, S. AND SCHAIBLE, S. (ed.), *Handbook of generalized convexity and generalized monotonicity*, Nonconvex Optimization and its Applications, vol. 76, Springer-Verlag, New York, 2005. MR 2098896 (2005i:90003)
- [34] HUANG, X. X. AND YANG, X. Q., *Nonlinear Lagrangian for multiobjective optimization and applications to duality and exact penalization*, SIAM J. Optim. **13** (2002), no. 3, 675–692 (electronic) (2003). MR 1972210 (2004c:90058)
- [35] JEYAKUMAR, V. AND LUC, D. T., *Nonsmooth vector functions and continuous optimization*, Springer Optimization and Its Applications, vol. 10, Springer, New York, 2008. MR 2354918 (2010b:90001)
- [36] JI, Y. AND LI, Y. AND ZHANG, K. AND ZHANG, XINLI, *A new nonmonotone trust-region method of conic model for solving unconstrained optimization*, J. Comput. Appl. Math. **233** (2010), no. 8, 1746–1754. MR 2564012
- [37] KAPLAN, A. AND TICHATSCHKE, R., *On inexact generalized proximal methods with a weakened error tolerance criterion*, Optimization **53** (2004), no. 1, 3–17. MR 2040631 (2004k:90095)
- [38] ———, *Bregman-like functions and proximal methods for variational problems with nonlinear constraints*, Optimization **56** (2007), no. 1-2, 253–265. MR 2288516 (2007k:90093)
- [39] ———, *Interior proximal method for variational inequalities on non-polyhedral sets*, Discuss. Math. Differ. Incl. Control Optim. **27** (2007), no. 1, 71–93. MR 2413806 (2009c:47113)
- [40] ———, *Bregman functions and auxiliary problem principle*, Optim. Methods Softw. **23** (2008), no. 1, 95–107. MR 2376773 (2008k:65117)
- [41] KIWIEL, K. C., *Proximal minimization methods with generalized Bregman functions*, SIAM J. Control Optim. **35** (1997), no. 4, 1142–1168. MR 1453294 (98e:90103)

- [42] LUC, D. T., *Theory of vector optimization*, Lecture Notes in Economics and Mathematical Systems, vol. 319, Springer-Verlag, Berlin, 1989. MR 1116766 (92e:90003)
- [43] MARTINET, B., *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Informat. Recherche Opérationnelle **4** (1970), no. Ser. R-3, 154–158. MR 0298899 (45 #7948)
- [44] ———, *Algorithmes pour la Résolution de problèmes d'optimisation et minimax*, Thèse d'état, Université de Grenoble, 1972.
- [45] MIETTINEN, K., *Nonlinear multiobjective optimization*, International Series in Operations Research & Management Science, 12, Kluwer Academic Publishers, Boston, MA, 1999. MR 1784937 (2001g:90005)
- [46] MOREAU, J.-J., *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France **93** (1965), 273–299. MR 0201952 (34 #1829)
- [47] PENG, Y. H., SHI, B. C. AND YAO, S. B., *Nonmonotone trust region algorithm for linearly constrained multiobjective programming*, J. Huazhong Univ. Sci. Technol. Nat. Sci. **31** (2003), no. 7, 113–114. MR 2025639
- [48] POLYAK, B. T., *Introduction to optimization*, Translations Series in Mathematics and Engineering, Optimization Software Inc. Publications Division, New York, 1987, Translated from the Russian, With a foreword by Dimitri P. Bertsekas. MR 1099605 (92b:49001)
- [49] ROCKAFELLAR, R. T., *Monotone operators and the proximal point algorithm*, SIAM J. Control Optimization **14** (1976), no. 5, 877–898. MR 0410483 (53 #14232)
- [50] ———, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Reprint of the 1970 original, Princeton Paperbacks. MR 1451876 (97m:49001)
- [51] ROCKAFELLAR, R. T. AND WETS, R. J.-B., *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317, Springer-Verlag, Berlin, 1998. MR 1491362 (98m:49001)
- [52] SHI, Z. J. AND GUO, J. H., *A new trust region method for unconstrained optimization*, J. Comput. Appl. Math. **213** (2008), no. 2, 509–520. MR 2392574 (2009m:90112)

- [53] TEBoulLE, M., *Entropic proximal mappings with applications to nonlinear programming*, Math. Oper. Res. **17** (1992), no. 3, 670–690. MR 1177730 (93h:90074)
- [54] ———, *Convergence of proximal-like algorithms*, SIAM J. Optim. **7** (1997), no. 4, 1069–1083. MR 1479615 (98k:65035)
- [55] VILLACORTA, K. D. AND OLIVEIRA, P. R., *An interior proximal method in vector optimization*, European J. Oper. Res., aceito para publicação, 2011.
- [56] XI, H. AND SHI, B. C., *A trust region method for multiobjective programming without constraints*, Math. Appl. (Wuhan) **13** (2000), no. 3, 67–69. MR 1774318
- [57] YAO, S. B. AND SHI, B. C. AND PENG, Y. H., *Nonmonotone trust region algorithms for multiobjective programming with linear constraints*, Math. Appl. (Wuhan) **15** (2002), no. suppl., 55–59. MR 1960175
- [58] YIGUI, O. AND QIAN, Z., *A nonmonotonic trust region algorithm for a class of semi-infinite minimax programming*, Appl. Math. Comput. **215** (2009), no. 2, 474–480. MR 2561504 (2010h:90147)
- [59] YU, Z. AND ZHANG, W. AND LIN, J., *A trust region algorithm with memory for equality constrained optimization*, Numer. Funct. Anal. Optim. **29** (2008), no. 5-6, 717–734. MR 2417951 (2009c:90115)