



Alireza Mohebi Ashtiani

OTIMIZAÇÃO GLOBAL DETERMINÍSTICA NO ESPAÇO-IMAGEM:  
PROBLEMAS MULTIPLICATIVOS E FRACIONÁRIOS

Campinas, SP

2012



UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

Alireza Mohebi Ashtiani

OTIMIZAÇÃO GLOBAL DETERMINÍSTICA NO ESPAÇO-IMAGEM:  
PROBLEMAS MULTIPLICATIVOS E FRACIONÁRIOS

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Automação.

Orientador: Prof. Dr. Paulo Augusto Valente Ferreira

Este Exemplar corresponde à versão final da tese defendida pelo aluno e orientado pelo Prof. Dr. Paulo Augusto Valente Ferreira.

---

Campinas, SP

2012

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

As370 Ashtiani, Alireza Mohebi  
Otimização global determinística no espaço–imagem:  
problemas multiplicativos e fracionários / Alireza  
Mohebi Ashtiani. – Campinas, SP: [s.n.], 2012.

Orientador: Paulo Augusto Valente Ferreira.  
Tese de doutorado - Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica e de  
Computação.

1. Otimização. 2. Otimização global. 3. Programação  
matemática 4. Algoritmo Branch and Bound. I.  
Ferreira, Paulo Augusto Valente, 1958-. II. Universidade  
Estadual de Campinas. Faculdade de Engenharia Elétrica  
e de Computação. III. Título.

Título em Inglês: Deterministic global optimization in image–space: multiplicative and  
fractional problems

Palavras-chave em Inglês: Optimization, Global optimization, Mathematical  
programming, Branch–and–Bound Algorithm

Área de concentração: Automação

Titulação: Doutor em Engenharia Elétrica

Banca Examinadora: Paulo Augusto Valente Ferreira, Christiano Lyra Filho, Ricardo  
Coração de Leão Fontoura de Oliveira, Rúbia Mara de Oliveira  
Santos, Alysson Machado Cosa

Data da defesa: 23-11-2012

Programa de Pós Graduação: Engenharia Elétrica

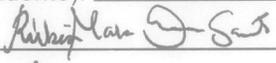
## COMISSÃO JULGADORA - TESE DE DOUTORADO

**Candidato:** Alireza Mohebi Ashtiani

**Data da Defesa:** 23 de novembro de 2012

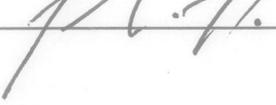
**Título da Tese:** "Otimização Global Determinística no Espaço-Imagem: Problemas Multiplicativos e Fracionários"

Prof. Dr. Paulo Augusto Valente Ferreira (Presidente): 

Profa. Dra. Rúbia Mara de Oliveira Santos: 

Prof. Dr. Alysson Machado Costa: 

Dr. Ricardo Coração de Leão Fontoura de Oliveira: 

Prof. Dr. Christiano Lyra Filho: 



# Dedicatória

*Aos meus pais, Aliakbar e Akram, pelo bom exemplo, pela honestidade, e pela educação que me deram, sempre com muita fé em Deus... .*

# Agradecimentos

Primeiramente, gostaria muito de expressar a minha sincera gratidão ao Professor Dr. Paulo Augusto Valente Ferreira pela excelente orientação constante, cuidadosa, exemplar e efetiva durante todo o processo, pela amizade, paciência, atenção, confiança, dedicação, e principalmente pela oportunidade oferecida de poder realizar este trabalho.

Agradeço também ...

- À minha família, em especial aos meus amados pais, Aliakbar e Akram, para os quais dedico este trabalho, pelo amor, dedicação, educação e incentivo que sempre me deram durante toda minha vida, sei que embora esteja morando muito longe deles estou sempre presente nas orações deles; aos meus queridos irmãos Reza e Hamid; à minha esposa Marcela pela convivência, paciência, companhia e incentivo durante esse período.
- Aos membros da Banca Examinadora, titulares e suplentes, Prof. Dr. Christiano Lyra Filho, Prof. Dr. Ricardo Coração de Leão Fontoura de Oliveira, Profa. Dra. Rúbia Mara de Oliveira Santos, Prof. Dr. Alysson Machado Cosa, Prof. Dr. João Bosco Ribeiro do Val, Prof. Dr. Akebo Yamakami e Prof. Dr. Ricardo Hiroshi Caldeira Takahashi que prontamente aceitaram ao convite.
- Aos meus amigos do Laboratórios do Departamento de Telemática (em ordem alfabética), Benito, Cristiano, Diego, Eduardo, Priscila, Rafael, Renato, Ricardo, Ricardo (Ceará), Taís, Tiago, Walter, pela amizade e por me receberem tão bem e terem me ajudado de diferentes maneiras.
- Ao Professor Dr. Cristiano Torezzan (FCA/UNICAMP) pela amizade e pelos conselhos ao longo dos últimos semestres como meu supervisor no Programa de Estágio Docente na Faculdade de Ciências Aplicadas da UNICAMP.
- Ao Professor Ali Fathi Ashtiani (ensino médio – Escola Shahid Mostafa Zeinali Ash-tiani), quem me ensinou Matemática, me ensinou como ensinar Matemática e me fez gostar de Matemática (um bom professor é sempre lembrado e pode marcar a trajetória dos seus alunos).
- À FEEC/UNICAMP que foram responsáveis pela minha formação.

- À CAPES, pelo apoio financeiro fundamental para a realização deste trabalho.
- Por fim, registro meus sinceros agradecimentos ao Brasil.

# Resumo

Muitos problemas práticos em Engenharia, Economia e Planejamento são modelados de maneira conveniente como problemas de Otimização Global. Esta tese tem como objetivo principal apresentar novas técnicas de Otimização Global com foco na resolução de duas importantes classes de problemas: problemas de Programação Multiplicativa Generalizada, os quais envolvem a minimização e a maximização de uma soma finita de produtos de funções convexas e côncavas, respectivamente, e problemas de Programação Fracionária Generalizada, os quais, por sua vez, envolvem a minimização e a maximização de uma soma finita de razões de funções convexa-côncava ou côncava-convexa, respectivamente. Na tese demonstra-se que todos estes problemas podem ser eficientemente resolvidos por um mesmo algoritmo de aproximação externa, a partir da reformulação dos problemas como problemas com infinitas restrições lineares de desigualdade. Um algoritmo baseado em enumeração de restrições e um algoritmo de aproximação externa combinado a uma técnica *branch-and-bound* são usados para resolver globalmente problemas de Programação Multiplicativa. Em seguida, as mesmas técnicas são empregadas na resolução de problemas de Programação Fracionária. Experiências computacionais atestam a viabilidade e a eficiência dos métodos de Otimização Global propostos, os quais também são facilmente programáveis a partir de pacotes de otimização disponíveis comercialmente.

**Palavras-chave:** Otimização, Otimização Global, Programação Matemática, Algoritmo *Branch-and-Bound*.



# Abstract

Many practical problems in Engineering, Economics and Planning are modeled in a convenient way by Global Optimization problems. The principal objective of this thesis is to introduce new global optimization techniques with focus on the resolution of two important classes of problems: Generalized Multiplicative Programming Problems, in which involve the minimization and maximization of a finite sum of products of convex and concave functions, respectively, and Generalized Fractional Programming Problems, in which, in turn, involve the minimization and maximization of a finite sum of convex–concave and concave–convex ratio functions, respectively. The thesis demonstrates that all these problems can be efficiently solved by the same outer approximation algorithm, from the reformulation of the problems as problems with infinite linear inequality constraints. An algorithm based on a constraint enumeration and an outer approximation algorithm together with a *branch-and-bound* technique are used to globally solve Multiplicative Programming problems. Then, the same techniques are employed in the resolution of Fractional Programming problems. Computational experiments certify the viability and efficiency of the proposed Global Optimization methods, which are also easily programmable through commercially available optimization packages.

**Keywords:** Optimization, Global Optimization, Mathematical Programming, *Branch-and-Bound* Algorithm.



# Lista de Símbolos e Siglas

Os símbolos e notações utilizadas nesta tese são bastante usuais. Sejam  $C$  e  $D$  dois conjuntos,  $A$  uma matriz  $m \times n$ ,  $f$  uma função,  $x$  e  $y$  dois vetores  $n$ -dimensionais reais,  $\mathcal{K}$  um cone. Então denota-se por

$\mathbb{N}$	– O conjunto dos números naturais
$\mathbb{R}^n$	– O conjunto dos vetores $n$ -dimensionais reais
$\mathbb{R}_+^n$	– O conjunto dos vetores não-negativos reais
$\mathbb{R}_{++}^n$	– O conjunto dos vetores positivos reais
$\mathbb{R}_-^n$	– O conjunto dos vetores não-positivos reais
$\mathbb{R}_{--}^n$	– O conjunto dos vetores negativos reais
$:=$	– Significa <i>igual por definição</i>
$x = y$	– Indica que $x_i = y_i, i = 1, 2, \dots, n$
$x \leq y$	– Indica que $x_i \leq y_i, i = 1, 2, \dots, n$
$x < y$	– Indica que $x_i < y_i, i = 1, 2, \dots, n$
$x \neq y$	– Indica que $x_i \neq y_i$ para algum $i \in \{1, 2, \dots, n\}$
$x \in C$	– $x$ pertence a $C$
$x \notin C$	– $x$ não pertence a $C$
$C \setminus \{x\}$	– Indica a subtração do elemento $x$ de $C$
$x^T$	– Indica a transposição de $x$
$\ x\ $	– A norma Euclidiana do vetor $x$
$\langle x, y \rangle$	– O produto escalar entre os vetores $x$ e $y$
$C \subseteq D$	– Indica que $C$ é um subconjunto de $D$
$C \subset D$	– Indica que $C$ é um subconjunto próprio de $D$
$C \cap D$	– A interseção dos conjuntos $C$ e $D$
$C \cup D$	– A união dos conjuntos $C$ e $D$
$C + D$	– A soma vetorial dos conjuntos $C$ e $D$
$C \times D$	– O produto cartesiano dos conjuntos $C$ e $D$
$\text{int } C$	– O interior do conjunto $C$
$\text{cl } C$	– O fecho do conjunto $C$
$\text{co}(C)$	– A casca convexa de $C$
$\text{cone}(C)$	– O fecho cônico de $C$
$\text{efi}(C)$	– O subconjunto dos pontos eficientes de $C$
$\text{comb}(C)$	– O conjunto de todas as combinações convexas de pontos de $C$
$\mathcal{N}_C$	– O cone normal de um conjunto convexo $C$
$\inf_{x \in C} f(x)$	– Indica o ínfimo ou maior limitante inferior de $f$ sobre $C$
$\sup_{x \in C} f(x)$	– Indica o supremo ou menor limitante superior de $f$ sobre $C$
$\partial(C)$	– A fronteira do conjunto $C$
$\mathcal{K} - \{v\}$	– Indica o conjunto $\{k - v : k \in \mathcal{K}\}$
$\mathcal{S}_+^n$	– O conjunto das matrizes simétricas semidefinidas positivas

$S_{++}^n$	- O conjunto das matrizes simétricas definidas positivas
$\{x^n\}$	- A sequência infinita de vetores $x^n, n = 0, 1, \dots$
$\lim_{x \rightarrow \infty} \{x^n\}$	- O limite de $x^n$
argmin	- Indica o argumento do mínimo
argmax	- Indica o argumento do máximo
$\text{epi}(f)$	- O epígrafo da função $f$
$\text{epi}_S(f)$	- O epígrafo restrito da função $f$
$\text{hipo}(f)$	- O hipografo da função $f$
$\text{dom}(f)$	- O domínio efetivo da função $f$
$\partial f(x)$	- O subdiferencial da função $f$
$f'(x; d)$	- A derivada direcional de $f$ em $x \in \mathbb{R}^n$ na direção $d \in \mathbb{R}^n$
$L(f; r)$	- O conjunto de nível da função $f$ (para $r \in \mathbb{R}$ )
$\nabla f$	- O gradiente (ou vector gradiente) da função $f$
$\nabla^2 f$	- A matriz Jacobiana da função $f$
$\text{im}(A)$	- Indica a imagem de $A$
$\text{ker}(A)$	- Indica o núcleo de $A$
$A^T$	- Indica a matriz transposta de $A$
$A > 0$	- Indica que a matriz $A$ é uma matriz simétrica definida positiva
$A \geq 0$	- Indica que a matriz $A$ é uma matriz simétrica semidefinida positiva
$\mathcal{A}(x^*)$	- O conjunto de índices das restrições de desigualdade ativas num ponto $x^*$

### Siglas (Algoritmos)

Algoritmo PMG-ENU	- Algoritmo de enumeração de restrições que resolve $(P_{MG})$
Algoritmo QMG-ENU	- Algoritmo de enumeração de restrições que resolve $(Q_{MG})$
Algoritmo PMG-BB	- Algoritmo <i>branch-and-bound</i> que resolve $(P_{MG})$
Algoritmo QMG-BB	- Algoritmo <i>branch-and-bound</i> que resolve $(Q_{MG})$
Algoritmo PFG-ENU	- Algoritmo de enumeração de restrições que resolve $(P_{FG})$
Algoritmo PR-BB	- Algoritmo <i>branch-and-bound</i> que resolve $(P_{FG})$ e $(Q_{FG})$

# Lista de Figuras

3.1	O mapeamento do espaço de decisões $\Omega$ no espaço–imagem $\mathcal{Y}$ por meio de uma função vetorial $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ . . . . .	42
3.2	Interpretação dos conceito de eficiência. . . . .	44
3.3	Todas as soluções eficientes são encontradas, uma vez que cada solução eficiente admite uma reta suporte no espaço–imagem – Exemplo 3.3.6. O método de soma ponderada permite encontrar as soluções suportadas (não–dominadas) do problema, isto é, pontos na casca convexa do conjunto–imagem. . . . .	52
3.4	Existem soluções eficientes que não são encontradas pelo método da soma ponderada, uma vez que nem todas as soluções eficientes do problema admitem uma reta suporte – Exemplo 3.3.7. Percebe-se que pelo método de soma ponderada não é possível identificar todas as soluções eficientes quando o espaço–imagem não é convexo. . . . .	53
3.5	Interpretação do problema $P_W$ . . . . .	54
3.6	Aproximação da função $\phi$ . . . . .	64
4.1	Árvore <i>branch–and–bound</i> para Exemplo 4.5.4. . . . .	102
5.1	Árvore <i>branch–and–bound</i> para o Exemplo 5.4.6. . . . .	140
5.2	Árvore de <i>branch–and–bound</i> para o Exemplo 5.4.9. . . . .	143



# Lista de Tabelas

4.1	Convergência do Algoritmo QMG-ENU – Exemplo 4.5.1 . . . . .	97
4.2	Convergência do Algoritmo PMG-BB – Exemplo 4.5.1 . . . . .	97
4.3	Convergência do Algoritmo QMG-ENU – Exemplo 4.5.2 . . . . .	98
4.4	Convergência do Algoritmo QMG-BB – Exemplo 4.5.2 . . . . .	99
4.5	Convergência do Algoritmo QMG-ENU – Exemplo 4.5.3 . . . . .	100
4.6	Convergência do Algoritmo PMG-BB – Exemplo 4.5.3 . . . . .	100
4.7	Convergência do Algoritmo QMG-ENU – Exemplo 4.5.4 . . . . .	101
4.8	Convergência do Algoritmo PMG-BB – Exemplo 4.5.4 . . . . .	101
4.9	Resultados dos Algoritmos PMG-ENU e PMG-BB para $(P_{MG})$ com $r = 1, 2, 3$ . .	104
4.10	Resultados dos Algoritmos QMG-ENU e QMG-BB para $(Q_{MG})$ com $r = 1, 2, 3$ . .	105
4.11	Crescimento dos tempos computacionais para $r = 1$ . . . . .	106
5.1	Convergência do Algoritmo PFG-ENU – Exemplo 5.4.1 . . . . .	137
5.2	Convergência do Algoritmo PFG-ENU – Exemplo 5.4.2 . . . . .	138
5.3	Convergência do Algoritmo PFG-ENU – Exemplo 5.4.3 . . . . .	138
5.4	Convergência do Algoritmo PFG-ENU – Exemplo 5.4.4 . . . . .	139
5.5	Convergência do Algoritmo para $(P_{FG})$ – Problemas 5.4.2–5.4.4. . . . .	145
5.6	Resultados dos Algoritmos PFG-ENU e PR-BB . . . . .	148



# Sumário

<b>Lista de Símbolos</b>	<b>xiv</b>
<b>Lista de Figuras</b>	<b>xvi</b>
<b>Lista de Tabelas</b>	<b>xviii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos da Tese . . . . .	5
1.2 Organização da Tese . . . . .	6
<b>2 Fundamentos de Otimização Global</b>	<b>9</b>
2.1 Elementos de Análise Convexa . . . . .	10
2.1.1 Conjuntos Convexos . . . . .	11
2.1.2 Funções Convexas . . . . .	16
2.1.3 Funções Convexas Diferenciáveis . . . . .	21
2.1.4 Funções Convexas não Diferenciáveis . . . . .	23
2.2 Propriedades Gerais de Problemas de Otimização . . . . .	24
2.2.1 Existência de Minimizadores Globais . . . . .	26
2.2.2 Otimização Irrestrita . . . . .	27
2.2.3 Otimização Restrita . . . . .	28
2.3 Otimização Global . . . . .	33
2.3.1 Uma Breve História de Otimização Global . . . . .	34
2.3.2 Alguns Algoritmos de Otimização Global . . . . .	35
2.4 Resumo . . . . .	37

---

<b>3</b>	<b>Otimização no Espaço–Imagem</b>	<b>39</b>
3.1	Otimização Multiobjetivo . . . . .	40
3.2	Conceitos Básicos e Definições . . . . .	41
3.3	Caracterização de Soluções Eficientes . . . . .	47
3.4	Abordagem por Projeção . . . . .	51
3.5	Algoritmo Básico . . . . .	60
3.6	Problema Max–Min . . . . .	62
3.7	Resumo . . . . .	65
<b>4</b>	<b>Programação Multiplicativa Generalizada</b>	<b>67</b>
4.1	Introdução . . . . .	68
4.1.1	Aplicações . . . . .	70
4.2	Programação Multiplicativa Generalizada . . . . .	73
4.2.1	Dualidade no Espaço–Imagem . . . . .	74
4.2.2	Formulações no Espaços Imagem . . . . .	75
4.3	Programação Multiplicativa Generalizada Côncava . . . . .	76
4.3.1	Problema Mestre — Algoritmo Global . . . . .	80
4.3.2	Resolução de $(P_{G^k})$ por Enumeração das Restrições . . . . .	82
4.4	Programação Multiplicativa Generalizada Convexa . . . . .	85
4.4.1	Projeção no Espaço–Imagem . . . . .	86
4.4.2	Algoritmo de Relaxação . . . . .	89
4.4.3	Algoritmo <i>Branch–and–Bound</i> para Resolução de $(P_{MG})$ . . . . .	90
4.4.4	Extensão para Problemas Multiplicativos Generalizados $(Q_{MG})$ . . . . .	94
4.5	Experimentos Numéricos . . . . .	96
4.5.1	Exemplos Aleatórios . . . . .	103

4.6	Resumo . . . . .	106
<b>5</b>	<b>Programação Fracionária Generalizada</b>	<b>107</b>
5.1	Introdução . . . . .	108
5.2	Programação Fracionária Convexa–Côncava . . . . .	110
5.2.1	Abordagem no Espaço–Imagem . . . . .	111
5.2.2	Gerando Problemas–Testes . . . . .	113
5.3	Programação Fracionária Côncava–Convexa . . . . .	118
5.3.1	Uma Abordagem de decomposição . . . . .	119
5.3.2	Projeção no Espaço–Imagem . . . . .	123
5.3.3	Algoritmo de Relaxação . . . . .	127
5.3.4	Algoritmo de Otimização Global para Resolução de $(Q_{FG})$ . . . . .	131
5.3.5	Extensão para Problemas Fracionários Generalizados $(P_{FG})$ . . . . .	134
5.4	Experimentos Numéricos . . . . .	135
5.4.1	Resultados Computacionais do Algoritmo PFG–ENU . . . . .	136
5.4.2	Resultados Computacionais dos Algoritmos PR–BB . . . . .	138
5.4.3	Exemplos Gerados (Pseudo–)Aleatoriamente . . . . .	145
5.5	Resumo . . . . .	146
<b>6</b>	<b>Conclusões</b>	<b>149</b>
6.1	Conclusões . . . . .	150
	<b>Bibliografia</b>	<b>152</b>





# 1

## Introdução

### Sumário

---

1.1	Objetivos da Tese . . . . .	5
1.2	Organização da Tese . . . . .	6

---

O termo otimização refere-se ao estudo de problemas nos quais se busca minimizar ou maximizar uma função, denominada de função objetivo, sujeito a um conjunto de restrições que limitam a escolha dos valores das variáveis de decisão do problema. Problemas de otimização surgem em inúmeras áreas do conhecimento, que incluem Ciências, Engenharia e Economia [PR02].

Em muitas situações, a convexidade da função objetivo ou da região viável do problema não pode ser facilmente verificada, e nesses casos é razoável supor que múltiplos ótimos locais existam. A Otimização Global trata de problemas de otimização não-linear com múltiplas soluções ótimas locais ([HPT95] e [HT93]). Verificar a existência de um ponto viável que satisfaça as condições de otimalidade de um problema não-convexo é um problema NP-difícil. Quando o problema tratado é não-convexo, os procedimentos de otimização clássicos não garantem convergência para uma solução ótima global do problema, eventualmente levando a uma solução inadequada (uma solução local, ou um ponto da região viável que não seja nem mesmo um ótimo local nem global) [Par93].

Em várias aplicações de otimização não-linear, uma soma de produtos de duas funções ou uma soma de funções fracionárias deve ser minimizada ou maximizada. Neste trabalho, propomos métodos de otimização global para estas classes de problemas de otimização globais, os chamados Problemas de Programação Multiplicativa Generalizada (PMG) e de Programação Fracionária Generalizada (PFG).

Aplicações de (PMG) e (PFG) surgem em geometria computacional [KYM91], projeto de circuitos integrados [Wat84], [MMH82] e [DS95], problemas de programação quadrática zero-um [PR87], problemas de programação linear zero-um [NW99], problemas de atribuição quadrática [PR87], problemas de economias de escala [FV93], problemas de regressão linear restrita [HPT95], problemas de complementaridade linear [FV93], problemas de análise de carteiras de investimento [HPT95], problemas de alocação [KTT97], jogos bimatriciais restritos [Man64], microeconomia [HQ71], otimização financeira [KSY93] e [MAF<sup>+</sup>97], otimização de árvores de decisões [Ben94], otimização robusta [MVZ95], mineração de dados e reconhecimento de padrões [BM94], otimização multicritérios [KR93] e em projetos de layouts [QG96], entre outros.

Sabe-se que o produto de funções convexas (côncavas) não é necessariamente uma função convexa (côncava), nem quase-convexa (côncava) [KK92]. Assim, tanto os problemas de Programação Multiplicativa nos quais se minimiza (maximiza) um produto de funções convexas (côncavas), quanto os problemas de Programação Multiplicativa Generalizada

nos quais se minimiza (maximiza) uma soma de produtos de duas funções convexas (côncavas), podem apresentar soluções ótimas locais que não são globais.

Uma técnica tradicional em Programação Multiplicativa ou em Programação Multiplicativa Generalizada é projetar o problema de interesse no *espaço–imagem*. O espaço–imagem é o espaço no qual as funções que constituem o objetivo global a otimizar têm as suas imagens. Diferentes abordagens têm sido propostas para resolver globalmente problemas de Programação Multiplicativa no espaço–imagem [KK95], [BB97], [OFo8] e [OF10]. Os problemas de programação multiplicativa generalizada convexas, descritos pela soma de produtos de duas funções convexas, e côncavos, descritos pela soma de produtos de duas funções côncavas, não são equivalentes, no sentido de que a partir da solução ótima de um destes problemas, uma solução ótima do problema restante não fica caracterizada por alguma transformação conhecida. Em [AF11b] e [AF10a], os autores mostram que estes problemas são equivalentes no espaço–imagem.

Nas últimas décadas, muitas técnicas de otimização global têm sido propostas para resolver globalmente problemas de programação multiplicativa, principalmente por Konno em co-autoria com Kuno e Yajima [KK90], [KK92], [KYK93], [KKY94] e [KK95]. Os algoritmos para resolver problemas multiplicativos podem ser classificados como algoritmos de aproximação externa [KKY94], [Tho91] e [OF10], métodos de enumeração de vértices [Par90], algoritmos *branch–and–bound* [MT92], [KK92], [JMT97], [Beno8] e [AFa], métodos de plano de corte no espaço–imagem dos objetivos [BB00], métodos baseados na análise do espaço–imagem [FP94], algoritmos simplex primal–dual [SS95], métodos heurísticos [BB97] e [LUY99], algoritmos de aproximação discreta [KK90], e algoritmos de simplex paramétricos [Swa66], [KK91], [KK92], [KKY92] e [SS95].

Em outras aplicações, as funções objetivos envolvem razões de funções. O presente trabalho também aborda os problemas de minimizar e maximizar uma soma de funções fracionárias sobre um conjunto convexo, os chamados problemas de Programação Fracionária Generalizada. No caso de minimização (maximização), cada fração é descrita como a razão entre uma função convexa (côncava) e uma função côncava (convexa), ambas positivas sobre na região viável do problema. Em [FJ01], os autores mostram que este problema é essencialmente um problema NP–difícil (apesar de sua estrutura especial) sob as hipóteses usuais sobre numeradores e denominadores. Este problema foi um dos problemas de Programação Fracionária menos pesquisados até cerca de 1990. Esta classe de problemas é atualmente tratada no contexto de otimização global.

Um dos primeiros modelos fracionários propostos na literatura (ainda não classificado desta forma) foi um modelo de equilíbrio para a expansão de uma economia, introduzido por von Neumann em 1937 ([vN37] e [vN45]). O modelo determina a taxa de crescimento de uma economia como o máximo da menor de várias razões entrada-saída. A maioria dos trabalhos em Programação Fracionária foi realizada a partir de 1960. Em 1962, Charnes e Cooper publicaram seu trabalho clássico [CC62], em que mostraram que um problema de programação linear fracionário com apenas uma razão pode ser reduzido a um problema linear, usando uma transformação de variáveis não-linear. Em [Mar64], os problemas fracionários lineares são resolvidos através de um procedimento baseado em pesquisa em vértices adjacentes, de forma análoga à utilizada pelo método simplex para resolver problemas de programação linear. O estudo dos problemas fracionários com apenas uma fração dominou a literatura sobre o assunto até a década de 1980. Muitos dos resultados, aplicações e algoritmos para problemas fracionários são abordados em [Sch78]. Desde então, foram publicadas duas outras revisões da literatura [Cra88] e [SM97] dedicadas exclusivamente à Programação Fracionária. Uma visão geral dos métodos para problemas de localização fracionários é apresentada em [Bar88].

A Programação Fracionária Generalizada têm sido freqüentemente estudada no contexto mais amplo da programação convexa generalizada [ADSZ88]. Razões de funções convexas e côncavas, bem como composições de tais funções, em geral não são convexas, mesmo no caso de razões lineares. Entretanto, a Programação Fracionária vem se beneficiando dos avanços no campo da Convexidade Generalizada [Mar75]. Uma revisão sobre problemas de programação fracionária com uma ou mais razões é apresentado em [Sch95], que contém a bibliografia mais extensa sobre a programação fracionária disponível até a data da sua publicação.

Problemas de Programação Fracionária surgem em problemas de transporte multietágios [AL70], em classificação [Rao71], otimização de carteiras de investimentos [KI88], problemas de corte de estoque [GG63], maximização do retorno de investimentos [Mje83], maximização de retorno/risco e teoria de portfólio [LM97] e [Sch95], problemas de minimização de custo/tempo [BFSZ96], [FDK98] e [Rad98], análise de sensibilidade de decisões [MR96], programação estocástica [CC63] e [Ber65], problemas de otimização geométrica [Ber65], física [Fal69], problemas de contratação governamental [CMW69], planejamento de transportes [AL70] e [FP92], financiamento e investimento [KI89], [KW96], [FP92] e [Sch95] e em muitas outras aplicações.

Nos últimos 40 anos houve um progresso significativo no desenvolvimento de algo-

ritmos determinísticos para problemas de otimização globais [HT93]. Mais recentemente, algoritmos específicos têm sido desenvolvidos, e entre estes, incluem-se os algoritmos para resolver globalmente problemas de Programação Fracionária Generalizada. A maioria destes algoritmos, no entanto, está limitada ao caso linear, ou seja, ao caso em que os numeradores e denominadores são funções afins e a região viável é um poliedro [CMS89], [KYM91], [FP94], [KA99], [KY99], [KFoo], [Kun02] e [PT03]. Em [Ben04], [Ben10] o autor desenvolve algoritmos para problemas em que os numeradores e denominadores são funções afins e a região viável é um conjunto compacto e convexo. Alguns algoritmos têm sido desenvolvidos para resolver globalmente programas fracionários não-lineares [Ben01], [Ben02a], [Ben02b], [DHT01] e [FJ01].

Vários algoritmos especializados têm sido propostos para resolver globalmente problemas de Programação Fracionária Generalizada: algoritmos do tipo simplex ou simplex paramétricos [CC62], [CMS89], [KY92], [Hir95], e [KW96], métodos de aproximação externa [KY99] e [Ben10], métodos heurísticos [KA99] e métodos de enumeração das restrições [AF10b]. No entanto, a abordagem que tem sido aplicada com mais sucesso a problemas de programação fracionária generalizada é baseada em *branch-and-bound* [KFoo], [DHT01], [Kun02], [Ben01], [Ben02a], [Ben02b], [Ben07a], [Ben07b], [Ben10], [AF10b], [AF12] e [AFb]. Para detalhes adicionais e uma revisão geral do problema, recomenda-se [SS03] e [FS04].

## 1.1 Objetivos da Tese

Este trabalho de tese tem como objetivo geral realizar um estudo com foco na resolução de problemas de Programação Multiplicativa e Programação Fracionária Generalizada, duas classes importantes de problemas de otimização global.

Propomos duas abordagens no espaço-imagem para resolver globalmente problemas multiplicativos generalizados. De acordo com a primeira abordagem, estes problemas de otimização não-convexos são reformulados como problemas de Programação Quadrática Indefinida com infinitas restrições lineares de desigualdade. Um algoritmo baseado em uma técnica de enumeração de restrições é usado para resolver o problema resultante. Na segunda abordagem, um procedimento *branch-and-bound* retangular é usado para resolver diretamente os problemas no espaço-imagem.

Em seguida, dois outros algoritmos de otimização global, baseados numa reformulação adequada do problema no espaço-imagem, são propostos para resolver globalmente pro-

blemas de Programação Fracionária Generalizada. Como no caso de problemas multiplicativos, um primeiro algoritmo encontra otimizadores (minimizadores ou maximizadores) globais como limites das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais, resolvidos por meio de um procedimento de enumeração de restrições, enquanto que um segundo algoritmo encontra minimizadores a partir de uma sequência de problemas lineares–fracionais especiais, resolvidos por meio de um procedimento *branch-and-bound* retangular.

## 1.2 Organização da Tese

A tese está organizada em seis capítulos com os seguintes conteúdos:

### Capítulo 1– Introdução

Neste capítulo fêz-se uma introdução geral à tese, foram apresentadas as formulações dos problemas a tratar, uma revisão da literatura sobre os principais problemas e os métodos disponíveis, bem como os objetivos a atingir. Por último, são apresentadas a estrutura e organização da tese.

### Capítulo 2– Fundamentos de Otimização Global

O Capítulo 2 é uma breve revisão de otimização global — conceitos básicos, dos principais modelos e algoritmos de otimização global. Neste capítulo descreve-se genericamente o algoritmo *branch-and-bound* usado na resolução de problemas de otimização global.

### Capítulo 3– Otimização no Espaço–Imagem

Neste capítulo introduz-se a idéia de otimização global no espaço–imagem. Discute-se o conceito de solução eficiente, a caracterização de soluções eficientes, a formulação do problema no espaço–imagem e um algoritmo básico, e, finalmente, alguns resultados sobre Programação Multiobjetivo.

### Capítulo 4– Programação Multiplicativa Generalizada

Neste capítulo são apresentados os problemas de Programação Multiplicativa Generalizada, isto é, os problemas de minimizar e maximizar uma soma finita de produtos de funções convexas e côncavas, respectivamente. Propomos duas novas abordagens

para resolver globalmente estes problemas. Minimizadores e maximizadores globais são obtidos como limites das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais resolvidos por meio de um procedimento de enumeração de restrições. Em seguida, por uma abordagem de otimização global que combina técnicas de aproximação externa e *branch-and-bound*.

### Capítulo 5– Programação Fracionária Generalizada

Este capítulo aborda os problemas de minimizar e maximizar uma soma finita de funções fracionárias, sendo cada fração descrita como a razão entre uma função convexa e uma função côncava, no caso de minimização, e de uma função côncava e uma função convexa, no caso de maximização. Dois algoritmos de otimização global baseados em reformulações adequadas do problema no espaço–imagem são propostos. De acordo com a primeira abordagem, os otimizadores globais do problema são obtidos como limites das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais, resolvidos por meio de um procedimento de enumeração de restrições. De acordo com a segunda abordagem, o problema é resolvido por um algoritmo de otimização que combina dois outros algoritmos: uma técnica *branch-and-bound* retangular, responsável por lidar com a não convexidade do espaço–imagem, e uma técnica de aproximação externa.

### Capítulo 6– Conclusões Gerais

O Capítulo 6 apresenta as conclusões gerais da tese e um resumo dos resultados obtidos nos capítulos anteriores.



# 2

## Fundamentos de Otimização Global

### Sumário

---

2.1	Elementos de Análise Convexa . . . . .	10
2.2	Propriedades Gerais de Problemas de Otimização . . . . .	24
2.3	Otimização Global . . . . .	33
2.4	Resumo . . . . .	37

---

Otimização é um ramo da matemática aplicada dedicado à caracterização do valor extremo de uma função, eventualmente sujeita a restrições sobre os valores dos seus argumentos. As técnicas de otimização surgiram associadas a problemas de logística de pessoal e gestão de transportes. Buscava-se o custo mínimo de operação sujeito a restrições, sendo que tanto a função de custo quanto as restrições eram funções lineares das variáveis de decisão. Problemas destes tipos são da classe de problemas de Programação Linear (PL). O algoritmo mais famosos para esses problemas são o Algoritmo Simplex, proposto por Dantzig no final dos anos 40, início dos 50 [Dan63] e os algoritmos de pontos interiores [Wri97]. A sua complexidade de pior caso é exponencial no número de variáveis do problema, mas o algoritmo se comporta de forma extremamente eficiente na maioria dos casos práticos.

Em muitos casos, a natureza do problema exige variáveis inteiras ou binárias. A imposição de tais restrições torna o problema mais difícil de resolver. Problemas de Programação Linear Inteira Mista (PLIM) são problemas em que parte das variáveis são inteiras, as demais contínuas; a função objetivo e as restrições são funções lineares das variáveis. Não obstante a linearidade da função objetivo do problema, o problema em si não é linear, uma vez que restrições de integralidade são não-lineares. Os dois algoritmos mais usados na resolução de PLIMs são os de planos de cortes e *branch-and-bound* e suas combinações.

Como a maioria dos métodos modernos de otimização global usam técnicas de otimização não-linear clássicas na sua fundamentação, uma visão geral sobre otimização não-linear é apresentada neste capítulo. Introduzem-se alguns conceitos básicos sobre análise convexa e condições necessárias e suficientes para otimalidade, local e global. Uma descrição do método *branch-and-bound* para problemas de otimização global genéricos também é apresentada.

## 2.1 Elementos de Análise Convexa

Convexidade é uma noção muito importante em Teoria de Otimização. Com hipóteses de convexidade e de qualificação de restrições, condições necessárias de otimalidade passam a ser suficientes. Em outras palavras, pode-se dizer que toda solução que satisfaça determinadas condições é uma solução ótima do problema. Todo minimizador local é global, e o conjunto de todos minimizadores globais é um conjunto convexo. A seguir apresentamos algumas noções básicas de Análise convexa utilizadas ao longo deste traba-

lho.

### 2.1.1 Conjuntos Convexos

**Definição 2.1.1.** Diz-se que um conjunto  $\Omega \subseteq \mathbb{R}^n$  é afim se a reta que passa por dois pontos distintos quaisquer em  $\Omega$  está contida inteiramente em  $\Omega$ .

**Definição 2.1.2.** Dados dois pontos quaisquer  $x, y \in \Omega$ , e um escalar real  $\lambda$ , o ponto  $\lambda x + (1 - \lambda)y$  é chamado de combinação afim de  $x$  e  $y$ .

Observe que um conjunto  $\Omega \subseteq \mathbb{R}^n$  é afim se contém a combinação afim de quaisquer dois pontos em  $\Omega$ . O conjunto vazio, o espaço  $\mathbb{R}^n$ , uma reta, os conjuntos formados por um único ponto, subespaços do  $\mathbb{R}^n$  e hiperplanos em  $\mathbb{R}^n$  são alguns exemplos de conjuntos afins.

**Lema 2.1.3.** Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto afim e  $x_0 \in \Omega$ . Então o conjunto

$$\mathcal{S} := \Omega - x_0 = \{x - x_0 \mid x \in \Omega\}.$$

é um subespaço do  $\mathbb{R}^n$ .

*Demonstração.* Suponha que  $x_1, x_2 \in \mathcal{S}$  e  $\lambda_1, \lambda_2 \in \mathbb{R}$ . Então  $x_1 + x_0 \in \Omega$  e  $x_2 + x_0 \in \Omega$ . Como  $\Omega$  é afim e  $\lambda_1 + \lambda_2 + (1 - \lambda_1 - \lambda_2) = 1$ , segue-se que

$$\lambda_1 x_1 + \lambda_2 x_2 + x_0 = \lambda_1(x_1 + x_0) + \lambda_2(x_2 + x_0) + (1 - \lambda_1 - \lambda_2)x_0 \in \Omega.$$

Conclui-se então que  $\lambda_1 x_1 + \lambda_2 x_2 \in \mathcal{S}$ , pois  $\lambda_1 x_1 + \lambda_2 x_2 + x_0 \in \Omega$ .

□

Pelo lema anterior, qualquer conjunto afim  $\Omega$  pode ser escrito como um subespaço mais um constante, na forma  $\Omega = \mathcal{S} + x_0$ ; a dimensão do conjunto afim  $\Omega$  é a dimensão do subespaço  $\mathcal{S} = \Omega - x_0$ , sendo  $x_0$  um elemento qualquer de  $\Omega$ . Um conjunto  $\Omega$  é afim se for uma translação de um espaço linear  $\mathcal{S}$ . Além disso, uma interseção de conjuntos afins ainda é um conjunto afim.

**Definição 2.1.4.** Um conjunto  $\Omega \subseteq \mathbb{R}^n$  é convexo se o segmento de reta que liga quaisquer dois pontos de  $\Omega$  estiver contido inteiramente em  $\Omega$ . Em outras palavras, tem-se  $\lambda x + (1 - \lambda)y \in \Omega$

para todos  $x, y \in \Omega$  e todo  $\lambda \in [0, 1]$ . O ponto  $\lambda x + (1 - \lambda)y$ ,  $\lambda \in [0, 1]$ , é chamado de combinação convexa dos pontos  $x$  e  $y$ .

De forma equivalente, um conjunto  $\Omega \subseteq \mathbb{R}^n$  é convexo se  $\lambda\Omega + (1 - \lambda)\Omega \subseteq \Omega$  para todo  $\lambda \in (0, 1)$ . Exemplos de conjuntos convexos são pontos, retas, segmentos de reta, o conjunto vazio, o espaço  $\mathbb{R}^n$ , qualquer subespaço do  $\mathbb{R}^n$ , vizinhanças, qualquer semi-espaço em  $\mathbb{R}^n$ , qualquer hiperplano em  $\mathbb{R}^n$ , uma bola em  $\mathbb{R}^n$ , um quadrado, uma esfera unitária e conjuntos afins.

É possível provar por indução que a definição de conjunto convexo se estende a uma combinação convexa arbitrária de pontos do conjunto  $\Omega$ .

**Teorema 2.1.5.** *Um conjunto  $\Omega \subseteq \mathbb{R}^n$  é convexo se e somente se  $\Omega$  contém todas as combinações convexas de qualquer número de pontos de  $\Omega$ : para quaisquer  $p \in \mathbb{N}$ ,  $x^i \in \Omega$  e  $\lambda_i \in [0, 1]$ ,  $i = 1, \dots, p$ , tais que  $\sum_{i=1}^p \lambda_i = 1$ , a combinação convexa  $\sum_{i=1}^p \lambda_i x^i$  pertence a  $\Omega$ .*

*Demonstração.* Veja [HPT95], Teorema 1.1. □

A seguinte proposição lista algumas operações com conjuntos que preservam convexidade.

**Proposição 2.1.6** (Propriedades de Conjuntos Convexos). *Interseção arbitrária, multiplicação por escalar, soma, e imagem inversa sob transformações afins preservam convexidade, ou seja*

- (a) *A intersecção  $\cap_{i \in I} \Omega_i$  de qualquer coleção  $\{\Omega_i \mid i \in I\}$  de conjuntos convexos, onde  $I$  é um conjunto qualquer (possivelmente infinito), é convexa;*
- (b) *A soma  $\Omega_1 + \Omega_2$  de dois conjuntos convexos  $\Omega_1$  e  $\Omega_2$  é convexa. Ademais, sejam  $\Omega_i \subseteq \mathbb{R}^n$ ,  $i = 1, 2$ , conjuntos convexos fechados e suponha que pelo menos um deles é limitado. Então o conjunto  $\Omega_1 + \Omega_2$  é convexo e fechado.*
- (c) *Para qualquer conjunto convexo  $\Omega$  e escalar  $\lambda$ , o conjunto  $\lambda\Omega$  é convexo. Além disso, se  $\Omega$  é um conjunto convexo e  $\lambda_1, \lambda_2$  são escalares positivos, tem-se que  $(\lambda_1 + \lambda_2)\Omega = \lambda_1\Omega + \lambda_2\Omega$ ;*
- (d) *O fecho e o interior de um conjunto convexo são convexos;*
- (e) *A imagem e a imagem inversa de um conjunto convexo sob transformações afins são conjuntos convexos.*

*Demonstração.* Veja Proposição 1.2.1 em [BNO03] e Proposição 3.2.3 em [IS05]. □

A caracterização do menor conjunto convexo que contém um dado subconjunto  $\Omega \subseteq \mathbb{R}^n$  é importante em certas discussões envolvendo condições de otimalidade e algoritmos de otimização não-linear.

**Definição 2.1.7.** *Dado um conjunto  $\Omega \subseteq \mathbb{R}^n$  qualquer, o menor conjunto convexo que contém  $\Omega$ , é chamado de casca convexa de  $\Omega$  e denotado por  $\text{co}(\Omega)$ . A casca convexa de  $\Omega$  é a interseção de todos os conjuntos convexos contendo  $\Omega$ , ou seja*

$$\text{co}(\Omega) = \bigcap \{C \mid C \text{ convexo em } \mathbb{R}^n \text{ e } C \supseteq \Omega\}.$$

Como a interseção de qualquer coleção de conjuntos convexos é um conjunto convexo,  $\text{co}(\Omega)$  é um conjunto convexo, qualquer que seja  $\Omega \subseteq \mathbb{R}^n$ . Obviamente, se  $\Omega \subseteq \mathbb{R}^n$  é convexo, então  $\text{co}(\Omega) = \Omega$ . Além disso, a casca convexa de um conjunto compacto é um conjunto compacto ([IS05], Proposição 3.25, página 81). Note que  $\text{co}(\Omega)$  pode não ser fechado mesmo que  $\Omega$  seja um conjunto fechado.

Seja  $\text{comb}(\Omega)$  o conjunto de todas as combinações convexas de pontos em  $\Omega$ . Claramente,  $\text{comb}(\Omega)$  é um conjunto convexo e  $\text{co}(\Omega) \subset \text{comb}(\Omega)$ . Como, por outro lado, um conjunto convexo contém todas as combinações convexas dos seus elementos (Teorema 2.1.5), devemos ter  $\text{comb}(\Omega) \subset \text{co}(\Omega)$ . Portanto, a casca convexa de um conjunto  $\Omega \subseteq \mathbb{R}^n$  qualquer é o conjunto de todas as combinações convexas de pontos de  $\Omega$ .

**Proposição 2.1.8.** *Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto qualquer. A casca convexa de  $\Omega$  é o conjunto de todas as combinações convexas de pontos de  $\Omega$ .*

A Proposição 2.1.8 nos diz que qualquer ponto de  $\text{co}(\Omega)$  pode ser escrito como uma combinação convexa dos pontos de  $\Omega$ . É interessante notar que, na representação de qualquer ponto  $x \in \text{co}(\Omega)$ , no máximo  $d + 1$  pontos de  $\Omega$  são necessários, sendo  $d$  é a dimensão de  $\Omega$ .

**Teorema 2.1.9** (Carathéodory). *Seja  $\Omega \subseteq \mathbb{R}^n$  com  $\dim \Omega = d$ . Então qualquer ponto  $x \in \text{co}(\Omega)$  pode ser escrito como uma combinação convexa de no máximo  $d + 1$  pontos em  $\Omega$ .*

**Definição 2.1.10.** *Um hiperplano  $\mathcal{H}(a, c)$  em  $\mathbb{R}^n$  é um conjunto da forma*

$$\mathcal{H}(a, c) := \{x \in \mathbb{R}^n \mid a^T x = c\},$$

onde  $a \in \mathbb{R}^n \setminus \{0\}$  e  $c \in \mathbb{R}$ . Da mesma forma, definimos os semi-espacos fechados

$$\mathcal{H}_+(a, c) := \{x \in \mathbb{R}^n \mid a^T x \geq c\}, \quad \mathcal{H}_-(a, c) := \{x \in \mathbb{R}^n \mid a^T x \leq c\}.$$

É fácil ver que  $\mathcal{H}(a, c)$ ,  $\mathcal{H}_+(a, c)$  e  $\mathcal{H}_-(a, c)$  são conjuntos convexos. Geometricamente, um hiperplano  $\mathcal{H}(a, c)$  divide o espaço em dois semi-espacos fechados,  $\mathcal{H}_+(a, c)$  e  $\mathcal{H}_-(a, c)$ . Os índices  $+$  e  $-$  significam que  $\mathcal{H}_+(a, c)$  é um semi-espaco fechado na direção  $a$  e  $\mathcal{H}_-(a, c)$  é um semi-espaco fechado na direção  $-a$ , respectivamente. Lembre-se que semi-espacos não são conjuntos afins.

**Definição 2.1.11.** Diz-se que um hiperplano  $\mathcal{H}(a, c)$  é um hiperplano separador de dois conjuntos convexos  $\Omega_1$  e  $\Omega_2$ , se  $\Omega_1 \subset \mathcal{H}_-(a, c)$  e  $\Omega_2 \subset \mathcal{H}_+(a, c)$ . Os conjuntos são estritamente separáveis se  $\Omega_1 \subset \mathcal{H}_-^o(a, c)$  e  $\Omega_2 \subset \mathcal{H}_+^o(a, c)$ , onde  $\mathcal{H}_+^o(a, c)$  e  $\mathcal{H}_-^o(a, c)$  são os semi-espacos abertos correspondente ao hiperplano  $\mathcal{H}(a, c)$ , definidos por

$$\mathcal{H}_+^o(a, c) := \{x \in \mathbb{R}^n \mid a^T x > c\}, \quad \mathcal{H}_-^o(a, c) := \{x \in \mathbb{R}^n \mid a^T x < c\}.$$

**Definição 2.1.12.** Diz-se que um hiperplano  $\mathcal{H}(a, c)$  é um hiperplano suporte a um conjunto convexo  $\Omega \subseteq \mathbb{R}^n$  se  $\Omega \subset \mathcal{H}_+(a, c)$  ou  $\Omega \subset \mathcal{H}_-(a, c)$ , e  $\mathcal{H}(a, c) \cap \partial\Omega \neq \emptyset$ . De maneira equivalente,  $\mathcal{H}(a, c)$  é hiperplano suporte de um conjunto convexo  $\Omega \subseteq \mathbb{R}^n$  se

$$\inf_{x \in \Omega} \langle a, x \rangle = c \quad \text{ou} \quad \sup_{x \in \Omega} \langle a, x \rangle = c.$$

**Definição 2.1.13.** Um conjunto  $\Omega \subseteq \mathbb{R}^n$  definido por um número finito de semi-espacos fechados em  $\mathbb{R}^n$  é chamado de poliedro. Um poliedro limitado não-vazio é chamado de politopo.

Poliedros e politopos são conjuntos convexos, muito importantes em otimização por descreverem geometricamente de uma região viável expressa por restrições lineares.

**Definição 2.1.14.** Diz-se que um conjunto  $\mathcal{K} \subseteq \mathbb{R}^n$  é um cone se  $\lambda\mathcal{K} \subseteq \mathcal{K}$  para todo  $\lambda \geq 0$ . Além disso, se  $\mathcal{K}$  é convexo, então  $\mathcal{K}$  é chamado de cone convexo.

Pela definição, se  $\mathcal{K}$  é um cone não-vazio, necessariamente  $0 \in \mathcal{K}$ . O conjunto  $\mathcal{K}$  é um cone com vértice no ponto  $\{v\}$  se  $\mathcal{K} - \{v\}$  é um cone com vértice na origem. Intuitivamente, um cone é um conjunto de direções. Exemplos de cones (convexos) são o espaco  $\mathbb{R}^n$ , qualquer subespaco do  $\mathbb{R}^n$ , o ortante não-negativo  $\mathbb{R}_+^n$ , semi-espacos fechados determinados por hiperplanos que passam pela origem, o cone formado por vetores com componentes decrescentes  $\mathbb{R}_\geq^n := \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n\}$ , o conjunto das matrizes simétricas semidefinidas positivas  $\mathcal{S}_+^n$ , e o cone de segunda ordem  $\{(t; x) \in \mathbb{R}^{n+1} \mid t \geq \|x\|\}$ .

**Definição 2.1.15.** Seja  $\mathcal{K} \subseteq \mathbb{R}^n$  um cone. O conjunto

$$\mathcal{K}^* := \{y \in \mathbb{R}^n \mid \langle y, d \rangle \leq 0, \quad \forall d \in \mathcal{K}\},$$

é chamado de cone dual de  $\mathcal{K}$ . Em termos geométricos,  $y \in \mathcal{K}^*$  se e somente se  $-y$  é normal a um hiperplano que suporta  $\mathcal{K}$  na origem. O cone  $\mathcal{K}^o = -\mathcal{K}^*$  chamado de cone polar de  $\mathcal{K}^*$ .

Cones duais também desempenham um papel muito importante em otimização. Em particular, temos o resultado a seguir.

**Proposição 2.1.16.** *Seja  $\emptyset \neq \mathcal{K} \subseteq \mathbb{R}^n$  um cone qualquer. Então o cone dual  $\mathcal{K}^*$  é um conjunto convexo fechado.*

**Definição 2.1.17.** *Um cone convexo  $\mathcal{K} \subseteq \mathbb{R}^n$  é poliedral se  $\mathcal{K}$  pode ser representado na forma*

$$\mathcal{K} = \{d \in \mathbb{R}^n \mid Ad \leq 0\},$$

ou

$$\mathcal{K} = \{Ad \mid d \in \mathbb{R}_+^n\},$$

para alguma matriz  $A$ .

O ortante não-negativo é um cone poliedral, mas nem o cone das matrizes simétricas semidefinidas positivas, nem o cone de segunda ordem são cones poliedrais.

**Proposição 2.1.18** (Propriedades de Cones).

- (a) A interseção  $\bigcap_{i \in I} \mathcal{K}_i$  de qualquer coleção  $\{\mathcal{K}_i \mid i \in I\}$  de cones, onde  $I$  é um conjunto qualquer (possivelmente infinito), é um cone;
- (b) O produto cartesiano  $\mathcal{K}_1 \times \mathcal{K}_2$  de dois cones  $\mathcal{K}_1$  e  $\mathcal{K}_2$  é um cone;
- (c) A soma  $\mathcal{K}_1 + \mathcal{K}_2$  de dois cones  $\mathcal{K}_1$  e  $\mathcal{K}_2$  é um cone;
- (d) Para qualquer cone  $\mathcal{K}$ , seu fecho  $\text{cl } \mathcal{K}$  é um cone. Além disso  $\mathcal{K}^* = (\text{cl } \mathcal{K})^* = (\text{co } \mathcal{K})^*$ . Se  $\mathcal{K}$  for um cone qualquer não-vazio, tem-se que  $(\mathcal{K}^*)^* = \text{cl co } \mathcal{K}$  e, em particular, se  $\mathcal{K}$  é convexo e fechado, então  $\mathcal{K} = (\mathcal{K}^*)^*$ .
- (e) A imagem e a imagem inversa de um cone sob transformações lineares e afins são cones.

**Definição 2.1.19.** *Para um conjunto  $\Omega \subseteq \mathbb{R}^n$  qualquer, o fecho cônico de  $\Omega$ , denotado por cone  $\Omega$ , é o menor cone convexo em  $\mathbb{R}^n$  que contém  $\Omega$ . Em outras palavras, o fecho cônico de um conjunto qualquer é a interseção de todos os cones convexos em  $\mathbb{R}^n$  que contêm  $\Omega$ .*

**Proposição 2.1.20** (Fecho Cônico). *Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo. Então*

$$\text{cone } \Omega = \{d \in \mathbb{R}^n \mid d = \lambda x, \ x \in \Omega, \ \lambda \in \mathbb{R}_+\}.$$

*Demonstração.* Uma prova para esta proposição pode ser encontrada em [ISO5], página 61.  $\square$

**Definição 2.1.21.** Diz-se que um cone  $\mathcal{K} \subseteq \mathbb{R}^n$  tem um número finito de geradores quando  $\mathcal{K}$  é o fecho cônico de um número finito de pontos em  $\mathbb{R}^n$ , ou seja, se existem  $p \in \mathbb{N}$  e  $a^i \in \mathbb{R}^n$ ,  $i = 1, \dots, p$ , tais que

$$\mathcal{K} = \text{cone} \{a^i : i = 1, \dots, p\}.$$

onde

$$\text{cone} \{a^i : i = 1, \dots, p\} = \left\{ x \in \mathbb{R}^n \mid x = \sum_{i=1}^p \lambda_i a^i, \lambda_i \geq 0, i = 1, \dots, p \right\}.$$

**Lema 2.1.22.** Qualquer cone com um número finito de geradores é um cone convexo e fechado.

**Teorema 2.1.23** (Teorema de Minkowski–Weyl). Um cone  $\mathcal{K} \subseteq \mathbb{R}^n$  tem um número finito de geradores se e somente se  $\mathcal{K}$  é um cone poliedral.

Uma outra noção importante utilizada na caracterização de condições de otimalidade de um problema com conjunto viável convexo é a de cone normal.

**Definição 2.1.24.** O cone normal de um conjunto convexo  $\Omega \subseteq \mathbb{R}^n$  num dado ponto  $\bar{x} \in \Omega$ , e em relação ao conjunto  $\Omega$ , é definido por

$$\mathcal{N}_{\Omega}(\bar{x}) := \{d \in \mathbb{R}^n \mid \langle d, x - \bar{x} \rangle \leq 0, \forall x \in \Omega\}.$$

### 2.1.2 Funções Convexas

Um dos conceitos mais importantes em Otimização é o conceito de função convexa. A importância desse conceito reside no fato de que qualquer mínimo local é também um mínimo global de um problema convexo, no sentido a ser definido.

**Definição 2.1.25.** Uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função linear se para todos  $x, y \in \mathbb{R}^n$  e  $\lambda_1, \lambda_2 \in \mathbb{R}$ , tem-se que

$$f(\lambda_1 x + \lambda_2 y) = \lambda_1 f(x) + \lambda_2 f(y).$$

**Definição 2.1.26.** Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo. Uma função  $f : \Omega \rightarrow \mathbb{R}$  é convexa quando para quaisquer  $x, y \in \Omega$  e  $\lambda \in [0, 1]$ , tem-se que

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

Se a desigualdade for estrita para todos  $x \neq y$  e  $\lambda \in (0, 1)$ , então  $f$  é uma função estritamente convexa. Diz-se ainda que uma função é fortemente convexa com módulo  $\gamma > 0$ , se para  $x, y \in \Omega$  e  $\lambda \in [0, 1]$ , tem-se que

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \gamma\lambda(1 - \lambda)\|x - y\|^2.$$

Se para quaisquer  $x, y \in \Omega$  e qualquer  $\lambda \in [0, 1]$ ,

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y),$$

então diz-se que a função  $f$  é côncava. Uma função  $f : \Omega \rightarrow \mathbb{R}$  é côncava se  $-f$  for uma função convexa.

Obviamente, uma função linear é uma função convexa (e côncava). Uma função fortemente convexa é uma função estritamente convexa e uma função estritamente convexa é uma função convexa. Geometricamente, para quaisquer pontos  $(x, f(x))$  e  $(y, f(y))$  do gráfico de uma função convexa (respectivamente, côncava)  $f$ , o segmento da reta unindo estes dois pontos localiza-se sempre acima (respectivamente, abaixo) da função.

**Exemplo 2.1.27.** As funções  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dadas por  $f(x) = \ln(e^{x_1} + \dots + e^{x_n})$ ,  $f(x) = \|x\|^p$ ,  $p \geq 1$  e  $f(x) = e^{\beta x^t A x}$ , onde  $A_{n \times n}$  é uma matriz simétrica semidefinida positiva e  $\beta$  é uma escalar positivo, são funções convexas.

**Exemplo 2.1.28.** A função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $f(x) = \|x - x_0\|^2$  para  $x_0 \in \mathbb{R}^n$  é uma função estritamente convexa.

**Exemplo 2.1.29.** A função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $f(x) = x_1^2 + \dots + x_n^2$  é um função fortemente convexa.

Muitas propriedades de funções convexas podem ser derivadas de propriedades de conjuntos convexos. Mais naturalmente associada a funções convexas é a definição de epígrafo de uma função convexa  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , que é o conjunto de pontos em  $\mathbb{R}^{n+1}$  sobre ou acima do gráfico de  $f$ .

**Definição 2.1.30.** Seja  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  uma função qualquer. O epígrafo de  $f$  é dado por

$$\text{epi}(f) := \left\{ (x, r) \in \Omega \times \mathbb{R} \mid f(x) \leq r \right\}.$$

Um conjunto relacionado é o epígrafo restrito de  $f$ , definido por

$$\text{epi}_S(f) := \{(x, r) \in \Omega \times \mathbb{R} \mid f(x) < r\}.$$

Em Análise Convexa é algumas vezes conveniente representar uma função  $f$  como [Roc70]

$$f(x) = \inf \{r \mid (x, r) \in \text{epi}(f)\},$$

e por definição,  $\inf\{\emptyset\} = \infty$ , o que só acontece se  $x$  não pertence ao domínio efetivo da função  $f$ ,  $\text{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$ . Observe então que  $\text{dom}(f)$  é não-vazio se e somente se  $\text{epi}(f)$  é não-vazio, e neste caso tem-se  $\text{dom}(f) = A(\text{epi}(f))$ , onde  $A$ , a projeção de  $\mathbb{R}^{n+1}$  em  $\mathbb{R}^n$ , é dada por  $A(x, r) = x$ . A conexão entre conjuntos convexos e funções convexas é por meio do epígrafo: uma função é convexa se e somente se o seu epígrafo é um conjunto convexo.

**Teorema 2.1.31** (Epígrafo de Funções Convexas). *Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo. Então uma função  $f : \Omega \rightarrow \mathbb{R}$  é convexa em  $\Omega$  se e somente se o epígrafo de  $f$  é um conjunto convexo em  $\mathbb{R}^{n+1}$ .*

Da maneira equivalente, uma função  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  é côncava se e somente se o seu hipografo, definido por

$$\text{hipo}(f) := \{(x, r) \in \Omega \times \mathbb{R} \mid r \leq f(x)\},$$

é um conjunto convexo. Outros conjuntos de interesse no contexto de funções convexas são os conjuntos de nível,  $L(f; r) := \{x \in \Omega \mid f(x) \leq r\}$ .

**Exemplo 2.1.32.** A função  $f : \mathbb{R}^n \times \mathcal{S}_{++}^n \rightarrow \mathbb{R}$  dada por  $f(x, Y) = x^T Y^{-1} x$  é convexa.

O epígrafo da função  $f$  é

$$\begin{aligned} \text{epi}(f) &= \{(x, Y, r) \in \mathbb{R}^n \times \mathcal{S}_{++}^n \times \mathbb{R} : Y > 0, x^T Y^{-1} x \leq r\}, \\ &= \{(x, Y, r) \in \mathbb{R}^n \times \mathcal{S}_{++}^n \times \mathbb{R} : \begin{bmatrix} Y & x \\ x^T & r \end{bmatrix} \geq 0, Y > 0\}, \end{aligned}$$

onde a segunda igualdade é obtida por meio do complemento de Schur [BVo4]. Como  $\text{epi}(f)$  é um conjunto convexo, conclui-se que  $f$  é uma função convexa.

**Teorema 2.1.33** (Convexidade de Conjuntos Nível). *Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo. Se*

$f : \Omega \rightarrow \mathbb{R}$  é convexa, então os conjuntos de nível  $L(f;r) := \{x \in \Omega \mid f(x) \leq r\}$  são convexos para todo  $r \in \mathbb{R}$ .

Note-se que a recíproca do teorema anterior não é verdadeira. Com efeito, existem funções não-convexas cujos conjuntos de nível são convexos. As funções que cujos conjuntos de nível são convexos são as chamadas funções quase-convexas.

**Definição 2.1.34.** *Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo e  $f : \Omega \rightarrow \mathbb{R}$  uma função definida em  $\Omega$ . A função  $f$  é quase-convexa em  $\Omega$  quando os conjuntos de nível  $L(f;r)$  são convexos para todo  $r \in \mathbb{R}$ .*

**Teorema 2.1.35.** *Seja  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo. Então  $f : \Omega \rightarrow \mathbb{R}$  é uma função quase-convexa em  $\Omega$  se, e somente se,*

$$f(\lambda x + (1 - \lambda)y) \leq \max \{f(x), f(y)\}, \quad \forall x, y \in \Omega, \quad \forall \lambda \in [0, 1].$$

*Demonstração.* Suponha-se que todos os conjuntos de nível de  $f$  são convexos. Então para todos  $x, y \in \Omega$ , fazendo  $r := \max\{f(x), f(y)\}$ , tem-se que  $x, y \in L(f;r)$ , e, conseqüentemente, pela convexidade de  $L(f;r)$ ,

$$f(\lambda x + (1 - \lambda)y) \leq r = \max \{f(x), f(y)\}, \quad \forall \lambda \in [0, 1].$$

Reciprocamente, suponha-se que para todos  $x, y \in \Omega$  e  $\lambda \in [0, 1]$  se verifica a desigualdade  $f(\lambda x + (1 - \lambda)y) \leq \tilde{r} = \max\{f(x), f(y)\}$ . Seja  $L(f;r)$  um conjunto de nível arbitrário de  $f$  e sejam  $x$  e  $y$  dois pontos quaisquer de  $L(f;r)$ . Então dado que  $f(x) \leq r$  e  $f(y) \leq r$ , pode-se concluir que  $\tilde{r} := \max\{f(x), f(y)\} \leq r$ , o que implica que  $x, y \in L(f;\tilde{r})$ . Como conseqüência,  $f(\lambda x + (1 - \lambda)y) \leq \tilde{r} \leq r$ . Logo  $\lambda x + (1 - \lambda)y \in L(f;r)$ . □

**Exemplo 2.1.36.** *A função não-convexa  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = x^3$  é uma função quase-convexa em  $\mathbb{R}$ .*

**Proposição 2.1.37.** *Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo e  $f : \Omega \rightarrow \mathbb{R}$ . Então a função  $f$  é quase-côncava em  $\Omega$  se  $-f$  for uma função quase-convexa. Uma função  $f$  simultaneamente quase-convexa e quase-côncava é chamada de quase-linear.*

**Exemplo 2.1.38.** *A função  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = e^{-x^2}$  e a função  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $f(x) = x_1 x_2$  são funções quase-côncavas, mas não côncavas.*

**Exemplo 2.1.39.** A função  $f = \log x$  é quase-linear em  $\mathbb{R}_+$  e a função fracional linear  $f(x) = \frac{a^T x + b}{c^T x + d}$  é quase-linear no semi-espaço  $c^T x + d > 0$ .

**Teorema 2.1.40.** Os resultados seguintes são bem conhecidos e suas demonstrações podem ser encontradas em [BNO03] e [Lue84], entre outros.:

- (a) Seja  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Então, a função  $\lambda f(x)$  é uma função convexa para todo  $\lambda \in \mathbb{R}_+$ ;
- (b) Sejam  $\Omega_i \subseteq \mathbb{R}^n$ ,  $i = 1, 2, \dots, p$  conjuntos convexos e  $f_i : \Omega_i \rightarrow \mathbb{R}$ ,  $i = 1, \dots, p$  funções convexas. Então, a função  $\sum_{i=1}^p \lambda_i f_i(x)$ ,  $\lambda_i \in \mathbb{R}_+$  é uma função convexa em  $\cap_{i=1}^p \Omega_i \subseteq \mathbb{R}^n$ ;
- (c) Sejam  $\Omega_i \subseteq \mathbb{R}^n$ ,  $i = 1, 2, \dots, p$  conjuntos convexos e  $f_i : \Omega_i \rightarrow \mathbb{R}$ ,  $i = 1, \dots, p$  funções convexas. Então,  $\max_{1 \leq i \leq p} f_i(x)$  é uma função convexa em  $\cap_{i=1}^p \Omega_i \subseteq \mathbb{R}^n$ ;
- (d) Sejam  $f : \Omega \subseteq \mathbb{R}^p \rightarrow \mathbb{R}$  uma função convexa e  $A : \mathbb{R}^n \rightarrow \mathbb{R}^p$  uma função afim. Então, a função  $f \circ A : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $(f \circ A)(x) = f(A(x))$  é uma função convexa;
- (e) Sejam  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e  $g : \mathbb{R} \rightarrow \mathbb{R}$  uma função convexa e não-decrescente. Então, a função  $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $(g \circ f)(x) = g(f(x))$  é uma função convexa;
- (f) Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Dados  $x_0, h \in \mathbb{R}^n$  e a função  $\tilde{f}_{x_0, h} : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $\tilde{f}_{x_0, h}(t) = f(x_0 + th)$ , então  $f$  é convexa se e somente se  $\tilde{f}_{x_0, h}$  é convexa para todos  $x_0, h \in \mathbb{R}^n$ ;
- (g) [Desigualdade de Jensen] Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo e  $f : \Omega \rightarrow \mathbb{R}$  uma função convexa. Então, para quaisquer  $p \in \mathbb{N}$ ,  $x^i \in \Omega$  e  $\lambda_i \in \mathbb{R}_+$ ,  $i = 1, \dots, p$  tais que  $\sum_{i=1}^p \lambda_i = 1$ , tem-se que

$$f\left(\sum_{i=1}^p \lambda_i x^i\right) \leq \sum_{i=1}^p \lambda_i f(x^i).$$

*Demonstração.* As provas de (a)–(b) são simples; a prova de (c) é baseada no fato de que o epígrafo do máximo é a interseção dos epígrafos das funções que definem o máximo, que pela Proposição 2.1.6 é um conjunto convexo. O Teorema 2.1.31 implica então na convexidade da função máximo. As provas de (d)–(e) seguem diretamente das definições. Vamos provar (f). Sejam  $f$  uma função convexa,  $x_0, h \in \mathbb{R}^n$  quaisquer e  $\lambda \in [0, 1]$ . Então, para  $t_1, t_2 \in \mathbb{R}$ , calculamos

$$\begin{aligned} \tilde{f}_{x_0, h}(\lambda t_1 + (1 - \lambda)t_2) &= f(x_0 + (\lambda t_1 + (1 - \lambda)t_2)h), \\ &= f(\lambda(x_0 + t_1 h) + (1 - \lambda)(x_0 + t_2 h)), \\ &\leq \lambda f(x_0 + t_1 h) + (1 - \lambda)f(x_0 + t_2 h), \\ &= \lambda \tilde{f}_{x_0, h}(t_1) + (1 - \lambda)\tilde{f}_{x_0, h}(t_2), \end{aligned}$$

implicando que  $\tilde{f}_{x_0, h}$  é convexa. Por outro lado, suponha que  $\tilde{f}_{x_0, h}$  é convexa para quaisquer  $x_0, h \in \mathbb{R}^n$ . Sejam  $x_1, x_2 \in \mathbb{R}^n$  e  $\lambda \in [0, 1]$ . Faça  $x_0 = x_1$  e  $h = x_2 - x_1$ . Neste caso,

$$\begin{aligned} f((1 - \lambda)x_1 + \lambda x_2) &= \tilde{f}_{x_0, h}(\lambda), \\ &= \tilde{f}_{x_0, h}(\lambda \cdot 1 + (1 - \lambda) \cdot 0), \\ &\leq \lambda \tilde{f}_{x_0, h}(1) + (1 - \lambda) \tilde{f}_{x_0, h}(0), \\ &= \lambda f(x_2) + (1 - \lambda) f(x_1), \end{aligned}$$

ou seja,  $f$  é uma função convexa. Para mostrar (g), usa-se a definição de epígrafo e o Teorema 2.1.31 com pontos  $(x^i, f(x^i)) \in \text{epi}(f)$ ,  $i = 1, \dots, p$ . □

A propriedade (c) do Teorema 2.1.40 pode ser generalizada para uma família infinita de funções convexas. Mas neste caso, as funções que definem o máximo (substituído pelo supremo) devem ser uniformemente limitadas superiormente para garantir que a função supremo tenha valores finitos no conjunto  $\Omega$ .

**Teorema 2.1.41** (Continuidade de Funções Convexas). *Seja  $\Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  um conjunto convexo e aberto e  $f : \Omega \rightarrow \mathbb{R}$  uma função convexa. Então  $f$  é localmente Lipschitz-contínua em  $\Omega$ . Em particular,  $f$  é contínua em  $\Omega$ .*

*Demonstração.* Veja [ISO5], página 136. □

### 2.1.3 Funções Convexas Diferenciáveis

Vamos supor que a função  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  seja diferenciável em um conjunto convexo e aberto  $\Omega$ , ou seja, em cada ponto  $x \in \Omega$  gradiente de  $f$ ,  $\nabla f(x)$ , existe. Neste caso, a convexidade de  $f$  admite várias caracterizações úteis.

**Teorema 2.1.42.** *Seja  $f : \Omega \rightarrow \mathbb{R}$  uma função diferenciável em um conjunto convexo e aberto  $\Omega \subseteq \mathbb{R}^n$ . Então  $f$  é uma função convexa se e somente se*

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

para todos  $x, y \in \Omega$ . Em palavras, a aproximação de primeira ordem de uma função convexa sempre está abaixo do gráfico da função em qualquer ponto do seu domínio. Também é possível afirmar que

$f$  é uma função estritamente convexa se a desigualdade acima é estrita para todos  $x, y \in \Omega$  tais que  $x \neq y$ .

Para função  $f$  duas vezes diferenciáveis em  $\Omega$ , temos a seguinte caracterização.

**Teorema 2.1.43.** *Seja  $f : \Omega \rightarrow \mathbb{R}$  uma função duas vezes diferenciável em um conjunto convexo e aberto  $\Omega \subseteq \mathbb{R}^n$ . Então  $f$  é uma função convexa se e somente se  $\nabla^2 f(x)$  é semidefinida positiva em todo ponto  $x \in \Omega$ , ou seja*

$$\langle \nabla^2 f(x)d, d \rangle \geq 0 \quad \forall x \in \Omega, \quad \forall d \in \mathbb{R}^n.$$

*Demonstração.* Suponha que  $f$  é duas vezes diferenciáveis em  $\Omega$  e sejam  $x_1, x_2 \in \Omega$ . Pelo Teorema de Taylor, tem-se que

$$f(x_2) = f(x_1) + (\nabla f(x_1))^T(x_2 - x_1) + \frac{1}{2}(x_2 - x_1)^T \nabla^2 f(x(\lambda))(x_2 - x_1),$$

onde  $x(\lambda) := x_1 + \lambda(x_2 - x_1) \in \Omega$  para algum  $\lambda \in [0, 1]$ . Se  $\nabla^2 f(x)$  é semidefinida positiva em todo ponto  $x \in \Omega$ , tem-se que  $(x_2 - x_1)^T \nabla^2 f(x(\lambda))(x_2 - x_1) \geq 0$ , concluindo-se  $f$  é uma função convexa (Teorema 2.1.42).

Por outro lado, vamos supor que exista  $x \in \Omega$  tal que  $\nabla^2 f(x)$  não é semidefinida positiva, ou seja, existe  $d \in \mathbb{R}^n$  tal que  $d^T \nabla^2 f(x)d < 0$ . Considere o ponto  $\tilde{x} := x + \epsilon d \in \Omega$  para algum  $\epsilon > 0$ . Para  $\epsilon > 0$  suficiente pequeno, o ponto  $x + \epsilon \lambda d = x + \lambda(\tilde{x} - x)$  estará próximo a  $x$  para todo  $\lambda \in [0, 1]$ . Como  $\nabla^2 f(\cdot)$  é contínua, tem-se que  $d^T \nabla^2 f(x + \epsilon \lambda d)d < 0$  para todo  $\lambda \in [0, 1]$ . Usando a expressão anterior, conclui-se que

$$f(\tilde{x}) < f(x) + (\nabla f(x))^T(\tilde{x} - x).$$

Pelo Teorema 2.1.42, segue-se que  $f$  não é convexa, o que é uma contradição. □

Note que o Teorema 2.1.42 implica que se  $\nabla^2 f(x)$  é definida positiva para todo  $x \in \Omega$ , então  $f$  é estritamente convexa. A recíproca não é verdadeira, em geral. Como exemplo,  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = x^4$  é estritamente convexa, mas  $\nabla^2 f(x)$  não é definida positiva no ponto  $x = 0$ .

**Exemplo 2.1.44.** *Seja  $f : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$  dada por  $f(x) = \log(\sum_{i=1}^n e^{x_i})$ . Mostremos que  $f$  é uma função convexa.*

As segundas derivadas parciais de  $f$  são

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \begin{cases} \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}} - \frac{e^{2x_i}}{(\sum_{i=1}^n e^{x_i})^2} & \text{se } i = j, \\ -\frac{e^{x_i+x_j}}{(\sum_{i=1}^n e^{x_i})^2} & \text{se } i \neq j. \end{cases}$$

Portanto,

$$\nabla^2 f(x) = \frac{1}{(e^T z)^2} ((e^T z) \text{diag}(z) - zz^t),$$

onde  $z := (e^{x_1}, \dots, e^{x_n})$ . Para todo  $d \in \mathbb{R}^n$  temos

$$\begin{aligned} d^T \nabla^2 f(x) d &= \frac{1}{(e^T z)^2} \left[ \left( \sum_{i=1}^n z_i \right) \left( \sum_{i=1}^n z_i d_i^2 \right) - \left( \sum_{i=1}^n z_i d_i \right)^2 \right], \\ &= \frac{1}{(e^T z)^2} \left[ \left( \sum_{i=1}^n (\sqrt{z_i})^2 \right) \left( \sum_{i=1}^n (\sqrt{z_i} d_i)^2 \right) - \left( \sum_{i=1}^n \sqrt{z_i} \cdot (\sqrt{z_i} d_i) \right)^2 \right], \\ &\geq 0, \end{aligned}$$

sendo a última passagem decorrente da desigualdade de Cauchy–Schwarz [Limo6]. Portanto,  $f$  é uma função convexa.

#### 2.1.4 Funções Convexas não Diferenciáveis

Mostramos que uma função diferenciável  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é convexa se e somente se, em cada ponto  $\bar{x} \in \mathbb{R}^n$ ,  $f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle$ . Entretanto, nem toda função convexa é diferenciável (e.g.,  $f = |x|$  é uma função convexa não-diferenciável).

A Otimização não Diferenciável atraiu a atenção de pesquisadores em Programação Matemática quando Held e Karp [HK70, HK71] propuseram um algoritmo eficiente para a solução do problema do caixeiro viajante. O algoritmo usava explicitamente como direção de busca o subgradiente de uma função linear por partes.

Quando  $\Omega$  é um conjunto aberto e  $f$  não é diferenciável em algum ponto  $x \in \Omega$ , é possível caracterizar a existência de um conjunto convexo limitado não-vazio de vetores com a propriedade a seguir.

**Definição 2.1.45.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Diz-se que um vetor  $s \in \mathbb{R}^n$  é um*

subgradiente de  $f$  no ponto  $x \in \mathbb{R}^n$  se

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

O conjunto de todos os subgradientes de  $f$  em  $x$  é o subdiferencial de  $f$  em  $x$ , que denotamos por  $\partial f(x)$ . Em outras palavras, qualquer subgradiente define uma aproximação linear para  $f$  cujo gráfico fica abaixo do gráfico de  $f$  e cujo valor coincide com o valor de  $f$  no ponto  $x$ .

**Proposição 2.1.46.** *Uma função convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é diferenciável no ponto  $x \in \mathbb{R}^n$  se e somente se o conjunto  $\partial f(x)$  contém um único elemento. Neste caso,  $\partial f(x) = \{\nabla f(x)\}$ .*

Uma função convexa admite em geral mais de um subgradiente. O subdiferencial  $\partial f(x)$  pode ser caracterizado da seguinte forma:

$$\begin{aligned} \partial f(x) &= \{s \in \mathbb{R}^n \mid f(y) \geq f(x) + \langle s, y - x \rangle, \quad \forall y \in \mathbb{R}^n\}, \\ &= \{s \in \mathbb{R}^n \mid f'(x; d) \geq \langle s, d \rangle, \quad \forall d \in \mathbb{R}^n\}, \end{aligned}$$

onde

$$f'(x; d) := \lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha},$$

define a derivada direcional de  $f$  em  $x \in \mathbb{R}^n$  na direção  $d \in \mathbb{R}^n$ .

**Teorema 2.1.47.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Então o conjunto  $\partial f(x)$  é convexo, compacto e não-vazio para todo  $x \in \mathbb{R}^n$ . Além disso, para toda direção  $d \in \mathbb{R}^n$ , tem-se que*

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

## 2.2 Propriedades Gerais de Problemas de Otimização

O objetivo em Programação Matemática é estudar propriedades e algoritmos para a resolução de problemas que exijam a determinação de valores mínimos ou máximos de funções sujeitos a restrições. Um problema de otimização genérico é um problema da forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in \Omega \subseteq \mathbb{R}^n, \end{aligned} \tag{2.2.1}$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo,  $x \in \mathbb{R}^n$  é o vetor das variáveis de decisão e  $\Omega$  o conjunto de todos os pontos viáveis do problema, também chamado de conjunto viável. Em geral, o conjunto viável  $\Omega$  é definido por restrições de igualdade e desigualdade, na forma

$$\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}, \quad (2.2.2)$$

onde  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ .

**Definição 2.2.1** (Minimizador Local). *Diz-se que um ponto  $x^* \in \Omega$  é um minimizador local do problema (2.2.1) se existe  $\epsilon > 0$  tal que  $f(x^*) \leq f(x)$  para todo  $x \in \Omega$  satisfazendo  $\|x - x^*\| \leq \epsilon$ .*

**Definição 2.2.2** (Minimizador Global). *Diz-se que um ponto  $x^* \in \Omega$  é um minimizador global do problema (2.2.1) se  $f(x^*) \leq f(x)$  para todo  $x \in \Omega$ .*

Nas Definições 2.2.1 e 2.2.2, os minimizadores são estritos quando as desigualdades são estritas para todo  $x \neq x^*$ ; todo minimizador global também é um minimizador local de  $f$  em  $\Omega$ . Problemas de maximização podem ser convertidos em problemas de minimização invertendo-se os sinais das funções objetivos.

**Definição 2.2.3.** *Diz-se que o problema (2.2.1) é problema de Programação Convexa quando  $\Omega$  é um conjunto convexo e  $f$  é uma função convexa.*

Problemas de Programação Linear formam uma classe especial de problemas de Programação Convexa, onde  $\Omega$  é poliedral e  $f$  é uma função linear. O problema (2.2.1) é não-linear se pelo menos uma das funções envolvidas no problema for não-linear.

**Definição 2.2.4.** *Quando  $\Omega = \mathbb{R}^n$ , diz-se que o problema (2.2.1) é irrestrito; quando  $\Omega \neq \mathbb{R}^n$ , diz-se que o problema é restrito.*

**Definição 2.2.5.** *O valor ótimo  $v^*$  do problema (2.2.1) é definido como*

$$v^* := \inf_{x \in \Omega} f(x).$$

*O conjunto de todas as soluções ótimas do problema (2.2.1), que denotamos por  $S_{opt}$ , é dado por  $S_{opt} := \{x \in \mathbb{R}^n : x \in \Omega, f(x) = v^*\}$ .*

### 2.2.1 Existência de Minimizadores Globais

Na Definição 2.2.5, se a função  $f$  não for limitada inferiormente no conjunto  $\Omega$ , então um minimizador global não existirá. Mas mesmo que o valor ótimo  $v^*$  exista e seja limitado, tem-se que  $v^* = \min_{x \in \Omega} f(x)$  se e somente se  $v^*$  for atingido num ponto  $x^* \in \Omega$ . O teorema a seguir garante a existência de minimizadores e maximizadores globais para funções contínuas.

**Teorema 2.2.6** (Teorema de Weierstrass). *Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto compacto não-vazio e  $f : \Omega \rightarrow \mathbb{R}$  uma função contínua em  $\Omega$ . Então existem pelo menos um minimizador e um maximizador global em  $\Omega$ .*

*Demonstração.* Uma prova deste teorema pode ser encontrada em [ISo5], página 10. □

**Corolário 2.2.7.** *Sejam  $\Omega \subseteq \mathbb{R}^n$  e  $f : \Omega \rightarrow \mathbb{R}$  uma função contínua em  $\Omega$ . Suponha que exista  $r \in \mathbb{R}$  tal que  $L(f; r)$  seja compacto e não-vazio. Então o problema (2.2.1) possui um minimizador global em  $\Omega$ .*

*Demonstração.* Pelo Teorema 2.2.6, o problema  $\min_{x \in L(f; r)} f(x)$  possui uma solução global, digamos  $x^* \in \Omega$ . Então, para todo  $x \in \Omega \setminus L(f; r)$  temos que  $f(x) \geq r \geq f(x^*)$ . Conclui-se então que  $x^*$  é um minimizador global do problema (2.2.1). □

A seguir enunciamos um teorema fundamental em Otimização Convexa.

**Teorema 2.2.8** (Teorema de Minimização Convexa). *Sejam  $\Omega \subseteq \mathbb{R}^n$  um conjunto convexo e  $f : \Omega \rightarrow \mathbb{R}$  uma função convexa. Então todo minimizador local é também um minimizador global do problema (2.2.1) e o conjunto de todos minimizadores globais do problema (2.2.1) é um conjunto convexo. Além disso, se  $f$  for uma função estritamente convexa, então o problema (2.2.1) possui no máximo um minimizador global.*

Nas duas seções a seguir descrevemos as condições necessárias e suficientes de otimalidades para problemas de otimização irrestrita e restrita nos quais as funções envolvidas são diferenciáveis.

### 2.2.2 Otimização Irrestrita

Considere o problema de otimização irrestrita

$$\text{minimizar } f(x) \text{ sujeito a } x \in \mathbb{R}^n, \quad (2.2.3)$$

no qual  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função diferenciável ou duas vezes diferenciável. O problema consiste em minimizar uma função de variáveis reais a um valor real sem que haja restrições aos valores das variáveis de decisão. A questão é: como verificarmos que um dado ponto  $x^* \in \mathbb{R}^n$  é um minimizador (local ou global) de  $f$ .

Quando a função objetivo  $f$  é uma função diferenciável ou duas vezes diferenciável, podemos utilizar as informações de derivadas de primeira e segunda ordem para testar se um dado ponto  $x^*$  pode ser um minimizador local (possivelmente estrito) do problema (2.2.3), examinando apenas o gradiente  $\nabla f(x^*)$  e a Hessiana  $\nabla^2 f(x^*)$ .

**Definição 2.2.9** (Ponto Estacionário). *Diz-se que um ponto  $x^* \in \mathbb{R}^n$  é um ponto estacionário do problema (2.2.3) se  $\nabla f(x^*) = 0$ .*

**Teorema 2.2.10** (Condições Necessárias de Primeira Ordem). *Se  $x^*$  é um minimizador local do problema (2.2.3) e  $f$  é diferenciável em uma vizinhança aberta de  $x^*$ , então  $\nabla f(x^*) = 0$ .*

Pelo Teorema 2.2.10, todo minimizador local deve ser um ponto estacionário do problema (2.2.3). Além disso, o Teorema 2.2.10 vale para um problema de minimização restrita da forma (2.2.1) quando o conjunto viável  $\Omega$  é aberto.

**Teorema 2.2.11** (Condições Necessárias de Segunda Ordem). *Se  $x^*$  é um minimizador local do problema (2.2.3) e  $f$  é duas vezes diferenciável em uma vizinhança aberta de  $x^*$ , então  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é semidefinida positiva.*

As condições necessárias de primeira e segunda ordem, Teoremas 2.2.10 e 2.2.11, são enunciadas supondo-se que  $x^*$  seja um minimizador local do problema (2.2.3), por meio de condições sobre  $\nabla f(x^*)$  e  $\nabla^2 f(x^*)$ . A seguir descrevemos condições que se satisfeitas por um ponto  $x^*$ , garantem que este ponto é um minimizador local de  $f$ , ou seja, enunciamos condições suficientes para que um ponto  $x^*$  seja um minimizador local de  $f$ .

**Teorema 2.2.12** (Condições Suficientes de Segunda Ordem). *Seja  $f$  uma função duas vezes diferenciável em uma vizinhança aberta de  $x^*$ . Se  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é definida positiva, então  $x^*$  é um minimizador local estrito do problema (2.2.3).*

Note que as condições do Teorema 2.2.12 garantem algo mais forte do que condições para minimizadores locais, já que garantem que o ponto será um minimizador local estrito. Note ainda que as condições do Teorema 2.2.12 não são necessárias no sentido de minimização estrita, ou seja, um ponto  $x^*$  pode ser um minimizador local estrito e não satisfazer as condições do Teorema 2.2.12, como por exemplo no caso  $f(x) = x^4$  e  $x^* = 0$ .

A seguir mostramos que quando a função objetivo é convexa, minimizadores locais e globais são simples de caracterizar.

**Teorema 2.2.13** (Condições Necessárias e Suficientes para um Problema Convexo Irrestrito). *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Então todo minimizador local é também um minimizador global do problema problema (2.2.3). Em particular, se  $f$  é diferenciável, então qualquer ponto estacionário de  $f$  é um minimizador global do problema (2.2.3).*

*Demonstração.* A primeira parte da demonstração do teorema pode ser encontrada por exemplo em [ISO5], Teorema 3.1.5. Vamos provar a segunda parte. Suponha que  $f$  é convexa e diferenciável. Neste caso, pelo Teorema 2.1.42 tem-se que

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

para todos  $x, y \in \mathbb{R}^n$ . Tomando  $x = x^*$ , e como  $x^*$  é um ponto estacionário, temos então que

$$f(y) \geq f(x^*), \quad \forall y \in \mathbb{R}^n,$$

o que completa a prova. □

### 2.2.3 Otimização Restrita

Na seção anterior estudamos as condições necessárias e suficientes de otimalidade para problemas de otimização irrestrita. Nesta seção estamos interessados em condições para problemas da forma

$$\text{minimizar } f(x) \text{ sujeito a } x \in \Omega, \tag{2.2.4}$$

no qual  $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$ , onde  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ .

Para desenvolver condições de otimalidade no caso restrito, precisamos de levar em

conta a estrutura do conjunto viável do problema. Como já mencionado na seção anterior, o Teorema 2.2.10 vale para um problema de minimização restrita da forma (2.2.4) quando o conjunto viável  $\Omega$  é aberto.

**Corolário 2.2.14.** *Seja  $x^* \in \text{int } \Omega$  um minimizador local do problema (2.2.4) e  $f$  diferenciável em  $x^*$ . Então  $\nabla f(x^*) = 0$ .*

Separaremos a discussão do caso geral em dois sub-problemas. No primeiro, consideramos problemas de otimização restrita com restrições de igualdade, ou seja, problemas da forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in \Omega = \{x \in \mathbb{R}^n \mid h(x) = 0\}, \end{aligned} \quad (2.2.5)$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ .

**Definição 2.2.15** (Condições de Qualificação de Restrições 1). *Dado um conjunto  $\Omega$  definido como no problema (2.2.5), dizemos que  $x^* \in \Omega$  atende uma condição de qualificação de restrições, se uma das seguintes condições é satisfeita: (1) o conjunto  $\{\nabla h_i(x^*), i = 1, \dots, l\}$  é um conjunto linearmente independente, ou seja,  $\text{rank } \nabla h(x^*) = l$ ; (2)  $h$  é uma função afim, ou seja,  $h(x) = Ax - b$ , onde  $A \in \mathbb{R}^{l \times n}$  e  $b \in \mathbb{R}^l$ . Em qualquer caso, dizemos que  $x^*$  é um ponto regular das restrições de (2.2.5).*

**Teorema 2.2.16** (Condição Necessária para Problemas Linearmente Restritos). *Sejam  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função diferenciável no ponto  $x^* \in \mathbb{R}^n$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  também diferenciável em  $x^*$ . Seja  $x^*$  um minimizador local do problema (2.2.5) e um ponto regular das restrições. Então*

$$\langle \nabla f(x^*), d \rangle = 0, \quad \forall d \in \ker \nabla h(x^*). \quad (2.2.6)$$

A expressão (2.2.6) pode ser reescrita na forma

$$-\nabla f(x^*) \in \text{im}(\nabla h(x^*))^T,$$

que é equivalente a existência de um vetor  $\lambda \in \mathbb{R}^l$  tal que

$$\nabla f(x^*) = (\nabla h(x^*))^T \lambda, \quad (2.2.7)$$

que é a conhecida condição clássica de Lagrange.

**Definição 2.2.17** (Função Lagrangiana para o Problema (2.2.5)). *Para o problema de otimização restrita (2.2.5), a função Lagrangeana  $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$  é dada por*

$$\mathcal{L}(x, \lambda) := f(x) + \sum_{i=1}^l \mu_i h_i(x).$$

A relação (2.2.7) leva a uma condição de otimalidade baseada na função Lagrangeana. A condição é de primeira ordem, pois utiliza apenas a informação dos gradientes da função objetivo e das restrições.

**Teorema 2.2.18** (Condição de Otimalidade de Lagrange para Problemas Linearmente Restritos). *Sejam  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função diferenciável no ponto  $x^* \in \mathbb{R}^n$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  também diferenciável em  $x^*$ . Seja  $x^*$  um minimizador local e um ponto regular do problema (2.2.5). Então existe  $\lambda^* \in \mathbb{R}^l$  tal que*

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0. \quad (2.2.8)$$

**Definição 2.2.19.** *Dizemos que um ponto  $x^* \in \mathbb{R}^n$  é um ponto estacionário do problema (2.2.5) se  $x^* \in \Omega$  e a equação (2.2.8) é satisfeita para algum  $\lambda^* \in \mathbb{R}^l$ . A este  $\lambda^*$  chamamos de multiplicador de Lagrange associado ao ponto estacionário  $x^*$ .*

Pelo Teorema 2.2.18, qualquer minimizador local que satisfaça a condição de qualificação de restrições é um ponto estacionário do problema (2.2.5).

**Teorema 2.2.20** (Condição Necessária de Segunda Ordem para Problemas Linearmente Restritos). *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  sejam funções diferenciáveis  $x^*$ , um minimizador local e ponto regular das restrições do problema (2.2.5). Então para  $\lambda^* \in \mathbb{R}^l$  que satisfaça (2.2.8),*

$$\langle \nabla_x^2 \mathcal{L}(x^*, \lambda^*) d, d \rangle \geq 0 \quad \forall d \in \ker \nabla h(x^*). \quad (2.2.9)$$

*Demonstração.* Veja Teorema 2.3.1 em [ISo5].

□

**Teorema 2.2.21** (Condição Suficiente de Segunda Ordem para Problemas Linearmente Restritos). *Sejam  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  funções diferenciáveis num ponto  $x^* \in \mathbb{R}^n$ . Se  $x^* \in \Omega$  e existe  $\lambda^* \in \mathbb{R}^l$  satisfazendo (2.2.8) tal que*

$$\langle \nabla_x^2 \mathcal{L}(x^*, \lambda^*) d, d \rangle > 0 \quad \forall d \in \ker \nabla h(x^*) \setminus \{0\}, \quad (2.2.10)$$

então  $x^*$  é um minimizador local estrito do problema (2.2.5).

*Demonstração.* Veja [ISO5], Teorema 2.3.2, para uma prova deste teorema. □

Vamos considerar agora um problema de otimização com restrições de igualdade e desigualdade:

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in \Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}, \end{array} \quad (2.2.11)$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . O tratamento de restrições de desigualdade requer os conceitos de restrição ativa e inativa.

**Definição 2.2.22.** Dado um ponto  $x \in \mathbb{R}^n$ , dizemos que a restrição de desigualdade  $g_i$ ,  $i \in \{1, \dots, m\}$  está ativa em  $x$  se  $g_i(x) = 0$  e inativa se  $g_i(x) < 0$ . O conjunto de índices das restrições de desigualdade ativas num ponto  $x^* \in \mathbb{R}^n$  será denotado por  $\mathcal{A}(x^*) = \{i \in \{1, \dots, m\} : g_i(x^*) = 0\}$ .

Vamos supor  $h$  e  $g$  diferenciáveis em  $x^* \in \Omega$ , e definir o cone obtido a partir dos gradientes das restrições de igualdade e de desigualdade no ponto  $x^*$  como

$$\mathcal{C}(x^*) := \{d \in \ker \nabla h(x^*) \mid \langle \nabla g(x^*), d \rangle \leq 0, \forall i \in \mathcal{A}(x^*)\}. \quad (2.2.12)$$

**Definição 2.2.23** (Condições de Qualificação de Restrições 2). Dado um conjunto  $\Omega$  definido como no problema (2.2.11), dizemos que  $x^* \in \Omega$  atende uma condição de qualificação de restrições, se uma das seguintes condições é satisfeita: (1)  $\{\nabla h_i(x^*), i = 1, \dots, l\} \cup \{\nabla g_i(x^*), i \in \mathcal{A}(x^*)\}$  é um conjunto linearmente independente; (2)  $h$  e  $g$  são funções afins; (3)  $\{\nabla h_i(x^*), i = 1, \dots, l\}$  é um conjunto linearmente independente e

$$\exists d^* \in \ker \nabla h_i(x^*) \quad \text{tal que} \quad \langle \nabla g(x^*), d^* \rangle < 0 \quad \forall i \in \mathcal{A}(x^*).$$

Em qualquer caso, dizemos que o ponto  $x^*$  é um ponto regular das restrições do problema (2.2.11).

**Teorema 2.2.24** (Condição Necessária de Primeira Ordem para Problemas Mistos). Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  sejam funções diferenciáveis num ponto  $x^* \in \mathbb{R}^n$ . Se  $x^*$  é um minimizador local e um ponto regular do problema (2.2.11), então

$$\langle \nabla f(x^*), d \rangle \geq 0 \quad \forall d \in \mathcal{C}(x^*), \quad (2.2.13)$$

ou seja

$$-f(x^*) \in (\mathcal{C}(x^*))^*.$$

*Demonstração.* Veja [ISO5], Teorema 4.1.3. □

**Teorema 2.2.25** (Condição Suficiente de Primeira Ordem para Problemas Mistos). *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  sejam funções diferenciáveis num ponto  $x^* \in \Omega$ . Se*

$$\langle \nabla f(x^*), d \rangle > 0 \quad \forall d \in \mathcal{C}(x^*) \setminus \{0\}, \quad (2.2.14)$$

então  $x^*$  é um minimizador local estrito do problema (2.2.11).

*Demonstração.* Veja [ISO5], Teorema 4.1.1. □

Combinando o Teorema 2.2.24 e o Lema de Farkas (ver [ISO5]) podemos afirmar que existem  $\lambda^* \in \mathbb{R}^l$  e  $\mu_i^* \in \mathbb{R}_+$ ,  $i \in \mathcal{A}(x^*)$ , tais que

$$-f(x^*) = \sum_{i=1}^l \lambda_i^* \nabla h_i(x^*) + \sum_{i \in \mathcal{A}(x^*)} \mu_i^* \nabla g_i(x^*). \quad (2.2.15)$$

**Definição 2.2.26** (Função Lagrangeana para o Problema (2.2.11)). *Para o problema de otimização restrita (2.2.11), a função Lagrangiana  $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^p \rightarrow \mathbb{R}$  é dada por*

$$\mathcal{L}(x, \lambda, \mu) := f(x) + \sum_{i=1}^l \lambda_i h_i(x) + \sum_{i=1}^p \lambda_i g_i(x).$$

A partir da expressão (2.2.15), chegamos ao celebrado Teorema de Karush–Kuhn–Tucker.

**Teorema 2.2.27** (Condição de Karush–Kuhn–Tucker). *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  sejam funções diferenciáveis num ponto  $x^* \in \mathbb{R}^n$ . Se  $x^*$  é um minimizador local e um ponto regular do problema (2.2.11), então existem  $\lambda^* \in \mathbb{R}^l$  e  $\mu^* \in \mathbb{R}_+^m$  tais que*

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0, \quad (2.2.16)$$

$$\mu_i^* g_i(x^*) = 0, \quad i = 1, \dots, m. \quad (2.2.17)$$

A condição (2.2.17), conhecida como condição de complementaridade, implica que os multiplicadores de Lagrange correspondentes às restrições de desigualdade inativas são iguais a zeros. Assim, podemos omitir os termos com índices  $i \notin \mathcal{A}(x^*)$  na condição (2.2.16).

**Definição 2.2.28.** Diz-se que um ponto  $x^* \in \Omega$  é ponto estacionário do problema (2.2.11), se existem  $\lambda^* \in \mathbb{R}^l$  e  $\mu^* \in \mathbb{R}_+^m$  tais que as condições (2.2.16) e (2.2.17) são satisfeitas. Os vetores  $\lambda^*$  e  $\mu^*$  são chamados de multiplicadores de Lagrange associados ao ponto estacionário  $x^*$ .

Pelo Teorema 2.2.27, qualquer minimizador local (regular) é um ponto estacionário do problema (2.2.11). As condições de Karush–Kuhn–Tucker, Teorema 2.2.27, são necessárias: um ponto estacionário não precisa ser um minimizador local do problema. Entretanto, se o problema for convexo, as condições tornam-se também suficientes.

**Teorema 2.2.29.** Seja  $x^* \in \mathbb{R}^n$  um ponto estacionário de um problema (2.2.11) no qual  $f$  e  $g$  são funções convexas e  $h$  é uma função afim (isto é, de um problema convexo). Então  $x^* \in \mathbb{R}^n$  é um minimizador local (e global) do problema (2.2.11).

## 2.3 Otimização Global

Uma modelagem mais realista dos problemas do mundo real muitas vezes exige a incorporação de não-convexidades no modelo, seja na função objetivo ou nas restrições. Problemas não-convexos apresentam muitos desafios, tanto no aspecto prático quanto no aspecto teórico. A presença de minimizadores locais que não são também globais exige a aplicação de técnicas de Otimização Global.

Um fator preponderante para a classificação de problemas de otimização entre “fáceis” e “difíceis” é o problema ser convexo ou não-convexo. Um problema é convexo se consiste na minimização de uma função convexa (ou maximização de uma função côncava) numa região convexa. Qualquer minimizador local de um problema convexo também é global (Teorema 2.2.8). Esta propriedade é muito importante em termos práticos, uma vez que os critérios de convergência aplicáveis a minimizadores locais levam a convergência a minimizadores globais.

O ramo da matemática aplicada dedicado à caracterização de extremos de problemas não-convexos é chamado de Otimização Global. Observe que se parte ou a totalidade

das variáveis de um problema estiver restrita a assumir apenas valores discretos, então o problema é não-convexo, mesmo que as funções envolvidas sejam lineares. Isso acontece porque a região viável do problema é um conjunto discreto de pontos, um conjunto não-convexo. Esta classe de problemas não-convexos (discretos) dispõe de algoritmos de otimização eficientes, reunidos sob a denominação genérica de Otimização Combinatória.

### 2.3.1 Uma Breve História de Otimização Global

Antes da introdução e uso de computadores eletrônicos, a exigência de otimização global não era contemplada devido ao enorme esforço de cálculo implicado. Otimalidade global era garantida apenas quando o problema era convexo, o que limitava os resultados obtidos. Os primeiros métodos utilizados em Otimização Global foram derivados de técnicas determinísticas, principalmente na técnica *dividir-e-conquistar*, introduzidos no final dos anos 1950.

Um algoritmo típico que utiliza o princípio de *dividir-e-conquistar* é o algoritmo *branch-and-bound*. Devido à natureza do algoritmo, pelo qual subproblemas são produzidos por uma ramificação do problema original em suas instâncias possíveis, o algoritmo *branch-and-bound* aplica-se muito bem a problemas discretos. Assim, as primeiras aplicações do algoritmo *branch-and-bound* foram dirigidas a problemas discretos, como o Problema do Caixeiro Viajante (TSP).

O primeiro artigo a tratar um problema de otimização global contínuo com uma técnica *branch-and-bound* (determinístico) foi publicado em 1969 [FS69]. Nas décadas de 1970 e 1980, trabalhos na área de otimização global determinística contínua ou inteira-mista eram escassos. A maioria dos trabalhos publicados neste período tratava de aplicações de Otimização Global a casos muito específicos, ou discutiam resultados teóricos sobre convergência.

A partir do final da década de 1980, trabalhos em otimização global determinística começaram a ser mais frequentes na literatura. Os trabalhos estavam concentrados em métodos iterativos aplicáveis a classes particulares de problemas [Par88], [MM88], [Hor88], [HDT88], [TH88], [KG88], [MM89], mas também surgiram estudos teóricos [HJL88] e [Hor89], e estudos sobre implementações paralelas [Par89], [Ge89] e [ES89].

No início da década de 1990, a maioria dos trabalhos ainda estava concentrada em aplicações de Otimização Global ou em algoritmos que comportavam-se bem em classes par-

ticulares de problemas. O primeiro algoritmo capaz de lidar diretamente com problemas não-convexos contínuos, foi o *branch-and-reduce* [RS95] e [RS96]. Pouco tempo depois, surgiu o algoritmo  $\alpha$ BB *branch-and-bound* pelo grupo de Floudas [AMF95], [AAMF96], [AAF97], [AAF98] e [ADFN98]. Posteriormente, uma série de algoritmos *branch-and-select* voltada para formulações não-convexas e para problemas não-lineares inteiros foram propostos [Smi96], [SP97], [SP99], [ZG99] e [KBoo].

### 2.3.2 Alguns Algoritmos de Otimização Global

Os métodos de otimização são geralmente divididos em duas categorias principais: *métodos determinísticos* e *métodos estocásticos*. Um algoritmo é determinístico se o algoritmo pode ser simulado por um autômato finito determinístico. Em cada fase de um algoritmo determinístico, sempre existe uma função de escolha que permite selecionar o próximo passo de maneira única; as provas de convergência para este tipo de algoritmo não envolvem a teoria da probabilidade. Um algoritmo estocástico, por outro lado, envolve um elemento aleatório na escolha do próximo passo da computação; as provas de convergência para este tipo de algoritmo envolvem o uso da teoria das probabilidades.

Neste trabalho, consideramos apenas algoritmos determinísticos, e dentre estes, algoritmos *branch-and-bound*. Estes algoritmos operam com limites superiores e inferiores do valor ótimo global e convergem para uma solução  $\epsilon$ -sub-ótima do problema. O algoritmo *branch-and-bound* adota a estratégia de divisão e conquista para problemas de Otimização Global. A estrutura geral do algoritmo é apresentada a seguir.

#### Algoritmo *Branch-and-Bound*

**Passo 0.** Inicialização do algoritmo;

**Passo 1.** Dividir um problema  $P$  em um conjunto de subproblemas menores  $\{P_k\}$  de forma que a solução de  $P$  possa ser obtida através da solução dos subproblemas  $P_k$ ;

**Passo 2.** Resolver os subproblemas  $P_k$ ;

**Passo 3.** Descartar subproblemas que não contenham a solução ótima;

**Passo 4.** Obter a solução do problema original a partir das soluções dos subproblemas.

Esta estratégia tem sido uma das mais utilizadas para encontrar a solução ótima de problemas de otimização NP-difíceis. Os algoritmos dividem o espaço das soluções; as divisões são feitas iterativamente, com a garantia de que os subproblemas sejam mais fáceis de resolver (de forma independente) do que o problema original. Além disso, procura-se

descartar subproblemas que não contenham uma solução ótima. A questão é como *quebrar*  $P$  em subproblemas menores, e depois os *recombinar* de maneira a obter uma solução ótima para o problema original.

O algoritmo de *branch-and-bound* que descrevemos aqui encontra um minimizador global de uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sobre um retângulo  $\mathcal{R}_{init}$  de dimensão  $m$ . Para um retângulo  $\mathcal{R} \subseteq \mathcal{R}_{init}$  definimos

$$\Phi_{min}(\mathcal{R}) = \min_{q \in \mathcal{R}} f(q).$$

Em seguida, o algoritmo estima  $\Phi_{min}(\mathcal{R}_{init})$  com uma precisão absoluta de  $\epsilon > 0$ , usando duas funções  $\Phi_{lb}(\mathcal{R})$  e  $\Phi_{ub}(\mathcal{R})$ , definidas em  $\{\mathcal{R} \mid \mathcal{R} \subseteq \mathcal{R}_{init}\}$  (cujos valores, presumivelmente, são mais fáceis de calcular do que  $\Phi_{min}(\mathcal{R})$ ). Estas duas funções satisfazem as seguintes condições:

**Condição 1.**  $\Phi_{lb}(\mathcal{R}) \leq \Phi_{min}(\mathcal{R}) \leq \Phi_{ub}(\mathcal{R})$ . As funções  $\Phi_{lb}$  e  $\Phi_{ub}$  fornecem um limite inferior e um limite superior para  $\Phi_{min}$ , respectivamente.

**Condição 2.** Quando o meio comprimento máximo dos lados de  $\mathcal{R}$ , denotado por  $\text{size}(\mathcal{R})$ , tende a zero, a diferença entre o limite superior e inferior converge *uniformemente* para zero, ou seja,

$$\forall \epsilon > 0, \exists \delta > 0 \text{ tal que} \\ \forall \mathcal{R} \subseteq \mathcal{R}_{init}, \quad \text{size}(\mathcal{R}) \leq \delta \implies \Phi_{ub}(\mathcal{R}) - \Phi_{lb}(\mathcal{R}) \leq \epsilon.$$

Qualitativamente falando, os limites  $\Phi_{lb}$  e  $\Phi_{ub}$  convergem para  $\Phi_{min}$  com a redução do retângulo inicial para um ponto.

Vamos agora descrever o algoritmo. Começamos calculando os valores  $\Phi_{lb}(\mathcal{R}_{init})$  e  $\Phi_{ub}(\mathcal{R}_{init})$ . Se  $\Phi_{ub}(\mathcal{R}_{init}) - \Phi_{lb}(\mathcal{R}_{init}) \leq \epsilon$ , então o algoritmo termina. Caso contrário, particionamos  $\mathcal{R}_{init}$  em sub-retângulos,  $\mathcal{R}_{init} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \dots \cup \mathcal{R}_N$ , e calculamos  $\Phi_{lb}(\mathcal{R}_i)$  e  $\Phi_{ub}(\mathcal{R}_i)$  para  $i = 1, 2, \dots, N$ . Desta forma,

$$\min_{1 \leq i \leq N} \Phi_{lb}(\mathcal{R}_i) \leq \Phi_{min}(\mathcal{R}_{init}) \leq \min_{1 \leq i \leq N} \Phi_{ub}(\mathcal{R}_i).$$

Se a diferença entre os novos limites for menor ou igual a  $\epsilon$ , então o algoritmo termina. Caso contrário, a partição de  $\mathcal{R}_{init}$  é refinada e os limites atualizados. Na descrição a se-

guir,  $k$  representa o índice de iteração;  $\mathcal{L}_k$  denota a lista de retângulos;  $L_k$  o limite inferior e  $U_k$ , o limite superior de  $\Phi_{min}(\mathcal{R}_{init})$  ao final de cada iteração  $k$ .

### Algoritmo Padrão *Branch-and-Bound*

$k = 0$ ;  
 $\mathcal{L}_0 = \{\mathcal{R}_{init}\}$ ;  
 $L_0 = \Phi_{lb}(\mathcal{R}_{init})$ ;  
 $U_0 = \Phi_{ub}(\mathcal{R}_{init})$ ;

**Enquanto**  $U_k - L_k > \epsilon$ , **faça**

- Escolher  $\mathcal{R} \in \mathcal{L}_k$  tal que o limite inferior em  $\mathcal{R}$  seja igual a  $L_k$ , isto é,  $\Phi_{lb}(\mathcal{R}) = L_k$ ;
- Dividir  $\mathcal{R}$  ao longo de um dos seus maiores lados em  $\mathcal{R}_I$  e  $\mathcal{R}_{II}$ ;
- Definir

$$\mathcal{L}_{k+1} := (\mathcal{L}_k - \{\mathcal{R}\}) \cup \{\mathcal{R}_I, \mathcal{R}_{II}\},$$

e  $L_{k+1}$  e  $U_{k+1}$  como os valores mínimos dos limites inferiores e superiores de todos retângulos  $\mathcal{R} \in \mathcal{L}_{k+1}$ , ou seja,

$$L_{k+1} := \min_{\mathcal{R} \in \mathcal{L}_{k+1}} \Phi_{lb}(\mathcal{R}),$$

$$U_{k+1} := \min_{\mathcal{R} \in \mathcal{L}_{k+1}} \Phi_{ub}(\mathcal{R});$$

- Fazer  $k := k + 1$ .

**fim enquanto.**

Uma análise mais detalhada deste algoritmo é realizada no Capítulo 4.

## 2.4 Resumo

Neste capítulo fez-se uma revisão de conceitos básicos de Análise Convexa, envolvendo definições e propriedades de conjuntos convexos e funções convexas. Também foram discutidas algumas propriedades gerais de problemas de otimização, a existência de soluções globais, e condições necessárias e suficientes de otimalidade de primeira e segunda ordens

para problemas de otimização restrita. O capítulo incluiu também uma introdução à otimização global e a alguns de seus algoritmos, como o algoritmo *branch-and-bound* na sua versão genérica. O conteúdo deste capítulo fundamenta os métodos de otimização a serem propostos nos Capítulos 4 e 5.

# 3

## Otimização no Espaço–Imagem

### Sumário

---

3.1	Otimização Multiobjetivo . . . . .	40
3.2	Conceitos Básicos e Definições . . . . .	41
3.3	Caracterização de Soluções Eficientes . . . . .	47
3.4	Abordagem por Projeção . . . . .	51
3.5	Algoritmo Básico . . . . .	60
3.6	Problema Max–Min . . . . .	62
3.7	Resumo . . . . .	65

---

Problemas de otimização que envolvem duas ou mais funções objetivos pertencem à classe de problemas de Otimização Multiobjetivo. A definição de eficiência introduzida por Pareto é utilizada: um ponto é Pareto-ótimo ou eficiente quando não é possível melhorar um objetivo sem piorar algum outro.

Existem várias abordagens disponíveis para tratar essa classe de problemas, as quais diferem na forma como o decisor articula as suas preferências em relação aos objetivos: antes, durante ou após do procedimento numérico de otimização. Neste capítulo introduzimos a idéia de otimização no espaço-imagem. Serão apresentadas a definição de solução eficiente (localmente eficiente e propriamente eficiente), caracterizações de soluções eficientes, o espaço-imagem do problema, a formulação do problema no espaço-imagem e um algoritmo básico a ser utilizado nos Capítulos 4 e 5.

### 3.1 Otimização Multiobjetivo

A otimização simultânea de várias funções objetivos é tratada em Otimização Multiobjetivo ou Otimização Vetorial. Recomenda-se [SD92], [Sta87] e [Sta88] para uma ampla discussão de temas e questões ligadas à Otimização Multiobjetivo.

A Otimização Multiobjetivo tem muitas aplicações em campos como Finanças, Administração, Teoria dos Jogos e Engenharia. No entanto, a resolução de problemas multiobjetivos não é uma tarefa fácil. A busca por todas as soluções ótimas de Pareto é um processo caro e demorado. Mesmo para problemas relativamente simples, determinar se um ponto pertence ao conjunto de Pareto é NP-difícil.

As soluções candidatas de um problema multiobjetivo são conhecidas como soluções Pareto-ótimas ou, menos freqüentemente, soluções Edgeworth-Pareto-ótimas em referência aos dois economistas, Edgeworth e Pareto, que desenvolveram a Teoria de Curvas de Indiferença no final do século 19. Edgeworth [Edg81] definiu otimalidade para problemas de tomada de decisão sob múltiplos objetivos no contexto de dois consumidores,  $P$  e  $\pi$ : “É necessário encontrar um ponto  $(x, y)$  tal que em qualquer direção em que dermos um passo infinitamente pequeno,  $P$  e  $\pi$  não aumentem juntas, mas que, quando a utilidade de um deles aumenta, a utilidade do outro diminua”.

Vilfredo Pareto [Paro6] foi um dos primeiros a analisar problemas de Economia por meio de ferramentas matemáticas. Em 1893, Pareto tornou-se presidente de Economia

Política na Universidade de Lausanne, Suíça, onde criou as suas duas teorias mais famosas, a Teoria da Circulação de Elites e a Teoria de Ótimo de Pareto. Stadler aplicou o conceito de Pareto-otimalidade aos campos da Engenharia e da Ciência em meados dos anos 1970 [Sta79], [Sta84]; as aplicações de Otimização Multiobjetivo em Engenharia cresceram ao longo das décadas seguintes [Ste86].

## 3.2 Conceitos Básicos e Definições

Um problema multiobjetivo genérico é formulado no *espaço das variáveis de decisão* como

$$P_x : \begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in \Omega, \end{array}$$

no qual  $\Omega \subseteq \mathbb{R}^n$  é um conjunto não-vazio e  $f : \Omega \rightarrow \mathbb{R}^m$ ,  $f := (f_1, f_2, \dots, f_m)$  é uma função vetorial, isto é, tal que cada ponto  $x \in \Omega$  é mapeado no espaço real  $m$ -dimensional.

Se todas as funções objetivos e as funções que descrevem as restrições forem lineares, então  $(P_x)$  é um problema de Otimização Multiobjetivo Linear. Se pelo menos uma das funções envolvidas não for linear,  $(P_x)$  é um problema de Otimização Multiobjetivo não-linear. Além disso, se as funções objetivos e o conjunto viável do problema forem convexos,  $(P_x)$  torna-se um problema de Otimização Multiobjetivo Convexo (e não-Convexo, caso contrário). A resolução de problemas multiobjetivos é dificultada pela frequente natureza conflitante dos objetivos: uma solução ótima segundo um objetivo pode não sê-lo de acordo com algum outro.

Uma diferença fundamental entre as otimizações escalar e multiobjetivo é que em Otimização Multiobjetivo cada solução viável  $x$  tem imagem em um ponto  $y$  no *espaço-imagem*. O mapeamento  $y = f(x)$  transforma um vetor  $n$ -dimensional em um vetor  $m$ -dimensional (ver Figura 3.1).

Como mencionamos anteriormente, existe um natural conflito entre as funções objetivos do problema. Uma consequência imediata deste fato é a perda do conceito usual de otimalidade, uma vez que a imagem de  $f$  é geralmente um conjunto parcialmente ordenado. Para introduzir o conceito de otimalidade aplicável ao caso, consideramos uma ordenação parcial num espaço linear real  $\mathcal{Z}$  como segue.

**Definição 3.2.1.** *Uma relação binária  $\leq$  em  $\mathcal{Z}$  é uma relação de ordem parcial em  $\mathcal{Z}$  se as seguintes*

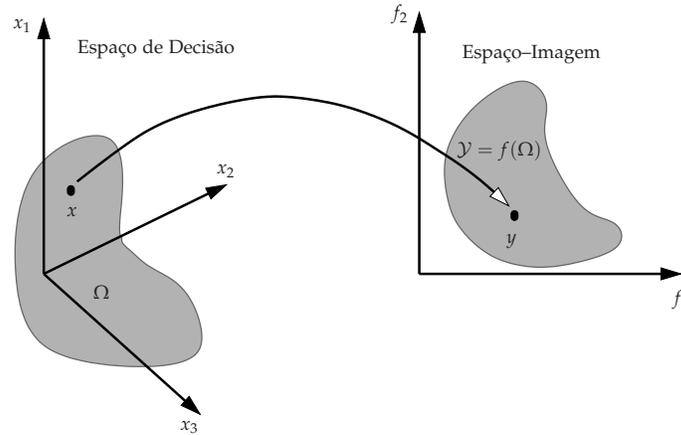


Figura 3.1: O mapeamento do espaço de decisões  $\Omega$  no espaço-imagem  $\mathcal{Y}$  por meio de uma função vetorial  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ .

propriedades são satisfeitas (para  $z_1, z_2, z_3, z_4 \in \mathcal{Z}$  e  $\lambda \in \mathbb{R}_+$ ):

- (i)  $z_1 \leq z_1$ ;
- (ii)  $z_1 \leq z_2, z_2 \leq z_3 \Rightarrow z_1 \leq z_3$ ;
- (iii)  $z_1 \leq z_2, z_4 \leq z_3 \Rightarrow z_1 + z_4 \leq z_2 + z_3$ ;
- (iv)  $z_1 \leq z_2 \Rightarrow \lambda z_1 \leq \lambda z_2$ ;

Além disso, a relação de ordem parcial  $\leq$  em  $\mathcal{Z}$  é anti-simétrica se

- (v)  $z_1 \leq z_2, z_2 \leq z_1 \Rightarrow z_1 = z_2$ .

**Definição 3.2.2.** Um espaço linear equipado com uma ordem parcial é chamado de espaço linear parcialmente ordenado.

A seguir apresentamos uma caracterização de uma ordem parcial num espaço linear real.

**Teorema 3.2.3.** (i) Se  $\leq$  é uma ordenação parcial em  $\mathcal{Z}$ , então o conjunto  $D := \{z \in \mathcal{Z} \mid 0 \leq z\}$  é um cone convexo; (ii) Se  $D$  é um cone convexo em  $\mathcal{Z}$ , então a relação binária  $\leq_D$ , definida por

$$z_1 \leq_D z_2 \Leftrightarrow z_2 - z_1 \in D,$$

é uma ordem parcial em  $\mathcal{Z}$ . É ainda possível afirmar que se  $\mathcal{Z}$  for um espaço linear real apontado, então  $\leq$  em  $\mathcal{Z}$  é anti-simétrica.

Convém destacar que

- (i) Um conjunto não-vazio  $\mathcal{K} \subseteq \mathbb{R}^n$  é um cone se  $\lambda x \in \mathcal{K}$  para todo  $x \in \mathcal{K}$  e  $\lambda \geq 0$ ;
- (ii) Um cone  $\mathcal{K}$  é apontado se  $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$ ;
- (iii) Um cone  $\mathcal{K}$  é sólido se  $\text{cor } \mathcal{K} \neq \emptyset$ , onde  $\text{cor } \mathcal{K}$  é o interior algébrico ou núcleo de  $\mathcal{K}$ .

Em particular, se  $\mathcal{K}$  é um cone convexo, então  $\text{cor } \mathcal{K} = \text{int } D$

**Exemplo 3.2.4.** O cone ordenado mais usual no  $\mathbb{R}^n$  é o ortante positivo  $\mathbb{R}_+^n$ . Este conjunto é um cone convexo apontado, fechado e sólido que define uma ordem parcial componente-a-componente em  $\mathbb{R}^n$ , também chamada de ordem de Pareto.

Um conjunto é parcialmente ordenado de acordo com  $\leq$  se satisfaz as propriedades (i)–(iv) da relação de ordem na Definição 3.2.1, mas não necessariamente todos os seus elementos são comparáveis no sentido indicado. O conjunto dos números reais  $\mathbb{R}$  é completamente ordenado; o  $\mathbb{R}^n$  apenas parcialmente. Uma das consequência deste fato é que problemas multiobjetivos são, em geral, caracterizados pela inexistência de soluções ótimas no sentido usual, isto é, em geral

$$\nexists x^* \in \Omega : f(x^*) \leq f(x), \quad \forall x \in \Omega \cap B(x^*, \epsilon),$$

onde  $B(x^*, \epsilon) := \{x \in \Omega : \|x - x^*\| < \epsilon\}$ , qualquer que seja  $\epsilon > 0$ . Em outras palavras, o conceito usual de otimalidade não é aplicável a problemas com múltiplos objetivos. É claro que implicitamente está se excluindo a possibilidade de ocorrência de soluções triviais, como a representada pela factibilidade da chamada *solução ideal* ou *utópica* do problema.

**Definição 3.2.5.** A *solução ideal*  $\underline{y}$  do problema é definida através de suas componentes como

$$\underline{y}_i = f_i(x^i), \quad i = 1, 2, \dots, m,$$

onde  $x^i := \underset{\mathbb{R}^n}{\text{argmin}} f_i(x)$ .

Note que se existisse  $x^* \in \Omega$  tal que  $\underline{y} = f(x^*)$ , então o problema estaria resolvido. Entretanto, a ocorrência de solução ideais em problemas multiobjetivos é bastante incomum.

Uma interpretação útil do problema pode ser realizada através do conceito de *dominância*. Para um ponto qualquer  $x^0 \in \Omega$ , defina os subconjuntos

$$\begin{aligned} \Omega_{<}(x^0) &:= \{x \in \Omega : f(x) \leq f(x^0) \text{ e } f(x) \neq f(x^0)\}, \\ \Omega_{\geq}(x^0) &:= \{x \in \Omega : f(x) \geq f(x^0)\}, \\ \Omega_{\sim}(x^0) &:= \{x \in \Omega : f(x) \not\leq f(x^0) \text{ e } f(x) \not\geq f(x^0)\}. \end{aligned}$$

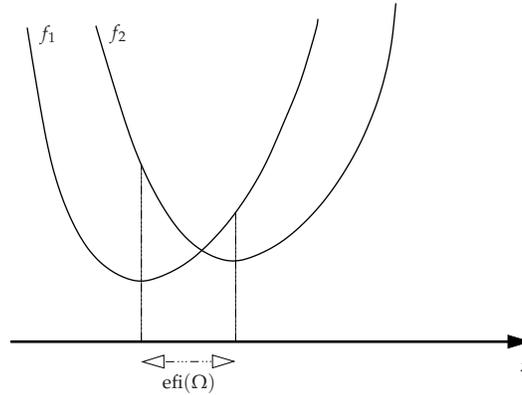


Figura 3.2: Interpretação dos conceito de eficiência.

Note que  $\Omega = \Omega_{<}(x^0) \cup \Omega_{\geq}(x^0) \cup \Omega_{\sim}(x^0)$ ,  $\forall x^0 \in \Omega$ . Os subconjuntos  $\Omega_{<}(x^0)$  e  $\Omega_{\geq}(x^0)$  contêm todos os pontos de  $\Omega$  que, respectivamente, dominam e são dominados pelo ponto  $x^0$  de acordo com a relação de ordem  $\leq$ . O subconjunto  $\Omega_{\sim}(x^0)$  representa todos os pontos não comparáveis a  $x^0$ . Lembrando que a formulação do problema busca a minimização simultânea de todos os objetivos e que  $x^0 \notin \Omega_{<}(x^0)$ , uma solução candidata  $x^* \in \Omega$  deverá ser tal que  $\Omega \cap \Omega_{<}(x^*) = \emptyset$ . Na verdade, em Otimização Multiobjetivo considera-se que haverá dois tipos de soluções, as soluções que devem ser descartadas (por serem dominadas) e as soluções que não podem ser descartadas automaticamente (soluções eficientes ou Pareto-ótimas). A determinação do conjunto de soluções eficientes, no todo ou em parte, é um dos problemas centrais da Otimização Multiobjetivo.

**Definição 3.2.6** (Dominância). *Diz-se que um ponto  $x \in \Omega$  domina outro ponto  $y \in \Omega$  se  $f_i(x) \leq f_i(y)$  para todo  $i = 1, 2, \dots, m$ , e  $f_j(x) < f_j(y)$  para algum  $j \in \{1, 2, \dots, m\}$ .*

**Definição 3.2.7** (Pareto Eficiente ou Pareto-Ótimo). *Um ponto  $x^* \in \Omega$  é Pareto-ótimo (ou eficiente ou não-dominado, ou não-inferior), se e somente se não existe nenhum outro ponto  $x \in \Omega$  tal que  $f_i(x) \leq f_i(x^*)$  para todo  $i = 1, 2, \dots, m$  e  $f_j(x) < f_j(x^*)$  para algum  $j \in \{1, 2, \dots, m\}$ .*

Uma solução eficiente (Pareto-ótima) não é dominada por qualquer outra solução viável do problema. Se  $x^* \in \Omega$  é uma solução eficiente, então, de acordo com a definição, qualquer alternativa  $x \in \Omega$  que proporcione um decréscimo em algum objetivo, relativo ao produzido por  $x^*$ , deve ao mesmo tempo levar ao acréscimo de pelo menos algum outro objetivo. Esta propriedade é ilustrada na Figura 3.2.

**Definição 3.2.8** (Conjunto Eficiente ou Pareto-Ótimo). *O conjunto de todas as soluções eficientes de um problema multiobjetivo é chamado de conjunto eficiente, e denotado por  $\text{efi}(\Omega)$ .*

Assumindo que a solução do problema multiobjetivo deva ser eficiente, o processo de otimização fica agora restrito ao subconjunto  $\text{efi}(\Omega)$ . Conforme será visto, este subconjunto pode ser caracterizado através de problemas de otimização escalares. Entretanto, os algoritmos utilizados para resolver problemas deste tipo convergem, em geral, para soluções ótimas locais. É natural então supor que qualquer técnica utilizada para gerar pontos de  $\text{efi}(\Omega)$  seja capaz de fornecer soluções eficientes locais.

**Definição 3.2.9** (Solução Localmente Eficiente). *Um ponto  $x^* \in \Omega$  é localmente eficiente, se e somente se existe  $\epsilon > 0$  tal que  $x^*$  é eficiente em  $\Omega \cap B(x^*, \epsilon)$ .*

Note que cada solução eficiente global é uma solução eficiente local. No entanto, o inverso não é verdadeiro a menos que algumas condições adicionais sejam verificadas.

**Teorema 3.2.10.** *Em um problema multiobjetivo convexo, toda solução eficiente local também é uma solução eficiente global.*

*Demonstração.* Suponha que  $x^* \in \Omega$  é uma solução eficiente local, mas não é uma solução eficiente global. Então existe  $y \in \Omega$  tal que  $f(y) \leq f(x^*)$  e  $f(y) \neq f(x^*)$ . Considere um escalar  $z := \lambda x^* + (1 - \lambda)y$  tal que  $z \in \Omega \cap B(x^*, \epsilon)$  e  $0 < \lambda < 1$ . Pela convexidade das funções objetivos, temos

$$f(z) \leq \lambda f(x^*) + (1 - \lambda)f(y) \leq f(x^*),$$

e

$$f(x^*) + (1 - \lambda)f(y) \neq f(x^*),$$

que demonstra que  $x^*$  não é uma solução eficiente local. □

O teorema a seguir estende o resultado acima.

**Teorema 3.2.11.** *Considere um problema multiobjetivo com um conjunto viável convexo e funções objetivos quase-convexas. Se pelo menos uma das funções objetivos é estritamente quase-convexa, então cada solução eficiente local do problema é também uma solução eficiente global.*

*Demonstração.* Para a prova, ver [RCRL95]. □

Antes de passar à caracterização de soluções eficientes, vamos introduzir definições adicionais.

**Definição 3.2.12.** Um ponto  $x^* \in \Omega$  é fracamente eficiente se não existe nenhum outro ponto  $x \in \Omega$  tal que  $f_i(x) < f_i(x^*)$  para todo  $i = 1, 2, \dots, m$ .

Um ponto é fracamente eficiente se não existe outro ponto que melhore *todas* as funções objetivos simultaneamente. Em contraste, um ponto é eficiente se não existe outro ponto que melhore *pelo menos* uma função objetivo sem prejuízo de qualquer outra função. Qualquer solução eficiente é também fracamente eficiente.

Soluções eficientes também podem ser classificadas como *próprias* ou *impróprias* [Geo68], [Yu85] e [Mie99].

**Definição 3.2.13.** Diz-se que  $x^* \in \Omega$  é uma solução propriamente eficiente se  $x^* \in \text{efi}(\Omega)$  e existe um escalar  $M > 0$  tal que, para cada  $i = 1, 2, \dots, m$  e cada  $x \in \Omega$  satisfazendo  $f_i(x) < f_i(x^*)$  existe no mínimo um  $j \neq i$  com  $f_j(x) > f_j(x^*)$  e

$$\frac{f_i(x) - f_i(x^*)}{f_j(x^*) - f_j(x)} \leq M.$$

Argumenta-se que, do ponto de vista do tomador de decisão, efeitos de saturação com os que poderiam ocorrer como soluções eficientes impróprias são indesejáveis ao se lidar com problemas práticos.

Encontrar todas as soluções eficientes muitas vezes é impraticável, uma vez que em geral existe um número infinito destas soluções. Além disso, e mesmo para problemas com dois objetivos, determinar se um ponto pertence ao conjunto eficiente é um problema NP-difícil [PY00]. Uma maneira de lidar com esse problema é introduzir o conceito de solução eficiente aproximada.

**Definição 3.2.14.** Dado um escalar  $\epsilon > 0$ , um conjunto eficiente  $\epsilon$ -aproximado, denotado por  $\mathcal{P}_\epsilon$ , é um subconjunto de  $\Omega$  tal que não existe outra solução  $y \in \Omega$  tal que  $(1 + \epsilon)f_i(y) \leq f_i(x)$  para todo  $x \in \mathcal{P}_\epsilon$  e para algum  $i$ .

Pela definição, o conjunto  $\mathcal{P}_\epsilon$  é um subconjunto de soluções viáveis que dominam, na prática, por um fator  $\epsilon > 0$ , qualquer outra solução viável do problema.

### 3.3 Caracterização de Soluções Eficientes

Um aspecto básico da abordagem de problemas multiobjetivos é a caracterização de soluções eficientes através de problemas escalares bem definidos. Assim sendo, o conceito de solução eficiente fica intrinsecamente associado às condições de otimalidade destes problemas. As relações entre soluções eficientes e problemas escalares passam a ser evidenciadas a partir do teorema seguinte.

**Teorema 3.3.1.**  $x^* \in \text{efi}(\Omega)$  se e somente se  $x^*$  resolve os  $m$  problemas escalares

$$\begin{aligned} P_i : \quad & \text{minimizar} \quad f_i(x) \\ & \text{sujeito a} \quad f_j(x) \leq f_j(x^*), \quad \forall j \neq i, \\ & \quad \quad \quad x \in \Omega. \end{aligned}$$

*Demonstração.* Suponha que  $x^*$  é eficiente e que portanto não existe  $x \in \Omega$  tal que  $f(x) \leq f(x^*)$  e  $f(x) \neq f(x^*)$ . Neste caso  $x^*$  resolve  $P_i$ ,  $i = 1, 2, \dots, m$ . Por outro lado, suponha que  $x^*$  resolve  $P_i$ ,  $i = 1, 2, \dots, m$ , mas  $x^* \notin \text{efi}(\Omega)$ . Então existe  $x \in \Omega$  tal que  $f(x) \leq f(x^*)$  e que para algum  $i$ ,  $f_i(x) < f_i(x^*)$ . Observe que nestas circunstâncias  $x^*$  não resolveria o problema  $P_i$ , contradizendo a hipótese inicial. □

Uma caracterização alternativa pode ser dada em termos da unicidade de solução de algum problema  $P_i$ .

**Teorema 3.3.2.** Se para algum  $i$ ,  $x^* \in \Omega$  é a solução ótima única de  $P_i$ , então  $x^* \in \text{efi}(\Omega)$ .

*Demonstração.* Suponha, por absurdo, que  $x^* \in \Omega$  é solução única de  $P_i$ , mas não é eficiente. Então existe  $x \in \Omega$  tal que  $f(x) \leq f(x^*)$  e  $f(x) \neq f(x^*)$  e, em particular,  $f_i(x) \leq f_i(x^*)$ , o que contraria a hipótese de unicidade de  $x^*$ . Portanto  $x^* \in \text{efi}(\Omega)$ . □

A caracterização de soluções eficientes por meio de problemas do tipo  $P_i$ ,  $i = 1, 2, \dots, m$  é interessante do ponto de vista teórico por induzir o desenvolvimento de condições analíticas para eficiência. Entretanto, do ponto de vista prático, esta caracterização tem pouca utilidade no sentido de gerar soluções eficientes. A geração de todo o conjunto  $\text{efi}(\Omega)$ , independentemente da natureza do problema considerado, é teoricamente possível tendo em vista o resultado a seguir.

**Teorema 3.3.3.** Se  $x^* \in \Omega$  é eficiente então existem um inteiro  $k \in I := \{1, 2, \dots, m\}$  e números reais  $\epsilon_j, j = 1, 2, \dots, m$  ( $j \neq k$ ) tais que  $x^*$  resolve

$$\begin{aligned} P_k(\epsilon) : \quad & \text{minimizar} \quad f_k(x) \\ & \text{sujeito a} \quad f_j(x) \leq \epsilon_j, \quad \forall j \neq k, \\ & \quad \quad \quad x \in \Omega, \end{aligned}$$

onde  $\epsilon$  está definido em  $\mathcal{E}_k := \{\epsilon = (\epsilon_1, \dots, \epsilon_{k-1}, \epsilon_{k+1}, \dots, \epsilon_m) : \Omega_k(\epsilon) \neq \emptyset\}$  e

$$\Omega_k(\epsilon) := \{x \in \Omega : f_j(x) \leq \epsilon_j, \quad \forall j \neq k\}.$$

*Demonstração.* Defina  $f_i^* = f_i(x^*), i = 1, 2, \dots, m$  e suponha que  $x^*$  não resolve o problema acima para nenhum  $k \in I$  e números reais  $\epsilon_j, j = 1, 2, \dots, m$  ( $j \neq k$ ). Neste caso, para  $k = 1$  e  $\epsilon_j = f_j^*, j = 2, 3, \dots, m$  deve existir  $x^0 \in \Omega$  tal que  $f_k(x^0) < f_k(x^*)$  e  $f_j(x^0) \leq f_j^*$ , o que contraliz a hipótese de que  $x^*$  é eficiente.  $\square$

Variando-se convenientemente  $\epsilon$  em  $\mathcal{E}_k$  é possível agora gerar todo o conjunto  $\text{efi}(\Omega)$ , porém nem toda solução do problema escalar  $P_k(\epsilon)$  é eficiente, o que é previsível dadas as similaridades entre os problemas  $P_k(\epsilon)$  e  $P_i$ . As condições suficientes para que, sendo solução do problema  $P_k(\epsilon)$ , uma decisão viável  $x^*$  seja eficiente estão resumidas no teorema seguinte.

**Teorema 3.3.4.** Dado  $\epsilon \in \mathcal{E}_k$ , a solução  $x^*$  de  $P_k(\epsilon)$  é eficiente se

- (i)  $x^*$  é a solução ótima única de  $P_k(\epsilon)$  para algum  $k$ ;
- (ii)  $x^*$  resolve  $P_k(\epsilon)$  para todo  $k \in I$ .

*Demonstração.* Seja  $x^*$  a solução ótima única de  $P_k(\epsilon)$  para algum  $k$ . Neste caso,  $x^*$  é também solução ótima única do problema  $P_k$

$$\begin{aligned} & \text{minimizar} \quad f_k(x) \\ & \text{sujeito a} \quad f_j(x) \leq f_j(x^*), \quad \forall j \neq k, \\ & \quad \quad \quad x \in \Omega, \end{aligned}$$

e em vista de resultados anteriores,  $x^*$  é eficiente. Por outro lado, suponha que  $x^*$  é eficiente, mas não resolve  $P_k(\epsilon)$  para  $k \in I$ . Assim sendo, existe um  $k^*$  tal que  $x^*$  não

resolve

$$\begin{aligned} P_{k^*}(\epsilon) : \quad & \text{minimizar} && f_{k^*}(x) \\ & \text{sujeito a} && f_j(x) \leq f_j(x^*), \quad \forall j \neq k^*, \\ & && x \in \Omega, \end{aligned}$$

contradizendo  $x^* \in \text{efi}(\Omega)$ . □

A caracterização de soluções eficientes por meio dos problemas  $P_k(\epsilon)$  possui algumas limitações de ordem prática. Verificar a otimalidade de  $x^*$  com respeito a  $m$  problemas de otimização pode exigir considerável esforço, principalmente devido a agregação de  $m - 1$  objetivos às restrições originais. Estas novas restrições podem modificar substancialmente a natureza do problema de otimização.

Neste aspecto, uma caracterização alternativa baseada na combinação linear dos objetivos mostra-se mais atrativa.

**Teorema 3.3.5.** *Seja  $x^* \in \Omega$  uma solução ótima de*

$$\begin{aligned} P_W : \quad & \text{minimizar} && \langle w, f(x) \rangle \\ & \text{sujeito a} && x \in \Omega, \end{aligned}$$

para  $w \in \mathbb{R}^m$ ,  $w \geq 0$  dado. Então  $x^*$  é uma solução eficiente do problema multiobjetivo se

1.  $x^*$  é a solução ótima única de  $P_W$ , ou
2.  $w_i > 0$ ,  $i = 1, 2, \dots, m$ .

*Demonstração.* Considere o primeiro caso. Se  $x^*$  é a solução ótima única de  $P_W$ , por definição

$$\langle w, f(x^*) - f(x) \rangle < 0, \quad \forall x \in \Omega.$$

Suponha que  $x^*$  não é eficiente, isto é, que exista  $x \in \Omega$  tal que  $f(x) \leq f(x^*)$  e  $f(x) \neq f(x^*)$ . Como  $w \geq 0$ , isto implicaria contradizer a unicidade da solução  $x^*$ , e portanto  $x^*$  é eficiente. O segundo caso é demonstrado utilizando-se argumentos similares. □

Como a solução de  $P_W$  não se altera quando a função objetivo é dividida por uma constante positiva pode-se considerar, sem perda de generalidade, que

$$w \in \mathcal{W} := \left\{ w \in \mathbb{R}^m : w \geq 0, \sum_{i=1}^m w_i = 1 \right\}.$$

Observe que o problema  $P_W$  não altera a classe de modelos tratados, isto é, a natureza das restrições da função objetivo de  $P_W$  é a mesma do problema multiobjetivo original. O mesmo não ocorre com o problema  $P_k(\epsilon)$ , uma vez que as  $\epsilon$ -restrições podem alterar a natureza do problema, caso as restrições originais sejam lineares e as funções objetivos não-lineares. Dentro deste contexto, a escolha de  $P_W$  é a mais natural no caso de problemas convexos, ficando o problema  $P_k(\epsilon)$  reservado para os casos em que a hipótese de convexidade não está presente.

**Exemplo 3.3.6.** *Considere o seguinte problema multiobjetivo*

$$P_x : \begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in \Omega, \end{array}$$

no qual  $\Omega := \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, -2 \leq x_2 \leq 2\}$  é um conjunto não-vazio em  $\mathbb{R}^2$  e  $f_1(x) = x_1$ ,  $f_2(x) = 1 + x_2^2 - x_1 - a \sin(b\pi x_1)$  são funções escalares

Pela abordagem baseada na soma ponderada, resolve-se

$$\min_{x \in \Omega} F(x) := w_1 f_1(x) + w_2 f_2(x), \quad w \in \mathcal{W},$$

O que é necessário para obter todas as soluções de um problema de programação multiobjetiva pela abordagem baseada na soma ponderada é a convexidade do espaço-imagem  $\mathcal{Y}$ . Como garantir a convexidade de  $\mathcal{Y}$ ? O espaço-imagem  $\mathcal{Y}$  pode ser não-convexo mesmo que o problema multiobjetivo seja convexo. Para que o espaço-imagem seja convexo, os parâmetros  $a$  e  $b$  são fixados em  $a = 0,2$  e  $b = 1$ .

Se o espaço-imagem do problema multiobjetivo for convexo, então todos os pontos em  $\text{efi}(\Omega)$  podem ser encontrados pela abordagem baseada na soma ponderada. Como a função objetivo  $F$  é uma combinação linear das funções objetivos  $f_1$  e  $f_2$ , a cada reta no espaço-imagem estará associada um valor de  $F$ . Os coeficientes angulares de retas estão relacionados aos pesos  $w_1, w_2$ ; a localização da reta depende do valor de  $F$ . O valor mínimo de  $F$  é encontrado em algum ponto. Pesos diferentes gerarão soluções ótimas diferentes, sempre na fronteira eficiente de Pareto. Todas as soluções eficientes (no sentido de minimização) podem se calculadas por meio da abordagem baseada na soma ponderada (ver Figura 3.3).

Em problemas multiobjetivo não-convexos, o método da soma ponderada pode não conseguir encontrar certas soluções eficientes. Nestes casos, apenas as soluções que se en-

contram na casca convexa do espaço–imagem podem ser determinadas. Essas soluções são chamadas *soluções suportadas*. Nem todas as soluções eficientes de um problema multiobjetivo não–convexo são suportadas.

**Exemplo 3.3.7.** *Considere novamente o problema multiobjetivo do Exemplo 3.3.6, onde os parâmetros  $a$  e  $b$  são fixados em valores  $a = 0,1$  e  $b = 3$ .*

Neste caso, o espaço–imagem é um conjunto não–convexo, e poderão existir soluções eficientes (no sentido de minimização) que não podem ser determinadas pelo método da soma ponderada (ver Figura 3.4).

### 3.4 Abordagem por Projeção

Os conceitos e resultados sobre otimização multiobjetivo apresentados até o momento foram discutidos a partir da formulação do problema no espaço das variáveis de decisão. Este tipo de formulação é natural e possui implicações teóricas e práticas fundamentais. Entretanto, interpretações geométricas importantes e resultados adicionais podem ser obtidos com a formulação do problema no *espaço–imagem*, ou *espaço dos objetivos* [Fer98]. Neste sentido, seja  $\mathcal{Y} := f(\Omega)$  a representação do mapeamento do subconjunto  $\Omega$  no espaço–imagem:

$$\mathcal{Y} := \{y \in \mathbb{R}^m : y = f(x), x \in \Omega\}.$$

O problema multiobjetivo pode ser formulado de forma equivalente como

$$P_{\mathcal{Y}} : \begin{array}{ll} \text{minimizar} & y \\ \text{sujeito a} & y \in \mathcal{Y}, \end{array}$$

e o conjunto de soluções eficientes associado é  $\text{efi}(\mathcal{Y}) = f(\text{efi}(\Omega))$ .

Ignorando momentaneamente que este último problema é apenas uma abstração matemática, a formulação no espaço–imagem poderia oferecer vantagens importantes sobre a formulação no espaço das variáveis de decisão no que diz respeito à dimensão e à natureza do espaço no qual o problema seria tratado. Observe que  $\Omega \subseteq \mathbb{R}^n$  e  $\mathcal{Y} \subseteq \mathbb{R}^m$  e, em geral, o número de objetivos  $m$  é pequeno. Neste caso, a abordagem do problema baseada no espaço–imagem traria benefícios computacionais consideráveis. Além disso, o conceito de solução eficiente fica muito bem caracterizado em termos geométricos por

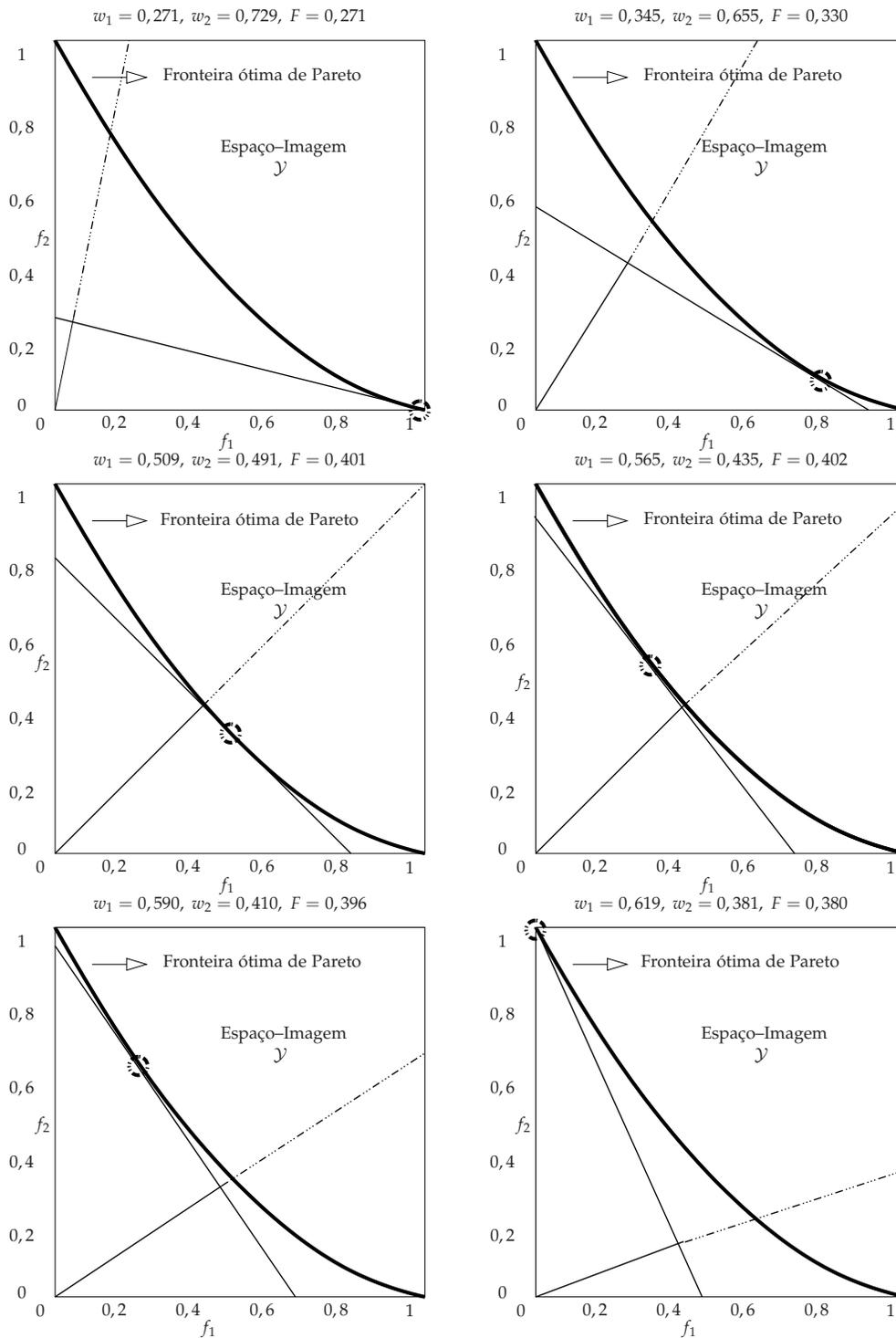


Figura 3.3: Todas as soluções eficientes são encontradas, uma vez que cada solução eficiente admite uma reta suporte no espaço-imagem – Exemplo 3.3.6. O método de soma ponderada permite encontrar as soluções suportadas (não-dominadas) do problema, isto é, pontos na casca convexa do conjunto-imagem.

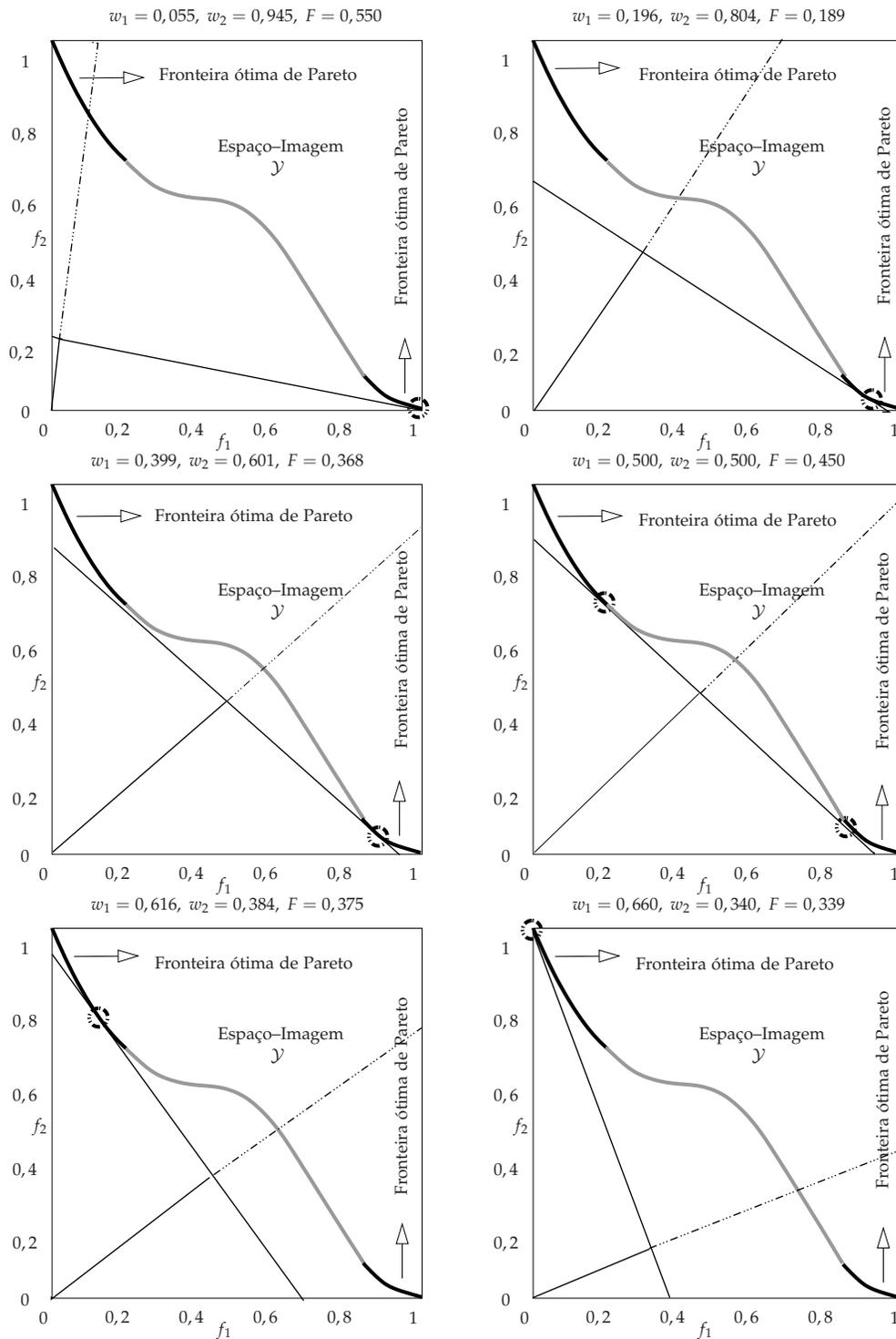


Figura 3.4: Existem soluções eficientes que não são encontradas pelo método da soma ponderada, uma vez que nem todas as soluções eficientes do problema admitem uma reta suporte – Exemplo 3.3.7. Percebe-se que pelo método de soma ponderada não é possível identificar todas as soluções eficientes quando o espaço-imagem não é convexo.

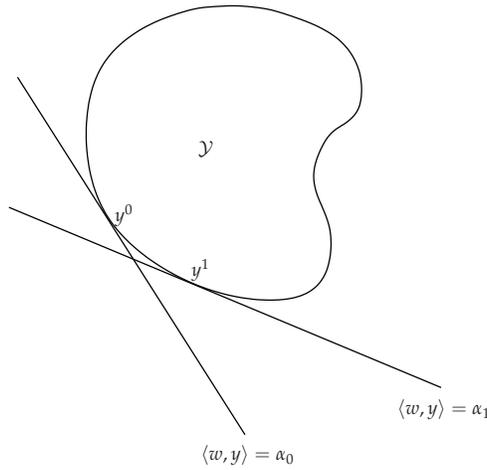


Figura 3.5: Interpretação do problema  $P_W$ .

meio da formulação no espaço-imagem.

**Proposição 3.4.1.** *Se  $y^* \in \text{efi}(\mathcal{Y})$ , então  $y^* \in \partial\mathcal{Y}$ .*

*Demonstração.* Veja por exemplo [AP96] ou [CSMSoo]. □

De fato, supor que  $y^* \in \text{int}(\mathcal{Y})$  permitiria definir uma vizinhança  $B(y^*, \epsilon)$  no entorno de  $y^*$ , e qualquer ponto  $y \in B(y^*, \epsilon)$  da reta que une  $y^*$  à origem do  $\mathbb{R}^m$  seria tal que  $y < y^*$ , contradizendo a hipótese de  $y^*$  ser eficiente. Conseqüentemente,  $y^* \in \partial\mathcal{Y}$ .

O problema  $P_W$  possui uma interpretação geométrica importante no espaço-imagem, conforme ilustra a Figura 3.5.

O problema pode ser reescrito como

$$P_y : \begin{array}{ll} \text{minimizar} & \langle w, y \rangle \\ \text{sujeito a} & y \in \mathcal{Y}. \end{array}$$

Seja  $y^0 \in \mathcal{Y}$  uma solução ótima do problema  $P_W$  e  $\alpha^0$  o valor mínimo associado. Observe então que  $\mathcal{H} := \{y : \langle w, y \rangle = \alpha^0\}$  define um hiperplano suporte a  $\mathcal{Y}$  no ponto  $y^0$ , uma vez que  $\langle w, y^0 \rangle = \alpha^0$  e  $\langle w, y \rangle \geq \alpha^0, \forall y \in \mathcal{Y}$ . Em outras palavras,  $\mathcal{Y}$  está contido no semi-espaço fechado  $\mathcal{H}_{\geq}$  determinado por  $\mathcal{H}$ . Se  $y^0$  é a única solução de  $P_W$  e/ou  $w \in \mathbb{R}_{++}^n$ , então  $y^0 \in \text{efi}(\mathcal{Y})$ .

Se qualquer solução eficiente admitisse um hiperplano suporte, então variando-se pa-

rametricamente  $w \in \mathcal{W}$ , todo o conjunto  $\text{efi}(\mathcal{Y})$  poderia ser gerado através de  $P_{\mathcal{W}}$ . Isto certamente ocorreria se  $\mathcal{Y}$  fosse um conjunto convexo. Infelizmente, nada se pode dizer quanto ao mapeamento  $\mathcal{Y} = f(\Omega)$  preservar a convexidade do problema original. Ainda assim, para problemas multiobjetivos convexos, demonstra-se que esta importante propriedade é válida.

**Teorema 3.4.2.** *Sejam  $f_i : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, m$ , funções convexas sobre um subconjunto convexo  $\Omega$ . Então  $\mathcal{Y} = f(\Omega)$  admite um hiperplano suporte em qualquer ponto de  $\text{efi}(\mathcal{Y})$ .*

*Demonstração.* A convexidade do conjunto convexo  $\text{co}(\mathcal{Y})$  implica na existência de um hiperplano suporte em qualquer ponto  $y^0 \in \text{efi}(\text{co}(\mathcal{Y}))$ . Neste caso, se  $y^0 \in \text{efi}(\text{co}(\mathcal{Y}))$  implicar em  $y^0 \in \text{efi}(\mathcal{Y})$ , o Teorema estará demonstrado. Suponha que  $y^0 \in \text{efi}(\text{co}(\mathcal{Y}))$ . Então  $y^0 \in \text{co}(\mathcal{Y})$ , e pelo Teorema de Carathéodory (Capítulo 2), pode-se representar  $y^0$  como

$$y^0 = \sum_{i=1}^{m+1} \lambda_i y^i, \quad \lambda_i \geq 0, \quad \sum_{i=1}^{m+1} \lambda_i = 1,$$

onde  $y^i = f(x^i) \in \mathcal{Y}$  para elementos apropriados  $x^i \in \Omega$ ,  $i = 1, 2, \dots, m+1$ . Neste caso, usando a convexidade das funções sobre o subconjunto convexo  $\Omega$ , obtém-se

$$y = f\left(\sum_{i=1}^{m+1} \lambda_i x^i\right) \leq \sum_{i=1}^{m+1} \lambda_i y^i = y^0.$$

Observe que  $y \in \mathcal{Y}$ , e que portanto  $y \in \text{co}(\mathcal{Y})$ . Como  $y^0 \in \text{efi}(\text{co}(\mathcal{Y}))$ , inclui-se que  $y \leq y^0$  ocorre sempre na igualdade, e que portanto  $y^0 \in \mathcal{Y}$ . Além disso, se  $y^0 \in \text{efi}(\text{co}(\mathcal{Y}))$ , então não existe qualquer outro  $y \in \text{co}(\mathcal{Y})$  tal que  $y \leq y^0$  e  $y \neq y^0$ . Como  $y^0 \in \mathcal{Y}$  e  $\mathcal{Y} \subseteq \text{co}(\mathcal{Y})$ , conclui-se também que  $y^0 \in \text{efi}(\mathcal{Y})$ .  $\square$

O problema multiobjetivo admite uma formulação convexa equivalente no espaço-imagem. Nesta nova formulação, a região de viabilidade para os vetores  $y$  é redefinida como

$$\mathcal{Y} + \mathcal{D} := \left\{ \xi \in \mathbb{R}^m : \xi = y + d, \quad y \in \mathcal{Y}, \quad d \in \mathcal{D} \right\},$$

no qual  $\mathcal{D}$  é o cone convexo

$$\mathcal{D} := \left\{ d \in \mathbb{R}^m : d \geq 0 \right\}.$$

O conjunto  $\mathcal{Y} + \mathcal{D}$  preserva a estrutura eficiente do conjunto  $\mathcal{Y}$ , como demonstrado a

seguir.

**Teorema 3.4.3.**  $\text{efi}(\mathcal{Y}) = \text{efi}(\mathcal{Y} + \mathcal{D})$ .

*Demonstração.* Suponha que  $y^* \in \text{efi}(\mathcal{Y})$ , mas  $y^* \notin \text{efi}(\mathcal{Y} + \mathcal{D})$ . Então existem elementos  $y \in \mathcal{Y}$  e  $d \in \mathcal{D}$  tais que  $y + d \leq y^*$  e  $y + d \neq y^*$ , o que implica  $y \leq y^*$  e  $y \neq y^*$ , isto é, que  $y^* \notin \text{efi}(\mathcal{Y})$ , contradizendo a suposição. Portanto,  $\text{efi}(\mathcal{Y}) \subseteq \text{efi}(\mathcal{Y} + \mathcal{D})$ . Por outro lado, assuma que  $y^* \in \text{efi}(\mathcal{Y} + \mathcal{D})$ , mas  $y^* \notin \text{efi}(\mathcal{Y})$ . Neste caso, existe  $y \in \mathcal{Y}$  tal que  $y \leq y^*$  e  $y \neq y^*$ , e, conseqüentemente, existe  $d \in \mathcal{D}$ ,  $d \neq 0$ , tal que  $y + d \leq y^*$ , isto é,  $y^* \notin \text{efi}(\mathcal{Y} + \mathcal{D})$ , o que contradiz mais uma vez a suposição. Logo  $\text{efi}(\mathcal{Y} + \mathcal{D}) \subseteq \text{efi}(\mathcal{Y})$ , e assim  $\text{efi}(\mathcal{Y}) = \text{efi}(\mathcal{Y} + \mathcal{D})$ .  $\square$

Além de preservar a estrutura eficiente de  $\mathcal{Y}$ , o conjunto  $\mathcal{Y} + \mathcal{D}$  possui a importante característica de ser convexo, o que não necessariamente ocorre com  $\mathcal{Y}$ .

**Teorema 3.4.4.** *Sejam  $f_i : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, m$ , funções convexas sobre um subconjunto convexo  $\Omega$ . Então o conjunto  $\mathcal{Y} + \mathcal{D}$  é convexo.*

*Demonstração.* Sejam  $y^1, y^2 \in \mathcal{Y} + \mathcal{D}$ . Especificamente,  $y^k = f(x^k) + d^k, k = 1, 2$ , com  $x^k \in \Omega$  e  $d^k \in \mathcal{D}$ . Então para qualquer  $\lambda \in [0, 1]$ ,

$$\begin{aligned} \lambda y^1 + (1 - \lambda)y^2 &= \lambda f(x^1) + (1 - \lambda)f(x^2) + \lambda d^1 + (1 - \lambda)d^2, \\ &\geq f(\lambda x^1 + (1 - \lambda)x^2) + \lambda d^1 + (1 - \lambda)d^2, \\ &= f(\lambda x^1 + (1 - \lambda)x^2) + \lambda d^1 + (1 - \lambda)d^2 + d_\lambda, \\ &= f(\lambda x^1 + (1 - \lambda)x^2) + \lambda(d_\lambda + d^1) + (1 - \lambda)(d_\lambda + d^2). \end{aligned}$$

A desigualdade surge da convexidade das funções e as igualdades subsequentes são obtidas para algum  $d_\lambda \geq 0$  apropriado. Como  $d_\lambda + d^1 \in \mathcal{D}$ ,  $d_\lambda + d^2 \in \mathcal{D}$  e  $\mathcal{D}$  é um conjunto convexo, segue que a combinação convexa de quaisquer dois pontos de  $\mathcal{Y} + \mathcal{D}$  pode ser sempre escrita na forma  $y + d$ , com  $y \in \mathcal{Y}$  e  $d \in \mathcal{D}$ . Portanto  $\mathcal{Y} + \mathcal{D}$  é um conjunto convexo.  $\square$

Como qualquer conjunto convexo,  $\mathcal{Y} + \mathcal{D}$  pode ser representado pela interseção de todos os semi-espacos fechados que o contém. O passo seguinte é substituir a região de viabilidade de  $P_y$  pelo conjunto  $\mathcal{Y} + \mathcal{D}$ . O problema resultante é convexo e será equivalente

a  $P_y$  se for possível garantir que pelo menos uma solução do problema original é eficiente (Teorema 3.4.3).

A minimização indicada em  $P_y$  sugere que melhores soluções são obtidas quando  $y$  apresenta menores valores. Uma maneira de expressar esta propriedade é definir uma função de preferência  $F : \mathcal{Y} \rightarrow \mathbb{R}$ , contínua e não-decrescente em relação a  $y \in \mathcal{Y}$ , isto é, dados dois pontos  $y^1, y^2 \in \mathcal{Y}$ , então  $F(y^1) \leq F(y^2)$  sempre que  $y^1 \leq y^2$ . O sentido da introdução da *função de preferência* é assegurar que se  $y^* \in \mathcal{Y}$  resolve  $P_y$  de acordo com algum critério adicional, então existe uma função  $F$  que quando minimizada sobre  $\mathcal{Y}$  fornece  $y^*$  como uma de suas possíveis soluções. Desta forma,  $P_y$  poderia ser substituído pelo problema

$$P_{\mathcal{F}} : \begin{array}{ll} \text{minimizar} & F(y) \\ \text{sujeito a} & y \in \mathcal{Y}, \end{array}$$

e o resultado de interesse demonstrado através do teorema seguinte.

**Teorema 3.4.5.** *Pelo menos uma solução de  $P_{\mathcal{F}}$  é eficiente.*

*Demonstração.* Suponha que  $y^*$  é uma solução ótima de  $P_{\mathcal{F}}$  e que  $y^0$  é uma solução ótima do problema auxiliar

$$\begin{array}{ll} \text{minimizar} & \sum_{i=1}^m y_i \\ \text{sujeito a} & y_j \leq y_j^*, \quad j = 1, 2, \dots, m, \\ & y \in \mathcal{Y}. \end{array}$$

Neste caso,  $y^0 \in \text{efi}(\mathcal{Y})$ . Caso contrário, a otimalidade com relação à soma das componentes de  $y$  seria falsa. Como  $y^0 \leq y^*$ , a solução factível  $y^0$  também resolve  $P_y$  pois

$$F(y^0) \leq F(y^*),$$

uma vez que  $F$  é uma função não-decrescente. □

Dado que pelo menos uma solução de  $P_{\mathcal{F}}$  é eficiente, pode-se *aumentar* a região viável

de  $P_{\mathcal{F}}$  de  $\mathcal{Y}$  para  $\mathcal{Y} + \mathcal{D}$ , obtendo-se o problema equivalente

$$P_{\mathcal{F}} : \begin{array}{ll} \text{minimizar} & F(y) \\ \text{sujeito a} & y \in \mathcal{Y} + \mathcal{D}, \end{array}$$

no qual a designação  $P_{\mathcal{F}}$  foi mantida por simplicidade.

A resolução de  $P_{\mathcal{F}}$  pode ser conduzida através de um método clássico de planos de corte. Se  $y^0 \notin \mathcal{Y} + \mathcal{D}$ , então existe um hiperplano  $\mathcal{H}$  do  $\mathbb{R}^m$  separando  $y^0$  de  $\mathcal{Y} + \mathcal{D}$ , e um dos semi-espacos gerados por  $\mathcal{H}$  contém  $\mathcal{Y} + \mathcal{D}$ . Suponha que um certo número destes semi-espacos, por simplicidade representados como  $\mathcal{H}^1, \mathcal{H}^2, \dots, \mathcal{H}^k$ , tenha sido determinado. Uma aproximação de ordem  $k$  para o problema seria

$$P_{\mathcal{F}^k} : \begin{array}{ll} \text{minimizar} & F(y) \\ \text{sujeito a} & y \in (\mathcal{Y} + \mathcal{D})^k := \bigcap_{i=1}^k \mathcal{H}^i. \end{array}$$

Seja  $y^k$  uma solução ótima de  $P_{\mathcal{F}^k}$ . Se  $y^k \in \mathcal{Y} + \mathcal{D}$ , então pode-se encerrar o processo. Caso contrário, se  $y^k \notin \mathcal{Y} + \mathcal{D}$ , obtém-se um semi-espaco adicional  $\mathcal{H}^{k+1}$ , no sentido de melhorar a aproximação anterior e a estimativa do valor ótimo da função  $F$ . Observe que  $\text{efi}(\mathcal{Y} + \mathcal{D}) \subseteq (\mathcal{Y} + \mathcal{D})^k$ , para todo  $k = 1, 2, \dots$ . Uma tarefa essencial no sentido de resolver  $P_{\mathcal{F}}$  através de aproximações por semi-espacos é verificar se  $y^k \in \mathcal{Y} + \mathcal{D}$ . Neste ponto, referências ao espaco das decisões são inevitáveis.

**Proposição 3.4.6.**  $\mathcal{Y} + \mathcal{D} = \mathcal{F}$ , onde

$$\mathcal{F} := \left\{ y \in \mathbb{R}^m : f(x) \leq y \text{ para algum } x \in \Omega \right\}.$$

*Demonstração.* Note que em  $\mathcal{F}$ , os vetores  $y$  podem se representados como  $y = f(x) + d$ , onde  $f(x) \in \mathcal{Y}$  e  $d \in \mathcal{D}$ . □

O teste de pertinência de um dado  $y$  em relação a  $\mathcal{F}$  é então viabilizado por meio de um resultado geral de otimização convexa.

**Lema 3.4.7.**  $y \in \mathcal{F}$  se e somente se  $y$  satisfaz o sistema de infinitas desigualdades lineares

$$\inf_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0, \quad \forall w \in \mathcal{W}.$$

Supõe-se que no lado esquerdo da desigualdade, o ínfimo é obtido em  $\Omega$  para cada

$w \in \mathcal{W}$ . As condições suficientes para que isto ocorra dependem das funções  $f_1, f_2, \dots, f_m$  e do conjunto  $\Omega \subseteq \mathbb{R}^n$ . Um modo prático para se testar a pertinência de um dado  $y \in \mathbb{R}^m$  em relação a  $\mathcal{F}$  é introduzido no teorema a seguir. Por conveniência, define-se

$$x(w) = \arg \min_{x \in \Omega} \langle w, f(x) - y \rangle.$$

**Teorema 3.4.8.**  $y \in \mathcal{F}$  se e somente se

$$\theta(y) := \sup_{w \in \mathcal{W}} \inf_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0. \quad (3.4.1)$$

*Demonstração.* Suponha que  $y \notin \mathcal{F}$ . Então existe um  $w^0 \in \mathcal{W}$  tal que

$$\langle w^0, f(x(w^0)) - y \rangle > 0,$$

e portanto

$$\theta(y) \geq \langle w^0, f(x(w^0)) - y \rangle > 0.$$

Por outro lado, se  $y \in \mathcal{F}$ , então aplica-se o resultado do Lema 3.4.7, em particular para  $w \in \mathcal{W}$  que maximiza o lado esquerdo da desigualdade, o que completa a prova.  $\square$

**Proposição 3.4.9.** A função  $\theta$  é uma função convexa em  $\mathbb{R}^m$ , e portanto é uma função contínua.

*Demonstração.* Dado quaisquer dois pontos  $y_1, y_2 \in \mathbb{R}^m$  e qualquer escalar  $\lambda \in [0, 1]$ , defina

$$\phi_y(w) := \inf_{x \in \Omega} \langle w, f(x) - y \rangle.$$

Tem-se então que

$$\begin{aligned} \theta(\lambda y_1 + (1 - \lambda)y_2) &= \sup_{w \in \mathcal{W}} \phi_{\lambda y_1 + (1 - \lambda)y_2}(w) \\ &= \sup_{w \in \mathcal{W}} \{ \lambda \phi_{y_1}(w) + (1 - \lambda) \phi_{y_2}(w) \} \\ &\leq \lambda \theta(y_1) + (1 - \lambda) \theta(y_2), \end{aligned}$$

sendo a segunda passagem decorrente da linearidade de  $\phi_y$  em relação a  $y$ . Portanto  $\theta$  é convexa em  $\mathbb{R}^m$ , e, por conseqüência, contínua.  $\square$

Na prática, as desigualdades de que trata o Lema 3.4.7 devem ser entendidas como restrições aos valores que a variável  $y$  poderá assumir. Em outras palavras, o conjunto de

soluções viáveis de  $P_{\mathcal{F}}$  passa a ser representado como

$$\langle w, y \rangle \geq \inf_{x \in \Omega} \langle w, f(x) \rangle, \quad w \in \mathcal{W}.$$

### 3.5 Algoritmo Básico

Problemas de otimização que envolvam um grande número de restrições são adequadamente resolvidos por meio de técnicas de relaxação. Assim, define-se um subconjunto inicial de restrições e resolve-se o problema relaxado. Caso o problema relaxado não possua solução viável, o mesmo ocorre com o problema original e encerra-se o procedimento com esta conclusão. Caso contrário, obtém-se uma solução ótima para o problema relaxado e verifica-se se a solução satisfaz todas as restrições ignoradas. Se satisfazer, uma solução ótima do problema original também terá sido encontrada. Caso contrário, determinam-se quais restrições foram violadas, agregando-as ao subconjunto de restrições a ser considerado numa próxima iteração. Este procedimento básico pode ser refinado do ponto de vista computacional removendo-se restrições inativas numa dada iteração, por exemplo.

No caso de problemas com infinitas restrições, como o problema multiobjetivo  $P_{\mathcal{F}}$ , o aspecto complicante é verificar se uma solução ótima do problema relaxado satisfaz todas as (infinitas) restrições ignoradas. Entretanto, quando

$$\theta(y) = \langle w^*, f(x(w^*)) - y \rangle > 0,$$

sabe-se que pelo menos uma restrição foi violada e então

$$\langle w^*, y \rangle \geq \langle w^*, f(x(w^*)) \rangle$$

terá sido a restrição mais violada. Um algoritmo básico que a cada iteração incorpora apenas a restrição mais violada ao subconjunto corrente de restrições do problema relaxado pode ser estabelecido como segue.

#### Algoritmo Básico

**Passo 0.** Faça  $k = 0$  e defina  $\mathcal{F}^k = \mathbb{R}^m$ ;

**Passo 1.** Resolva o problema multiobjetivo relaxado

$$\begin{array}{ll} \text{minimizar} & F(y) \\ \text{sujeito a} & y \in \mathcal{F}^k, \end{array}$$

obtendo  $y \in \mathcal{F}^k$ .

**Passo 2.** Encontre  $\theta(y^k) = \langle w^k, f(x(w^k)) - y^k \rangle$ , sendo  $(x(w^k), w^k)$  a solução do problema representado no lado esquerdo de (3.4.1) para  $y = y^k$ . Se  $\theta(y^k) < \epsilon$ , onde  $\epsilon > 0$  é uma tolerância pré-especificada, então  $y^k$  resolve o problema no espaço-imagem e  $x(w^k) \in \Omega$  resolve o problema no espaço das decisões. Caso contrário, defina

$$\mathcal{F}^{k+1} := \left\{ y \in \mathcal{F}^k : \langle w^k, y^k \rangle \geq \langle w^k, f(x(w^k)) \rangle \right\},$$

faça  $k \leftarrow k + 1$  e volte ao **Passo 1**.

Na prática é sempre conveniente considerar  $\mathcal{F}^k := \{y \in \mathbb{R}^m : \underline{y} \leq y \leq \bar{y}\}$ , onde  $\underline{y}$  é a solução utópica do problema e  $\bar{y}$  é um limitante superior suficientemente grande para não excluir possíveis soluções eficientes do problema.

A prova de que a sequência  $\{y^k\}$  gerada pelo algoritmo converge para um ponto limite  $y^*$  pode ser conduzida por meio de argumentos gerais aplicáveis à convergência dos métodos de planos de corte. A demonstração de que o ponto limite de  $\{y^k\}$  resolve o problema multiobjetivo é apresentada a seguir.

**Teorema 3.5.1.** *Seja  $y^*$  o ponto limite da sequência  $\{y^k\}$  gerada pelo Algoritmo Básico. Então  $y^*$  resolve  $P_{\mathcal{F}}$ .*

*Demonstração.* Seja  $\{y^{k_l}\}$ ,  $l \in \mathbb{N}$ , qualquer subsequência de  $\{y^k\}$  convergente para  $y^*$ . Numa iteração arbitrária  $k_l$ , a última restrição incorporada a  $\mathcal{F}^{k_l}$  pode ser escrita como

$$0 \geq \langle w^{k_l}, f(x^{k_l}) - y^{k_l} \rangle + \langle w^{k_l}, y^{k_l} - y \rangle, \quad y \in \mathcal{F}^{k_l},$$

ou

$$0 \geq \theta(y^{k_l}) + \langle w^{k_l}, y^{k_l} - y \rangle, \quad y \in \mathcal{F}^{k_l},$$

e para  $k_{l'}$  com  $l' > l$ , a inequação anterior deve ser satisfeita para  $y^{k_{l'}}$ , isto é

$$\theta(y^{k_l}) \leq \langle w^{k_l}, y^{k_l} - y \rangle,$$

implicando que  $\theta(y^*) \leq 0$ , uma vez que  $\theta$  é contínua,  $\|w^{k_l}\| < 1$ ,  $\forall l \in \mathbb{N}$  e  $\|y^{k_{l'}} - y^{k_l}\| \rightarrow 0$  quando  $l \rightarrow \infty$ . Em outras palavras,  $y^* \in \mathcal{F}$ . Seja  $F^*$  o valor ótimo de  $F$ . Como  $\mathcal{F} \subseteq \mathcal{F}^k$ ,  $\forall k \in \mathbb{N}$ , então  $F(y^k) \leq F^*$ ,  $\forall k \in \mathbb{N}$ . Assim, pela continuidade de  $F$ , tem-se que  $F(y^*) \leq F^*$ , e portanto  $y^*$  resolve  $P_{\mathcal{F}}$ .  $\square$

### 3.6 Problema Max–Min

No Passo 2 do Algoritmo Básico, o valor da função  $\theta$  é obtido resolvendo-se o problema do tipo max–min

$$\begin{aligned} & \text{maximizar} && \phi(w) \\ & \text{sujeito a} && w \in \mathcal{W}, \end{aligned}$$

onde

$$\phi(w) := \inf_{x \in \Omega} \langle w, f(x) \rangle, \quad w \in \mathcal{W}.$$

Por sua vez, o ínfimo sobre  $x \in \Omega$  é obtido resolvendo-se o problema paramétrico

$$\begin{aligned} P_{\mathcal{W}} : & \text{minimizar} && \langle w, f(x) \rangle \\ & \text{sujeito a} && w \in \mathcal{W}, \quad x \in \Omega. \end{aligned}$$

Um método clássico de planos de corte pode ser usado para resolver o problema max–min e assim obter-se o valor  $\theta(y)$ . O método baseia-se em propriedades importantes da função  $\phi$ .

**Proposição 3.6.1.** *A função  $\phi(w) := \inf_{x \in \Omega} \langle w, f(x) \rangle$  é côncava sobre o conjunto  $\mathcal{W}$ .*

*Demonstração.* Seja  $w = \lambda w^1 + (1 - \lambda)w^2$  com  $w^1, w^2 \in \mathcal{W}$  e  $\lambda \in [0, 1]$ . Então,

$$\begin{aligned} \phi(\lambda w^1 + (1 - \lambda)w^2) &= \inf_{x \in \Omega} \langle \lambda w^1 + (1 - \lambda)w^2, f(x) - y \rangle, \\ &= \inf_{x \in \Omega} \left\{ \lambda \langle w^1, f(x) - y \rangle + (1 - \lambda) \langle w^2, f(x) - y \rangle \right\}, \\ &\geq \lambda \inf_{x \in \Omega} \langle w^1, f(x) - y \rangle + (1 - \lambda) \inf_{x \in \Omega} \langle w^2, f(x) - y \rangle, \\ &\geq \lambda \phi(w^1) + (1 - \lambda) \phi(w^2). \end{aligned}$$

$\square$

A função  $\phi$  é côncava, mas não necessariamente diferenciável sobre  $\mathcal{W}$ . O teorema a seguir indica como obter os subgradientes de  $\phi$  necessários à implementação de um método de planos de corte.

**Teorema 3.6.2.** *Seja  $\partial\phi(w^0)$  o subdiferencial de  $\phi$ , e  $\Omega(w^0)$  o conjunto solução de  $P_{\mathcal{W}}$  no ponto  $w^0 \in \mathcal{W}$ , respectivamente. Então*

$$f(x) - y \in \partial\phi(w^0), \quad \forall x \in \Omega(w^0).$$

*Demonstração.* Por definição

$$\phi(w) \leq \phi(w^0) + (\phi^{sg})^T(w - w^0), \quad \forall w, w^0 \in \mathcal{W},$$

onde  $\phi^{sg} \in \mathbb{R}^m$  é um subgradiente de  $\phi$  no ponto  $w^0$ . Por outro lado,

$$\phi(w) \leq \langle w, f(x) - y \rangle, \quad \forall x \in \Omega$$

e

$$\phi(w^0) \leq \langle w^0, f(x) - y \rangle, \quad \forall x \in \Omega(w^0).$$

Tomando  $x \in \Omega(w^0)$  e subtraindo as duas últimas expressões, obtém-se,

$$\begin{aligned} \phi(w) &\leq \phi(w^0) + \langle f(x) - y, w - w^0 \rangle, \\ &= \phi(w^0) + (f(x) - y)^T (w - w^0), \end{aligned}$$

e portanto  $f(x) - y \in \partial\phi(w^0)$ . □

Lembrando que cada subgradiente determina uma função suporte linear no ponto considerado, pode-se construir uma aproximação linear por partes para  $\phi$  na forma

$$\phi^k(w) = \min_{1 \leq j \leq k} \left\{ \langle w, f(x(w^j)) - y \rangle \right\},$$

onde  $k$  é o número de pontos considerados. A Figura 3.6 ilustra o processo de aproximação de  $\phi$  através de funções (hiperplanos) suportes.

O problema de maximização de  $\phi^k$  sobre  $w \in \mathcal{W}$  é então substituído pelo problema

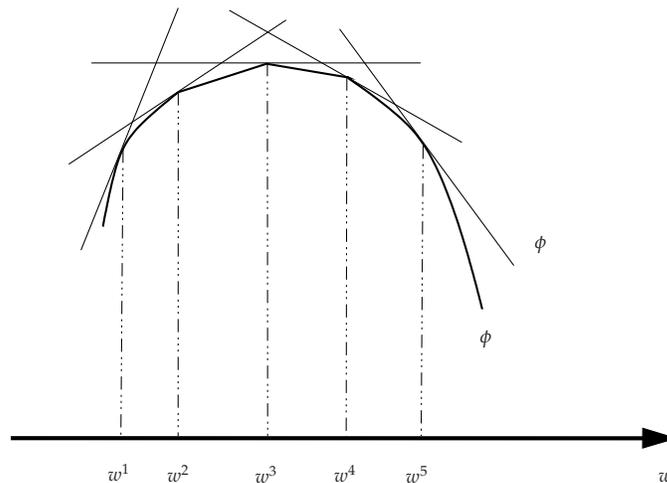


Figura 3.6: Aproximação da função  $\phi$ .

linear nas variáveis  $(w, \sigma)$ :

$$\begin{aligned} &\text{maximizar} && \sigma \\ &\text{sujeito a} && \sigma \leq \langle w, f(x(w^i)) - y \rangle, \quad i = 1, 2, \dots, k, \\ &&& w \in \mathcal{W}, \quad \sigma \in \mathbb{R}. \end{aligned}$$

Seja  $(w^{k+1}, \sigma^{k+1})$  a solução encontrada. O problema de minimização é resolvido para  $w = w^{k+1}$ , e então a restrição linear correspondente é agregada ao problema anterior com  $k$  substituído por  $k + 1$ . Uma vez que  $\sigma^{k+1}$  é um limite superior para  $\phi(w^{k+1})$ , o algoritmo convergirá quando  $\sigma^{k+1} - \phi(w^{k+1}) < \epsilon$ , para algum  $k$  e uma tolerância pré-especificada  $\epsilon > 0$ .

Um aspecto a ser destacado é que, por meio da abordagem por projeção, a solução do problema multiobjetivo pode ser obtida a partir de uma seqüência de problemas formulados no  $\mathbb{R}^m$ , independentemente da natureza das funções  $f_1, f_2, \dots, f_m$  e do subconjunto  $\Omega \subseteq \mathbb{R}^n$  sobre o qual estas funções estão definidas. As características particulares de cada classe de problemas são relevantes apenas no que diz respeito à solução do problema paramétrico

$$\begin{aligned} P_{\mathcal{W}} : & \text{minimizar} && \langle w, f(x) \rangle \\ & \text{sujeito a} && w \in \mathcal{W}, \quad x \in \Omega, \end{aligned}$$

uma vez que cada classe de problemas impõe condições de otimalidade específicas e for-

nece métodos numéricos apropriados para se obter soluções ótimas para  $P_{\mathcal{W}}$ .

## 3.7 Resumo

Neste capítulo estudamos conceitos e técnicas de otimização multiobjetivo. Analizamos formulações de problemas de otimização multiobjetivos tanto no espaço dos variáveis de decisão quanto no espaço–imagem, e discutimos maneiras diferentes de caracterizar e gerar soluções eficientes (ou Pareto–ótimas), que constituem o conjunto de soluções candidatas a resolver problemas multiobjetivos. Os conceitos e métodos analisados — em particular o Algoritmo Básico (Seção 3.5) — serão empregados nos capítulos 4 e 5 a seguir, nos quais problemas multiplicativos e fracionários generalizados são vistos como problemas multiobjetivos especiais.



# 4

## Programação Multiplicativa Generalizada

### Sumário

---

4.1	Introdução . . . . .	68
4.2	Programação Multiplicativa Generalizada . . . . .	73
4.3	Programação Multiplicativa Generalizada Côncava . . . . .	76
4.4	Programação Multiplicativa Generalizada Convexa . . . . .	85
4.5	Experimentos Numéricos . . . . .	96
4.6	Resumo . . . . .	106

---

Neste capítulo propomos uma abordagem no espaço–imagem para a resolução de problemas multiplicativos generalizados, que envolvem a maximização e a minimização de somas finitas de produtos de funções côncavas e convexas, respectivamente, sobre um conjunto convexo compacto e não–vazio. Veremos que, no espaço–imagem, estes dois problemas podem ser eficientemente resolvidos por um mesmo algoritmo de aproximação externa. Mostramos como reformular esses problemas não–convexos como problemas de Programação Quadrática Indefinida com infinitas restrições lineares de desigualdade. Dois algoritmos com a estrutura do Algoritmo Básico do capítulo anterior são propostos, o primeiro baseado em uma técnica de enumeração de restrições, o segundo em uma técnica *branch–and–bound*. O desempenho computacional dos algoritmos é avaliado com base em problemas–teste da literatura.

## 4.1 Introdução

Neste capítulo abordaremos duas variantes do problema de otimizar uma soma de produtos de funções escalares positivas sobre um conjunto viável convexo, problema que tem sido referido na literatura de otimização global como Problema Multiplicativo Generalizado. A generalização em relação ao problema multiplicativo tradicional [KK95], no qual um único produto é otimizado, baseia-se em permitir que a função objetivo seja expressa como uma soma de produtos, em particular, como uma soma de produtos de duas funções [KKY94], [KFoo], [Beno8] e [OF10], que é a formulação adotada na presente tese. Na última década, muitos algoritmos têm sido propostos para resolver globalmente o problema multiplicativo generalizado, especialmente por Konno e co-autores [KK92], [KYM91] e [KSY93].

Os problemas de programação multiplicativa generalizada considerados apresentam as seguintes formas gerais:

$$\min_{x \in \Omega} \left[ f_1(x) + \sum_{i=1}^r f_{2i}(x) f_{2i+1}(x) \right], \quad (4.1.1)$$

e

$$\max_{x \in \Omega} \left[ g_1(x) + \sum_{i=1}^r g_{2i}(x) g_{2i+1}(x) \right], \quad (4.1.2)$$

nos quais  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções convexas e  $g_k : \mathbb{R}^n \rightarrow \mathbb{R}$  são côncavas para  $k = 1, 2, \dots, 2r + 1$ . Assume-se que  $\Omega$  é um subconjunto convexo compacto e não–vazio do  $\mathbb{R}^n$ , e que as funções  $f_k$  e  $g_k$  são positivas em  $\Omega$  para  $k = 1, 2, \dots, 2r + 1$ . Em otimização

global, os problemas (4.1.1) e (4.1.2) são conhecidos como problema multiplicativo generalizado convexo [KKY94], [OF10] e problema multiplicativo generalizado côncavo [Beno8], [AF11b], respectivamente, embora as funções objetivos de ambos os problemas sejam não-convexas, em geral.

Diferentes abordagens têm sido propostas para resolver globalmente problemas dos tipos (4.1.1) e (4.1.2). Em [KKY94], os autores apresentam uma técnica de aproximação externa, que estende a abordagem para programação multiplicativa, introduzida em [KYK93], para a minimização de uma soma de produtos de duas funções convexas. Os autores reduzem o problema (4.1.1) a um problema de minimização côncava  $2r$ -dimensional equivalente, introduzindo variáveis auxiliares. Em seguida, o problema resultante é resolvido por meio de um algoritmo de planos de corte. A transformação paramétrica proposta em [KYK93] é estendida em [JMT97] para a minimização da soma de uma função convexa e um produto de funções convexas. Este problema particular de programação multiplicativa generalizada é reescrito como um problema de minimização quase-côncava e resolvido por um algoritmo *branch-and-bound* cônico.

[KFoo] discute um algoritmo *branch-and-bound* para resolver problemas de programação multiplicativa generalizada lineares da forma (4.1.1), na qual as funções  $f_k$  são lineares e positivas em  $\Omega \subseteq \mathbb{R}^n$ ,  $r$  é um número pequeno e  $\Omega$  é um politopo. Os autores mostram que este problema pode ser resolvido de maneira eficiente adaptando-se um algoritmo *branch-and-bound*, proposto por em [AMF95], para problemas não-convexos contendo produtos de duas variáveis. Esse problema é, por sua vez, um problema de Programação Quadrática Indefinida, conhecido por ser NP-difícil [Mat96].

Existem vários algoritmos eficientes para o problema correspondente a  $r = 1$ . Algoritmos Simplex paramétricos são propostos em [KKY92] e [Swa66]. O algoritmo Simplex paramétrico proposto em [KKY92] emprega não mais do que o dobro do tempo de computação necessário para resolver problemas de programação lineares. Problemas multiplicativos generalizados lineares ( $r \geq 2$ ) são muito mais difíceis de resolver. Algoritmos têm sido propostos [KK91], e [PAT95], mas o esforço computacional cresce exponencialmente à medida que  $r$  aumenta. Em [PAT95], uma implementação cuidadosa de um algoritmo *branch-and-bound* foi capaz de resolver problemas multiplicativos generalizados lineares com  $r = 10$ .

Em [RS03] desenvolve-se um algoritmo de otimização global especializado para problemas de programação multiplicativa generalizada lineares baseado em subestimadores

convexos [RS01] e em uma nova técnica de ramificação para algoritmos *branch-and-bound* retangulares. Cada termo na função objetivo é definido como sendo o produto de um número qualquer de funções lineares, irrestritas em sinal, em oposição à formulação convencional, que envolve a soma de produtos de duas funções positivas ([KK95] e [SS95]).

[Beno8] apresentou o primeiro algoritmo para resolver globalmente o problema (4.1.2). O algoritmo proposto em [Beno8] implementa um algoritmo *branch-and-bound*, aplicado a um problema é equivalente ao problema (4.1.2). Durante o processo, os limites superiores exigidos são calculados por meio da resolução de problemas de programação convexa. O algoritmo *branch-and-bound* descrito em [Beno8] opera no espaço  $\mathbb{R}^r$  em vez do espaço  $\mathbb{R}^n$ , uma característica que tende a reduzir o esforço computacional necessário.

Um algoritmo para problemas de programação multiplicativa generalizada do tipo (4.1.1), baseado em elementos de Análise Convexa, foi proposto em [OF10]. Neste trabalho, os autores estendem a abordagem proposta em [OF08] para resolver globalmente problemas de programação convexa multiplicativa. Os aspectos centrais do algoritmo proposto em [OF08] são a redução do problema de programação convexa multiplicativa à minimização de uma função quase-côncava sobre um politopo, realizada por meio de um procedimento de enumeração de vértices.

A extensão proposta em [OF10] implica no uso de uma técnica mais geral para resolver problema no espaço-imagem, representado por um problema quadrático indefinido de pequena dimensão e estrutura particular. Usando resultados de Análise Convexa, o problema (4.1.1) é reformulado no espaço-imagem como um problema de Programação Quadrática Indefinita equivalente, com infinitas restrições lineares de desigualdade. Em seguida, um algoritmo que combina relaxação com uma técnica de enumeração de restrições é usado para resolver o problema no espaço-imagem.

### 4.1.1 Aplicações

Problemas de programação multiplicativa generalizada (4.1.1) e (4.1.2) atraem considerável atenção na literatura por suas aplicações em várias áreas de estudo, como Microeconomia, Otimização Financeira, Otimização de Portfólios, Otimização Robusta, Geometria Computacional, Projetos de Leiaute [KYM91], [KSY93], [AF10a], [AF11a], [AF11c]. Muitas

aplicações dos problemas (4.1.1) e (4.1.2) envolvem a formulação

$$\begin{aligned} \min/\text{maximizar} \quad & h(x) + \sum_{i=1}^r \left[ \langle a^i, x \rangle + \alpha_i \right] \left[ \langle b^i, x \rangle + \beta_i \right] \\ \text{sujeito a} \quad & x \in \Omega, \end{aligned} \tag{4.1.3}$$

onde  $h$  é uma função convexa/côncava sobre  $\Omega$ , e nenhuma restrição de sinal é imposta a  $\langle a^i, x \rangle + \alpha_i$  ou  $\langle b^i, x \rangle + \beta_i$ ,  $i = 1, 2, \dots, r$ ,  $a^i, b^i \in \mathbb{R}^n$  e  $\alpha_i, \beta_i \in \mathbb{R}$ . Para mostrar que a hipótese de positividade de  $\langle a^i, x \rangle + \alpha_i$  e  $\langle b^i, x \rangle + \beta_i$ ,  $i = 1, 2, \dots, r$ , não é restritiva, para cada  $i = 1, 2, \dots, r$  considere

$$p_i < \min \left\{ \min_{x \in \Omega} \left[ \langle a^i, x \rangle + \alpha_i \right], \min_{x \in \Omega} \left[ \langle b^i, x \rangle + \beta_i \right] \right\}.$$

Em seguida, nos problemas (4.1.1) e (4.1.2), defina

$$\begin{aligned} f_1(x) &:= h(x) + \sum_{i=1}^r \left[ p_i \langle a^i, x \rangle + p_i \langle b^i, x \rangle + p_i(\alpha_i + \beta_i) - p_i^2 \right], \\ f_{2i}(x) &:= \langle a^i, x \rangle + (\alpha_i - p_i), \quad i = 1, 2, \dots, r, \\ f_{2i+1}(x) &:= \langle b^i, x \rangle + (\beta_i - p_i), \quad i = 1, 2, \dots, r, \end{aligned}$$

e

$$\begin{aligned} g_1(x) &:= h(x) + \sum_{i=1}^r \left[ p_i \langle a^i, x \rangle + p_i \langle b^i, x \rangle + p_i(\alpha_i + \beta_i) - p_i^2 \right], \\ g_{2i}(x) &:= \langle a^i, x \rangle + (\alpha_i - p_i), \quad i = 1, 2, \dots, r, \\ g_{2i+1}(x) &:= \langle b^i, x \rangle + (\beta_i - p_i), \quad i = 1, 2, \dots, r. \end{aligned}$$

Demonstra-se então que os problemas (4.1.1) e (4.1.2) são equivalentes aos problema (4.1.3), e todas as hipóteses exigidas são satisfeitas. As formulações (4.1.1)–(4.1.2) são representativas de problemas importantes de otimização. Um problema geral de Programação Quadrática pode ser escrito como

$$\begin{aligned} \min_{x \in \Omega} \quad & \frac{1}{2} x^T Q x + \langle q, x \rangle + \gamma \\ \Omega &:= \{x \in \mathbb{R}^n \mid Ax \leq b, \quad x \geq 0\}, \end{aligned}$$

onde  $Q \in \mathbb{R}^{n \times n}$  é uma matriz simétrica de posto  $p \leq n$ ,  $q \in \mathbb{R}^n$  e  $\gamma \in \mathbb{R}$ . [Tuy98]

fornece um método construtivo simples para encontrar vetores linearmente independentes  $v^1, v^2, \dots, v^r \in \mathbb{R}^n$  e vetores  $w^1, w^2, \dots, w^r \in \mathbb{R}^n$  tais que, para todo  $x \in \mathbb{R}^n$ ,

$$\frac{1}{2}x^T Qx = \frac{1}{2} \sum_{i=1}^r \langle v^i, x \rangle \langle w^i, x \rangle.$$

Uma classe importante de problemas de programação fracionária generalizada (estudado no Capítulo 5) é também representada pela formulação (4.1.1). Suponha que  $f_0, f_1, \dots, f_r$  são funções convexas e positivas e  $g_1, g_2, \dots, g_r$  são funções côncavas e positivas em  $\Omega \subseteq \mathbb{R}^n$ . Então cada razão  $f_i/g_i(x)$  é uma função convexa e positiva em  $\Omega$ , e o problema de programação fracionária generalizada

$$\min_{x \in \Omega} f_0(x) + \sum_{i=1}^r \frac{f_i(x)}{g_i(x)} \tag{4.1.4}$$

reduz-se a (4.1.1).

Assim, as formulações (4.1.1)–(4.1.2) representam todas as aplicações de Programação Quadrática Indefinida e Programação Fracionária Generalizada

As formulações (4.1.3) também permitem descrever problemas de Programação Bilinear como casos especiais. Sejam  $f \in \mathbb{R}^q, u \in \mathbb{R}^r, F \in \mathbb{R}^{q \times s}$  e  $U \in \mathbb{R}^{r \times t}$ . Suponha que

$$Y := \{y \in \mathbb{R}^s \mid Fy \leq f, y \geq 0\},$$

e

$$Z := \{z \in \mathbb{R}^t \mid Uz \leq u, z \geq 0\},$$

são politopos compactos não-vazios. O problema geral de Programação Bilinear é dado por

$$\begin{aligned} &\text{minimizar} && \langle a, y \rangle + \langle b, z \rangle + y^T C z \\ &\text{sujeito a} && y \in Y, \quad z \in Z, \end{aligned}$$

onde  $a \in \mathbb{R}^s, b \in \mathbb{R}^t$  e  $C^{s \times t}$  é uma matriz de posto  $r$ . Em [KY] os autores mostram, por meio de um processo construtivo, que o problema de Programação Bilinear pode ser reescrito na forma

$$\begin{aligned} &\text{maximizar} && \langle -a, y \rangle + \langle -b, z \rangle + \sum_{i=1}^r \langle v^i, y \rangle \langle w^i, y \rangle \\ &\text{sujeito a} && y \in Y, \quad z \in Z, \end{aligned}$$

no qual  $v^i \in \mathbb{R}^s, w^i \in \mathbb{R}^t, i = 1, 2, \dots, r$ . Este último problema é um caso especial da formulação (4.1.3) com  $x^T = [y^T, z^T]$ ,  $h(x) = \langle -a, y \rangle + \langle -b, z \rangle$ ,  $(a^i)^T = [(v^i)^T, 0^T]$  e  $(b^i)^T = [0^T, (w^i)^T]^T, i = 1, 2, \dots, r, \alpha_i = \beta_i = 0, i = 1, 2, \dots, r$ , e

$$\Omega = \left\{ (y, z) \in \mathbb{R}^{s+t} \mid Fy \leq f, Uz \leq u, y, z \geq 0 \right\}.$$

Portanto, entre as aplicações do problema (4.1.1) estão todas as aplicações de problemas de Programação Bilinear [KT97], [Man64], [Fre79], [Fal73].

Um problema de Programação Linear Zero–Um pode ser escrito como

$$\begin{aligned} & \text{maximizar} && \langle c, x \rangle \\ & \text{sujeito a} && x \in \Omega, \\ & && x_i \in \{0, 1\}, \quad i = 1, 2, \dots, n, \end{aligned} \tag{4.1.5}$$

onde  $c \in \mathbb{R}^n$  e  $\Omega \subseteq \mathbb{R}^n$  é um politopo não–vazio. Em [Rag69], demonstra-se que para  $\epsilon > 0$  suficientemente pequeno,  $x^*$  é uma solução ótima para problema (4.1.5) se e somente se  $x^*$  é uma solução ótima para o problema

$$\begin{aligned} & \text{maximizar} && \langle c, x \rangle + \frac{1}{\epsilon} \sum_{i=1}^n (-x_i)(1 - x_i) \\ & \text{sujeito a} && x \in \Omega, \\ & && 0 \leq x_i \leq 1, \quad i = 1, 2, \dots, n. \end{aligned} \tag{4.1.6}$$

Como o problema (4.1.6) é um caso especial do problema (4.1.3), segue-se que todas as numerosas aplicações de Programação Linear Zero–Um estão incluídas entre as aplicações dos problemas (4.1.1) e (4.1.2). Veja por exemplo [NW98], para a discussão de algumas dessas aplicações.

## 4.2 Programação Multiplicativa Generalizada

Considere os problemas de otimização

$$P_{MG} : \min_{x \in \Omega} \left[ v_1 = f_1(x) + \sum_{i=1}^r f_{2i}(x) f_{2i+1}(x) \right],$$

e

$$Q_{MG} : \max_{x \in \Omega} \left[ v_2 = g_1(x) + \sum_{i=1}^r g_{2i}(x)g_{2i+1}(x) \right],$$

nos quais  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $g_k : \mathbb{R}^n \rightarrow \mathbb{R}$ , para  $k = 1, 2, \dots, 2r + 1$ , e  $\Omega$  satisfazem as condições impostas no começo da Seção 4.1. Como as funções objetivos  $v_1$  e  $v_2$  não são, em geral, funções convexas e côncavas, respectivamente, os problemas  $(P_{MG})$  e  $(Q_{MG})$  podem ter ótimos locais que não são globais.

Os pressupostos que cada função participante na soma seja positiva sobre o conjunto viável do problema, e, adicionalmente, de que cada função seja convexa, se o objetivo é minimizar, ou côncava, se o objetivo é maximizar, fornecem alguma estrutura para o problema e não excluem aplicações importantes. Apesar da positividade e das hipóteses de convexidade/concavidade, os problemas multiplicativos generalizados  $(P_{MG})$  e  $(Q_{MG})$  são ambos não-convexos: o produto de funções convexas positivas não é necessariamente uma função convexa ou quase-convexa, e embora o produto de funções côncavas positivas seja uma função quase-côncava, a soma das funções quase-côncavas não é uma função quase-côncava [HPT95], em geral.

**Exemplo 4.2.1.** Como um caso especial do problema  $(Q_{MG})$ , considere

$$\left| \begin{array}{ll} \text{maximizar} & (x_1 + x_2) + 4(-x_1)(x_2) \\ & x_1, x_2 \\ \text{sujeito a} & 2x_1 + x_2 \geq 12, \\ & x_1 + 2x_2 \geq 12, \\ & x_1, x_2 \leq 6. \end{array} \right.$$

É possível mostrar que  $x^0 = (6, 3)$  e  $x^1 = (3, 6)$  são soluções ótimas locais para este problema, cada uma fornecendo  $-63$  para o valor da função objetivo, mas que  $x^0$  e  $x^1$  não são soluções ótimas globais. O valor ótimo  $-56$  é atingido em  $x^* = (4, 4)$ .

### 4.2.1 Dualidade no Espaço-Imagem

Uma das conclusões deste capítulo é que os problemas  $(P_{MG})$  e  $(Q_{MG})$  são análogos do ponto de vista da abordagem proposta em [OF10]. A relação entre os problemas  $(P_{MG})$  e  $(Q_{MG})$  torna-se aparente quando suas funções objetivos são reescritas como as composições  $v_1(x) = u(f(x))$  e  $v_2(x) = u(g(x))$ , onde  $f := (f_1, f_2, \dots, f_m)$ ,  $g := (g_1, g_2, \dots, g_m)$ , e

$u : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $m = 2r + 1$ , é dada por

$$u(\xi) := \xi_1 + \sum_{i=1}^r \xi_{2i} \xi_{2i+1} = \frac{1}{2} \xi^T Q \xi + c^T \xi, \quad (4.2.1)$$

com definições adequadas para a matriz  $Q \in \mathbb{R}^{m \times m}$  e o vetor de zeros e uns  $c \in \mathbb{R}^m$ ,  $m = 2r + 1$ .

Deve-se notar que os problemas  $(P_{MG})$  e  $(Q_{MG})$  não são equivalentes no sentido de que, a partir de uma solução ótima de um deles, uma solução ótima para o outro é facilmente determinada. No entanto, os mesmos princípios utilizados em [OF10] para resolver globalmente o problema de minimização  $(P_{MG})$  podem ser aplicados, com as modificações necessárias, para resolver globalmente o problema de maximização  $(Q_{MG})$ .

#### 4.2.2 Formulações no Espaços Imagem

Aos problemas  $(P_{MG})$  e  $(Q_{MG})$  podem-se associar os problemas de minimizar e maximizar as funções objetivos vetoriais  $f := (f_1, f_2, \dots, f_m)$  e  $g := (g_1, g_2, \dots, g_m)$ , respectivamente, em  $\Omega$ . A função  $u$  definida por (4.2.1) pode ser vista como uma função de agregação para o problema de minimizar  $f$  em  $\Omega$ , ou para o problema de maximizar  $g$  em  $\Omega$  [Yu85].

Sejam  $\mathcal{Y} := f(\Omega)$  e  $\mathcal{Z} := g(\Omega)$  as imagens de  $\Omega$  pelas funções  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , respectivamente. As projeções dos problemas  $(P_{MG})$  e  $(Q_{MG})$  nos espaços imagem  $\mathcal{Y}$  e  $\mathcal{Z}$  são

$$\min_{y \in \mathcal{Y}} u(y), \quad (4.2.2)$$

e

$$\max_{z \in \mathcal{Z}} u(z). \quad (4.2.3)$$

A representação (4.2.2), proposta em [OF10], se apoia em duas propriedades básicas: a positividade de  $f(\Omega)$  e sua  $\mathbb{R}_+^m$ -convexidade [Yu74], que significa que  $f(\Omega) + \mathbb{R}_+^m$  é um conjunto convexo. Este último resultado tem um papel central em Otimização Multiobjetivo [Yu85], e em Análise do Espaço-Imagem [Gia05], onde  $f(\Omega) + \mathbb{R}_+^m$  é conhecida como uma *extensão cônica* de  $f(\Omega)$ . Estas mesmas propriedades fundamentais (positividade e  $\mathbb{R}_+^m$ -convexidade) estão presentes na abordagem do problema (4.2.3) proposta nesta tese. A positividade da função vetorial côncava e a  $\mathbb{R}_-^m$ -convexidade de  $\mathcal{Z}$  formam a base para o desenvolvimento de um algoritmo de Otimização Global com propriedades similares às

apresentadas pelo algoritmo proposto em [OF10].

A equivalência entre o problema (4.2.2) e a sua reformulação,

$$\min_{y \in (\mathcal{Y} + \mathbb{R}_+^m)} u(y), \quad (4.2.4)$$

é estabelecida usando-se o fato de que a função de agregação  $u$ , definida por (4.2.1), é estritamente crescente em  $\mathcal{Y} + \mathbb{R}_+^m$ , como uma consequência direta da positividade de  $f$  em  $\Omega$ . Intuitivamente, ao resolver (4.2.4), o caráter estritamente crescente de  $u$  não admitirá incrementos não-negativos nos vetores de  $\mathcal{Y}$ , tornando evidente que um minimizador global de (4.2.4) também é um minimizador global de (4.2.2). O mesmo raciocínio se aplica na direção inversa, ou seja, qualquer minimizador global de (4.2.2) também é um minimizador global de (4.2.4).

**Teorema 4.2.2.** [OF10] *Seja  $y^*$  uma solução ótima para o problema (4.2.4). Então  $y^*$  também é uma solução ótima para o problema (4.2.2). Reciprocamente, se  $y^*$  resolve (4.2.2), então  $y^*$  também resolve (4.2.4).*

A monotonicidade de  $u$  e a concavidade de  $g$  podem ser usadas para mostrar que (4.2.3) é equivalente, no sentido descrito acima, ao problema

$$\max_{z \in (\mathcal{Z} - \mathbb{R}_+^m) \cap \mathbb{R}_{++}^m} u(z). \quad (4.2.5)$$

### 4.3 Programação Multiplicativa Generalizada Côncava

Considere o problema (4.2.3) e lembre que  $\mathcal{Z} := g(\Omega)$ . A continuidade (implicada pela concavidade) de  $g$  e a compactidade de  $\Omega$  implicam na compactidade de  $\mathcal{Z}$ , mas não a sua convexidade. Entretanto,  $\mathcal{Z}$  é cone-convexo, ou seja,  $\mathcal{G} := \mathcal{Z} - \mathbb{R}_+^m$  é um conjunto convexo [Yu85].

**Proposição 4.3.1.** *Defina  $\mathcal{G} := \mathcal{Z} - \mathbb{R}_+^m$ . Então*

$$\mathcal{G} := \{z \in \mathbb{R}^m : g(x) \geq z \text{ para algum } x \in \Omega\}.$$

*Demonstração.* Pela definição de  $\mathcal{G}$ , qualquer vetor  $z \in \mathcal{G}$  pode ser escrito como  $z := g(x) - d$ , sendo que  $g(x) \in \mathcal{Z}$ ,  $x \in \Omega$ , e  $d \in \mathbb{R}_+^m$ .  $\square$

As abordagens que visam a resolução do problema ( $Q_{PM}$ ) por meio da resolução do seu problema equivalente (4.2.3) no espaço–imagem diferem basicamente na forma de representar o conjunto (geralmente) não–convexo  $\mathcal{Z}$ . O problema da representação de  $\mathcal{G}$  em termos práticos é abordada por meio de um resultado fundamental em Análise Convexa, já apresentado no Capítulo 3.

**Lema 4.3.2.** *Dado  $z \in \mathbb{R}^m$ , então  $z \in \mathcal{G}$  se e somente se  $z$  satisfaz o sistema de infinitas desigualdades lineares*

$$\langle w, z \rangle \leq \max_{x \in \Omega} \langle w, g(x) \rangle \quad \text{para todo } w \in \mathcal{W}, \quad (4.3.1)$$

O problema (4.2.3) admite uma formulação equivalente com uma região viável convexa.

**Teorema 4.3.3.** *Seja  $z^*$  uma solução ótima para*

$$\max_{z \in \mathcal{G} \cap \mathbb{R}_{++}} u(z). \quad (4.3.2)$$

*Então  $z^*$  também é uma solução ótima para o problema (4.2.3). Reciprocamente, se  $z^*$  resolve (4.2.3), então  $z^*$  também resolve (4.3.2).*

*Demonstração.* Como qualquer  $z = g(x)$ ,  $x \in \Omega$ , é viável para (4.3.2), o conjunto viável de (4.3.2) contém o conjunto viável de (4.2.3), e o valor ótimo de (4.3.2) é um limite superior para o valor ótimo de (4.2.3). Se  $z^*$  resolve (4.3.2), então pelo Lema 4.3.2 existe  $x^* \in \Omega$  tal que  $g(x^*) \geq z^*$  e, na prática,  $g(x^*) = z^*$ . Caso contrário, (ou seja, se  $g(x^*) \geq z^*$  e  $g(x^*) \neq z^*$ ), a viabilidade de  $g(x^*)$  para (4.3.2) e o caráter estritamente crescente de  $u$  em  $\mathcal{G}$  estaria em contradição com a otimalidade de  $z^*$ . Como  $g(x^*)$  é viável para (4.2.3), conclui-se que  $z^*$  também resolve (4.2.3). A afirmação inversa é provada usando argumentos semelhantes.  $\square$

O problema (4.3.2) é um problema de Otimização Semi–infinita devido ao sistema de desigualdades semi–infinito (4.3.1) (parametrizado por  $w \in \mathcal{W}$ ), que descreve a sua região viável. Algoritmos de relaxação têm sido utilizados para resolver problemas semi–infinitos. O algoritmo de relaxação adotado neste trabalho se desenvolve por meio da determinação de  $z^k$ , um maximizador global de  $u$  sobre uma aproximação externa  $\mathcal{G}^k$  de  $\mathcal{G}$ , descrita por um subconjunto das restrições de desigualdade (4.3.1). Caso o problema relaxado não possua uma solução viável, o mesmo ocorre com o problema original e encerra-se o procedimento com esta conclusão. Caso contrário, obtém-se uma solução ótima para o problema relaxado e verifica-se se a solução satisfaz todas as restrições ignoradas. Se

satisfazer, uma solução ótima do problema original também terá sido encontrada. Caso contrário, determina-se e agrega-se a  $\mathcal{G}^k$  a restrição mais violada por  $z^k$ . Este procedimento básico pode ser refinado do ponto de vista computacional removendo-se restrições inativas numa dada iteração, por exemplo. A restrição mais violada é determinada pelo cálculo de

$$\theta(z^k) := \min_{w \in \mathcal{W}} \phi_{z^k}(w), \quad (4.3.3)$$

sendo

$$\phi_{z^k}(w) := \max_{x \in \Omega} \langle w, g(x) - z^k \rangle. \quad (4.3.4)$$

**Lema 4.3.4.**  $z^k \in \mathcal{G}$  se e somente se  $\theta(z^k) \geq 0$ .

*Demonstração.* Suponha que  $z^k \in \mathcal{G}$ , ou seja,  $g(x) \geq z^k$  para algum  $x \in \Omega$ . Seja  $w^* \in \mathcal{W}$  um maximizador de  $\max_{x \in \Omega} \langle w, g(x) - z^k \rangle$ . Obviamente, isto implica que  $\theta(z^k) \geq 0$ . Reciprocamente, assumamos que  $z^k \notin \mathcal{G}$ , ou seja,  $g_j(x) < z_j^k$  para todo  $x \in \Omega$  e algum  $j$ . Neste caso, existe algum  $w^0 \in \mathcal{W}$  tal que

$$\theta(z^k) \leq \max_{x \in \Omega} \langle w^0, g(x) - z^k \rangle < 0,$$

o que completa a prova. □

Um ponto  $z^k$  satisfaz o sistema de desigualdades semi-infinito se e somente se  $\theta(z^k) \geq 0$ . Se  $\theta(z^k) < 0$ , então a solução ótima do Problema Min-Max (4.3.3)–(4.3.4), denotada como  $(w^k, x^k)$ , caracteriza a restrição de desigualdade mais violada. Problema Min-Max são encontrados frequentemente em áreas como Otimização, Engenharia, Controle Ótimo, Microeconomia, Teoria dos Jogos, entre outras.

**Teorema 4.3.5.** Dado  $z \in \mathbb{R}^m$ ,  $\theta(z)$  é o valor ótimo do problema de otimização convexa

$$\begin{aligned} & \text{maximizar} && \sigma \\ & \text{sujeito a} && g(x) \geq \sigma e + z, \quad x \in \Omega, \end{aligned} \quad (4.3.5)$$

onde  $\sigma \in \mathbb{R}$  e  $e \in \mathbb{R}^m$  é o vetor  $m$ -dimensional de uns ( $e = (1, 1, \dots, 1)$ ).

*Demonstração.* O problema dual Lagrangeano de (4.3.5) é

$$\text{minimizar}_{w \in \mathbb{R}_+^m} \max_{x \in \Omega, \sigma \in \mathbb{R}} \{ \sigma + \langle w, g(x) - \sigma e - z \rangle \}, \quad (4.3.6)$$

no qual  $w \in \mathbb{R}_+^m$  é o vetor de variáveis duais associadas às restrições de desigualdade. A existência do máximo impõe que  $\sum_{i=1}^m w_i = 1$ , e o problema (4.3.6) reduz-se a

$$\text{minimizar } \max_{w \in \mathcal{W}} \langle w, g(x) - z \rangle, \quad (4.3.7)$$

cujo valor ótimo, por definição, é  $\theta(z)$ . Como o problema primal (4.3.5) é convexo, sob condições de qualificação das restrições [Ber95], não há gap de dualidade e os problemas (4.3.5) e (4.3.7) têm o mesmo valor ótimo,  $\theta(z)$ .  $\square$

O Teorema 4.3.5 sugere a seguinte interpretação geométrica para  $\theta(z)$ , dado  $z \in \mathbb{R}^m$ . Seja  $(x^*, w^*)$  soluções primal e dual ótimas de problema (4.3.5). Pelo Teorema 4.3.5,  $g(x^*) \geq \theta(z)e + z$ , e se  $w^* \in \mathbb{R}_{++}^m$ , a desigualdade torna-se uma igualdade. O caso  $\theta(z) < 0$  é mais relevante para a análise, porque o algoritmo de relaxação proposto na próxima seção gera uma seqüência de pontos inviáveis ( $z \notin \mathcal{G}$ ,  $\theta(z) < 0$ ) convergindo para uma solução ótima de (4.3.2). Neste caso,  $z_i - g_i(x^*) = -\theta(z) > 0$  para  $i = 1, 2, \dots, m$ . Portanto, uma condição do tipo  $-\theta(z) < \epsilon$ ,  $\epsilon > 0$ , implicará que a violação das desigualdades  $g(x) \geq z$  se dará por no máximo  $\epsilon$ .

A seguir, apresentamos algumas propriedades geométricas adicionais para  $\phi_z$  e  $\theta$ , nas quais são análogas às obtidas no Capítulo 3 para problemas de Otimização Multiobjetivo. Veja também [AF11b].

**Teorema 4.3.6.** *Sejam  $\phi_z$  e  $\theta$  definidas por (4.3.4) e (4.3.3), respectivamente. Então*

- (i) *Dado  $y \in \mathbb{R}^m$ , a função  $\phi_z$  é convexa em  $\mathcal{W}$ ;*
- (ii)  *$(g(x(w)) - z)$  é um subgradiente de  $\phi_z$  em  $w \in \mathcal{W}$ ;*
- (iii) *A função  $\theta$  é côncava em  $\mathbb{R}^m$ .*

Em cada ponto  $w \in \mathcal{W}$ , um subgradiente  $g(x(w)) - z$  determina um hiperplano suporte ao epígrafo de  $\phi_z$ , o que nos permite construir aproximações lineares por partes para  $\phi_z$ . A  $l$ -ésima aproximação de  $\phi_z$  seria

$$\phi_z^l(w) = \max_{1 \leq i \leq l} \{ \langle w, g(x(w^i)) - z \rangle \}. \quad (4.3.8)$$

Em vez da minimização convexa em (4.3.3), consideramos o problema de minimizar  $\phi_z^l$  em  $\mathcal{W}$ , que por sua vez, pode ser descrito como um problema de Programação Linear. Especificamente, para algum  $l$  fixo e  $w^l \in \mathcal{W}$ , encontramos  $x(w^l)$  resolvendo o problema

de Programação Côncava  $\max_{x \in \Omega} \langle w, g(x) - z \rangle$ , e, em seguida, obtemos  $\sigma^{l+1}$  e  $w^{l+1}$  como o valor ótimo e uma solução ótima do seguinte problema de Programação Linear

$$\begin{aligned} & \text{minimizar} && \sigma \\ & \text{sujeito a} && \sigma \geq \langle w, g(x(w^i)) - z \rangle, \quad i = 1, 2, \dots, l, \\ & && w \in \mathcal{W}, \sigma \in \mathbb{R}. \end{aligned} \tag{4.3.9}$$

Este procedimento já foi discutido no Capítulo 3, e será reintroduzido a seguir por conveniência.

**Algoritmo MIN-MAX:** (Viabilidade de  $z \in \mathbb{R}^m$ )

**Passo 0.** Faça  $l := 0$  e escolha  $w^l \in \mathcal{W}$ ;

**Passo 1.** Resolva o problema de Programação Côncava

$$P_W : \max_{x \in \Omega} \langle w^l, g(x) \rangle$$

obtendo  $x(w^l)$ .

**Passo 2.** Resolva o problema de Programação Linear nas variáveis  $w$  e  $\sigma$ :

$$\begin{aligned} P_L : & \text{minimizar} && \sigma \\ & \text{sujeito a} && \sigma \geq \langle w, g(x(w^i)) - z \rangle, \quad i = 1, 2, \dots, l, \\ & && w \in \mathcal{W}, \sigma \in \mathbb{R}, \end{aligned} \tag{4.3.10}$$

obtendo  $\sigma^{l+1}, w^{l+1}$  e  $\phi_z(w^{l+1})$ . Se  $\phi_z(w^{l+1}) - \sigma^{l+1} < \epsilon_1$ , sendo  $\epsilon_1 > 0$  uma pequena tolerância, declare  $\theta(z) := \sigma^{l+1}$  e pare. Caso contrário, faça  $l \leftarrow l + 1$  e retorne ao **Passo 1**.

### 4.3.1 Problema Mestre — Algoritmo Global

A seguir, propomos duas técnicas de otimização para resolver globalmente os problemas de programação multiplicativa generalizada ( $Q_{MG}$ ). Primeiramente, apresenta-se uma modificação do algoritmo proposto em [OF10] para resolver ( $Q_{MG}$ ). Em seguida, um algoritmo *branch-and-bound* voltado à resolução dos problemas ( $P_{MG}$ ) e ( $Q_{MG}$ ) será apresentado.

Considere o politopo inicial

$$\mathcal{G}^0 := \{z \in \mathbb{R}^m : 0 < \underline{z} \leq z \leq \bar{z}\}, \quad (4.3.11)$$

em que  $\underline{z}$  e  $\bar{z}$  são limitantes inferior e superior para  $z$  em  $\Omega$ , respectivamente. Os valores  $\underline{z}$  e  $\bar{z}$  podem ser definidos como

$$\underline{z}_i = \min_{x \in \Omega} g_i(x), \quad \bar{z}_i = \max_{x \in \Omega} g_i(x), \quad i = 1, 2, \dots, m,$$

por meio da resolução de  $m$  minimizações côncavas e  $m$  maximizações côncavas, respectivamente. Na prática,  $\underline{z}$  pode ser tomado como um vetor positivo arbitrariamente pequeno.

Inicialmente, o máximo de  $u$  (expressão (4.2.1)) em  $\mathcal{G}^0$  é alcançado em  $z^0 = \bar{z}$ . O ponto utópico  $z^0$  raramente satisfaz o sistema de desigualdades (4.3.1), ou seja  $\theta(z^0) < 0$ , em geral. Seja  $w^0 \in \mathcal{W}$  um minimizador ótimo correspondente em (4.3.3). Verificamos então que  $z^0$  não pertence ao semi-espço suporte

$$\mathcal{H}_-^0 = \{z \in \mathbb{R}^m : \langle w^0, z \rangle \leq \langle w^0, g(x(w^0)) \rangle\}. \quad (4.3.12)$$

Uma aproximação externa melhorada para  $\mathcal{G}$  é  $\mathcal{G}^1 = \mathcal{H}_-^0 \cap \mathcal{G}^0$ . Se  $z^1$  que maximiza  $u$  em  $\mathcal{G}^1$  também é tal que  $\theta(z^1) < 0$ , então um novo semi-espço suporte  $\mathcal{H}_-^1$  é determinado, e a região viável de (4.3.2) é aproximada de uma forma melhor por  $\mathcal{G}^2 = \mathcal{G}^1 \cap \mathcal{H}_-^1$ . Repetindo-se este procedimento, em uma iteração arbitrária  $k$  do algoritmo de relaxação, o seguinte problema relaxado é resolvido:

$$P_{\mathcal{G}^k} : \max_{z \in \mathcal{G}^k} u(z).$$

Numa iteração genérica  $k$  do procedimento iterativo, após a geração de  $k$  hiperplanos de corte, o problema se resume na obtenção de uma solução ótima global do problema ( $P_{\mathcal{G}^k}$ ), que pode ser reformulado como o Problema Quadrático Indefinido

$$\begin{aligned} &\text{maximizar} && u(z) = \frac{1}{2}z^T Qz + c^T z \\ &\text{sujeito a} && A^{(k)}z \leq b^{(k)}, \\ &&& \underline{z} \leq z \leq \bar{z}, \end{aligned} \quad (4.3.13)$$

em que  $A^{(k)} \in \mathbb{R}^{k \times m}$ ,  $b^{(k)} \in \mathbb{R}^k$ ,  $\underline{z} \in \mathbb{R}^m$  e  $\bar{z} \in \mathbb{R}^m$  caracterizam a representação matricial do problema ( $P_{\mathcal{G}^k}$ ).

### 4.3.2 Resolução de $(P_{\mathcal{G}^k})$ por Enumeração das Restrições

No problema (4.3.13), a ser resolvido em uma iteração arbitrária  $k$  do algoritmo,

$$Q = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}_{(2r+1) \times (2r+1)}, \quad c^T := \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (4.3.14)$$

e  $A^{(k)} \in \mathbb{R}^{k \times m}$ ,  $b^{(k)} \in \mathbb{R}^k$ ,  $\underline{z} \in \mathbb{R}^m$  e  $\bar{z} \in \mathbb{R}^m$  caracterizam a representação matricial do problema  $(P_{\mathcal{G}^k})$ .

O polinômio característico de  $Q$

$$\det(\lambda I - Q) = \lambda \underbrace{(\lambda^2 - 1) \cdots (\lambda^2 - 1)}_{r \text{ vezes}} = 0,$$

possui exatamente  $r$  autovalores positivos iguais a 1,  $r$  autovalores negativos iguais a  $-1$ , e um autovalor igual a zero. Portanto, a matriz  $Q$  é uma matriz indefinida, e por consequência  $(P_{\mathcal{G}^k})$  representa um problema de Programação Quadrática Indefinida.

Considere um problema de otimização restrita na forma geral

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in \Omega := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}, \end{aligned} \quad (4.3.15)$$

na qual  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ , e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ . A função Lagrangiana associada a (4.3.15) é dada por  $\mathcal{L}(x, \lambda) := f(x) + g(x)^T \mu + h(x)^T \lambda$ . Do Capítulo 2 sabe-se que se  $x^*$  é um minimizador local de (4.3.15) e satisfaz qualquer uma das condições de regularidade das restrições, então existem  $\lambda^* \in \mathbb{R}^l$  e  $\mu^* \in \mathbb{R}_+^m$  tais que  $\nabla \mathcal{L}_x(x^*, \lambda^*, \mu^*) = 0$ , e

$$d^T \nabla_x^2 \mathcal{L}(x^*, \lambda^*, \mu^*) d \geq 0, \quad \text{sempre que} \quad \nabla g_{\mathcal{I}(x^*)}(x) d = 0, \quad \text{e} \quad \nabla h(x) d = 0, \quad (4.3.16)$$

onde  $g_{\mathcal{I}(x^*)}$  denota as componentes de  $g$  ativas em  $x^*$ . Note-se que uma condição de regularidade possível é que os gradientes das restrições ativas avaliados em  $x^*$  sejam linearmente independentes.

O Teorema estabelece que a Hessiana da função Lagrangeana possui  $s$  autovalores negativos em um minimizador local que satisfaça as condições necessárias de otimalidade de segunda ordem.

**Teorema 4.3.7.** [HPRS91] *Sejam  $f, g$  e  $h$  funções contínuas e duas vezes diferenciáveis em um ponto  $x$ . Suponha que a Hessiana  $\nabla_x^2 \mathcal{L}(x, \lambda, \mu)$  possua  $s$  autovalores negativos (contando multiplicidades), e que (4.3.16) se verifica. Então a dimensão do espaço gerado pelos gradientes das restrições ativas é no mínimo  $s$ . Conseqüentemente, pelo menos  $s$  restrições de (4.3.15) estão ativas em  $x$ .*

**Proposição 4.3.8.** *Pelo menos  $r$  restrições estão ativas em qualquer solução ótima global (local) de  $(P_{\mathcal{G}^k})$ .*

*Demonstração.* A prova é uma conseqüência direta do Teorema 4.3.7. □

Pelos resultados anteriores, sabe-se agora que uma solução ótima de  $(P_{\mathcal{G}^k})$  ocorre na fronteira de  $\mathcal{G}^k$  e pode ser encontrada por *enumeração de restrições* [HPT95]. O número de restrições de  $(P_{\mathcal{G}^k})$  é  $2m + k$ , sendo  $m = 2r + 1$ , o número de funções que compõem a função objetivo de  $(Q_{MG})$ , e  $k$  o número corrente de semi-espacos gerados pelo algoritmo.

Pelas condições de Karush–Kuhn–Tucker, qualquer solução ótima do problema  $(P_{\mathcal{G}^k})$  deve satisfazer as seguintes condições:

$$\begin{aligned} Hz &\leq h, \\ Qz + c - H^T \mu &= 0, \\ \mu^T (Hz - h) &= 0, \\ \mu &\geq 0, \end{aligned}$$

sendo

$$H := \begin{bmatrix} A^{(k)} \\ \dots \\ I \\ \dots \\ -I \end{bmatrix}, \quad H \in \mathbb{R}^{(k+2m) \times m}, \quad \text{e} \quad h := \begin{bmatrix} b^{(k)} \\ \dots \\ \bar{z} \\ \dots \\ -\underline{z} \end{bmatrix}, \quad h \in \mathbb{R}^{k+2m}.$$

Como uma solução ótima do problema  $(P_{\mathcal{G}^k})$  ocorre na fronteira da região viável do problema,  $\{z \in \mathbb{R}^m : Hz \leq h\}$ , existe um conjunto de índices  $\mathcal{I} \subseteq \{1, 2, \dots, k + 2m\}$  tal

que  $H_{\mathcal{I}}z \leq h_{\mathcal{I}}$ , onde  $(H_{\mathcal{I}}, h_{\mathcal{I}})$  representa o subconjunto de linhas de  $(H, h)$  indexado por  $\mathcal{I}$ . Os multiplicadores de Lagrange correspondentes a restrições inativas são iguais a zero. Portanto, as condições KKT reduzem-se ao sistema

$$\begin{aligned} H_{\mathcal{I}}z &= h_{\mathcal{I}}, \\ Qz + c - H^T\mu &= 0, \\ \mu_{\bar{\mathcal{I}}} &= 0, \end{aligned}$$

sendo  $\bar{\mathcal{I}} = \{1, 2, \dots, k + 2m\} \setminus \mathcal{I}$ , o conjunto complementar de  $\mathcal{I}$ .

Uma abordagem adequada para resolver problemas do tipo  $(P_{\mathcal{G}^k})$  consiste em enumerar todos os possíveis conjuntos  $\mathcal{I}$ , resolver os sistemas lineares associados,  $H_{\mathcal{I}}z = h_{\mathcal{I}}$ , checar se a solução  $z \in \mathbb{R}^m$  é viável para  $(P_{\mathcal{G}^k})$  e se  $\mu \geq 0$ , e então, dentre as soluções viáveis encontradas, determinar aquela que fornece o maior valor para a função objetivo de  $(P_{\mathcal{G}^k})$ .

Para cada escolha possível de  $\mathcal{I}$ , encontrar as soluções  $(z, \mu)$  do sistema associado, e checar se a solução é viável e se  $\mu \geq 0$ , exige  $\mathcal{O}(2^{k+2m}(k+3m)^3)$  operações, uma vez que existem  $2^{k+2m}$  maneiras de se selecionar subconjuntos de  $\{1, 2, \dots, k + 2m\}$  e cada sistema de equações é resolvido e, no máximo  $\mathcal{O}((k+3m)^3)$  operações aritméticas [HPT95]. Vale lembrar que como as restrições de limitação  $\underline{z}_i \leq z_i$  e  $z_i \leq \bar{z}_i$  não podem estar simultaneamente ativas se os limites forem diferentes, o número de subconjuntos a serem explorados é significativamente menor.

Suponha que, numa iteração qualquer  $k$ , o problema  $(P_{\mathcal{G}^k})$  é resolvido por enumeração de restrições, obtendo-se a solução ótima global  $z^k$ . O Algoritmo MIN-MAX testa  $z^k$ , retornando com uma nova restrição de desigualdade,  $z^k \notin \mathcal{G}^k$ . Na iteração  $k + 1$  deve-se trabalhar com o conjunto de restrições

$$\{1, 2, \dots, k + 2m\} \cup \{\text{nova restrição}\},$$

mas subconjuntos de  $\{1, 2, \dots, k + 2m\}$  não precisam ser explorados, porque levariam à mesma solução  $z^k$  anterior, que se revelou inviável; é suficiente explorar as combinações de  $\{1, 2, \dots, k + 2m\}$  com a nova restrição. Se o algoritmo global convergir após  $l$  iterações, este procedimento terá evitado  $2^l$  combinações desnecessárias [OF10].

O Algoritmo QMG-ENU a seguir resolve globalmente o problema de programação multiplicativa generalizada côncava  $(Q_{MG})$  por meio de aproximação externa. As propriedades de convergência infinita e finita do Algoritmo QMG-ENU são semelhantes às exibidas pelo

algoritmo de [OF10], o algoritmo proposto para resolver o problema de programação multiplicativa generalizada do tipo  $(P_{MG})$ .

**Algoritmo QMG-ENU:** Resolução de  $(Q_{MG})$  por meio de Enumeração de Restrições

**Passo 0:** Defina  $\mathcal{G}^0$  e faça  $k := 0$ ;

**Passo 1:** Resolva o problema de programação generalizada côncava

$$P_{\mathcal{G}^k} : \max_{z \in \mathcal{G}^k} \frac{1}{2} z^T Q z + c^T z,$$

por meio de Enumeração de Restrições, obtendo  $z^k$ ;

**Passo 2:** Encontre  $\theta(z^k)$  resolvendo o subproblema minmax (4.3.3)–(4.3.4). Se  $-\theta(z^k) < \epsilon_{rel}$ , sendo  $\epsilon_{rel} > 0$  é uma tolerância pré-especificada, pare:  $z^k$  e  $x^k$  são soluções  $\epsilon_{rel}$ -ótimas de (4.3.2) and  $(Q_{MG})$ , respectivamente. Caso contrário, defina

$$\mathcal{G}^{k+1} := \{z \in \mathcal{G}^k : \langle w^k, z \rangle \leq \langle w^k, g(x^k) \rangle\},$$

faça  $k \leftarrow k + 1$  e volte ao **Passo 1**.

**Teorema 4.3.9.** *Qualquer ponto de acumulação  $z^*$  da seqüência gerada pelo Algoritmo QMG-ENU é uma solução ótima do problema de programação convexa multiplicativa generalizada  $(Q_{MG})$ .*

O teste de convergência no Passo 2 do Algoritmo QMG-ENU é baseado na interpretação do Teorema 4.3.5. A tolerância  $\epsilon_{rel}$  é uma medida da violação da restrição  $z \in \mathcal{G}$ .

Na seção a seguir propomos uma abordagem alternativa para os problemas  $(P_{MG})$  e  $(Q_{MG})$ . Embora baseada em diferentes argumentos, a nova abordagem resulta em um algoritmo global similar ao algoritmo QMG-ENU, com uma resolução do problema-mestre por meio de um algoritmo *branch-and-bound*.

## 4.4 Programação Multiplicativa Generalizada Convexa

Considere o problema  $(P_{MG})$

$$P_{MG} : \min_{x \in \Omega} \left[ f_1(x) + \sum_{i=1}^r f_{2i}(x) f_{2i+1}(x) \right].$$

Começamos lembrando que a função objetivo de  $(P_{MG})$  pode ser vista como uma função de agregação especial para o problema de minimizar a função objetivo vetorial  $f := (f_1, f_2, \dots, f_m)$  em  $\Omega$ . Uma solução  $x^* \in \Omega$  é uma solução eficiente para este problema se não existe outra solução  $x \in \Omega$  tal que  $f(x) \leq f(x^*)$  e  $f(x) \neq f(x^*)$ . A positividade de  $f$  sobre  $\Omega$  implica que a função de objetivo de  $(P_{MG})$  é uma função monótona estritamente crescente em cada uma das componentes de  $f$ , e, em consequência, uma solução ótima de  $(P_{MG})$  deve ser eficiente. O Teorema 4.4.1 a seguir é uma consequência imediata da interpretação do problema  $(P_{MG})$  com um problema de programação multiobjetivo.

**Teorema 4.4.1.** *Seja  $x^* \in \Omega$  qualquer solução ótima do problema de  $(P_{MG})$  e considere o simplex  $m$ -dimensional*

$$w \in \mathcal{W} := \left\{ w \in \mathbb{R}_+^m : \sum_{i=1}^m w_i = 1 \right\}.$$

*Então existe algum  $w \in \mathcal{W}$  (em geral, dependente da solução ótima  $x^*$ ) tal que  $x^*$  também resolve o problema convexo*

$$\left| \begin{array}{ll} \underset{x}{\text{minimizar}} & \sum_{i=1}^m w_i f_i(x) \\ \text{sujeito a} & x \in \Omega. \end{array} \right. \quad (4.4.1)$$

A caracterização das soluções ótimas de  $(P_{MG})$  como soluções eficientes da sua versão multiobjetivo tem sido explorada na literatura [Ben99], [OFo8]. Tal caracterização está também presente na abordagem proposta nesta seção.

#### 4.4.1 Projeção no Espaço-Imagem

Sejam  $y^L$  e  $y^U$  vetores  $2m$ -dimensionais ( $m = 2r + 1$ ) tais que  $y^L \leq f(x) \leq y^U$  para todo  $x \in \Omega$ , e considere o politopo inicial

$$\mathcal{Y}^0 := \left\{ y \in \mathbb{R}^m : 0 < y^L \leq y \leq y^U \right\}. \quad (4.4.2)$$

Vetores  $y^L$  e  $y^U$  podem ser definidos como

$$y_i^L = \min_{x \in \Omega} f_i(x), \quad y_i^U = \max_{x \in \Omega} f_i(x), \quad i = 1, 2, \dots, m.$$

Ambos os vectores são facilmente obtidos quando  $f$  é uma função vetorial linear. No

caso não-linear geral, encontrar  $y^L$  envolve a resolução  $m$  problemas convexos. Um procedimento prático para obter  $y^U$  tal que  $f(x) \leq y^U$  para todo  $x \in \Omega$  é sugerido em [Ben99].

Considere o problema de otimização alternativo

$$\left| \begin{array}{l} \text{minimizar}_{x,y} \quad y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{sujeito a} \quad f(x) \leq y, \\ \quad \quad \quad x \in \Omega, y \in \mathcal{Y}^0. \end{array} \right. \quad (4.4.3)$$

Tendo em vista a definição de  $\mathcal{Y}^0$ , o problema  $(P_{MG})$  pode ser reescrito como

$$\left| \begin{array}{l} \text{minimizar}_{x,y} \quad y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{sujeito a} \quad f(x) = y, \\ \quad \quad \quad x \in \Omega, y \in \mathcal{Y}^0, \end{array} \right.$$

o que mostra que o problema (4.4.3) é uma relaxação do problema  $(P_{MG})$ . Como a função de objetivo comum de  $(P_{MG})$  e (4.4.3) é uma função monótona estritamente crescente, qualquer solução ótima  $(x^*, y^*)$  de (4.4.3) deve ser tal que  $f(x^*) = y^*$ , e então  $x^*$  também resolveria o problema  $(P_{MG})$ .

O problema (4.4.3) tem  $n + m$  variáveis de decisão, uma soma de produtos elementares como função objetivo e uma região viável conjuntamente convexa nas variáveis  $x$  e  $y$ . Uma representação alternativa da sua região viável leva a um problema equivalente com apenas  $m$  variáveis.

A projeção do problema (4.4.3) em  $\mathbb{R}^m$  é

$$\left| \begin{array}{l} \text{minimizar}_y \quad v(y) \\ \text{sujeito to} \quad y \in \mathcal{V} \cap \mathcal{Y}^0, \end{array} \right. \quad (4.4.4)$$

onde

$$\begin{aligned} v(y) := \min_x \quad & y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{sujeito a} \quad & f(x) \leq y, \\ & x \in \Omega, \end{aligned} \quad (4.4.5)$$

e

$$\mathcal{V} := \left\{ (y) \in \mathbb{R}^m : f(x) \leq y, \text{ para algum } x \in \Omega \right\}. \quad (4.4.6)$$

Adotando a convenção de que o valor do problema (4.4.5) é  $+\infty$  se  $y \notin \mathcal{V}$  (isto é, quando o problema 4.4.5 é inviável), obtemos

$$v(y) = \begin{cases} y_1 + \sum_{i=1}^r y_{2i} y_{2i+1}, & \text{se } y \in \mathcal{V}, \\ +\infty, & \text{caso contrário,} \end{cases}$$

isto é, o valor de  $v$  é definido pelo *existência* de um  $x \in \Omega$  que satisfaz o sistema de desigualdades  $f(x) \leq y$ .

O problema (4.4.4) é, por sua vez, equivalente a

$$\left| \begin{array}{l} \text{minimizar}_y \quad y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{sujeito a} \quad \sum_{i=1}^m w_i y_i \geq \min_{x \in \Omega} \sum_{i=1}^m w_i f_i(x) \quad \text{para todo } w \in \mathcal{W}, \\ \quad \quad \quad y \in \mathcal{Y}^0. \end{array} \right. \quad (4.4.7)$$

no qual o conjunto  $\mathcal{V}$  é substituído por sua representação como um sistema de desigualdades lineares semi-infinito [Geo70], [OFo8], [OF10]. O problema (4.4.7) pertence à categoria de problemas de otimização semi-infinita linearmente restrito. Adotando a mesma estratégia da seção anterior, expressamos a restrição semi-infinita como uma única restrição,

$$\theta(y) \leq 0, \quad (4.4.8)$$

onde

$$\theta(y) := \max_{w \in \mathcal{W}} \psi_y(w) \quad (4.4.9)$$

e

$$\psi_y(w) := \min_{x \in \Omega} \left\{ \sum_{i=1}^m w_i (f_i(x) - y_i) \right\}. \quad (4.4.10)$$

Um ponto  $y \in \mathbb{R}^m$  satisfaz o sistema de infinitas desigualdades do problema (4.4.7) se e somente se  $\theta(y) \leq 0$ .

A função  $\psi_y$  é côncava em  $\mathcal{W}$ , e portanto  $\theta(y)$  em (4.4.8) é calculado por meio da

resolução de um problema côncavo. Além disso,  $\theta$  é convexa (e portanto, contínua) em  $\mathbb{R}^m$ .

por meio da aplicação da dualidade Lagrangeana, demonstra-se (como no Teorema 4.3.5) que  $\theta(y)$  é o valor ótimo do problema convexo

$$\left| \begin{array}{l} \text{minimizar}_{x,\sigma} \quad \sigma \\ \text{sujeito a} \quad f(x) \leq \sigma e + y, \\ \quad \quad \quad x \in \Omega, \end{array} \right. \quad (4.4.11)$$

no qual  $e \in \mathbb{R}^m$  é o vetor  $m$ -dimensional de uns e  $\sigma \in \mathbb{R}$ .

No caso de violação ( $\theta(y) > 0$ ), cada desigualdade em  $f(x) \leq y$  será violada por no máximo  $\theta(y)$ . A maximização de  $\phi_y$  em  $\mathcal{W}$  é realizada pelo algoritmo MIN-MAX discutido na Seção 4.3, reescrito como um algoritmo MIN-MAX.

#### 4.4.2 Algoritmo de Relaxação

Na iteração  $k$  do algoritmo, o problema relaxado assume a forma

$$\left| \begin{array}{l} \text{minimizar}_y \quad y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{sujeito a} \quad \sum_{i=1}^m w_i^s y_i \geq \sum_{i=1}^m w_i^s f_i(x^s), \quad s = 0, 1, \dots, k, \\ \quad \quad \quad y \in \mathcal{Y}^0, \end{array} \right. \quad (4.4.12)$$

na qual  $(x^s, w^s)$ ,  $s = 0, 1, \dots, k$ , são as soluções ótimas obtidas pelo algoritmo MIN-MAX que encontra  $\theta(y^s)$ ,  $s = 0, 1, \dots, k$ . Seja  $y^{k+1}$  uma solução ótima do problema relaxado (4.4.12). Se  $\theta(y^{k+1}) \leq 0$ , então  $y^{k+1}$  é também uma solução ótima do problema semi-infinito (4.4.7). Caso contrário, ( $\theta(y^{k+1}) > 0$ ), pelo menos a seguinte restrição de desigualdade é violada por  $y^{k+1}$ :

$$\sum_{i=1}^m w_i^{k+1} y_i \geq \sum_{i=1}^m w_i^{k+1} f_i(x^{k+1}). \quad (4.4.13)$$

Entre todas as desigualdades lineares que constituem a restrição semi-infinita, (4.4.13) é a mais violada por  $y^{k+1}$ . Em termos geométricos, tal desigualdade produz um *corte mais profundo* no retângulo  $\mathcal{Y}^0$ . Isto é, o semi-espaco (4.4.13) não contém  $y^{k+1}$  e a sua intersecção com  $\mathcal{Y}^0$  fornece uma aproximação externa melhorada de  $\mathcal{Y}^0$  pois, ao remover  $y^{k+1}$ , uma

quantidade máxima de  $\mathcal{Y}^0$  é também removida pelo plano de corte definido por

$$\sum_{i=1}^m w_i^{k+1} y_i = \sum_{i=1}^m w_i^{k+1} f_i(x^{k+1}). \quad (4.4.14)$$

A desigualdade (4.4.13) é incorporada ao conjunto de restrições do problema (4.4.12) e repete-se o procedimento. A prova de convergência infinita do algoritmo de relaxação descrito acima para um minimizador global de (4.4.7) é essencialmente a mesma fornecida no Capítulo 3 para a convergência do Algoritmo Básico.

### 4.4.3 Algoritmo *Branch-and-Bound* para Resolução de $(P_{MG})$

O problema relaxado (4.4.12) possui  $m$  variáveis,  $k + 2m$  restrições lineares e uma soma de produtos elementares como função objetivo a ser minimizada. Algoritmos *branch-and-bound* são técnicas de otimização global voltadas para problemas não-convexos [LW66] e [Moog1]. São algoritmos determinísticos, no sentido de que operam com limites superiores e inferiores para o valor ótimo do problema, e terminam com um certificado que atesta que a solução encontrada é  $\epsilon$ -sub-ótima. Por outro lado, algoritmos *branch-and-bound* podem ser (e muitas vezes são) de lenta convergência, e no pior caso exigem um esforço computacional que cresce exponencialmente com o tamanho do problema.

Uma estratégia bem sucedida em otimização global para lidar com produtos elementares tem sido substituir cada ocorrência de um produto elementar pelo seu envelope convexo (envelope côncavo, no caso de maximização), definida como a função convexa que melhor subestima o produto em um dado intervalo fechado. Certas classes de problemas não-convexos podem ser abordados por meio de técnicas de otimização convexa quando se utiliza o conceito de envelope convexo de uma função.

**Definição 4.4.2** (Envelope Convexo). *A função convexa  $\phi$  que melhor se ajusta a uma dada função não-convexa numa região retângular  $\mathcal{R} \subseteq \mathbb{R}^n$  é chamada de envelope convexo de  $f$ , e deve satisfazer as seguintes propriedades:*

- (i)  $\phi$  é convexa em  $\mathcal{R}$ ;
- (ii)  $f(x) \geq \phi(x)$  para todo  $x \in \mathcal{R}$ ;
- (iii) Para todas as funções  $g$  que satisfazem (i) e (ii),  $\phi(x) \geq g(x)$  para todo  $x \in \mathcal{R}$ .

Infelizmente, não existe nenhum método geral para determinar o envelope convexo de uma função qualquer em uma dada região [MF95].

Produtos de funções monovariáveis  $f_i : \mathbb{R} \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, N$ ,

$$p(x) := \prod_{i=1}^N f_i(x_i), \quad p(x) : \mathbb{R}^n \rightarrow \mathbb{R},$$

são funções não-convexas, em geral, mesmo quando as funções monovariáveis  $f_i$  são convexas. Um grande número de não-linearidades presentes em problemas de Engenharia pode ser descrito como produtos de funções monovariáveis.

**Teorema 4.4.3.** *Sejam  $f$  e  $g$  funções monovariáveis duas vezes diferenciáveis definidas dentro do retângulo  $[x^L, x^U] \times [y^L, y^U]$  e*

$$l(x, y) := \max \left\{ \phi(f^L g(y)) + \phi(g^L f(x)) - f^L g^L, \phi(f^U g(y)) + \phi(g^U f(x)) - f^U g^U \right\},$$

onde

$$\begin{aligned} f^L &= \inf_{x^L \leq x \leq x^U} f(x), \\ f^U &= \sup_{x^L \leq x \leq x^U} f(x), \\ g^L &= \inf_{y^L \leq y \leq y^U} g(y), \\ g^U &= \sup_{y^L \leq y \leq y^U} g(y), \end{aligned}$$

e  $\phi(f^L g(y))$ ,  $\phi(g^L f(x))$ ,  $\phi(f^U g(y))$  e  $\phi(g^U f(x))$  são os envelopes convexos das funções monovariáveis  $f^L g(y)$ ,  $g^L f(x)$ ,  $f^U g(y)$ , e  $g^U f(x)$ , respectivamente. Então

- (i)  $l(x, y)$  é convexa em  $[x^L, x^U] \times [y^L, y^U]$ ;
- (ii)  $f(x)g(y) \geq l(x, y)$ ,  $\forall (x, y) \in [x^L, x^U] \times [y^L, y^U]$ .

*Demonstração.* Veja [MF95]. □

A partir do Teorema 4.4.3, a função  $l(x, y)$  pode ser utilizada como um envelope convexo do produto de duas funções duas vezes diferenciáveis. É possível mostrar que sob certas condições  $l(x, y)$  corresponde ao envelope convexo do produto  $f(x)g(y)$ . Estas condições são estabelecidas a seguir.

**Teorema 4.4.4.** *Se (i) as funções monovariáveis  $g^L f(x)$ ,  $g^U f(x)$  e  $f^L g(y)$ ,  $f^U g(y)$  são côncavas em  $[x^L, x^U]$  e  $[y^L, y^U]$ , respectivamente, e (ii) as funções  $f(x)$  e  $g(y)$  são monotônicas, então  $l(x, y)$*

é um envelope convexo de  $f(x)g(y)$ , ou seja

$$\phi(f(x)g(y)) = l(x, y), \quad \forall (x, y) \in [x^L, x^U] \times [y^L, y^U].$$

*Demonstração.* Veja [MF95]. □

Pelos Teoremas 4.4.3 e 4.4.4, os envelopes convexos e côncavos para os termos bilineares  $y_{2i}y_{2i+1}$  do problema (4.4.7) podem ser representados, a partir das definições  $f(y_{2i}) := y_{2i}$ ,  $g(y_{2i+1}) := y_{2i+1}$  e  $f(y_{2i}) := -y_{2i}$ ,  $g(y_{2i+1}) := y_{2i+1}$

$$y_{2i}y_{2i+1} \geq \max \left\{ y_{2i}^L y_{2i+1} + y_{2i+1}^L y_{2i} - y_{2i}^L y_{2i+1}^L, \right. \\ \left. y_{2i}^U y_{2i+1} + y_{2i+1}^U y_{2i} - y_{2i}^U y_{2i+1}^U \right\}, \quad (4.4.15)$$

ou

$$-y_{2i}y_{2i+1} \geq \max \left\{ -y_{2i}^U y_{2i+1} - y_{2i+1}^L y_{2i} + y_{2i}^U y_{2i+1}^L, \right. \\ \left. -y_{2i}^L y_{2i+1} - y_{2i+1}^U y_{2i} + y_{2i}^L y_{2i+1}^U \right\}, \quad (4.4.16)$$

para todo  $[y_{2i}^L, y_{2i}^U] \times [y_{2i+1}^L, y_{2i+1}^U]$ . Os lados direitos de (4.4.15) e (4.4.16) são os envelopes convexo e côncavo de  $y_{2i}y_{2i+1}$  em  $[y_{2i}^L, y_{2i}^U] \times [y_{2i+1}^L, y_{2i+1}^U]$ .

Substituindo cada produto elementar que aparece na função objetivo de (4.4.7) pelo seu envelope convexo, obtemos o seguinte problema de otimização linear semi-infinito:

$$\left| \begin{array}{l} \text{minimizar}_{y, \lambda} \quad \lambda_1 + \sum_{i=1}^r \lambda_{i+1} \\ \text{sujeito a} \quad \sum_{i=1}^m w_i y_i \geq \min_{x \in \Omega} \sum_{i=1}^m w_i f_i(x) \quad \text{para todo } w \in \mathcal{W}, \\ \lambda_1 \geq y_1, \\ \lambda_{i+1} \geq y_{2i}^L y_{2i+1} + y_{2i+1}^L y_{2i} - y_{2i}^L y_{2i+1}^L, \quad i = 1, 2, \dots, r, \\ \lambda_{i+1} \geq y_{2i}^U y_{2i+1} + y_{2i+1}^U y_{2i} - y_{2i}^U y_{2i+1}^U, \quad i = 1, 2, \dots, r, \\ y \in \mathcal{Y}^0. \end{array} \right. \quad (4.4.17)$$

Devido à subestimação de produtos introduzidos pelos envelopes convexos, o valor ótimo do problema (4.4.17) é menor ou igual ao valor ótimo do problema (4.4.7), que por sua vez, é igual ao valor ótimo do problema original ( $P_{MG}$ ). Sendo um subestimador do

problema (4.4.7), o problema (4.4.17) permite a resolução do problema (4.4.7) por meio de um algoritmo *branch-and-bound* [HPT95].

A seguir, desenvolvemos um algoritmo *branch-and-bound* para resolver globalmente o problema (4.4.7) (e, portanto, o problema  $(P_{MG})$ ), por meio do problema (4.4.17). Em termos gerais, a região viável do problema (4.4.17) é sucessivamente subdividida em partes mais refinadas (*branching*), sobre as quais os limitantes superiores e inferiores do valor ótimo do problema (4.4.17) são determinados (*bounding*). Partes da região viável com limitantes superiores menores do que o melhor limitante inferior encontrado até a iteração corrente do algoritmo são eliminadas de considerações posteriores (*pruning*), pois estas partes não podem conter uma solução ótima do problema (4.4.17).

Em algoritmos *branch-and-bound* retangulares, a região viável do problema é dividida em sub-retângulos. Um sub-retângulo  $\mathcal{Y}$  de  $\mathcal{Y}^0$  é um conjunto da forma (4.4.2) e limitantes  $y^L(\mathcal{Y})$  e  $y^U(\mathcal{Y})$  satisfazendo  $y^L(\mathcal{Y}) \geq y^L$  e  $y^U(\mathcal{Y}) \leq y^U$ , com o entendimento de que  $y^L(\mathcal{Y}^0) = y^L$  e  $y^U(\mathcal{Y}^0) = y^U$ .

Por conveniência, os problemas (4.4.17) e (4.4.7) serão também referidos como problemas  $(P_L)$  e  $(P_S)$ , respectivamente. Sejam  $P_L(\mathcal{Y})$  e  $P_S(\mathcal{Y})$  problemas das formas  $(P_L)$  e  $(P_S)$  quando  $\mathcal{Y}^0$  é substituído por  $\mathcal{Y}$ . Seja  $(y(\mathcal{Y}), \lambda(\mathcal{Y}))$  qualquer solução ótima de  $(P_L(\mathcal{Y}))$  determinada pelo algoritmo de relaxação. O valor ótimo de  $(P_S(\mathcal{Y}))$  é menor ou igual ao limitante superior  $\mu(\mathcal{Y}) := y_1(\mathcal{Y}) + \sum_{i=1}^r y_{2i}(\mathcal{Y})y_{2i+1}(\mathcal{Y})$ , porque  $y(\mathcal{Y})$  é viável para  $(P_S(\mathcal{Y}))$ , e maior ou igual ao limitante inferior  $\gamma(\mathcal{Y}) := \lambda_1(\mathcal{Y}) + \sum_{i=1}^r \lambda_{i+1}(\mathcal{Y})$ , porque  $(P_L(\mathcal{Y}))$  é uma subestimação de  $(P_S(\mathcal{Y}))$ .

No algoritmo *branch-and-bound* apresentado a seguir, dois novos sub-retângulos são produzidos seccionando-se um dado retângulo no ponto médio de um dos seus lados.

**Algorithm** PMG-BB: Algoritmo *branch-and-bound* para Resolver  $(P_{MG})$

**Passo 0:** Resolva o problema  $(P_L(\mathcal{Y}^0))$ , obtendo uma solução ótima  $(y(\mathcal{Y}^0), \lambda(\mathcal{Y}^0))$ . Faça  $\mu^0 := \mu(\mathcal{Y}^0)$ ,  $\gamma^0 := \gamma(\mathcal{Y}^0)$ ,  $\mathcal{L}^0 := \{\mathcal{Y}^0\}$  e  $q = 0$ ;

**Passo 1:** Se  $\mu^q = \gamma^q$ , então pare. A solução incumbente  $y^q$  é uma solução ótima para  $(P_S)$ ;

**Passo 2:** Encontre  $\mathcal{Y} \in \mathcal{L}^q$  tal que  $\gamma(\mathcal{Y}) = \gamma^q$ . Divida  $\mathcal{Y}$  em sub-retângulos  $\mathcal{Y}^I$  e  $\mathcal{Y}^{II}$  e defina

$$\mathcal{L}^{q+1} := (\mathcal{L}^q \setminus \{\mathcal{Y}\}) \cup \{\mathcal{Y}^I, \mathcal{Y}^{II}\}.$$

Compute  $\mu(\mathcal{Y}^I)$ ,  $\gamma(\mathcal{Y}^I)$ ,  $\mu(\mathcal{Y}^{II})$  e  $\gamma(\mathcal{Y}^{II})$ . Elimine todos os sub-retângulos  $\mathcal{Y} \in \mathcal{L}^{q+1}$  tais que  $(P_L(\mathcal{Y}))$  seja inviável ou  $\mu(\mathcal{Y}) > \gamma^q$ ;

**Passo 3:** Encontre  $\mathcal{Y}^* \in \arg \min_{\mathcal{Y} \in \mathcal{L}^{q+1}} \mu(\mathcal{Y})$  e faça  $y^{q+1} := y(\mathcal{Y}^*)$ ,  $\mu^{q+1} := \mu(\mathcal{Y}^*)$ ,  $\gamma^{q+1} := \min_{\mathcal{Y} \in \mathcal{L}^{q+1}} \gamma(\mathcal{Y})$ ,  $q := q + 1$  e volte ao **Passo 1**.

As regras de branching, bounding e pruning do Algoritmo PMG-BB preenchem todas as condições necessárias para a convergência de algoritmos *branch-and-bound* [HPT95]. Assim, qualquer ponto de acumulação  $y^*$  da sequência  $\{y^q\}$  gerada pelo Algoritmo PMG-BB resolve o problema  $(P_S)$ . Resultados de convergência também garantem que para  $q$  suficientemente grande,

$$\mu^q - \gamma^q \leq \epsilon_{bb},$$

no qual  $\epsilon_{bb} > 0$  é qualquer tolerância (geralmente pequena) para a convergência finita do Algoritmo PMG-BB.

Um aspecto muito importante da implementação de algoritmo *branch-and-bound* por meio do algoritmo de relaxação é a propagação e reutilização de cortes mais profundos por meio da árvore *branch-and-bound*.

Quando  $\mathcal{Y}$  é dividido em  $\mathcal{Y}^I$  e  $\mathcal{Y}^{II}$ , os problemas  $P_L(\mathcal{Y}^I)$  e  $P_L(\mathcal{Y}^{II})$  herdam os cortes mais profundos que caracterizam a solução ótima de  $P_L(\mathcal{Y})$ , e então o algoritmo de relaxação gera apenas os cortes mais profundos adicionais necessários para encontrar suas soluções ótimas – a solução ótima de  $P_L(\mathcal{Y})$  é geralmente um bom ponto de partida para  $(P_L(\mathcal{Y}^I))$  ou  $(P_L(\mathcal{Y}^{II}))$ . Esta estratégia geralmente acelera os procedimentos de bounding e pruning do algoritmo *branch-and-bound* PMG-BB, como ilustrado na seção dos resultados computacionais.

#### 4.4.4 Extensão para Problemas Multiplicativos Generalizados $(Q_{MG})$

O algoritmo de otimização global proposto acima é facilmente estendido para problemas da forma  $(Q_{MG})$ . O problema  $(Q_{MG})$  é equivalente a

$$\left\{ \begin{array}{ll} \underset{x,z}{\text{maximizar}} & z_1 + \sum_{i=1}^r z_{2i} z_{2i+1} \\ \text{sujeito a} & g(x) \geq z, \\ & x \in \Omega, z \in \mathcal{Z}^0. \end{array} \right. \quad (4.4.18)$$

no qual

$$\mathcal{Z}^0 := \left\{ z \in \mathbb{R}^m : 0 < z^L \leq z \leq z^U \right\}, \quad (4.4.19)$$

onde os vetores  $z^L$  e  $z^U$  são definidos como

$$z_i^L = \min_{x \in \Omega} g_i(x), \quad z_i^U = \max_{x \in \Omega} g_i(x), \quad i = 1, 2, \dots, m.$$

Usando técnicas de projeção e relaxação análogas às utilizadas para o problema  $(P_{MG})$ , o problema  $(Q_{MG})$  é reformulado e resolvido no espaço  $z$ . O problema projetado assume a forma

$$\left| \begin{array}{l} \text{maximizar}_z \quad z_1 + \sum_{i=1}^r z_{2i} z_{2i+1} \\ \text{sujeito a} \quad \sum_{i=1}^m w_i z_i \leq \max_{x \in \Omega} \sum_{i=1}^m w_i g_i(x) \quad \text{para todo } w \in \mathcal{W}, \\ \quad \quad \quad z \in \mathcal{Z}^0. \end{array} \right. \quad (4.4.20)$$

e  $z \in \mathcal{Z}^0$  satisfaz a desigualdade semi-infinita se e somente se

$$\beta(z) := \min_{w \in \mathcal{W}} \max_{x \in \Omega} \sum_{i=1}^m w_i (g_i(x) - z_i) \geq 0.$$

A convergência finita do algoritmo de relaxação é estabelecida quando

$$-\beta(z^k) \leq \epsilon_{rel},$$

onde  $\epsilon_{rel} > 0$  é uma pequena tolerância.

Os envelopes côncavos para os termos bilineares são dados por

$$z_{2i} z_{2i+1} \leq \min \left\{ z_{2i}^U z_{2i+1} + z_{2i+1}^L z_{2i} - z_{2i}^U z_{2i+1}^L, \right. \\ \left. z_{2i}^L z_{2i+1} + z_{2i+1}^U z_{2i} - z_{2i}^L z_{2i+1}^U \right\}, \quad (4.4.21)$$

para todo  $[z_{2i}^L, z_{2i}^U] \times [z_{2i+1}^L, z_{2i+1}^U]$ . O lado direito de (4.4.21) é envelope côncavo de  $z_{2i} z_{2i+1}$  em  $[z_{2i}^L, z_{2i}^U] \times [z_{2i+1}^L, z_{2i+1}^U]$ . Assim, obtemos o seguinte problema de otimização linear semi-

infinito:

$$\begin{array}{l}
 \text{maximizar}_{z, \zeta} \quad \zeta_1 + \sum_{i=1}^r \zeta_{i+1} \\
 \text{sujeito a} \quad \sum_{i=1}^m w_i z_i \leq \max_{x \in \Omega} \sum_{i=1}^m w_i g_i(x) \quad \text{para todo } w \in \mathcal{W}, \\
 \zeta_1 \leq z_1, \\
 \zeta_{i+1} \leq z_{2i}^U z_{2i+1} + z_{2i+1}^L z_{2i} - z_{2i}^U z_{2i+1}^L, \quad i = 1, 2, \dots, r, \\
 \zeta_{i+1} \leq z_{2i}^L z_{2i+1} + z_{2i+1}^U z_{2i} - z_{2i}^L z_{2i+1}^U, \quad i = 1, 2, \dots, r, \\
 z \in \mathcal{Z}^0.
 \end{array} \tag{4.4.22}$$

Um algoritmo *branch-and-bound*, denotado por QMG-BB, com características semelhantes ao do algoritmo PMG-BB, é então obtido.

## 4.5 Experimentos Numéricos

O desempenho computacional dos algoritmos de otimização global propostos foram investigados com base em problemas de programação multiplicativa generalizada selecionados a partir da literatura. No caso dos algoritmos *branch-and-bound*, o número de bissecções e o número de cortes mais profundos geradas pelos algoritmos são os parâmetros mais importantes a serem analisados.

Todos os algoritmos foram implementados em MATLAB versão 7.0.1/*Optimization Toolbox* versão 4, executando sobre uma plataforma computacional composta por um computador Intel Core Duo, 3.33 GHz com 4GB de memória RAM, 32 bits. As tolerâncias adotadas para as convergências finitas dos algoritmos foram fixadas em  $\epsilon_{en} = 0,00001$  (para o algoritmo QMG-ENU),  $\epsilon_{rel} = 0,00001$  (para o algoritmo de relaxação) e  $\epsilon_{bb} = 0,005$  (para os algoritmos *branch-and-bound*), a menos de especificações em contrário.

**Exemplo 4.5.1.** Considere inicialmente o exemplo ilustrativo discutido em [CFSYPP12], [JLC12],

[SS95], onde algoritmos alternativos para problemas convexos multiplicativos são propostos:

$$\begin{array}{l} \text{minimizar} \quad (x_1 + 1) + (2x_1 - 3x_2 + 13)(x_1 + x_2 - 1) \\ \text{sujeito a} \quad -x_1 + 2x_2 \leq 8, \\ \quad \quad \quad -x_2 \leq -3, \\ \quad \quad \quad x_1 + 2x_2 \leq 12, \\ \quad \quad \quad x_1 - 2x_2 \leq -5, \\ \quad \quad \quad x_1 \geq 0, \quad x_2 \geq 0. \end{array}$$

As funções  $f_1, f_2$  e  $f_3$  são convexas e positivas em  $\Omega$ ;  $\Omega$  é um conjunto compacto e não-vazio. Considerando  $\underline{y} = (1; 1; 2)$  e  $\bar{y} = (4,50; 7,25; 6,75)$ , obtêm-se os resultados apresentados nas Tabelas 4.1 e 4.2.

Tabela 4.1: Convergência do Algoritmo QMG-ENU – Exemplo 4.5.1

$k$	$y^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(1,0000; 1,0000; 2,0000)	(0,0000; 0,2500; 0,7500)	(0,0000; 4,0000)	0,7500
Valor $\epsilon_{en}$ -ótimo = 4,0000				

O Algoritmo QMG-ENU (modificado para resolver  $(P_{MG})$ ) convergiu após 2 iterações para a solução  $\epsilon_{en}$ -ótima global  $x^* = (0,0000; 4,0000)$ . O valor ótimo da função multiplicativa generalizada foi  $v^*(P_{MG}) = 4,0000$ .

Tabela 4.2: Convergência do Algoritmo PMG-BB – Exemplo 4.5.1

$k$	$y^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(1,0000; 1,0000; 1,0000; 2,0000)	(0,0000; 0,0000; 0,2500; 0,7500)	(0; 4)	0,7500
	(1,0000; 1,0000; 1,0000; 3,0000)	(1,0000; 0,0000; 0,0000; 0,0000)	(0; 4)	0,0000
Valor $\epsilon_{bb}$ -ótimo = 4,0000				

No Passo 0 do algoritmo *branch-and-bound* PMG-BB obtivemos  $\mu^0 - \gamma^0 = 4,0000 - 4,0000 = 0,0000$ , significando que o algoritmo convergiu na iteração  $q = 0$  sem realização de uma única ramificação. Os cortes mais profundos gerados pelo algoritmo de relaxação ao resolver o problema são exibidos na Tabela 4.2. O algoritmo de relaxação convergiu na segunda iteração para a mesma solução encontrada pelo algoritmo de enumeração de restrições (Tabela 4.1). Dois cortes mais profundos (introduzido no retângulo inicial) foram necessários. A mesma solução é relatada em [CFSYPP12], [JLC12] e [SS95] após duas (com  $\epsilon = 0,001$ ), duas (com  $\epsilon = 0,000001$ ) e três (com  $\epsilon$  não indicado) iterações, respectivamente.

**Exemplo 4.5.2.** Como um segundo exemplo, considere o problema de programação multiplicativa côncava utilizado em [Beno8], [AF10a], [AF11b]:

$$\begin{array}{l}
 \text{maximizar} \quad (x_1 - x_2 + 4) + (5 - 0.25x_1^2)(0.125x_2 + 1) \\
 \quad \quad \quad + (0.25x_1 + 1)(4 - 0.125x_2^2), \\
 \text{sujeito a} \quad 5x_1 - 8x_2 \geq -24, \\
 \quad \quad \quad 5x_1 + 8x_2 \leq 44, \\
 \quad \quad \quad 6x_1 - 3x_2 \leq 15, \\
 \quad \quad \quad 4x_1 + 5x_2 \geq 10, \\
 \quad \quad \quad x_1 \geq 0.
 \end{array}$$

As funções  $g_1, g_2, g_3, g_4$  e  $g_5$  são positivas em  $\Omega$  e os limitantes inferiores e superiores de  $z$  são  $\underline{z} = (1; 1; 1; 1; 1,7422)$  e  $\bar{z} = (6, 5; 5; 1, 5312; 2; 4)$ , respectivamente. Pela Tabela 4.3, o Algoritmo QMG-ENU convergiu após 9 iterações para a solução  $\epsilon_{en}$ -ótima global  $x^* = (2, 5000; 0, 0000)$  do problema. O valor ótimo da função multiplicativa generalizada foi  $v^*(Q_{MG}) = 16, 4375$ .

Tabela 4.3: Convergência do Algoritmo QMG-ENU – Exemplo 4.5.2

$k$	$z^k$	$w^k$	$x(w^k)$	$-\theta(z^k)$
0	(6,5000; 5,0000; 1,5312; 2,0000; 4)	(0,3530; 0,6470; 0,0000; 0,0000; 0)	(1,9647; 0,4282)	0,9645
1	(6,5000; 3,5092; 1,5312; 2,0000; 4)	(0,1024; 0,0746; 0,8230; 0,0000; 0)	(2,7676; 0,5697)	0,4412
2	(6,4359; 3,5442; 1,0000; 2,0000; 4)	(0,0000; 0,1574; 0,0000; 0,8426; 0)	(2,6641; 0,3750)	0,3316
3	(6,5000; 3,1215; 1,0303; 1,6855; 4)	(0,2000; 0,0000; 0,0000; 0,8000; 0)	(4,0000; 3,0000)	0,0484
4	(6,5000; 3,4452; 1,0010; 1,6250; 4)	(0,4094; 0,5906; 0,0000; 0,0000; 0)	(2,4916; 0,0067)	0,0045
5	(6,4962; 3,4401; 1,0019; 1,6259; 4)	(0,1020; 0,0817; 0,8163; 0,0000; 0)	(2,4995; 1,9395)	0,0014
6	(6,4962; 3,4401; 1,0002; 1,6259; 4)	(0,0000; 0,1673; 0,0000; 0,8327; 0)	(2,4785; 0,0619)	0,0012
7	(6,5000; 3,4375; 1,0000; 1,6250; 4)	(0,4095; 0,5905; 0,0000; 0,0000; 0)	(2,4915; 0,0068)	2,6788e-006
8	(6,5000; 3,4375; 1,0000; 1,6250; 4)	(0,0000; 0,1668; 0,0000; 0,8332; 0)	(2,5000; 0,0000)	7,6611e-007

Valor  $\epsilon_{en}$ -ótimo = 16,4375

Pela Tabela 4.4, no Passo 0 do Algoritmo *branch-and-bound* QMG-BB obtivemos  $\mu^0 - \gamma^0 = 16, 4375 - 16, 4373 = 0, 0002$  que é menor do que a tolerância  $\epsilon_{bb} = 0, 005$ , significando que com  $\epsilon_{bb} = 0, 0002$ , o algoritmo convergiu na iteração  $q = 0$  sem realização de uma única ramificação. Os cortes mais profundos gerados pelo algoritmo de relaxação ao resolver o problema são exibidos na Tabela 4.4. O algoritmo de relaxação convergiu na 15ª iteração para a mesma solução encontrada pelo algoritmo de enumeração de restrições. A mesma solução é encontrada em [Beno8] com um número bem maior de iterações (67 iterações). Em [Beno8], um algoritmo alternativo para problemas multiplicativos côncavos que

combina *branch-and-bound* com técnicas baseadas em aproximações externas é proposto.

Tabela 4.4: Convergência do Algoritmo QMG-BB – Exemplo 4.5.2

$k$	$z^k$	$w^k$	$x(w^k)$	$-\beta(z^k)$
	(1; 6,5000; 5,0000; 1,5312; 2,0000; 4)	(0; 0,3534; 0,6466; 0,0000; 0,0000; 0,0000)	(1,9629; 0,4297)	0,9645
	(1; 6,5000; 3,5083; 1,5312; 2,0000; 4)	(0; 0,1024; 0,0745; 0,8231; 0,0000; 0,0000)	(2,7812; 0,5633)	0,4411
	(1; 6,4370; 3,5428; 1,0000; 2,0000; 4)	(0; 0,0000; 0,1584; 0,0000; 0,8416; 0,0000)	(2,6750; 0,4200)	0,3313
	(1; 6,5000; 1,9959; 1,1323; 1,8974; 4)	(0; 0,2000; 0,0000; 0,0000; 0,8000; 0,0000)	(4,0000; 3,0000)	0,2179
	(1; 6,5000; 3,4436; 1,0011; 1,6250; 4)	(0; 0,4093; 0,5907; 0,0000; 0,0000; 0,0000)	(2,4885; 0,0092)	0,0036
	(1; 6,4970; 3,4396; 1,0019; 1,6258; 4)	(0; 0,1020; 0,0818; 0,8162; 0,0000; 0,0000)	(2,4964; 1,9416)	0,0014
	(1; 6,4970; 3,4396; 1,0002; 1,6258; 4)	(0; 0,0000; 0,1673; 0,0000; 0,8327; 0,0000)	(2,4785; 0,0621)	9,8366e-004
o	(1; 6,4272; 3,3468; 1,0182; 1,6432; 4)	(0; 0,0000; 0,0000; 0,0000; 0,2123; 0,7877)	(2,5645; 0,1289)	0,0021
	(1; 6,4664; 3,3956; 1,0084; 1,6334; 4)	(0; 0,0000; 0,0000; 0,0000; 0,1020; 0,8980)	(2,5243; 0,0487)	5,0334e-004
	(1; 6,4858; 3,4198; 1,0035; 1,6286; 4)	(0; 0,0000; 0,0000; 0,0000; 0,0409; 0,9591)	(2,5071; 0,0142)	9,6701e-005
	(1; 6,4951; 3,4314; 1,0012; 1,6262; 4)	(0; 0,0000; 0,0000; 0,0000; 0,0153; 0,9847)	(2,5000; 0,0000)	1,8919e-005
	(1; 6,4933; 3,4292; 1,0017; 1,6267; 4)	(0; 0,0000; 0,1663; 0,0000; 0,8337; 0,0000)	(2,5097; 0,0195)	8,9758e-006
	(1; 6,4994; 3,4368; 1,0001; 1,6251; 4)	(0; 0,0000; 0,1665; 0,0000; 0,8335; 0,0000)	(2,5059; 0,0117)	1,4427e-006
	(1; 6,4981; 3,4351; 1,0005; 1,6255; 4)	(0; 0,0000; 0,0000; 0,0114; 0,0002; 0,9884)	(2,5059; 0,0117)	5,6424e-006
	(1; 6,4995; 3,4368; 1,0001; 1,6251; 4)	(0; 0,0000; 0,0000; 0,0105; 0,0000; 0,9895)	(2,5000; 0,0000)	1,3638e-006
Valor $\epsilon_{bb}$ -ótimo = 16,4375				

**Exemplo 4.5.3.** Considere o problema de programação multiplicativa convexa discutido em [OF10] e [KKY94]:

$$\begin{array}{l}
 \text{minimizar} \quad (3x_1 - 4x_2 + 15) + (x_1 + 2x_2 - 1.5)(2x_1 - x_2 + 4) \\
 \quad \quad \quad + (x_1 - 2x_2 + 8.5)(2x_1 + x_2 - 1), \\
 \text{sujeito a} \quad 5x_1 - 8x_2 \geq -24, \\
 \quad \quad \quad 5x_1 + 8x_2 \leq 44, \\
 \quad \quad \quad 6x_1 - 3x_2 \leq 15, \\
 \quad \quad \quad 4x_1 + 5x_2 \geq 10, \\
 \quad \quad \quad x_1 \geq 0.
 \end{array}$$

Novamente, as funções  $f_1, f_2, f_3, f_4$  e  $f_5$  são positivas em  $\Omega$ . Os limitantes inferiores e superiores de  $y$  são  $\underline{y} = (3; 1; 1; 2; 2)$  e  $\bar{y} = (22, 5; 9; 9; 11; 11)$ , respectivamente. O algoritmo QMG-ENU convergiu após 3 iterações para a solução  $\epsilon_{en}$ -ótima global  $x^* = (0, 0000; 3, 0000)$  do problema, com o valor ótimo igual a  $v^*(P_{MG}) = 12, 5000$ . A mesma solução é relatada em [KKY94], no qual uma técnica de aproximação externa alternativa é proposta, sem indicar o número de iterações necessárias.

A convergência do algoritmo PMG-BB (com  $\epsilon_{bb} = 0, 05$ ) é detalhada na Tabela 4.6. No Passo 0 do algoritmo obtivemos  $\mu^0 - \gamma^0 = 12, 5042 - 12, 5000 = 0, 0058$ , significando que o

Tabela 4.5: Convergência do Algoritmo QMG-ENU – Exemplo 4.5.3

$k$	$y^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(3; 1,0000; 1; 2,0000; 2)	(0,3333; 0,6667; 0,0000; 0,0000; 0,0000)	(0,0000; 2,0000)	2,3333
1	(3; 4,5000; 1; 2,0000; 2)	(0,0000; 0,0000; 0,0000; 0,9130; 0,0870)	(2,0000; 4,2500)	0,4565
2	(3; 4,5000; 1; 2,5000; 2)	(0,5370; 0,2535; 0,1205; 0,0891; 0,0000)	(0,0000; 3,0000)	2,8897e-009
Valor $\epsilon_{en}$ -ótimo = 12,5000				

algoritmo convergiu na iteração  $q = 0$  sem realização de uma única ramificação. Os cortes mais profundos gerados pelo algoritmo de relaxação ao resolver o problema são exibidos na Tabela 4.6. O algoritmo de relaxação convergiu na segunda iteração para a mesma solução encontrada pelo algoritmo de enumeração de restrições (Tabela 4.5). Dois cortes mais profundos (introduzido no retângulo inicial) foram necessários.

Tabela 4.6: Convergência do Algoritmo PMG-BB – Exemplo 4.5.3

$k$	$y^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(1; 3; 1,0000; 1; 2,0000; 2)	(0,0000; 0,3333; 0,6667; 0,0000; 0; 0,0000)	(0; 2)	2,3333
	(1; 3; 4,5000; 1; 2,5021; 2)	(0,0005; 0,5935; 0,0494; 0,2132; 0; 0,1434)	(0; 3)	2,8342e-009
Valor $\epsilon_{bb}$ -ótimo = 12,5000				

**Exemplo 4.5.4.** Como o último exemplo ilustrativo escolhido da literatura considere o problema de programação multiplicativa côncava discutido em [AF11b]:

$$\begin{array}{l}
 \text{maximizar} \quad (3x_1 - 4x_2 + 15) + (x_1 + 2x_2 - 1.5)(2x_1 - x_2 + 4) \\
 \quad \quad \quad + (x_1 - 2x_2 + 8.5)(2x_1 + x_2 - 1), \\
 \text{sujeito a} \quad 5x_1 - 8x_2 \geq -24, \\
 \quad \quad \quad 5x_1 + 8x_2 \leq 44, \\
 \quad \quad \quad 6x_1 - 3x_2 \leq 15, \\
 \quad \quad \quad 4x_1 + 5x_2 \geq 10, \\
 \quad \quad \quad x_1 \geq 0.
 \end{array}$$

A solução  $\epsilon_{en}$ -ótima global  $x^* = (4,0000; 3,0000)$  fornecendo o valor ótimo  $v^*(Q_{MG}) = 156,5000$  foi encontrada pelo algoritmo QMG-ENU após 5 iterações. A Tabela 4.7 resume os dados produzidos pelo QMG-ENU ao resolver o problema.

Com  $\epsilon_{bb} = 0,05$ , o algoritmo QMG-BB convergiu para a mesma solução encontrada pelo algoritmo QMG-ENU. Os cortes mais profundos gerados pelo algoritmo de relaxação ao resolver o problema são exibidos na Tabela 4.8.

Tabela 4.7: Convergência do Algoritmo QMG-ENU – Exemplo 4.5.4

$k$	$z^k$	$w^k$	$x(w^k)$	$-\theta(z^k)$
0	(22,5; 9,0000; 9; 11,0000; 11)	(0,5000; 0,5000; 0,0000; 0,0000; 0,0000)	(4; 3,0000)	4,0000
1	(14,5; 9,0000; 9; 11,0000; 11)	(0,0000; 0,2045; 0,0000; 0,6006; 0,1949)	(4; 3,0000)	3,0000
2	(14,5; 9,0000; 9; 6,0053; 11)	(0,0000; 0,0000; 0,0000; 0,0000; 1,0000)	(4; 3,0000)	1,0000
3	(14,5; 9,0000; 9; 6,3298; 10)	(0,0433; 0,9563; 0,0004; 0,0000; 0,0000)	(2; 4,2500)	0,4565
4	(15,0; 8,5000; 9; 6,0000; 10)	(0,1456; 0,3848; 0,2625; 0,0107; 0,1964)	(4; 3,0000)	5,1230e-011

Valor  $\epsilon_{\eta}$ -ótimo = 156,5000

Tabela 4.8: Convergência do Algoritmo PMG-BB – Exemplo 4.5.4

$q$	$k$	$z^k$	$w^k$	$x(w^k)$	$-\beta(z^k)$
	0	(1; 22,5000; 9,0000; 9; 11,0000; 11,0)	(0; 0,5000; 0,5000; 0,0000; 0,0000; 0,0000)	(2,5; 0)	4,0000
	1	(1; 14,5000; 9,0000; 9; 11,0000; 11,0)	(0; 0,0000; 0,3750; 0,0000; 0,6250; 0,0000)	(4,0; 3)	3,0000
0	2	(1; 14,5000; 9,0000; 9; 6,2000; 11,0)	(0; 0,0000; 0,0000; 0,0000; 0,0000; 1,0000)	(4,0; 3)	1,0000
	3	(1; 14,5000; 9,0000; 9; 6,2000; 10,0)	(0; 0,0435; 0,9565; 0,0000; 0,0000; 0,0000)	(2,0; 4)	0,4565
	4	(1; 15,0000; 8,5000; 9; 6,5000; 10,0)	(1; 0,0000; 0,0000; 0,0000; 0,0000; 0,0000)	(4,0; 3)	2,1242e-009
0-I	0	(1; 12,7500; 8,6023; 9; 6,4386; 10,0)	(0; 0,0000; 0,9130; 0,0870; 0,0000; 0,0000)	(4,0; 3)	0,0934
	1	(1; 12,7500; 8,5000; 9; 6,5000; 10,0)	(1; 0,0000; 0,0000; 0,0000; 0,0000; 0,0000)	(4,0; 3)	3,8723e-012
0-II	-	-	-	-	-
1-I	-	-	-	-	-
1-II	0	(1; 17,6250; 5,8750; 9; 8,0750; 10,0)	(0; 0,4444; 0,0000; 0,0000; 0,0000; 0,5556)	(2,5; 0)	1,1667
	1	(1; 17,6250; 5,8750; 9; 8,0750; 7,9)	(1; 0,0000; 0,0000; 0,0000; 0,0000; 0,0000)	(4,0; 3)	1,3113e-011
2-I	-	-	-	-	-
2-II	-	-	-	-	-
3-I	-	(1; 7,8750; 8,5000; 9; 6,5000; 10,0)	(1; 0,0000; 0,0000; 0,0000; 0,0000; 0,0000)	(4,0; 3)	3,0093e-014
3-II	-	-	-	-	-
4-I	-	Problema Infactível			
4-II	-	-	-	-	-

Valor  $\epsilon_{bb}$ -ótimo = 156,5000

A Figura 4.1 ilustra a árvore *branch-and-bound* produzida pelo algoritmo QMG-BB. O número de cortes mais profundos adicionais exigidos em cada subproblema (I ou II) é informado entre parênteses, precedido pela número da iteração. Cinco cortes mais profundos foram produzidos no Passo 0 do algoritmo QMG-BB. A bisseção do retângulo inicial deu origem aos problemas 0-I e 0-II, que herderam 5 cortes do Passo 0. Dois cortes adicionais foram necessários para resolver o problema 0-I; nenhum corte adicional foi necessário para resolver o problema 0-II.

A bisseção do retângulo associado ao problema 0-II deu origem aos problemas 1-I e 1-II, que herderam os 5 cortes do problema 0-II. Dois cortes adicionais foram necessários para resolver o problema 1-II; nenhum corte adicional foi necessário para resolver o problema

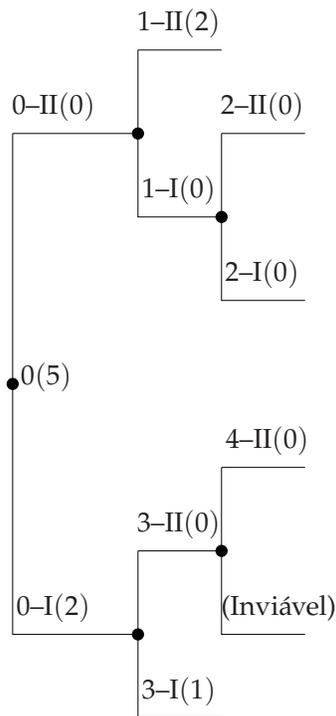


Figura 4.1: Árvore *branch-and-bound* para Exemplo 4.5.4.

1-I. Três outras bisseções ocorreram até a convergência do algoritmo QMG-BB. Um corte adicional foi necessário para resolver o problema 3-I. A solução do problema global exigiu um total de 10 cortes mais profundos.

Nos Exemplos (4.5.1), (4.5.2) e (4.5.3), a combinação de envelopes e cortes mais profundos foi suficiente para resolver os problemas multiplicativos generalizados no Passo 0 dos algoritmos *branch-and-bound*, sem bisseções do retângulo inicial. No Exemplo (4.5.4), 5 bisseções foram realizadas, mas poucos cortes mais profundos foram necessários depois Passo 0 do algoritmo. A propagação e reutilização de cortes mais profundos por meio da árvore *branch-and-bound* minimizou o número de cortes mais profundos necessários para a convergência dos algoritmos *branch-and-bound* e manteve a otimização (linear) em cada sub-retângulo a mais simples possível. Um comportamento semelhante foi observado durante a aplicação dos algoritmos *branch-and-bound* propostos para outros exemplos extraídos da literatura.

A maior parte do esforço computacional requerido pelos algoritmos *branch-and-bound* propostos concentra-se na geração de cortes mais profundos por meio do algoritmo MIN-MAX ou MAX-MIN, conforme o caso. Apesar de cada iteração deste algoritmo requerer a resolu-

ção de um problema de otimização convexa, e em seguida de um problema linear simples, experiências numéricas indicam que a carga computacional exigida pela geração de cortes mais profundos é compensada por uma convergência mais rápida dos algoritmos *branch-and-bound* propostos.

### 4.5.1 Exemplos Aleatórios

Além desses exemplos ilustrativos acima, relatamos a seguir os resultados de experimentos mais exaustivos com os algoritmos apresentados nas seções anteriores. Consideramos duas subclasses de problemas multiplicativos generalizados, nos quais as funções objetivos a serem otimizadas são descritas como a soma de uma função linear com a soma de produtos de funções lineares:

$$P_{MG} \left\{ \begin{array}{l} \text{minimizar} \quad \langle c^1, x \rangle + \sum_{i=1}^r \langle c^{2i}, x \rangle \langle c^{2i+1}, x \rangle \\ \text{sujeito a} \quad Ax \geq b, \quad x \in \mathbb{R}_+^n, \end{array} \right.$$

e

$$Q_{MG} \left\{ \begin{array}{l} \text{maximizar} \quad \langle c^1, x \rangle + \sum_{i=1}^r \langle c^{2i}, x \rangle \langle c^{2i+1}, x \rangle \\ \text{sujeito a} \quad Ax \leq b, \quad x \in \mathbb{R}_+^n, \end{array} \right.$$

nos quais  $A \in \mathbb{R}^{p \times n}$ ,  $b \in \mathbb{R}^p$ ,  $c^i \in \mathbb{R}^n$ ,  $i = 1, 2, \dots, m$  são matrizes e vetores constantes com elementos gerados aleatoriamente com distribuição uniforme no intervalo  $[0, 100]$ .

Como em [OF10], as seguintes quantidades foram definidas (valores totais):  $W$ , número de problemas  $P_W$  resolvidos;  $C$ , números de planos de corte necessários para convergência e  $T$ , tempo de CPU (em segundos). Dez problemas com combinações selecionadas de  $n$  (número de variáveis),  $p$  (número de restrições) e  $r$  (número de produtos) foram resolvidos.

Novamente, todos os algoritmos foram implementados em MATLAB versão 7.0.1/*Optimization Toolbox* versão 4, executando sobre uma plataforma computacional composta por um computador Intel Core Duo, 3.33 GHz com 4GB de memória RAM, 32 bits e as tolerâncias adotadas para as convergências finitas desses algoritmos foram fixadas em  $\epsilon_{en} = 0,00001$ ,  $\epsilon_{rel} = 0,00001$  e  $\epsilon_{bb} = 0,05$ . As Tabelas a seguir apresentam a média e desvio padrão (em parênteses) dos valores de  $C$ ,  $W$  e  $T$ .

As Tabelas 4.9 e 4.10 apresentam os resultados obtidos pelas abordagens propostas neste

capítulo ao resolver globalmente problemas multiplicativos generalizados ( $P_{MG}$ ) e ( $Q_{MG}$ ), respectivamente, com  $r = 1$ ,  $r = 2$  e  $r = 3$ .

Tabela 4.9: Resultados dos Algoritmos PMG-ENU e PMG-BB para ( $P_{MG}$ ) com  $r = 1, 2, 3$ .

$r$	1	1	1	1	1	1	1
$n$	20	20	50	50	100	100	150
$p$	10	30	30	70	70	130	130
ALGORITMO PMG-ENU							
C	3,80 (1,23)	4,00 (0,93)	6,67 (1,50)	7,22 (2,23)	6,70 (0,95)	8,40 (0,84)	8,80 (1,30)
W	11,00 (3,86)	14,60 (4,74)	16,36 (3,70)	16,23 (7,15)	17,40 (4,61)	22,90 (1,17)	26,20 (3,08)
T	0,75 (0,29)	0,99 (0,34)	2,90 (1,12)	5,91 (1,86)	9,92 (1,62)	25,04 (5,83)	27,48 (10,10)
ALGORITMO PMG-BB							
C	3,71 (0,32)	4,21 (0,48)	6,56 (1,95)	7,39 (0,84)	7,28 (1,05)	8,94 (1,18)	9,18 (1,08)
W	10,29 (4,94)	17,71 (2,80)	17,98 (4,41)	22,09 (7,14)	29,32 (6,15)	31,11 (6,66)	34,14 (5,78)
T	0,73 (0,11)	0,81 (0,16)	2,22 (0,62)	5,98 (2,13)	8,03 (1,91)	11,33 (4,47)	14,82 (4,73)
$r$	2	2	2	2	3	3	3
$n$	20	20	50	50	20	50	50
$p$	10	30	30	70	30	30	70
ALGORITMO PMG-ENU							
C	5,29 (1,67)	6,90 (1,11)	8,15 (1,23)	8,09 (0,52)	6,85 (1,05)	7,36 (1,32)	8,88 (1,49)
W	24,24 (3,20)	36,09 (3,62)	39,98 (4,48)	47,99 (8,27)	42,49 (5,58)	41,04 (9,47)	46,99 (9,09)
T	3,45 (1,08)	6,82 (0,92)	14,31 (6,92)	28,86 (4,32)	37,77 (10,10)	50,09 (17,34)	66,61 (15,57)
ALGORITMO PMG-BB							
C	7,41 (0,78)	7,78 (1,04)	9,00 (0,48)	9,27 (1,17)	9,65 (0,49)	10,19 (1,64)	11,84 (1,65)
W	28,19 (7,08)	38,55 (4,41)	42,24 (6,03)	45,38 (8,51)	48,99 (8,10)	54,58 (10,78)	56,21 (11,24)
T	2,99 (0,38)	4,35 (1,29)	8,50 (2,52)	16,96 (3,67)	19,22 (4,04)	28,72 (8,31)	30,23 (4,87)

De acordo com as Tabelas 4.9 e 4.10, os valores de  $n$  e  $p$  têm pouca influência no valor de  $C$ , o número de planos de corte mais profundos necessários para convergência, e uma influência significativa no valor de  $W$ , o número de problemas  $P_W$  resolvidos, e conseqüentemente no valor de  $T$ , tempo de CPU, porque  $W$  está relacionado à resolução de problemas de otimização no  $\mathbb{R}^n$ . As Tabelas 4.9 e 4.10 também mostram que identificar os cortes mais profundos se torna cada vez mais difícil, uma vez que a razão  $W/C$  aumenta à medida que  $r$  cresce. Entretanto, o crescimento da razão  $W/C$  é suave nas experiências computacionais realizadas.

Uma técnica de aproximação externa alternativa é proposta em [KKY94] para resolver globalmente problemas multiplicativos generalizados convexos ( $P_{MG}$ ). Uma vez que os resultados relatados na Tabela 4.9 e os fornecidos em [KKY94] foram obtidos usando

Tabela 4.10: Resultados dos Algoritmos QMG-ENU e QMG-BB para  $(Q_{MG})$  com  $r = 1, 2, 3$ .

$r$	1	1	1	1	1	1	1
$n$	20	20	50	50	100	100	150
$p$	10	30	30	70	70	130	130
ALGORITMO QMG-ENU							
C	2,10 (0,85)	2,60 (1,43)	4,30 (1,34)	5,40 (1,84)	7,10 (2,02)	6,70 (2,75)	6,90 (2,42)
W	6,70 (0,94)	11,20 (1,38)	14,90 (3,65)	15,80 (1,98)	19,14 (5,61)	17,90 (4,27)	19,80 (7,55)
T	0,73 (0,31)	1,03 (0,47)	2,79 (2,00)	3,96 (1,48)	7,02 (2,87)	14,61 (6,41)	19,14 (4,67)
ALGORITMO QMG-BB							
C	2,90 (1,62)	2,80 (1,92)	6,70 (2,22)	7,10 (0,95)	6,50 (0,33)	7,10 (1,30)	7,60 (0,75)
W	9,80 (2,47)	9,60 (2,23)	15,70 (5,21)	14,90 (4,44)	19,80 (3,67)	20,90 (5,49)	24,60 (7,69)
T	0,60 (0,26)	0,53 (0,22)	2,03 (0,73)	3,90 (1,23)	6,22 (1,71)	9,50 (2,10)	11,70 (4,56)
$r$	2	2	2	2	3	3	3
$n$	20	20	50	50	20	50	50
$p$	10	30	30	70	30	30	70
ALGORITMO QMG-ENU							
C	4,30 (2,63)	4,80 (1,69)	6,90 (2,56)	8,30 (2,21)	5,00 (1,25)	4,60 (1,90)	6,90 (0,99)
W	27,80 (14,24)	30,90 (9,69)	37,30 (12,88)	60,40 (16,35)	32,00 (13,65)	34,80 (16,64)	35,50 (8,55)
T	3,72 (2,99)	4,89 (1,47)	15,06 (9,54)	21,01 (9,28)	30,95 (15,29)	41,57 (22,54)	50,42 (19,11)
ALGORITMO QMG-BB							
C	7,11 (1,31)	9,00 (0,45)	10,60 (0,90)	12,10 (0,23)	11,80 (0,15)	11,90 (0,32)	12,20 (1,11)
W	31,02 (3,62)	35,70 (5,63)	37,00 (8,24)	39,10 (5,89)	43,00 (5,32)	48,80 (7,11)	50,30 (9,27)
T	2,69 (0,67)	2,71 (0,58)	7,11 (1,12)	10,54 (3,25)	13,80 (5,88)	19,30 (6,14)	28,20 (5,22)

diferentes recursos computacionais, a seguinte medida de desempenho relativo é adotada:

$$\tau_{i,j} = \frac{\text{tempo médio para } n = i \text{ e } p = j}{\text{tempo médio para } n = 20 \text{ e } p = 10} \quad (4.5.1)$$

O crescimento dos tempos computacionais necessários dos algoritmos propostos nesta tese e o algoritmo proposto em [KKY94] são apresentados na Tabela 4.11. Para o problema  $(Q_{MG})$  não se encontra disponível uma experiência numérica extensa, realizada com base em problemas aleatórios. Realizamos apenas uma análise comparativa entre os algoritmos QMG-ENU e QMG-BB propostos nesta tese.

Os números médios de vértices gerados pelo método de [KKY94] são progressivamente maiores, o que dificulta a aplicação do algoritmo para  $r > 3$ . Por outro lado, os algoritmos propostos nesta tese geram números relativamente pequenos de cortes no espaço-imagem. Com isso, estes dois algoritmos, mesmo sendo de natureza exponencial, conseguem bom desempenho nos casos testados. Além disso, pela análise das Tabelas 4.9, 4.10 e 4.11,

Tabela 4.11: Crescimento dos tempos computacionais para  $r = 1$ .

Problema ( $P_{MG}$ )	$\tau_{20,30}$	$\tau_{50,30}$	$\tau_{50,70}$	$\tau_{100,70}$	$\tau_{100,130}$	$\tau_{150,130}$
Algoritmo proposto em [KKY94]	3,80	17,40	55,80	167,00	576,40	964,40
Algoritmo PMG-ENU	1,32	3,87	7,88	13,23	33,39	36,64
Algoritmo PMG-BB	1,11	3,04	8,19	11,00	15,30	20,30
Problema ( $Q_{MG}$ )	$\tau_{20,30}$	$\tau_{50,30}$	$\tau_{50,70}$	$\tau_{100,70}$	$\tau_{100,130}$	$\tau_{150,130}$
Algoritmo QMG-ENU	1,41	3,82	5,42	9,62	20,01	26,22
Algoritmo QMG-BB	0,88	3,38	4,83	10,37	15,83	19,50

percebe-se que os algoritmos PMG-BB e QMG-BB, apesar de gerar em um número maior de cortes no espaço-imagem, têm tempos computacionais menores quando comparado aos algoritmos PMG-ENU e QMG-ENU, o que pode ser explicado pela natureza enumerativa (explícita) destes últimos algoritmos.

## 4.6 Resumo

Neste capítulo propusemos duas abordagens diferentes para resolver globalmente problemas multiplicativos generalizados convexos e côncavos. Mostramos que estes problemas podem ser reformulados como problemas de Programação Quadrática Indefinida com infinitas restrições lineares de desigualdade. Uma técnica de enumeração de restrições e uma técnica *branch-and-bound* foram usadas para resolver globalmente os problemas. Experiências computacionais demonstraram o bom desempenho dos algoritmos propostos.

# 5

## Programação Fracionária Generalizada

### Sumário

---

5.1	Introdução . . . . .	108
5.2	Programação Fracionária Convexa–Côncava . . . . .	110
5.3	Programação Fracionária Côncava–Convexa . . . . .	118
5.4	Experimentos Numéricos . . . . .	135
5.5	Resumo . . . . .	146

---

O problema de otimizar uma ou mais funções fracionárias é conhecido como problema de Programação Fracionária. Um dos problemas fracionários mais difíceis é o que envolve a soma de frações. De fato, este problema foi um dos problemas de programação fracionária menos pesquisados até cerca de 1990. Em [FJ01], os autores mostram que, apesar de sua estrutura especial sob hipóteses usuais sobre numeradores e denominadores, o problema é essencialmente um problema NP-difícil. Este capítulo propõe novas abordagens para os problemas de minimização e maximização de uma soma de funções fracionárias sobre um conjunto convexo.

## 5.1 Introdução

A otimização global de uma soma de funções fracionárias lineares ou não-lineares sobre um conjunto convexo tem atraído o interesse de vários pesquisadores desde a década de 1970. Até aproximadamente 1990 este problema foi um dos menos pesquisados na área de Programação Fracionária, a despeito de suas várias aplicações. Vários algoritmos especializados têm sido propostos e desenvolvidos para resolver globalmente estes problemas reconhecidamente não-convexos

Razões de funções convexas ou côncavas, bem como composições de tais razões, não são convexas em geral, mesmo em caso de funções lineares. Mas freqüentemente são convexas generalizadas em algum sentido. Desde o início, a programação fracionária tem se beneficiado de avanços no campo da Convexidade Generalizada [Mar75]. De fato, os anais dos seis simpósios internacionais sobre Convexidade Generalizada já realizados contêm contribuições em Programação Fracionária [SZ81], [Singh & Dass, 1989], [Cambini et al., 1990], [Komlosi et al., 1994], [Crouzeix et al., 1998] e [Hadjisavvas et al., 2001].

Inicialmente restrita a problemas com apenas uma fração, o campo da Programação Fracionária ampliou seu escopo nas últimas décadas ao incorporar a análise de problemas multi-frações [Sch81]. O estudo de problemas com apenas uma fração dominou grande parte da literatura até cerca de 1980. Muitos dos resultados conhecidos até então são apresentados na primeira monografia sobre Programação Fracionária, publicada por Schaible em 1978 [Sch78]; veja também [Sch81] e [SI83]. Desde então, outros dois trabalhos dedicados exclusivamente à programação fracionária foram publicados [Cra88], [SM97]. Cada um deles inclui um capítulo sobre problemas fracionários generalizados.

Sejam  $p$ ,  $q$  e  $h_k$ ,  $k = 1, 2, \dots, m$  funções reais definidas em um conjunto  $C$  do espaço

euclidiano  $n$ -dimensional  $\mathbb{R}^n$ . Considere

$$f(x) := \frac{p(x)}{q(x)},$$

e

$$\Omega := \left\{ x \in C : h_k(x) \leq 0, k = 1, 2, \dots, m \right\},$$

sendo  $q(x) > 0$  para todo  $x \in \Omega$ .

Os problemas não-lineares

$$\begin{aligned} \min_{x \in \Omega} f(x), \\ \max_{x \in \Omega} f(x), \end{aligned} \tag{5.1.1}$$

são chamados de problemas fracionários de razão única. Em algumas aplicações, mais do que uma razão aparece na função objetivo:

$$\begin{aligned} \min_{x \in \Omega} \max_{1 \leq i \leq r} f_i(x), \\ \max_{x \in \Omega} \min_{1 \leq i \leq r} f_i(x), \end{aligned} \tag{5.1.2}$$

nos quais  $f_i(x) := p_i(x)/q_i(x)$ . Os problemas (5.1.2) são referidos como problemas fracionários generalizados, ou problemas fracionários min-max e max-min, respectivamente [SI83]. Quando as funções  $p$  e  $q$  são afins (lineares mais uma constante) e  $\Omega$  é um poliedro convexo, os problemas (5.1.1) são chamados de problemas fracionários lineares assumindo as formas

$$\begin{aligned} \min \left\{ \frac{c^T x + \alpha}{d^T x + \beta} : Ax \leq b, x \geq 0 \right\}, \\ \max \left\{ \frac{c^T x + \alpha}{d^T x + \beta} : Ax \leq b, x \geq 0 \right\}, \end{aligned} \tag{5.1.3}$$

nos quais  $c, d \in \mathbb{R}^n$ ,  $\alpha, \beta \in \mathbb{R}$ ,  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ .

Problemas com uma única razão surgem em problemas de tomada de decisões. Podem também ocorrer indiretamente, como resultado de algum tipo de modelagem [Sch78]. Em algumas, maximizar a de um sistema leva a um problema fracionário. Para uma exposição de problemas fracionários com uma única razão, sugerimos o capítulo sobre a programação fracionária do *Handbook of Global Optimization* [Sch95].

Os problemas de programação fracionária generalizada são conhecidos por possuírem soluções locais que não são globais. O objetivo deste capítulo é fornecer novas abordagens

para resolver globalmente problemas de programação fracionária generalizados. Os problemas de programação fracionária generalizada (PFG) considerados neste trabalho apresentam as seguintes formas gerais:

$$P_{FG} : \min_{x \in \Omega} \sum_{i=1}^r \frac{p_i(x)}{q_i(x)},$$

e

$$Q_{FG} : \max_{x \in \Omega} \sum_{i=1}^r \frac{q_i(x)}{p_i(x)}.$$

nos quais as funções  $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são convexas e as funções  $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são côncavas para  $i = 1, 2, \dots, r$ . Assume-se que  $\Omega$  é um subconjunto convexo, compacto e não-vazio de  $\mathbb{R}^n$ , e que as funções  $p_i$  e  $q_i$  são positivas em  $\Omega$  para  $i = 1, 2, \dots, r$ . Em otimização global, os problemas  $(P_{FG})$  e  $(Q_{FG})$  têm sido referidos como os problemas fracionários generalizados convexo-côncavo e côncavo-convexo, respectivamente [SS03].

## 5.2 Programação Fracionária Convexa-Côncava

Esta seção aborda o problema

$$P_{FG} : \min_{x \in \Omega} \sum_{i=1}^r \frac{p_i(x)}{q_i(x)},$$

que consiste em minimizar uma soma de razões, sendo cada termo da soma uma razão entre uma função convexa e uma função côncava, ambas positivas sobre a região viável do problema. Dois algoritmos de otimização global baseados numa reformulação adequada do problema no espaço-imagem são propostos. Minimizadores globais são obtidos como o limite das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais, resolvidos por meio de um procedimento de enumeração de restrições, de acordo com o primeiro algoritmo, e de uma sequência de problemas lineares-fracionais especiais, resolvidos por meio de um procedimento *branch-and-bound* retangular, de acordo com o segundo algoritmo. Ambos os algoritmos exploram o relativamente pequeno número de semi-espaços necessários para aproximar o problema original no espaço-imagem.

### 5.2.1 Abordagem no Espaço–Imagem

Os problemas de programação fracionária generalizada convexa–côncava são formulados como

$$P_{FG} : \min_{x \in \Omega} \sum_{i=1}^r \frac{p_i(x)}{q_i(x)},$$

no qual  $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções convexas e  $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções côncavas para  $k = 1, 2, \dots, r$ . Assume-se que  $\Omega$  é um subconjunto convexo, compacto e não–vazio do  $\mathbb{R}^n$ , e que as funções  $p_i$  e  $q_i$  são positivas em  $\Omega$  para  $i = 1, 2, \dots, r$ .

A abordagem no espaço–imagem apresentada a seguir é análoga à proposta no Capítulo 4 para resolver problemas multiplicativos generalizados das formas  $(P_{MG})$  e  $(Q_{FG})$ . O Lema abaixo garante a equivalência entre os problemas  $(P_{FG})$  e  $(P_{MG})$ .

**Lema 5.2.1.** *A função  $1/g$  é convexa e positiva se  $g$  é uma função côncava e positiva.*

Definindo  $\tilde{q}_i(x) := 1/q_i(x)$ , para  $i = 1, 2, \dots, r$ , o problema  $(P_{FG})$  pode ser reduzido ao problema de programação multiplicativa generalizada convexa

$$\tilde{P}_{MG} : \min_{x \in \Omega} \sum_{i=1}^r p_i(x) \tilde{q}_i(x).$$

Ao problema  $(\tilde{P}_{MG})$  também pode-se associar o problema de minimizar a função objetivo vetorial  $\tilde{f} := (p_1, \tilde{q}_1, p_2, \tilde{q}_2, \dots, p_r, \tilde{q}_r)$  em  $\Omega$ . A função objetivo em  $(\tilde{P}_{MG})$  pode ser escrita como a composição  $\tilde{u}(\tilde{f}(x))$  em que  $\tilde{u} : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $m = 2r$ , é definida como

$$\tilde{u}(y) := \sum_{i=1}^r y_{p_i} y_{\tilde{q}_i}.$$

A função  $\tilde{u}$  pode ser vista como uma função de agregação para o problema de minimizar  $\tilde{f}$  em  $\Omega$ . A projeção do problema  $(\tilde{P}_{MG})$  assume a forma

$$\tilde{P}_{\mathcal{F}} : \min_{y \in \mathcal{Y}} \tilde{u}(y),$$

sendo  $\mathcal{Y} := \{y \in \mathbb{R}^m : y = \tilde{f}(\Omega), x \in \Omega\}$  o espaço–imagem.

Como foi discutido no capítulo anterior, o problema se resume na obtenção da solução

ótima global do problema mestre

$$\tilde{P}_{\mathcal{F}^k} : \min_{y \in \mathcal{F}^k} \tilde{u}(y),$$

numa iteração genérica  $k$ , isto é, após a geração de  $k$  restrições de desigualdade (hiperplanos de corte), responsáveis pela aproximação corrente de  $\mathcal{F}$ . O conjunto  $\mathcal{F}^k$  é descrito por  $k$  restrições de desigualdade e pelas restrições de limitação  $\underline{y}_i \leq y_i \leq \bar{y}_i$ ,  $i = 1, 2, \dots, m$ . Na forma matricial, o problema  $(\tilde{P}_{\mathcal{F}^k})$  é equivalente ao seguinte problema de programação quadrática linearmente restrito:

$$\left| \begin{array}{ll} \text{minimizar} & \tilde{u}(y) = \frac{1}{2}y^T Q y \\ \text{sujeito a} & C^{(k)}y \geq d^{(k)}, \\ & \underline{y} \leq y \leq \bar{y}, \end{array} \right. \quad (5.2.1)$$

no qual  $C^{(k)} \in \mathbb{R}^{k \times m}$ ,  $d^{(k)} \in \mathbb{R}^k$ ,  $\underline{y} \in \mathbb{R}^m$  e  $\bar{y} \in \mathbb{R}^m$  caracterizam a representação matricial de  $\mathcal{F}^k$ . A matriz  $Q \in \mathbb{R}^{m \times m}$  tem a seguinte forma

$$Q = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad Q \in \mathbb{R}^{m \times m}.$$

Dada a estrutura bloco–diagonal de  $Q$ , seu polinômio característico pode ser escrito como

$$\det(\lambda I - Q) = \underbrace{(\lambda^2 - 1) \cdots (\lambda^2 - 1)}_{r \text{ vezes}} = 0,$$

e os autovalores de  $Q$  são 1 e  $-1$ ; portanto  $Q$  é indefinida. Logo, o problema de programação quadrática (5.2.1) é um problema de programação quadrática indefinida.

Pelo menos  $r$  restrições estarão ativas em qualquer solução ótima do problema (5.2.1). Devido ao pequeno número de restrições geradas pelo algoritmo de relaxação, o problema (5.2.1) pode ser eficientemente resolvido por meio da técnica de enumeração de restrições [HPT95]. Além disso, somente as combinações de pelo menos  $r$  restrições que incluem a mais violada por  $y^k$  precisam ser consideradas na iteração  $k + 1$  do processo (para maiores detalhes, veja a Seção 4.3.2).

Com isso e com pequenas modificações necessárias, o algoritmo PMG–ENU proposto no Capítulo 4 é capaz de resolver globalmente o problema de programação fracionária generalizada convexa–côncava ( $P_{FG}$ ); todas as propriedades derivadas anteriormente, como convergência para uma solução ótima global, permanecem válidas. O algoritmo PFG–ENU resolve globalmente o problema de programação fracionária generalizada convexa–côncava ( $P_{FG}$ ) por meio de enumeração de restrições.

### 5.2.2 Gerando Problemas–Testes

Nesta seção será descrito um método para construir uma classe de problemas–testes que envolvem a minimização de soma de razões lineares sobre um conjunto convexo, compacto e não–vazio. Dado um conjunto convexo, compacto e não–vazio, o método determina uma função que é a soma de razões lineares e atinge um mínimo global sobre o conjunto em um ponto que pode ser encontrado por meio de programação convexa e busca unidimensional. Geralmente a função também terá mínimos locais que não serão globais. O método baseia-se no seguinte resultado.

**Teorema 5.2.2.** *Seja  $i \in \{1, 2, \dots, r\}$  e considere*

$$Z := \left\{ x \in \Omega : a(\langle c_i, x \rangle + \alpha_i) + b(\langle d_i, x \rangle + \beta_i) = K \right\},$$

com

$$\Omega \subset \left\{ x \in \mathbb{R}^n : x \geq 0, x \neq 0 \right\},$$

sendo  $a, b$  e  $K$  escalares e  $a$  e  $b$  ambos nulos. Então, a função

$$f_i(x) := \frac{\langle c_i, x \rangle + \alpha_i}{\langle d_i, x \rangle + \beta_i},$$

é uma função convexa em  $Z$  sempre que  $a \neq 0$  e  $K/a \geq 0$ .

*Demonstração.* Dada a convexidade de  $\Omega$ , o conjunto  $Z$  também é um conjunto convexo. Fixe  $i \in \{1, 2, \dots, r\}$  e assumamos  $a \neq 0$  e  $K/a \geq 0$ . Sejam  $z^1, z^2 \in Z$  e suponhamos que  $\lambda$  é um escalar que satisfaz  $0 \leq \lambda \leq 1$ . Como  $a \neq 0$  e  $z^1, z^2 \in Z$ , para  $j = 1, 2$

$$\langle c_i, z^j \rangle + \alpha_i = (K/a) - (b/a)(\langle d_i, z^j \rangle + \beta_i). \quad (5.2.2)$$

Portanto,

$$\begin{aligned}
 f_i[\lambda z^1 + (1 - \lambda)z^2] &= \frac{\langle c_i, \lambda z^1 + (1 - \lambda)z^2 \rangle + \alpha_i}{\langle d_i, \lambda z^1 + (1 - \lambda)z^2 \rangle + \beta_i} \\
 &= \frac{\lambda(\langle c_i, z^1 \rangle + \alpha_i) + (1 - \lambda)(\langle c_i, z^2 \rangle + \alpha_i)}{\lambda(\langle d_i, z^1 \rangle + \beta_i) + (1 - \lambda)(\langle d_i, z^2 \rangle + \beta_i)} \\
 &= \frac{K}{a} \left[ \frac{1}{\lambda(\langle d_i, z^1 \rangle + \beta_i) + (1 - \lambda)(\langle d_i, z^2 \rangle + \beta_i)} \right] - \frac{b}{a} \\
 &\leq \frac{K}{a} \left[ \frac{\lambda(\langle d_i, z^2 \rangle + \beta_i) + (1 - \lambda)(\langle d_i, z^1 \rangle + \beta_i)}{(\langle d_i, z^1 \rangle + \beta_i)(\langle d_i, z^2 \rangle + \beta_i)} \right] - \frac{b}{a} \\
 &= \frac{K}{a} \left[ \frac{\lambda}{\langle d_i, z^1 \rangle} + \frac{1 - \lambda}{\langle d_i, z^2 \rangle} \right] - \frac{b}{a} \\
 &= \frac{\lambda(K/a)}{\langle d_i, z^1 \rangle} - \frac{b}{a}\lambda + \frac{(1 - \lambda)(K/a)}{\langle d_i, z^2 \rangle} - \frac{b}{a}(1 - \lambda) \\
 &= \lambda f_i(z^1) + (1 - \lambda)f_i(z^2),
 \end{aligned}$$

e portanto  $f_i$  é uma função convexa. □

Os passos do procedimento para a construção de problemas–testes em programação fracionária generalizada linear são descritos a seguir.

**Procedimento PT-PC:** Construção de Problemas–Testes ( $P_{FGL}$ )

**Passo 1.** Escolha um valor inteiro  $r \geq 2$ . Escolha vetores positivos  $c_1 \in \mathbb{R}^n$  e  $d_i \in \mathbb{R}^n$ ,  $i = 1, 2, \dots, r$ .

**Passo 2.** Escolha números positivos  $\alpha_1 \in \mathbb{R}$  e  $\beta_i \in \mathbb{R}$ ,  $i = 1, 2, \dots, r$ .

**Passo 3.** Para cada  $i = 2, 3, \dots, r$ , defina  $c_i \in \mathbb{R}^n$  e  $\alpha_i \in \mathbb{R}$  por

$$\begin{aligned}
 c_i &= c_1 + d_1 - d_i, \\
 \alpha_i &= \alpha_1 + \beta_1 - \beta_i.
 \end{aligned}$$

O procedimento PT-PC constrói um problema fracionário linear de modo que para cada  $i = 1, 2, \dots, r$ ,

$$\langle d_i, x \rangle + \beta_i > 0, \quad \text{para todo } x \in \Omega,$$

e, sem perda de generalidade, podemos supor que o numerador  $(\langle c_i, x \rangle + \alpha_i)$  é não-negativo para todo  $x \in \Omega$ .

Um problema–teste construído pelo procedimento PT–PC em geral possui mínimos locais que não são mínimos globais. Apesar disso, qualquer problema–teste pode ser globalmente resolvido por meio de otimização convexa e unidimensional. Suponha que o problema foi construído pelo procedimento PT–PC. Sejam

$$K_{max} := \max_{x \in \Omega} \left( \langle c_1 + d_1, y \rangle + \alpha_1 + \beta_1 \right),$$

e

$$K_{min} := \min_{x \in \Omega} \left( \langle c_1 + d_1, y \rangle + \alpha_1 + \beta_1 \right).$$

Note que  $K_{max}$  e  $K_{min}$  são números reais positivos, e que cada um deles é determinado por meio da resolução de um problema convexo. Para cada  $K$  tal que

$$K_{min} \leq K \leq K_{max},$$

seja  $g(K)$  definido por

$$g(K) := \min_{x \in \Omega} \sum_{i=1}^r \frac{\langle c_i, x \rangle + \alpha_i}{\langle d_i, x \rangle + \beta_i} \quad (5.2.3)$$

$$\text{sujeito a } \langle c_1 + d_1, x \rangle + \alpha_1 + \beta_1 = K,$$

Por construção, para cada  $x \in \Omega$  existe um  $K \in [K_{min}, K_{max}]$  tal que  $x$  satisfaz a restrição  $\langle c_1 + d_1, x \rangle + \alpha_1 + \beta_1 = K$  do problema (5.2.3). Conclui-se que o valor ótimo global  $v$  do problema fracionário linear é dado por

$$v = \inf g(K) \quad \text{sujeito a } K \in [K_{min}, K_{max}]. \quad (5.2.4)$$

Note que, para cada  $K \in [K_{min}, K_{max}]$ , pelo Teorema 5.2.2, a função

$$f_1(x) := \frac{\langle c_1, x \rangle + \alpha_1}{\langle d_1, x \rangle + \beta_1},$$

é convexa na região viável do problema (5.2.3). Se  $K \in [K_{min}, K_{max}]$ , pelo Passo 3 do Procedimento PT–PC, para cada  $i = 1, 2, \dots, r$ , a região definida pelas condições

$$\langle c_i + d_i, x \rangle + \alpha_i + \beta_i = K \quad \text{e } x \in \Omega,$$

é equivalente à região definida pelas restrições do problema (5.2.3).

Para cada  $K \in [K_{min}, K_{max}]$ , o problema (5.2.3) envolve a minimização de uma função convexa e contínua sobre um conjunto convexo, compacto e não-vazio. Além disso, como a função  $g$  também é convexa no intervalo  $[K_{min}, K_{max}]$ , o ínfimo em (5.2.4) é atingido.

A partir da discussão acima, um problema gerado pelo Procedimento PT-PC, pode ser resolvido globalmente da seguinte maneira. Primeiro, resolve-se dois problemas de programação convexa para obter os valores  $K_{min}$  e  $K_{max}$ . Em seguida, aplicamos qualquer procedimento de busca unidimensional para encontrar o valor  $K^*$  de  $K$  que globalmente minimiza a função convexa  $g(K)$  no intervalo  $[K_{min}, K_{max}]$ ; o valor de  $g(K)$  é determinado por meio da resolução do problema de programação convexa (5.2.3). O valor ótimo global  $v$  do problema é igual a  $g(K^*)$ , e qualquer solução ótima para o problema de programação convexa (5.2.3) com  $K = K^*$  é uma solução ótima global do problema original.

### Alguns Problemas Testes

A seguir, aplicamos o procedimento PT-PC para gerar alguns problemas-testes de programação fracionária generalizada linear. Suponha que o conjunto  $\Omega$  é dado por

$$\Omega := \left\{ x \in \mathbb{R}^2 : 3x_1^2 + x_2^2 \leq 48, x_1 + x_2 \geq 1, x_1, x_2 \geq 0 \right\}.$$

**Exemplo 5.2.3.** Seguindo os passos 1–3 do procedimento PT-PC, temos

**Passo 1.** Escolha  $r = 2$ , e

$$c_1^T = (1, 0; 3, 0), \quad d_1^T = (4, 0; 1, 0), \quad d_2^T = (1, 0; 1, 0).$$

**Passo 2.** Escolha  $\alpha_1 = 2, 0$ ,  $\beta_1 = 3, 0$  e  $\beta_2 = 4, 0$ .

**Passo 3.** Construa  $c_2 \in \mathbb{R}^2$  e  $\alpha_2 \in \mathbb{R}$  da seguinte maneira

$$c_2 = \begin{pmatrix} 1, 0 \\ 3, 0 \end{pmatrix} + \begin{pmatrix} 4, 0 \\ 1, 0 \end{pmatrix} - \begin{pmatrix} 1, 0 \\ 1, 0 \end{pmatrix} = \begin{pmatrix} 4, 0 \\ 3, 0 \end{pmatrix},$$

$$\alpha_2 = 2, 0 + 3, 0 - 4, 0 = 1, 0.$$

O problema resultante é dado por

$$\begin{aligned} \min \quad & \frac{x_1 + 3x_2 + 2}{4x_1 + x_2 + 3} + \frac{4x_1 + 3x_2 + 1}{x_1 + x_2 + 4} \\ \text{sujeito a} \quad & x \in \Omega := \left\{ x \in \mathbb{R}^2 : 3x_1^2 + x_2^2 \leq 48, x_1 + x_2 \geq 1, x_1, x_2 \geq 0 \right\}. \end{aligned}$$

Como resultado, temos

$$\begin{aligned} K_{max} &= \max_{x \in \Omega} 5x_1 + 4x_2 + 5 = 39,176, \\ K_{min} &= \min_{x \in \Omega} 5x_1 + 4x_2 + 5 = 9,000. \end{aligned}$$

Para cada  $K \in [K_{min}, K_{max}]$ ,  $g(K)$  é dada por

$$g(K) := \min_{x \in \Omega} \frac{x_1 + 3x_2 + 2}{4x_1 + x_2 + 3} + \frac{4x_1 + 3x_2 + 1}{x_1 + x_2 + 4}, \quad (5.2.5)$$

$$\text{sujeito a } 5x_1 + 4x_2 + 5 = K.$$

Ao resolver globalmente o problema–teste, podemos empregar qualquer procedimento de busca unidimensional para encontrar o valor  $K^*$  que minimiza  $g(K)$  no intervalo  $[K_{min}, K_{max}]$ . Por meio de aplicação de uma busca do tipo bisseção, por exemplo, encontramos, após 17 bisseções,

$$K^* = 10,0000,$$

$$g(K^*) = 1,4285,$$

e uma solução ótima para o problema de programação convexa (5.2.5) com  $K = 10,0000$  é dada por

$$(x^*)^T = (1,0000; 0,0000).$$

Como resultado,  $(x^*)^T = (1,0000; 0,0000)$  resolve globalmente o problema–teste, e o valor ótimo global do problema é igual a 1,4285.

**Exemplo 5.2.4.** Escolha  $r = 2$ , e

$$\begin{aligned} c_1^T &= (2, 0; 4, 0), \quad d_1^T = (2, 0; 2, 0), \quad d_2^T = (3, 0; 2, 0), \\ \alpha_1 &= 1, 0 \quad \beta_1 = 2, 0 \quad \beta_2 = 2, 0. \end{aligned}$$

Neste caso, o problema seria

$$\begin{array}{ll} \text{minimizar} & \frac{2x_1 + 4x_2 + 1}{2x_1 + 2x_2 + 2} + \frac{x_1 + 4x_2 + 1}{3x_1 + 2x_2 + 2} \\ \text{sujeito a} & x \in \Omega := \left\{ x \in \mathbb{R}^2 : 3x_1^2 + x_2^2 \leq 48, \quad x_1 + x_2 \geq 1, \quad x_1, x_2 \geq 0 \right\}. \end{array}$$

**Exemplo 5.2.5.** Escolha  $r = 3$ , e

$$\begin{aligned} c_1^T &= (7, 0; 5, 0), \quad d_1^T = (1, 0; 2, 0), \quad d_2^T = (5, 0; 3, 0), \quad d_3^T = (1, 0; 5, 0) \\ \alpha_1 &= 10, 0 \quad \beta_1 = 1, 0 \quad \beta_2 = 3, 0 \quad \beta_3 = 7, 0. \end{aligned}$$

Usando o procedimento PT-PC, podemos ver que o seguinte problema de programação fracionária generalizada linear será gerado

$$\begin{array}{ll} \text{minimizar} & \frac{3x_1 + 3x_2 + 3}{x_1 + 2x_2 + 1} + \frac{3x_1 + 4x_2 + 8}{5x_1 + 5x_2 + 3} + \frac{7x_1 + 2x_2 + 4}{x_1 + 5x_2 + 7} \\ \text{sujeito a} & x \in \Omega := \left\{ x \in \mathbb{R}^2 : 3x_1^2 + x_2^2 \leq 48, \quad x_1 + x_2 \geq 1, \quad x_1, x_2 \geq 0 \right\}. \end{array}$$

### 5.3 Programação Fracionária Côncava–Convexa

A maioria das técnicas de otimização global que contemplam problemas de programação fracionária generalizada exploraram o conceito de espaço–imagem. O espaço–imagem é definido por um mapeamento que associa uma variável adicional a cada razão [FP94], [DHT01], ou variáveis adicionais ao numerador e ao denominador de cada razão [Beno2b], [Kuno2]. Há vantagens teóricas e práticas em considerar problemas no espaço–imagem. Primeiramente, existem correspondências entre as soluções globais dos problemas de interesse e as soluções Pareto-ótimas (ou soluções eficientes) de certos problemas de otimização multiobjetivo associados. Além disso, os problemas de otimização tornam-se mais tratáveis no espaço–imagem.

A abordagem proposta na seção a seguir para problemas de programação fracionária generalizada côncava–côncava (e também convexa–côncava) explora conceitos e métodos de programação multiobjetivo, o conceito de espaço–imagem, o conceito de envelope convexo, e uma técnica *branch-and-bound* retangular.

### 5.3.1 Uma Abordagem de decomposição

Nesta seção será proposto um algoritmo para resolver globalmente problemas de programação fracionária generalizada côncava–convexa ( $P_{FG}$ ). Ao projetar o problema no espaço–imagem, obtemos o problema equivalente de maximizar uma soma de razões elementares sujeito a uma restrição de desigualdade linear semi-infinita. Então mostramos que o problema equivalente é globalmente solúvel por um algoritmo de otimização que combina dois outros algoritmos: uma técnica *branch-and-bound* retangular, responsável por lidar com a não–concavidade do problema no espaço–imagem, e uma técnica de relaxação (aproximação externa).

A abordagem proposta nesta seção é aplicável tanto ao problema de programação fracionária generalizada côncava–convexa

$$Q_{FG} : \max_{x \in \Omega} \sum_{i=1}^r \frac{q_i(x)}{p_i(x)}.$$

como ao problema de programação fracionária generalizada convexa–côncava

$$P_{FG} : \min_{x \in \Omega} \sum_{i=1}^r \frac{p_i(x)}{q_i(x)},$$

nos quais  $\Omega \subset \mathbb{R}^n$  um conjunto convexo compacto e não–vazio, as funções  $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, r$  são convexas e positivas em  $\Omega$  e as funções  $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, r$  são côncavas e positivas em  $\Omega$ .

Apesar de todas as hipóteses anteriores sobre a convexidade, a côncavidade e a positividade das funções envolvidas na função objetivo, os problemas de programação fracionária generalizada ( $Q_{FG}$ ) e ( $P_{FG}$ ) são ambos não–convexos, isto é, podem apresentar uma solução ótima local que não seja uma solução ótima global. Além disso, as soluções ótimas de ( $Q_{FG}$ ) e ( $P_{FG}$ ) não estão relacionadas por qualquer propriedade conhecida, embora possamos dizer algo sobre seus valores ótimos.

**Teorema 5.3.1.** *Os valores ótimos de ( $Q_{FG}$ ) e ( $P_{FG}$ ) satisfazem a desigualdade*

$$v^*(Q_{FG})v^*(P_{FG}) \geq r^2, \tag{5.3.1}$$

nos quais  $v^*(Q_{FG})$  e  $v^*(P_{FG})$  são os valores ótimos de ( $Q_{FG}$ ) e ( $P_{FG}$ ), respectivamente.

*Demonstração.* Usando a positividade de  $p_i(x)$ ,  $q_i(x)$ ,  $i = 1, 2, \dots, r$  em  $\Omega$ , e invocando a desigualdade entre as médias aritmética e harmônica, obtemos

$$\frac{1}{r} \sum_{i=1}^r \frac{q_i(x)}{p_i(x)} \geq \frac{r}{\sum_{i=1}^r \frac{p_i(x)}{q_i(x)}}, \quad (5.3.2)$$

para todo  $x \in \Omega$ . Então, a desigualdade (5.3.1) é estabelecida, tomando o máximo nos dois lados de (5.3.1) sobre  $\Omega$ . □

O Teorema (5.3.1) fornece um limitante inferior para o valor ótimo de  $(P_{FG})$  quando o valor ótimo de  $(Q_{FG})$  é conhecido, e vice-versa. É claro que  $v^*(Q_{FG})v^*(P_{FG}) = 1$  quando  $r = 1$ . Como exemplo, considere o seguinte problema de programação fracionária generalizada linear:

$$\left| \begin{array}{l} \text{minimizar}_{x_1, x_2} \quad \frac{x_1}{2x_1 + x_2} + \frac{x_2}{2} \\ \text{sujeito a} \quad 2x_1 + x_2 \leq 6, \\ \quad \quad \quad 3x_1 + x_2 \leq 8, \\ \quad \quad \quad -x_1 + x_2 \geq -1, \\ \quad \quad \quad x_1, x_2 \geq 1. \end{array} \right. \quad (5.3.3)$$

O valor ótimo de (5.3.3) é  $v^*(P_{FG}) = 0.83$ , atingido em  $x^* = (1; 1)$  usando o algoritmo de enumeração de restrições proposto na seção anterior. O limitante inferior para o valor ótimo de  $(Q_{FG})$  é  $4,82 = \frac{4}{0,83} \leq v^*(Q_{FG})$ , e o seu valor ótimo é  $v^*(Q_{FG}) = 6,5$  (atingido em  $x^* = (1; 4)$  usando o algoritmo proposto nesta seção). Embora experiências numéricas mostrem que a qualidade do limitante inferior varia de acordo com os dados do problema, a desigualdade (5.3.1) é útil por indicar o que esperar se mudar de um problema para o outro.

Nesta seção, desenvolvemos uma abordagem de otimização global para o problema de programação fracionária generalizada côncava-convexa  $(Q_{FG})$ . Um desenvolvimento análogo pode ser feito em termos de problema de programação fracionária generalizada convexa-côncava  $(P_{FG})$ , essencialmente trocando maximizações por minimizações e concavidades por convexidades.

Intrinsecamente associada a qualquer solução ótima de  $(Q_{FG})$  está a definição de solução Pareto-ótima ou solução eficiente [Yu85]. Esta associação deriva da interpretação do problema  $(Q_{FG})$  como uma instância paramétrica do problema de programação fracionária

multiobjetivo

$$\left| \begin{array}{l} \underset{x}{\text{maximizar}} \quad \left( \frac{q_1(x)}{p_1(x)}, \frac{q_2(x)}{p_2(x)}, \dots, \frac{q_r(x)}{p_r(x)} \right) \\ \text{sujeito a} \quad x \in \Omega. \end{array} \right. \quad (5.3.4)$$

Um ponto viável  $x^0 \in \Omega$  é uma solução eficiente do problema (5.3.4) se não existe qualquer outro ponto  $x \in \Omega$  tal que  $q_i(x)/p_i(x) \geq q_i(x^0)/p_i(x^0)$ ,  $i = 1, 2, \dots, r$ , com pelo menos uma desigualdade estrita. Sabe-se que qualquer solução ótima de  $(Q_{FG})$  é uma solução eficiente do problema de programação fracionária multiobjetivo (5.3.4) [DHT01].

Nossa abordagem baseia-se numa caracterização alternativa da eficiência das soluções ótimas de problema  $(Q_{FG})$ : qualquer solução ótima de  $(Q_{FG})$  é uma solução eficiente para o problema de programação multiobjetivo convexo

$$\left| \begin{array}{l} \underset{x}{\text{maximizar}} \quad (q_1(x), q_2(x), \dots, q_p(x), -p_1(x), -p_2(x), \dots, -p_r(x)) \\ \text{sujeito a} \quad x \in \Omega. \end{array} \right. \quad (5.3.5)$$

De fato, suponha que  $x^* \in \Omega$  é uma solução ótima para  $(Q_{FG})$ , mas não é uma solução eficiente de (5.3.5). Neste caso, existe um outro ponto  $x^0 \in \Omega$  tal que

$$q_i(x^0) \geq q_i(x^*), \quad -p_i(x^0) \geq -p_i(x^*), \quad i = 1, 2, \dots, r,$$

com pelo menos uma desigualdade estrita, o que contradiz a otimalidade de  $x^*$  por  $(Q_{FG})$  (porque o valor fornecido por  $x^0$  seria maior do que o valor fornecido por  $x^*$ ). O Teorema a seguir é uma consequência imediata da associação de problema  $(Q_{FG})$  com o problema de programação multiobjetiva convexo (5.3.5).

**Teorema 5.3.2.** *Seja  $x^* \in \Omega$  qualquer solução ótima do problema de  $(Q_{FG})$  e considere o simplex  $2r$ -dimensional*

$$w \in \mathcal{W} := \left\{ w \in \mathbb{R}_+^{2r} : \sum_{k=1}^{2r} w_k = 1 \right\}.$$

*Então existe algum  $w \in \mathcal{W}$  (em geral, dependente da solução ótima  $x^*$ ) tal que  $x^*$  também*

resolva o problema convexo

$$\left\{ \begin{array}{l} \underset{x}{\text{maximizar}} \quad \left\{ \sum_{i=1}^r w_i q_i(x) - \sum_{j=r+1}^{2r} w_j p_{j-r}(x) \right\} \\ \text{sujeito a} \quad x \in \Omega. \end{array} \right. \quad (5.3.6)$$

*Demonstração.* Qualquer solução ótima do problema  $(Q_{FG})$  é uma solução eficiente do problema (5.3.5). Cada solução eficiente do problema (5.3.5) é, por sua vez, uma solução ótima do problema (5.3.6) para algum  $w \in \mathcal{W}$  [Yu85].  $\square$

É um pouco surpreendente que qualquer solução ótima do problema de programação não-convexa  $(Q_{FG})$  também seja uma solução ótima de um problema convexo. O ingrediente-chave para isso é a consideração de uma função vetorial aumentada no problema de programação multiobjetiva convexo (5.3.5).

Devido à forma especial de sua função objetivo côncava, o problema (5.3.6) é mais fácil de resolver do que o problema  $(Q_{FG})$ . Em particular, o problema (5.3.6) é linear quando  $(Q_{FG})$  é um problema de programação fracionária generalizada linear, o que implica que qualquer solução ótima do problema  $(Q_{FG})$  se encontra na fronteira (mas não necessariamente em um vértice) do politopo  $\Omega$ .

Se  $r = 1$ , então o problema (5.3.6) reduz-se a

$$\left\{ \begin{array}{l} \underset{x}{\text{maximizar}} \quad w_1 q_1(x) - (1 - w_1) p_1(x) \\ \text{sujeito a} \quad x \in \Omega, \end{array} \right.$$

que (para  $w_1 > 0$ ) é equivalente ao problema paramétrico

$$\left\{ \begin{array}{l} \underset{x}{\text{maximizar}} \quad q_1(x) - \lambda p_1(x), \quad \lambda := (1 - w_1)/w_1 \\ \text{sujeito a} \quad x \in \Omega. \end{array} \right.$$

O problema anterior tem sido muito estudado em conexão com o problema de programação fracionária de uma única razão [SS03].

O Teorema 5.3.2 garante a existência, mas não sugere um procedimento para determinar  $w \in \mathcal{W}$  que levaria a uma solução ótima de  $(Q_{FG})$  por meio da solução do problema convexo (5.3.6). Em essência, o algoritmo de otimização global desenvolvido nesta seção encontra iterativamente  $w \in \mathcal{W}$  com esta propriedade.

### 5.3.2 Projeção no Espaço–Imagem

Abordagens de otimização global baseadas nas imagens das funções vetoriais  $p := (p_1, p_2, \dots, p_r)$  e  $q := (q_1, q_2, \dots, q_r)$  que caracterizam as funções envolvidas na soma das razões têm sido relatadas na literatura de programação fracionária. Uma representação diferente do espaço–imagem do problema ( $Q_{FG}$ ) é proposta nesta seção.

Sejam  $y^L := (y_1^L, y_2^L, \dots, y_r^L)$ ,  $y^U := (y_1^U, y_2^U, \dots, y_r^U)$  ( $0 < y^L \leq y^U$ ),  $z^L := (z_1^L, z_2^L, \dots, z_r^L)$  and  $z^U := (z_1^U, z_2^U, \dots, z_r^U)$  ( $0 < z^L \leq z^U$ ) vetores  $r$ –dimensional tais que  $y^L \leq p(x) \leq y^U$  and  $z^L \leq q(x) \leq z^U$  para todo  $x \in \Omega$ . Os vetores  $y^L$ ,  $y^U$ ,  $z^L$  e  $z^U$  podem ser definidos como

$$y_i^L := \min_{x \in \Omega} p_i(x), \quad y_i^U := \max_{x \in \Omega} p_i(x), \quad z_i^L := \min_{x \in \Omega} q_i(x), \quad z_i^U := \max_{x \in \Omega} q_i(x),$$

para  $i = 1, 2, \dots, r$ . Todos estes valores são facilmente obtidos no caso de problemas de programação fracionária generalizada lineares. No caso geral de problemas de programação fracionária generalizada não–lineares,  $y^L$  e  $z^U$  são os valores ótimos de  $2r$  problemas convexos, enquanto  $y^U$  e  $z^L$  exigem a resolução de  $2r$  problemas não–convexos. Um procedimento prático (também baseado em otimização convexa) para encontrar os valores de  $y^U$  e  $z^L$  tais que  $p(x) \leq y^U$  and  $z^L \leq q(x)$  para todos  $x \in \Omega$  é sugerido em [Ben99].

Considere, então, o retângulo

$$\mathcal{R}^0 := \{(y, z) \in \mathbb{R}^r \times \mathbb{R}^r : y^L \leq y \leq y^U, z^L \leq z \leq z^U\}, \quad (5.3.7)$$

e o problema de soma de razões

$$\left| \begin{array}{l} \text{maximizar}_{x, y, z} \quad \sum_{i=1}^r \frac{z_i}{y_i} \\ \text{sujeito a} \quad p(x) \leq y, \\ \quad \quad \quad q(x) \geq z, \\ \quad \quad \quad x \in \Omega, (y, z) \in \mathcal{R}^0. \end{array} \right. \quad (5.3.8)$$

O resultado seguinte é essencial para o desenvolvimento da abordagem dedicado ao problema ( $Q_{FG}$ ) no espaço imagen.

**Teorema 5.3.3.** *Se  $(x^*, y^*, z^*)$  é uma solução ótima para o problema (5.3.8), então  $y^* = p(x^*)$ ,  $z^* = q(x^*)$ , e  $x^*$  é uma solução ótima para o problema ( $Q_{FG}$ ). Reciprocamente, se  $x^*$  é uma solução ótima para o problema ( $Q_{FG}$ ), então  $(x^*, y^*, z^*)$  é uma solução ótima para o problema (5.3.8), com*

$y^* := p(x^*)$  e  $z^* := q(x^*)$ .

*Demonstração.* Seja  $(x^*, y^*, z^*)$  uma solução ótima para o problema (5.3.8). Então  $q(x^*) \geq z^* \geq z^L \geq 0$  e  $y^* \geq p(x^*) \geq y^L > 0$ . Isso implica que para cada  $i = 1, 2, \dots, r$

$$\frac{q_i(x^*)}{p_i(x^*)} \geq \frac{z^*}{y^*},$$

que por sua vez implica que

$$v^*(Q_{FG}) \geq \sum_{i=1}^r \frac{z^*}{y^*}. \quad (5.3.9)$$

Sejam  $\tilde{z}_i := q_i(x^*)$  e  $\tilde{y}_i := p_i(x^*)$  para  $i = 1, 2, \dots, r$ . Portanto,  $(x^*, \tilde{z}, \tilde{y})$  é viável para  $(Q_{FG})$ , e como  $(x^*, y^*, z^*)$  é uma solução ótima global  $(Q_{FG})$ , teremos

$$v^*(Q_{FG}) \leq \sum_{i=1}^r \frac{z^*}{y^*}. \quad (5.3.10)$$

Usando (5.3.9) e (5.3.10), obtemos

$$v^*(Q_{FG}) = \sum_{i=1}^r \frac{z^*}{y^*}, \quad (5.3.11)$$

implicando que  $q(x^*) = z^*$  e  $p(x^*) = y^*$ . Para provar que  $x^*$  é uma solução ótima global de  $(Q_{FG})$ , suponha que  $\hat{x}$  é uma solução viável para  $(Q_{FG})$ . Podemos verificar facilmente que  $(\hat{x}, \hat{z}, \hat{y})$ , com  $\hat{z} := q(\hat{x})$  e  $\hat{y} := p(\hat{x})$ , é uma solução viável do problema (5.3.8), de modo que  $\sum_{i=1}^r q_i(\hat{x})/p_i(\hat{x}) \leq \sum_{i=1}^r z_i^*/y_i^*$ , que implica que  $x^*$  é uma solução ótima global de  $(Q_{FG})$ .

Por outro lado, suponha que  $x^*$  é uma solução ótima global para o problema  $(Q_{FG})$ ,  $z^* := q(x^*)$  e  $y^* := p(x^*)$ . Por definição  $(x^*, y^*, z^*)$  é uma solução viável para (5.3.8), e

$$v^*(Q_{FG}) = \sum_{i=1}^r \frac{z^*}{y^*}.$$

Seja  $(\bar{x}, \bar{z}, \bar{y})$  uma solução viável do problema (5.3.8). Então,  $v^*(Q_{FG}) \geq \sum_{i=1}^r q_i(\bar{x})/p_i(\bar{x})$  e  $\sum_{i=1}^r q_i(\bar{x})/p_i(\bar{x}) \geq \sum_{i=1}^r \bar{z}_i/\bar{y}_i$ . Consequentemente,

$$\sum_{i=1}^r \frac{z_i^*}{y_i^*} = v^*(Q_{FG}) \geq \sum_{i=1}^r \frac{\bar{z}_i}{\bar{y}_i},$$

o que completa a prova da segunda parte do teorema.  $\square$

O problema (5.3.8) tem  $n + 2r$  variáveis de decisão, uma soma de razões elementares como função objetivo e uma região viável convexa nas variáveis  $x$ ,  $y$  e  $z$ . Problemas de soma de razões na forma equivalente (5.3.8) têm sido resolvidos por meio de técnicas *branch-and-bound* baseadas em relaxações côncavas da sua função objetivo em partições (subdivisões) de  $\mathcal{R}^0$  [DHT01], [Kuno2], [Benozb].

Observamos que, apesar das operações típicas de *branch-and-bound* poderiam ser aplicadas somente às variáveis  $y$  e  $z$ , o problema (5.3.8) está formulado no espaço  $\mathbb{R}^{n+2r}$ . Uma representação alternativa da sua região viável leva a um problema equivalente no espaço  $\mathbb{R}^{2r}$ .

A projeção do problema (5.3.8) no  $\mathbb{R}^{2r}$  é

$$\left\{ \begin{array}{l} \underset{y,z}{\text{maximizar}} \quad v(y,z) \\ \text{sujeito a} \quad (y,z) \in \mathcal{V} \cap \mathcal{R}^0, \end{array} \right. \quad (5.3.12)$$

sendo que

$$\begin{aligned} v(y,z) := \underset{x}{\text{maximizar}} \quad & \sum_{i=1}^r \frac{z_i}{y_i} \\ \text{sujeito a} \quad & p(x) \leq y, \\ & q(x) \geq z, \\ & x \in \Omega, \end{aligned} \quad (5.3.13)$$

e

$$\mathcal{V} := \left\{ (y,z) \in \mathbb{R}^{2r} : p(x) \leq y, q(x) \geq z \text{ para algum } x \in \Omega \right\}. \quad (5.3.14)$$

Adotando a convenção de que o valor do problema (5.3.13) é  $-\infty$  se  $(y,z) \notin \mathcal{V}$  (isto é, quando o problema 5.3.13 é inviável), obtemos

$$v(y,z) = \begin{cases} \sum_{i=1}^r \frac{z_i}{y_i}, & \text{se } (y,z) \in \mathcal{V}, \\ -\infty, & \text{caso contrário,} \end{cases}$$

isto é, o valor de  $v$  é definido pela existência de um  $x \in \Omega$  que satisfaz o sistema de desigualdades  $f(x) \leq y$ ,  $g(x) \geq z$ ;  $v$  não depende do valor que  $x \in \Omega$  assume.

**Teorema 5.3.4.** *Se  $(x^*, y^*, z^*)$  é uma solução ótima para o problema (5.3.8), então  $(y^*, z^*)$  é uma solução ótima para o problema (5.3.12). Reciprocamente, se  $(y^*, z^*)$  é uma solução ótima para o*

problema (5.3.12), então existe  $x^* \in \Omega$  tal que  $y^* = p(x^*)$ ,  $z^* = q(x^*)$ , e  $(x^*, y^*, z^*)$  é uma solução ótima para o problema (5.3.8).

*Demonstração.* A equivalência entre problemas das formas (5.3.8) e (5.3.12) é estabelecida sob hipóteses mais gerais em [Geo70]. Além disso, observamos que se  $(y^*, z^*)$  é uma solução ótima do problema (5.3.12), então, em princípio,  $y^* \geq f(x^*)$  e  $z^* \leq g(x^*)$  para algum  $x^* \in \Omega$ . No entanto, a otimalidade de  $(y^*, z^*)$  seria facilmente refutada por  $(y^0, z^0)$ ,  $y^0 := f(x^*)$  e  $z^0 := g(x^*)$ , se  $y^* \neq f(x^*)$  ou  $z^* \neq g(x^*)$ .  $\square$

A projeção do problema (5.3.8) no espaço–imagem se materializa quando o conjunto  $\mathcal{V}$  é explicitamente representado.

A representação de  $\mathcal{V}$  já foi objeto de análise nos Capítulos 3 e 4: existe  $x \in \Omega$  tal que  $\phi(x) \leq t$  se e somente se

$$\sum_{i=1}^m w_i t_i \leq \max_{x \in \Omega} \sum_{i=1}^m w_i \phi_i(x) \quad \text{para todos } w \in \mathcal{W}. \quad (5.3.15)$$

Fazendo  $m := 2r$ ,  $\phi := (q, -p)$  and  $t := (z, -y)$ , o problema (5.3.12) assume a seguinte forma equivalente no espaço–imagem:

$$(P_S) \left\{ \begin{array}{l} \text{maximizar}_{y,z} \quad \sum_{i=1}^r \frac{z_i}{y_i} \\ \text{sujeito a} \quad \sum_{i=1}^r w_i z_i - \sum_{j=r+1}^{2r} w_j y_{j-r} \leq b(w) \quad \text{para todo } w \in \mathcal{W}, \\ (y, z) \in \mathcal{R}^0. \end{array} \right. \quad (5.3.16)$$

no qual

$$b(w) := \max_{x \in \Omega} \left\{ \sum_{i=1}^r w_i q_i(x) - \sum_{j=r+1}^{2r} w_j p_{j-r}(x) \right\}.$$

O problema  $(P_S)$  consiste em maximizar uma soma de razões elementares sujeito a uma restrição de desigualdade linear semi–infinita e as restrições delimitadoras originais. Seguindo a mesma linha de desenvolvimento dos Capítulos 3 e 4, expressamos a restrição semi–infinita como uma única restrição,

$$\theta(y, z) \geq 0, \quad (5.3.17)$$

onde

$$\theta(y, z) := \min_{w \in \mathcal{W}} \psi_{y,z}(w) \quad (5.3.18)$$

e

$$\psi_{y,z}(w) := \max_{x \in \Omega} \left\{ \sum_{i=1}^r w_i (q_i(x) - z_i) - \sum_{j=r+1}^{2r} w_j (p_{j-r}(x) - y_{j-r}) \right\}. \quad (5.3.19)$$

Os subíndices em  $\psi$  lembram de que esta função de  $w \in \mathcal{W}$  é parametrizada por  $y$  e  $z$ .

Seja  $(\bar{x}, \bar{w})$  uma solução ótima do problema min-max que define o valor de  $\theta$  no lado esquerdo de (5.3.17). Se  $\theta(y, z) < 0$ , então pelo menos a seguinte restrição de desigualdade é violada por  $(y, z)$ :

$$\sum_{i=1}^r \bar{w}_i z_i - \sum_{j=r+1}^{2r} \bar{w}_j y_{j-r} \leq \sum_{i=1}^r \bar{w}_i q_i(\bar{x}) - \sum_{j=r+1}^{2r} \bar{w}_j p_{j-r}(\bar{x}). \quad (5.3.20)$$

Dentre todas as desigualdades lineares – semi-espacos suportes para  $\mathcal{V}$  que constituem a restrição semi-infinita, a desigualdade (5.3.20) é a mais violada por  $(y, z)$ . Em termos geométricos, tal desigualdade produz um *corte mais profundo* no retângulo  $\mathcal{R}^0$ .

### 5.3.3 Algoritmo de Relaxação

A seguir apresentamos um algoritmo de relaxação para a resolução do problema  $(P_S)$ . O índice da  $l$ -éssima desigualdade mais violada é denotado por  $w^l$ ,  $l \geq 0$ .

**Algoritmo PS-REL:** Algoritmo de Relaxação para  $(Q_{FG})$

**Passo 0:** Defina  $\mathcal{R}^0$  e faça  $k := 0$ ;

**Passo 1:** Resolva o problema relaxado

$$(P_R) \left\{ \begin{array}{l} \text{maximizar}_{y,z} \quad \sum_{i=1}^r \frac{z_i}{y_i} \\ \text{sujeito a} \quad \sum_{i=1}^r w_i^l z_i - \sum_{j=r+1}^{2r} w_j^l y_{j-r} \leq b(w^l), \quad l = 0, 1, \dots, k-1, \\ (y, z) \in \mathcal{R}^0, \end{array} \right.$$

obtendo  $(y^k, z^k)$ ;

**Passo 2:** Encontre  $\theta(y^k, z^k)$ , obtendo  $(x^k, w^k)$ . Se  $\theta(y^k, z^k) \geq 0$ , então pare:  $(y^k, z^k)$  é uma solução ótima para o problema (5.3.16). Caso contrário,

$$\sum_{i=1}^p w_i^k z_i - \sum_{j=p+1}^{2p} w_j^k y_{j-p} \leq b(w^k),$$

é a restrição mais violada. Faça  $k \leftarrow k + 1$  e volte ao **Passo 1**.

A solução ótima do problema  $(P_R)$  no Passo 1 do Algoritmo PS-REL será tratada na próxima seção. A seguir, discutiremos o Passo 2 e a convergência do Algoritmo PS-REL.

As funções  $\theta$  e  $\psi_{y,z}$  (definidas em (5.3.18) e (5.3.19), respectivamente), apresentam propriedades importantes. O valor de  $\psi_{y,z}$  é calculado por meio da resolução de um problema convexo. Além disso, sendo o máximo pontual de funções lineares (indexadas por  $x \in \Omega$ ),  $\psi_{y,z}$  é uma função convexa [BVo4]. Portanto,  $\theta(y, z)$  é o valor ótimo de um problema convexo. Algumas propriedades adicionais de  $\theta$  são apresentados a seguir.

**Teorema 5.3.5.** *A função  $\theta$  definida por (5.3.18) é côncava e o seu valor ótimo é igual ao valor ótimo do problema convexo*

$$\left. \begin{array}{l} \text{maximizar } \sigma \\ \text{sujeito a } \end{array} \right\} \begin{array}{l} p(x) \leq y - \sigma e, \\ q(x) \geq z + \sigma e, \\ x \in \Omega, \end{array} \quad (5.3.21)$$

no qual  $e \in \mathbb{R}^m$  é o vetor  $m$ -dimensional de uns.

*Demonstração.* Dados quaisquer dois pontos  $(y^1, z^1), (y^2, z^2) \in \mathbb{R}^r \times \mathbb{R}^r$  e qualquer escalar  $\lambda \in [0, 1]$ , temos

$$\begin{aligned} \theta(\lambda(y^1, z^1) + (1 - \lambda)(y^2, z^2)) &= \min_{w \in \mathcal{W}} \psi_{\lambda y^1 + (1-\lambda)y^2, \lambda z^1 + (1-\lambda)z^2}(w) \\ &= \min_{w \in \mathcal{W}} \{ \lambda \psi_{y^1, z^1}(w) + (1 - \lambda) \psi_{y^2, z^2}(w) \} \\ &\geq \lambda \theta(y^1, z^1) + (1 - \lambda) \theta(y^2, z^2), \end{aligned}$$

onde a segunda igualdade vem do fato de que  $\psi_{y,z}$  é uma função linear em  $(y, z)$ . Conseqüentemente,  $\theta$  is côncava.

O problema dual Lagrangeano de (5.3.21) é

$$\min_{w \in \mathbb{R}_+^{2r}} \max_{x \in \Omega, \sigma \in \mathbb{R}} \left\{ \sigma + \sum_{i=1}^r w_i (q_i(x) - z_i - \sigma) + \sum_{j=r+1}^{2r} w_j (-p_{j-r}(x) + y_{j-r} - \sigma) \right\}, \quad (5.3.22)$$

no qual  $w = (w_1, w_2, \dots, w_{2r}) \in \mathbb{R}_+^{2r}$  é o vetor de variáveis duais associadas às restrições de desigualdade de (5.3.21). A existência do máximo impõe que

$$\sum_{k=1}^{2r} w_k = 1,$$

o que reduz o problema dual (5.3.22) ao problema min–max

$$\min_{w \in \mathcal{W}} \max_{x \in \Omega} \left\{ \sum_{i=1}^r w_i (q_i(x) - z_i) - \sum_{j=r+1}^{2r} w_j (p_{j-r}(x) - y_{j-r}) \right\}, \quad (5.3.23)$$

cujo valor ótimo é, por definição,  $\theta(y, z)$ . Como o problema primal (5.3.21) é convexo, sob condições de qualificação de restrições [Ber95], não há gap de dualidade e os problemas (5.3.21) e (5.3.23) têm o mesmo valor ótimo,  $\theta(y, z)$ . □

A cônicaidade de  $\theta$  implica na sua continuidade, que é um ingrediente importante da demonstração da convergência do Algoritmo PS–REL a um solução ótima do problema  $(P_S)$  (e, portanto, de  $(Q_{FG})$ ).

**Teorema 5.3.6.** *Qualquer ponto de acumulação  $(y^*, z^*)$  da sequência  $\{(y^k, z^k)\}$  gerada pelo Algoritmo PS–REL é uma solução ótima do problema de programação fracionária  $(P_S)$ .*

*Demonstração.* O problema relaxado  $(P_R)$  sempre tem uma solução ótima e a última desigualdade incorporada a  $(P_R)$  é

$$\sum_{i=1}^r w_i^k (q_i(x^k) - z_i) - \sum_{j=r+1}^{2r} w_j^k (p_{j-r}(x^k) - y_{j-r}) \geq 0,$$

e pode ser reescrita como

$$\begin{aligned} \sum_{i=1}^r w_i^k (z_i - z_i^k) - \sum_{j=r+1}^{2r} w_j^k (y_{j-r} - y_{j-r}^k) \\ \leq \sum_{i=1}^r w_i^k (q_i(x^k) - z_i^k) - \sum_{j=r+1}^{2r} w_j^k (p_{j-r}(x^k) - y_{j-r}^k) = \theta(y^k, z^k). \end{aligned} \quad (5.3.24)$$

Em qualquer iteração subsequente  $r > k$  do algoritmo, devemos ter

$$\begin{aligned} -\theta(y^k, z^k) &\leq \sum_{i=1}^r w_i^k (z_i^k - z_i^r) - \sum_{j=r+1}^{2r} w_j^k (y_{j-r}^k - y_{j-r}^r) \\ &\leq \|w^k\| \left\| \begin{bmatrix} z^k - z^r \\ -y^k + y^r \end{bmatrix} \right\| \\ &\leq \left\| \begin{bmatrix} z^k - z^r \\ -y^k + y^r \end{bmatrix} \right\|, \end{aligned}$$

porque  $\|w^k\| \leq 1$  para todo  $w^k \in \mathcal{W}$ . Como  $k \rightarrow \infty$ , obtemos  $(y^k, z^k) \rightarrow (y^*, z^*)$ ,  $(y^r, z^r) \rightarrow (y^*, z^*)$  e a continuidade de  $\theta$  em  $(y^*, z^*)$  fornece  $\theta(y^*, z^*) \geq 0$ . Portanto,  $(y^*, z^*) \in \mathcal{V} \cap \mathcal{R}^0$ , isto é,  $(y^*, z^*)$  é uma solução viável de  $(P_S)$ . Como a região viável do problema  $(P_R)$  é uma aproximação externa da região viável do problema  $(P_S)$  para todo  $k$ , concluímos que  $v^*(P_S) \leq \sum_{i=1}^r z_i^* / y_i^*$ . Consequentemente,  $(y^*, z^*)$  é uma solução ótima de  $(P_S)$ .  $\square$

Pelo Teorema 5.3.6, segue-se que para  $k$  suficientemente grande,

$$-\theta(y^k, z^k) \leq \epsilon_{rel}, \quad (5.3.25)$$

no qual  $\epsilon_{rel} > 0$  é qualquer tolerância (pequena) para a convergência finita do algoritmo de relaxação PS-REL.

O Teorema 5.3.5 fornece a racionalidade do critério de convergência (5.3.25). Para ver como, seja  $x^k$  uma solução ótima do problema (5.3.21) para  $y = y^k$  e  $z = z^k$  em qualquer iteração  $k$  do Algoritmo PS-REL. Pelo Teorema 5.3.5,  $p(x^k) \leq y^k - \theta(y^k, z^k)e$ ,  $q(x^k) \geq z^k + \theta(y^k, z^k)e$ , como a implementação de um procedimento de aproximação externa, o Algoritmo PS-REL gera uma sequência de pontos inviáveis  $\{(y^k, z^k)\}$  (isto é,  $\theta(y^k, z^k) < 0$  para todo  $k$ ) convergindo a uma solução viável (e então ótima) do problema  $(P_S)$ . Neste caso, o critério de convergência finita (5.3.25) garante que  $x^k$  viola cada desigualdade do

problema (5.3.21) por, no máximo,  $\epsilon_{rel}$ .

Toda a informação necessária no Passo 2 do Algoritmo PS-REL pode ser obtida resolvendo-se o problema primal (5.3.21) ou o seu problema dual Lagrangeano (5.3.23). As experiências numéricas apresentadas no fim deste capítulo estão baseadas na resolução do problema dual (5.3.23) por meio do método de planos de corte já discutido no Capítulo 4.

### 5.3.4 Algoritmo de Otimização Global para Resolução de $(Q_{FG})$

O problema  $(P_R)$  no Passo 1 do Algoritmo PS-REL possui  $2r$  variáveis,  $k + 4r$  restrições lineares de desigualdade e uma soma de razões elementares como função objetivo a ser maximizada. No entanto, problema  $(P_R)$  ainda reflete o caráter não-convexo do problema original  $(Q_{FG})$ .

Uma estratégia bem sucedida em otimização global para lidar com razões elementares tem sido substituir cada razão elementar por (no caso de maximização) seu envelope côncavo, definido como a função que melhor sobreestima a razão em um conjunto convexo e compacto. Para termos fracionários da forma  $z_i/y_i$ , temos

$$\frac{z_i}{y_i} \leq \min \left\{ \frac{z_i}{y_i^L} - \frac{z_i^L y_i}{y_i^L y_i^U} + \frac{z_i^L}{y_i^U}, \frac{z_i}{y_i^U} - \frac{z_i^U y_i}{y_i^L y_i^U} + \frac{z_i^U}{y_i^L} \right\} \quad (5.3.26)$$

para todo  $(z_i, y_i) \in [z_i^L, z_i^U] \times [y_i^L, y_i^U]$ . O lado direito de (5.3.26) é um envelope côncavo de  $z_i/y_i$  em  $[z_i^L, z_i^U] \times [y_i^L, y_i^U]$ .

Substituindo cada razão elementar que aparece na função objetivo de  $(P_S)$  por seu en-

velope côncavo, obtemos o seguinte problema de otimização linear semi-infinito:

$$(P_L) \quad \left\{ \begin{array}{l} \text{maximizar}_{y,z,\lambda} \quad \sum_{i=1}^r \lambda_i \\ \text{sujeito a} \quad \sum_{i=1}^r w_i z_i - \sum_{j=r+1}^{2r} w_j y_{j-r} \leq b(w) \quad \text{para todo } w \in \mathcal{W}, \\ \lambda_i \leq \frac{z_i}{y_i^L} - \frac{z_i^L y_i}{y_i^L y_i^U} + \frac{z_i^L}{y_i^U}, \quad i = 1, 2, \dots, r, \\ \lambda_i \leq \frac{z_i}{y_i^U} - \frac{z_i^U y_i}{y_i^L y_i^U} + \frac{z_i^U}{y_i^L}, \quad i = 1, 2, \dots, r, \\ (y, z) \in \mathcal{R}^0. \end{array} \right.$$

Após substituir cada razão elementar pelo seu envelope côncavo, observa-se que  $v^*(P_L) \geq v^*(P_S) = v^*(Q_{FG})$ . Note também que qualquer relaxação (após  $k$  iterações do Algoritmo PS-REL) do problema  $(P_L)$  é um problema de otimização linear com  $3r$  variáveis e  $k + 6r$  restrições de desigualdade.

A seguir desenvolvemos um algoritmo *branch-and-bound* para resolver globalmente o problema  $(P_S)$  (e portanto, o problema  $(Q_{FG})$ ) por meio de problema  $(P_L)$ . Em termos gerais, a região viável de  $(P_L)$  é sucessivamente subdividida em partes mais refinadas (*branching*), sobre as quais os limitantes superiores e inferiores do valor ótimo de  $(P_S)$  são determinados (*bounding*). Partes da região viável com limitantes inferiores excedendo o melhor limitante superior encontrado até a iteração atual do algoritmo são eliminadas de considerações posteriores (*pruning*), pois estas partes não podem conter uma solução ótima de  $(P_S)$ .

Em algoritmos *branch-and-bound* retangular, a região viável do problema é particionada em sub-retângulos. Um sub-retângulo  $\mathcal{R}$  de  $\mathcal{R}^0$  é um conjunto da forma (5.3.7) e limitantes  $y^L(\mathcal{R})$ ,  $y^U(\mathcal{R})$ ,  $z^L(\mathcal{R})$  e  $z^U(\mathcal{R})$  satisfazendo  $y^L(\mathcal{R}) \geq y^L$ ,  $y^U(\mathcal{R}) \leq y^U$ ,  $z^L(\mathcal{R}) \geq z^L$  e  $z^U(\mathcal{R}) \leq z^U$  (com o entendimento de que  $y^L(\mathcal{R}^0) = y^L$ , e assim por diante).

Sejam  $(P_L(\mathcal{R}))$  e  $(P_S(\mathcal{R}))$  problemas das formas  $(P_L)$  e  $(P_S)$  quando  $\mathcal{R}^0$  é substituído por  $\mathcal{R}$ ; seja  $(y(\mathcal{R}), z(\mathcal{R}), \lambda(\mathcal{R}))$  qualquer solução ótima do  $(P_L(\mathcal{R}))$ . O valor ótimo do  $(P_S(\mathcal{R}))$  é maior igual do que o limitante inferior  $\gamma(\mathcal{R}) := \sum_{i=1}^r z_i(\mathcal{R})/y_i(\mathcal{R})$ , porque  $(y(\mathcal{R}), z(\mathcal{R}))$  é viável para  $P_S(\mathcal{R})$ , e menor igual do que o limitante superior  $\mu(\mathcal{R}) := \sum_{i=1}^r \lambda_i(\mathcal{R})$ , pois  $(P_L(\mathcal{R}))$  é uma majoração de  $(P_S(\mathcal{R}))$ .

No algoritmo *branch-and-bound* proposto a seguir, Algoritmo PR-BB, sub-retângulos

são produzidos bisseccionando-se um dado retângulo no ponto médio de um dos seus lados (ao longo do seu maior lado).

**Algoritmo PR-BB:** Algoritmo *branch-and-bound* para Resolver  $(P_R)$

**Passo 0:** Resolva o problema  $(P_L(\mathcal{R}^0))$ , obtendo uma solução ótima  $(y(\mathcal{R}^0), z(\mathcal{R}^0), \lambda(\mathcal{R}^0))$ .

Faça  $\gamma^0 := \gamma(\mathcal{R}^0)$ ,  $\mu^0 := \mu(\mathcal{R}^0)$ ,  $\mathcal{L}^0 := \{\mathcal{R}^0\}$  e  $q = 0$ ;

**Passo 1:** Se  $\mu^q = \gamma^q$ , então pare. A solução incumbente  $(y^q, z^q)$  é uma solução ótima para  $(P_S)$ ;

**Passo 2:** Encontre  $\mathcal{R} \in \mathcal{L}^q$  tal que  $\mu(\mathcal{R}) = \mu^q$ . Divida  $\mathcal{R}$  em sub-retângulos  $\mathcal{R}^I$  e  $\mathcal{R}^{II}$  e defina

$$\mathcal{L}^{q+1} := (\mathcal{L}^q \setminus \{\mathcal{R}\}) \cup \{\mathcal{R}^I, \mathcal{R}^{II}\}.$$

Compute  $\gamma(\mathcal{R}^I)$ ,  $\mu(\mathcal{R}^I)$ ,  $\gamma(\mathcal{R}^{II})$  e  $\mu(\mathcal{R}^{II})$ . Elimine todos os sub-retângulos  $\mathcal{R} \in \mathcal{L}^{q+1}$  tais que  $(P_L(\mathcal{R}))$  seja inviável ou  $\gamma(\mathcal{R}) > \mu^q$ ;

**Passo 3:** Encontre  $\mathcal{R}^* \in \arg \max_{\mathcal{R} \in \mathcal{L}^{q+1}} \gamma(\mathcal{R})$  e faça  $(y^{q+1}, z^{q+1}) := (y(\mathcal{R}^*), z(\mathcal{R}^*))$ ,  $\gamma^{q+1} := \gamma(\mathcal{R}^*)$ ,  $\mu^{q+1} := \max_{\mathcal{R} \in \mathcal{L}^{q+1}} \mu(\mathcal{R})$ ,  $q := q + 1$  e volte ao **Passo 1**.

As regras de branching, bounding e pruning do Algoritmo PR-BB preenchem todas as condições necessárias para a convergência de algoritmos de *branch-and-bound* [HPT95]. Assim, qualquer ponto de acumulação  $(y^*, z^*)$  da sequência  $\{(y^q, z^q)\}$  gerada pelo Algoritmo PR-BB resolve o problema de programação fracionária  $(P_S)$ . Resultados de convergência também garantem que para  $q$  suficientemente grande,

$$\mu^q - \gamma^q \leq \epsilon_{bb},$$

no qual  $\epsilon_{bb} > 0$  é qualquer tolerância (pequena) para a convergência finita do Algoritmo PR-BB.

Para a discussão de um aspecto importante da implementação do algoritmo PR-BB por meio do algoritmo de relaxação PS-REL da seção anterior, considere o problema linear  $(P_L(\mathcal{R}))$ .

Quando  $\mathcal{R}$  é dividido em  $\mathcal{R}^I$  e  $\mathcal{R}^{II}$ , os problemas  $(P_L(\mathcal{R}^I))$  e  $(P_L(\mathcal{R}^{II}))$  herdam as restrições geradas para obter uma solução ótima de  $(P_L(\mathcal{R}))$ , e então o Algoritmo PS-REL gera apenas os cortes adicionais necessários para encontrar suas soluções ótimas – a solução

ótima de  $(P_L(\mathcal{R}))$  é geralmente um bom ponto de partida para  $(P_L(\mathcal{R}^I))$  ou  $(P_L(\mathcal{R}^{II}))$ . Esta estratégia acelera os procedimentos de bounding e pruning do Algoritmo PR-BB, como ilustrado na Seção 5.4.

### 5.3.5 Extensão para Problemas Fracionários Generalizados $(P_{FG})$

A técnica de otimização global proposta é extensível a problemas de programação fracionária generalizada convexa-côncava da forma  $(P_{FG})$ . Em particular, o problema  $(P_{FG})$  é equivalente a

$$\left\{ \begin{array}{l} \text{minimizar}_{x,y,z} \quad \sum_{i=1}^p \frac{y_i}{z_i} \\ \text{sujeito a} \quad p(x) \leq y, \\ \quad \quad \quad q(x) \geq z, \\ \quad \quad \quad x \in \Omega, (y, z) \in \mathcal{R}^0. \end{array} \right. \quad (5.3.27)$$

Usando técnicas análogas de projeção e de relaxação, o problema  $(P_{FG})$  é reformulado e resolvido no espaço-imagem. O problema no espaço-imagem é

$$(P_T) \left\{ \begin{array}{l} \text{minimizar}_{y,z} \quad \sum_{i=1}^r \frac{y_i}{z_i} \\ \text{sujeito a} \quad \sum_{i=1}^r w_i y_i - \sum_{j=r+1}^{2r} w_j z_{j-r} \geq b(w) \quad \text{para todo } w \in \mathcal{W}, \\ \quad \quad \quad (y, z) \in \mathcal{R}^0. \end{array} \right. \quad (5.3.28)$$

no qual

$$b(w) := \min_{x \in \Omega} \left\{ \sum_{i=1}^r w_i p_i(x) - \sum_{j=r+1}^{2r} w_j q_{j-r}(x) \right\},$$

e  $(y, z) \in \mathcal{R}^0$  satisfaz a desigualdade semi-infinita se e somente se

$$\beta(y, z) := \max_{w \in \mathcal{W}} \min_{x \in \Omega} \left\{ \sum_{i=1}^r w_i (p_i(x) - y_i) - \sum_{j=r+1}^{2r} w_j (q_{j-r}(x) - z_i) \right\} \leq 0.$$

A convergência finita do algoritmo de relaxação é estabelecida quando

$$\beta(y, z) \leq \epsilon_{rel}$$

onde  $\epsilon_{rel} > 0$  é uma pequena tolerância.

Os envelopes convexos para os termos fracionários  $y_i/z_i$  da função objetivo de  $(P_T)$  assumem a forma

$$\frac{y_i}{z_i} \geq \max \left\{ \frac{y_i^L}{z_i} + \frac{y_i}{z_i^U} - \frac{y_i^L}{z_i^U}, \frac{y_i^U}{z_i} + \frac{y_i}{z_i^L} - \frac{y_i^U}{z_i^L} \right\} \quad (5.3.29)$$

para todo  $(y_i, z_i) \in [y_i^L, y_i^U] \times [z_i^L, z_i^U]$ . O lado direito de (5.3.29) é um envelope convexo de  $y_i/z_i$  em  $[y_i^L, y_i^U] \times [z_i^L, z_i^U]$ . O algoritmo *branch-and-bound* PR-BB pode ser facilmente modificado para resolver globalmente o problema  $(P_T)$ . Um limitante inferior para o valor ótimo de  $(P_T)$  é encontrado resolvendo-se o problema convexo semi-infinito com  $3r$  variáveis

$$(P_C) \quad \left\{ \begin{array}{l} \text{minimizar}_{y,z,\lambda} \quad \sum_{i=1}^r \lambda_i \\ \text{sujeito a} \quad \sum_{i=1}^r w_i y_i - \sum_{j=r+1}^{2r} w_j z_{j-r} \geq b(w) \quad \text{para todo } w \in \mathcal{W}, \\ \lambda_i \geq \frac{y_i^L}{z_i} + \frac{y_i}{z_i^U} - \frac{y_i^L}{z_i^U}, \quad i = 1, 2, \dots, r, \\ \lambda_i \geq \frac{y_i^U}{z_i} + \frac{y_i}{z_i^L} - \frac{y_i^U}{z_i^L}, \quad i = 1, 2, \dots, r, \\ (y, z) \in \mathcal{R}^0 \end{array} \right.$$

para o qual cada razão elementar foi substituída pelo seu envelope convexo. Observa-se que  $v^*(P_C) \leq v^*(P_T) = v^*(P_{FG})$ .

## 5.4 Experimentos Numéricos

O desempenho computacional dos algoritmos de otimização global propostos neste capítulo para resolver globalmente os problemas de programação fracionária generalizada convexa-côncava e côncava-convexa  $(P_{FG})$  e  $(Q_{FG})$ , respectivamente, foram avaliados com base em problemas selecionados da literatura de programação fracionária.

Os Algoritmos PFG-ENU, PS-REL, PR-BB e o procedimento PT-PC (para problemas-testes) foram implementados em MATLAB versão 7.0.1/*Optimization Toolbox* versão 4, sobre uma plataforma computacional composta por um computador Intel Core Duo, 3,33 GHz

com 4GB de memória RAM, 32 bits. As tolerâncias adotadas para as convergências finitas desses algoritmos foram fixadas em  $\epsilon_{en} = 0,00001$ ,  $\epsilon_{rel} = 0,00001$  e  $\epsilon_{bb} = 0,005$ , respectivamente, a menos de especificações em contrário.

### 5.4.1 Resultados Computacionais do Algoritmo PFG-ENU

Vale lembrar que para problemas de programação fracionária ( $P_{FG}$ ), além dos exemplos encontrados na literatura de programação fracionária, temos uma ampla escolha de problemas-testes que podem ser gerados pelo procedimento PT-PC.

**Exemplo 5.4.1.** *O problema a seguir foi utilizado para ilustrar os algoritmo alternativos propostos em [FP94], [Beno7a].*

$$\left\{ \begin{array}{l} \underset{x_1, x_2}{\text{minimizar}} \quad \frac{-x_1 + 2,000x_2 + 2,000}{3,000x_1 - 4,000x_2 + 5,000} + \frac{4,000x_1 - 3,000x_2 + 4,000}{-2,000x_1 + x_2 + 3,000} \\ \text{sujeito a} \quad x_1 + x_2 \leq 1,5000, \\ \quad \quad \quad x_1 - x_2 \leq 0,0000, \\ \quad \quad \quad 0,0000 \leq x_1, x_2 \leq 1,0000. \end{array} \right. \quad (5.4.1)$$

As funções  $p_1, \tilde{q}_1, p_2$  e  $\tilde{q}_2$  são positivas em  $\Omega$ . Os limitantes inferiores e superiores de  $(y_p, y_{\tilde{q}})$  são:

$$\begin{aligned} (\underline{y}_p, \underline{y}_{\tilde{q}}) &= (2,0000; 0,2000; 1,0000; 0,25000), \\ (\bar{y}_p, \bar{y}_{\tilde{q}}) &= (4,0000; 1,0000; 4,75000; 0,4444), \end{aligned}$$

respectivamente. O Algoritmo PFG-ENU convergiu após 9 iterações para a solução  $\epsilon_{en}$ -ótima  $x^* = (0,0000; 0,2829)$ , com valor ótimo  $v^* = 1,6232$ . A Tabela 5.1 apresenta os dados produzidos pelo Algoritmo PFG-ENU ao resolver o problema (5.4.1).

As soluções relatadas em [FP94] e [Beno7a] são  $x^* = (0,0000; 0,2839)$  (com o valor ótimo igual a  $v^* = 1,6240$ ) e  $x^* = (0,0000; 0,2679)$  (com o valor ótimo igual a  $v^* = 1,6236$ , encontrada na 18ª iteração do algoritmo alternativo proposto com  $\epsilon = 0,0001$ ), respectivamente.

Como outras experiências numéricas, considere os seguintes problemas-testes gerados pelo Procedimento PT-PC proposto.

Tabela 5.1: Convergência do Algoritmo PFG-ENU – Exemplo 5.4.1

$k$	$y_{p_1}^k, y_{\tilde{q}_1}^k, y_{p_2}^k, y_{\tilde{q}_2}^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(2,0000; 0,2000; 1,0000; 0,2500)	(0,6000; 0,0000; 0,4000; 0,0000)	(0,0000; 1,0000)	1,2000
1	(4,0000; 0,2000; 1,0000; 0,2500)	(0,0000; 0,6513; 0,3487; 0,0000)	(0,0000; 0,8554)	0,4336
2	(3,1709; 0,2000; 2,2436; 0,2500)	(0,0000; 0,8597; 0,1403; 0,0000)	(0,0000; 0,5354)	0,1499
3	(2,4589; 0,2000; 3,3117; 0,2500)	(0,0450; 0,0000; 0,0000; 0,9550)	(0,0000; 0,2585)	0,0569
4	(2,4589; 0,2000; 3,3117; 0,3096)	(0,0000; 0,9278; 0,0722; 0,0000)	(0,0000; 0,2153)	0,0417
5	(2,6945; 0,2724; 2,9582; 0,2985)	(0,0000; 0,9075; 0,0925; 0,0000)	(0,0000; 0,3459)	0,0041
6	(2,5701; 0,2579; 3,1448; 0,3044)	(0,0000; 0,9179; 0,0821; 0,0000)	(0,0000; 0,2846)	0,0011
7	(2,5723; 0,2594; 3,1415; 0,3043)	(0,0442; 0,0000; 0,0000; 0,9558)	(0,0000; 0,2884)	2,1818e-005
8	(2,5661; 0,2585; 3,1508; 0,3046)	(0,0000; 0,9182; 0,0818; 0,0000)	(0,0000; 0,2829)	6,4854e-007

Valor  $\epsilon_{en}$ -ótimo = 1,6232

**Exemplo 5.4.2.**

$$\begin{array}{l}
 \text{minimizar}_{x_1, x_2} \quad \frac{x_1 + 3x_2 + 2}{4x_1 + x_2 + 3} + \frac{4x_1 + 3x_2 + 1}{x_1 + x_2 + 4} \\
 \text{sujeito a} \quad 3x_1^2 + x_2^2 \leq 48, \\
 \quad \quad \quad x_1 + x_2 \geq 1, \\
 \quad \quad \quad x_1, x_2 \geq 0.
 \end{array} \tag{5.4.2}$$

**Exemplo 5.4.3.**

$$\begin{array}{l}
 \text{minimizar}_{x_1, x_2} \quad \frac{2x_1 + 4x_2 + 1}{2x_1 + 2x_2 + 2} + \frac{x_1 + 4x_2 + 1}{3x_1 + 2x_2 + 2} \\
 \text{sujeito a} \quad 3x_1^2 + x_2^2 \leq 48, \\
 \quad \quad \quad x_1 + x_2 \geq 1, \\
 \quad \quad \quad x_1, x_2 \geq 0.
 \end{array} \tag{5.4.3}$$

**Exemplo 5.4.4.**

$$\begin{array}{l}
 \text{minimizar}_{x_1, x_2} \quad \frac{3x_1 + 3x_2 + 3}{x_1 + 2x_2 + 1} + \frac{3x_1 + 4x_2 + 8}{5x_1 + 5x_2 + 3} + \frac{7x_1 + 2x_2 + 4}{x_1 + 5x_2 + 7} \\
 \text{sujeito a} \quad 3x_1^2 + x_2^2 \leq 48, \\
 \quad \quad \quad x_1 + x_2 \geq 1, \\
 \quad \quad \quad x_1, x_2 \geq 0.
 \end{array} \tag{5.4.4}$$

O conjunto  $\Omega \subseteq \{x \in \mathbb{R}^2 : x \geq 0, x \neq 0\}$  nos Exemplos 5.4.2–5.4.4 é um conjunto convexo, compacto e não-vazio. As funções  $p_1, \tilde{q}_1, p_2, \tilde{q}_2$ , nos Exemplos 5.4.2–5.4.3, e as funções  $p_1, \tilde{q}_1, p_2, \tilde{q}_2, p_3, \tilde{q}_3$ , no Exemplo 5.4.3, são todas positivas em  $\Omega$ . Por meio de aplicação de uma busca do tipo bisseção, encontramos  $x^* = (1,0000; 0,0000)$  para o Exemplo 5.4.2 (após 17 bisseções),  $x^* = (1,0000; 0,0000)$  para o Exemplo 5.4.3 (após 14 bisseções) e  $x^* = (0,0000; 6,9276)$  para o Exemplo 5.4.4 (após 19 bisseções). As Tabelas 5.2–5.4 apresen-

tam os dados produzidos pelo Algoritmo PFG-ENU ao resolver os Exemplos 5.4.2-5.4.4.

Tabela 5.2: Convergência do Algoritmo PFG-ENU – Exemplo 5.4.2

$k$	$y_{p_1}^k, y_{q_1}^k, y_{p_2}^k, y_{q_2}^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(3,0000; 0,0489; 4,0000; 0,0833)	(0,3333; 0,0000; 0,6667; 0,0000)	(0,0000; 1,0000)	0,6667
1	(3,0000; 0,0489; 5,0000; 0,0833)	(0,0043; 0,0000; 0,0085; 0,9872)	(1,1505; 0,0000)	0,1152
2	(5,0000; 0,0489; 4,0000; 0,2000)	(0,0000; 0,8708; 0,1292; 0,0000)	(0,1580; 0,8420)	0,1724
3	(3,0000; 0,0489; 5,3349; 0,1971)	(0,0682; 0,9318; 0,0000; 0,0000)	(1,0840; 0,0000)	0,0871
4	(4,2787; 0,0489; 5,3349; 0,1916)	(0,0000; 0,9822; 0,0178; 0,0000)	(1,1183; 0,0000)	0,0858
5	(3,0010; 0,1424; 4,9995; 0,2000)	(0,0000; 0,9413; 0,0587; 0,0000)	(0,9590; 0,0410)	4,7564e-004
6	(3,0000; 0,1425; 5,0064; 0,1999)	(0,0744; 0,9256; 0,0000; 0,0000)	(1,0272; 0,0000)	3,7073e-004
7	(3,0000; 0,1429; 5,0000; 0,2000)	(0,0383; 0,0000; 0,0000; 0,9617)	(1,0000; 0,0000)	1,3745e-006
Valor $\epsilon_{en}$ -ótimo = 1,4286				

Tabela 5.3: Convergência do Algoritmo PFG-ENU – Exemplo 5.4.3

$k$	$y_{p_1}^k, y_{q_1}^k, y_{p_2}^k, y_{q_2}^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(3,0000; 0,0556; 2,0000; 0,0492)	(0,0474; 0,9526; 0,0000; 0,0000)	(1,5134; 0,0000)	0,1852
1	(6,8512; 0,0556; 2,0000; 0,0492)	(0,0000; 0,8997; 0,1003; 0,0000)	(1,0000; 0,0000)	0,1749
2	(6,8512; 0,0556; 3,7368; 0,0492)	(0,0000; 0,9702; 0,0298; 0,0000)	(2,2554; 0,0000)	0,0808
3	(6,8512; 0,0556; 6,2604; 0,0492)	(0,0214; 0,9786; 0,0000; 0,0000)	(1,8392; 0,0000)	0,0716
4	(9,9921; 0,0556; 6,2604; 0,0492)	(0,0044; 0,9956; 0,0000; 0,0000)	(4,0000; 0,0000)	0,0399
5	(7,7059; 0,1055; 4,6359; 0,0492)	(0,0086; 0,0000; 0,0000; 0,9914)	(2,5599; 0,0000)	0,0400
6	(3,0000; 0,2471; 2,0192; 0,1232)	(0,0357; 0,0000; 0,0000; 0,9643)	(2,0350; 0,0000)	0,0740
7	(3,0000; 0,2471; 2,0192; 0,1908)	(0,0450; 0,0000; 0,0000; 0,9550)	(1,0000; 0,0000)	0,0088
8	(3,0000; 0,2471; 2,0192; 0,1974)	(0,0534; 0,9466; 0,0000; 0,0000)	(1,0000; 0,0000)	5,8001e-006
Valor $\epsilon_{en}$ -ótimo = 1,1500				

Observe que o Algoritmo PFG-ENU convergiu após 8, 9 e 15 iterações para as soluções  $\epsilon_{en}$ -ótimas  $x^* = (1,0000; 0,0000)$ ,  $x^* = (1,0000; 0,0000)$  e  $x^* = (0,0000; 6,9282)$  dos problemas 5.4.2, 5.4.3 e 5.4.4, respectivamente, as mesmas soluções encontradas por meio da aplicação da busca por bisseção.

### 5.4.2 Resultados Computacionais dos Algoritmos PR-BB

Nesta seção são apresentadas algumas experiências numéricas que permitem avaliar o desempenho computacional da técnica *branch-and-bound* proposta neste capítulo, os algoritmos PR-BB, para resolver globalmente problemas de programação fracionária generalizada ( $Q_{FG}$ ) e ( $P_{FG}$ ). O número de bisseções e o número de cortes mais profundos gerados pelos algoritmos são os parâmetros mais importantes da análise.

Tabela 5.4: Convergência do Algoritmo PFG-ENU – Exemplo 5.4.4

$k$	$y_{p_1}^k, y_{q_1}^k, y_{p_2}^k, y_{q_2}^k, y_{p_3}^k, y_{q_3}^k$	$w^k$	$x(w^k)$	$\theta(y^k)$
0	(6,0000 ; 0,0649; 11,0000; 0,0233; 45; 0,0239)	(0,0010; 0,8515;0,1521;0;0;0)	(0,6397; 0,3603)	0,3628
1	(6,0000 ; 0,0649; 13,3542; 0,0233; 45; 0,0239)	(0,0623; 0,9377; 0; 0; 0; 0)	(0,0000; 1,0841)	0,2508
2	(10,0237; 0,0649; 13,3842; 0,0233; 45; 0,0239)	(0; 0,9668;0,0332; 0; 0; 0)	(0,0000; 1,3945)	0,1990
3	(10,0237; 0,0649; 19,2916; 0,0233; 45; 0,0239)	(0,0199; 0,9799; 0; 0; 0; 0)	(0,0000; 2,3765)	0,1081
4	(15,5325; 0,0649; 19,2916; 0,0233; 45; 0,0239)	(0,0649; 0,9351; 0; 0; 0; 0)	(0,0000; 2,8482)	4,8101e-004
5	(15,5325; 0,0649; 26,8026; 0,0233; 45; 0,0239)	(0,0157; 0,9843; 0; 0; 0; 0)	(0,0000; 4,2802)	2,7612e-005
6	(21,1003; 0,0649; 26,8026; 0,0233; 45; 0,0239)	(0,0223; 0,0005;0,9772;0;0;0)	(0,0000; 4,6094)	8,2117e-004
7	(21,1003; 0,0649; 33,5988; 0,0233; 45; 0,0239)	(0,0039; 0,9961; 0; 0; 0; 0)	(0,0000; 6,0382)	0,0115
8	(17,9158; 0,0889; 28,3752; 0,0233; 45; 0,0239)	(0; 0,9901; 0;0,0099; 0; 0)	(0,8314; 4,3237)	7,4596e-004
9	(24,0721; 0,0649; 33,5983; 0,0233; 45; 0,0239)	(0,0016; 0,9983; 0; 0; 0; 0)	(0,0000; 6,0372)	1,0744e-004
10	(23,6338; 0,0666; 35,8368; 0,0233; 45; 0,0239)	(0,0012; 0; 0;0,9988; 0; 0)	(0,0000; 6,8784)	0,0035
11	(24,0731; 0,0649; 36,4792; 0,0262; 45; 0,0239)	(0,0014; 0,9516;0;0;0,0469)	(0,1276; 6,9247)	0,0019
12	(23,2768; 0,0680; 35,3148; 0,0272; 45; 0,0239)	(0,0312; 0,9688; 0; 0; 0; 0)	(0,0000; 6,9282)	5,9506e-005
13	(23,7829; 0,0673; 35,5892; 0,0266; 45; 0,0239)	(0,0019; 0,9981; 0; 0; 0; 0)	(0,0000; 6,8155)	2,8593e-004
14	(23,7817; 0,0673; 35,7112; 0,0266; 45; 0,0239)	(0,0010; 0; 0; 0; 0; 0,9990)	(0,0000; 6,9282)	9,4066e-006

Valor  $\epsilon_{en}$ -ótimo = 2,9786

**Exemplo 5.4.5.** (Exemplo 6.1 em [Benozza])

$$\begin{array}{l}
 \text{maximizar}_{x_1, x_2} \quad \frac{2x_1 + x_2}{x_1} + \frac{2}{x_2} \\
 \text{sujeito a} \quad 2x_1 + x_2 \leq 6, \\
 \quad \quad \quad 3x_1 + x_2 \leq 8, \\
 \quad \quad \quad -x_1 + x_2 \geq -1, \\
 \quad \quad \quad x_1, x_2 \geq 1.
 \end{array} \tag{5.4.5}$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema simples de soma de razões após apenas uma bisseção; um único corte mais profundo (introduzido no retângulo inicial  $\mathcal{R}^0$ ) foi necessário. O Algoritmo PR-BB convergiu para a solução  $x^* = (1,0000; 4,0000)$ ,  $v^*(Q_{FG}) = 6,5000$  e, na convergência,  $\mu^1 - \gamma^1 = 6,5000 - 6,5000 = 0,0000 \leq \epsilon_{bb}$ . A mesma solução é relatada em [Benozza] com  $\epsilon_{bb} = 0,05$ , após 11 iterações.

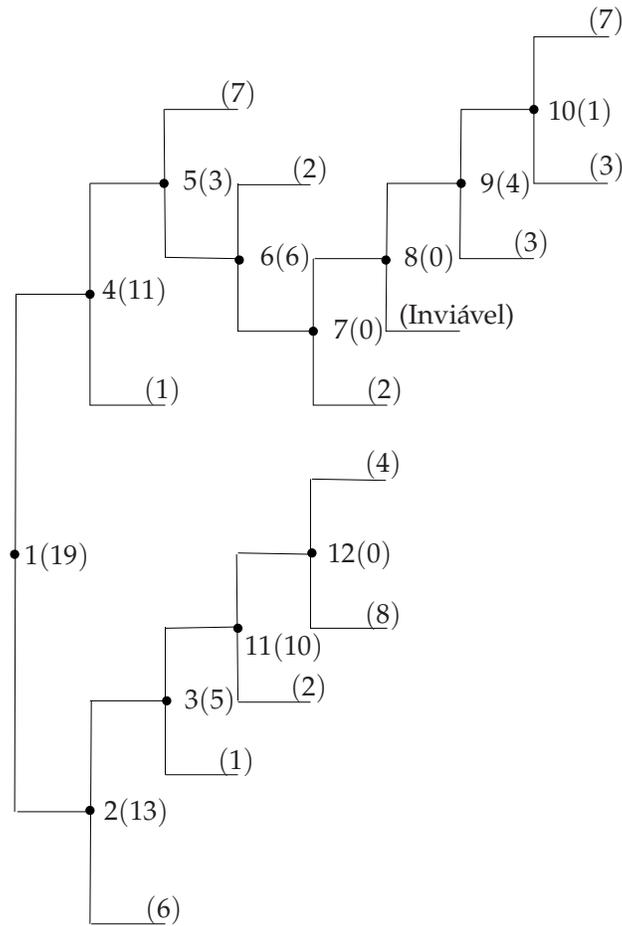


Figura 5.1: Árvore *branch-and-bound* para o Exemplo 5.4.6.

**Exemplo 5.4.6.** (Exemplo 6.2 em [Benozzi]).

$$\begin{array}{l}
 \left. \begin{array}{l}
 \text{maximizar}_{x_1, x_2} \quad \frac{-x_1^2 + 3x_1 - x_2^2 + 3x_2 + 3.5}{x_1 + 1} + \frac{x_2}{x_1^2 - 2x_1 + x_2^2 - 8x_2 + 20} \\
 \text{sujeito a} \quad 2x_1 + x_2 \leq 6, \\
 \quad \quad \quad 3x_1 + x_2 \leq 8, \\
 \quad \quad \quad -x_1 + x_2 \geq -1, \\
 \quad \quad \quad x_1, x_2 \geq 1.
 \end{array} \right\} \quad (5.4.6)
 \end{array}$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada não-linear após 12 bisseções, como ilustrado na Figura 5.1. O número de cortes mais profundos adicionais necessários em cada subproblema é informado entre parênteses precedido pela ordem de bisseção.

Após os 19 cortes gerados no Passo 0 do Algoritmo PR-BB, o número de cortes adicionais tende a diminuir à medida que partições mais refinadas são necessárias. O algoritmo convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (1,0000; 1,7413)$ ,  $v^*(Q_{FG}) = 4,0608$ , um pouco melhor do que a solução encontrada em [Benoz], ( $x^* = (1,0020; 1,7100)$ ,  $v^*(Q_{FG}) = 4,0319$ , obtida com  $\epsilon_{bb} = 0,05$ , após 27 iterações). A solução  $\epsilon_{bb}$ -ótima foi descoberta pelo Algoritmo PR-BB após 11 iterações, e na convergência  $\mu^{11} - \gamma^{11} = 4,0619 - 4,0606 = 0,0013 \leq \epsilon_{bb}$ .

**Exemplo 5.4.7.** (Exemplo 3.1 em [Benoz]).

$$\left\{ \begin{array}{l} \text{maximizar}_{x_1, x_2} \quad \frac{-x_1 + 2x_2 + 2}{3x_1 - 4x_2 + 5} + \frac{4x_1 - 3x_2 + 4}{-2x_1 + x_2 + 3} \\ \text{sujeito a} \quad 2x_1 + x_2 \leq 6, \\ \quad \quad \quad x_1 + x_2 \leq 1,5, \\ \quad \quad \quad x_1 - x_2 \leq 0, \\ \quad \quad \quad 0 \leq x_1, x_2 \leq 1. \end{array} \right. \quad (5.4.7)$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada linear após uma única bisseção; três cortes mais profundos foram necessários. O Algoritmo PR-BB convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (0,0000; 1,0000)$ ,  $v^*(Q_{FG}) = 4,2500$ , e na convergência  $\mu^1 - \gamma^1 = 4,2500 - 4,2500 = 0,0000 \leq \epsilon_{bb}$ . A mesma solução é relatada em [Benoz] com  $\epsilon_{bb}$  não explicitamente indicada, após 11 iterações.

**Exemplo 5.4.8.** ([DHT01], página 463).

$$\left\{ \begin{array}{l} \text{maximizar}_{x_1, x_2, x_3} \quad \frac{3x_1 + x_2 - 2x_3 + 0.8}{2x_1 - x_2 + x_3} + \frac{4x_1 - 2x_2 + x_3}{7x_1 + 3x_2 - x_3} \\ \text{sujeito a} \quad -x_1 - x_2 + x_3 \geq -1, \\ \quad \quad \quad x_1 - x_2 + x_3 \geq 1, \\ \quad \quad \quad -12x_1 - 5x_2 - 12x_3 \geq -34.8, \\ \quad \quad \quad 12x_1 - 12x_2 - 7x_3 \geq -29.1, \\ \quad \quad \quad 6x_1 - x_2 - x_3 \geq 4.1, \\ \quad \quad \quad x_1, x_2, x_3 \geq 0. \end{array} \right. \quad (5.4.8)$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada linear após 17 iterações; treze cortes mais profundos foram necessários. O algoritmo convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (1,0000; 0,0000; 0,0000)$ ,



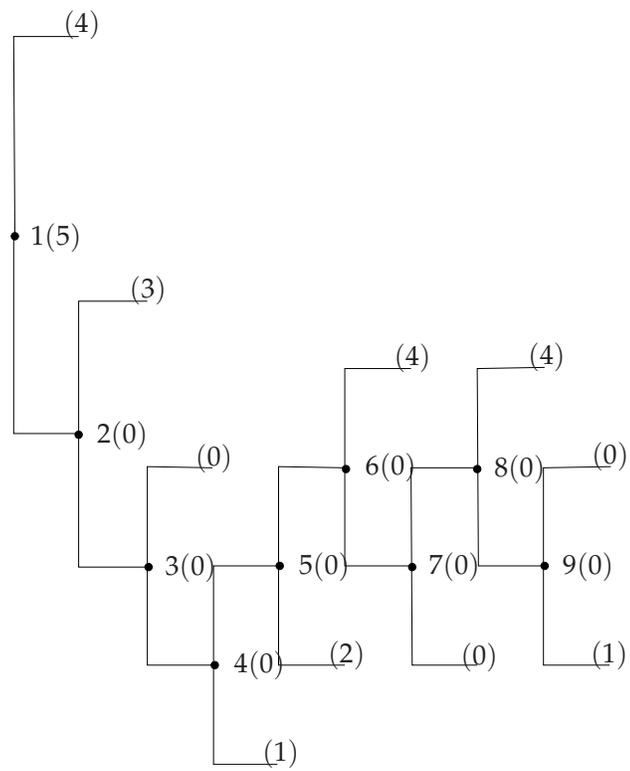


Figura 5.2: Árvore de *branch-and-bound* para o Exemplo 5.4.9.

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada linear na iteração 74; quatro cortes mais profundos foram necessários. O algoritmo convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (3,0000; 4,0000)$ ,  $v^*(Q_{FG}) = 5,0000$ . A mesma solução é relatada em [WSo8] com  $\epsilon_{bb} = 0,0001$  após 32 iterações. Na convergência,  $\mu^{74} - \gamma^{74} = 5,0031 - 4,9984 = 0,0047 \leq \epsilon_{bb}$ ; três cortes mais profundos foram gerados.

**Exemplo 5.4.11.** (Exemplo 1 em [Kuno2]).

$$\begin{array}{l}
 \left. \begin{array}{l}
 \text{maximizar}_{x_1, x_2, x_3} \quad \frac{4x_1 + 3x_2 + 3x_3 + 50}{3x_2 + 3x_3 + 50} + \frac{3x_1 + 4x_3 + 50}{4x_1 + 4x_2 + 5x_3 + 50} \\
 \quad + \frac{x_1 + 2x_2 + 5x_3 + 50}{x_1 + 5x_2 + 5x_3 + 50} + \frac{x_1 + 2x_2 + 4x_3 + 50}{5x_2 + 4x_3 + 50} \\
 \text{sujeito a} \quad 2x_1 + x_2 + 5x_3 \leq 10, \\
 \quad x_1 + 6x_2 + 3x_3 \leq 10, \\
 \quad 5x_1 + 9x_2 + 2x_3 \leq 10, \\
 \quad 9x_1 + 7x_2 + 3x_3 \leq 10, \\
 \quad x_1, x_2, x_3 \geq 0.
 \end{array} \right\} \quad (5.4.11)
 \end{array}$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada linear na iteração 10, após a geração de doze cortes mais profundos. O algoritmo convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (1,1111; 0,0000; 0,0000)$ ,  $v^*(Q_{FG}) = 4,0907$ , um aumento significativo no que diz respeito à solução encontrada em [Kuno2], ( $x^* = (0,0000; 1,1111; 0,0000)$ ,  $v^*(Q_{FG}) = 3,7984$ , obtida com  $\epsilon_{bb} = 0,001$ , após 31 iterações). Na convergência;  $\mu^{10} - \gamma^{10} = 4,0949 - 4,0907 = 0,0042 \leq \epsilon_{bb}$ ; dez cortes mais profundos foram necessários.

**Exemplo 5.4.12.** (Exemplo 2 em [Benozb]).

$$\begin{array}{l}
 \left. \begin{array}{l}
 \text{maximizar}_{x_1, x_2, x_3} \quad \frac{-x_1^2 + 4x_1 - 2x_2^2 + 8x_2 - 3x_3^2 + 12x_3 + 56}{-2x_1^2 + 16x_1 - x_2^2 + 8x_2 + 2} \\
 \quad + \frac{x_1^2 - 2x_1 + x_2^2 - 2x_2 + x_3 + 20}{2x_1 + 4x_2 + 6x_3} \\
 \text{sujeito a} \quad x_1 + x_2 + x_3 \leq 10, \\
 \quad -x_1 - x_2 + x_3 \leq 4, \\
 \quad x_1, x_2, x_3 \geq 1.
 \end{array} \right\} \quad (5.4.12)
 \end{array}$$

O Algoritmo PR-BB convergiu para uma solução  $\epsilon_{bb}$ -ótima deste problema de programação fracionária generalizada linear na iteração 59, após a geração de 250 cortes mais profundos. O algoritmo convergiu para a solução  $\epsilon_{bb}$ -ótima  $x^* = (1,8196; 0,0000; 0,0000)$ ,  $v^*(Q_{FG}) = 6,1198$ . Na convergência,  $\mu^{59} - \gamma^{59} = 6,1247 - 6,1198 = 0,0049 \leq \epsilon_{bb}$ ; 219 cortes mais profundos foram necessários. A mesma solução é relatada em [Benozb] para  $\epsilon_{bb} = 0,01$ , após 24 iterações.

A seguir são apresentados alguns experimentos numéricos envolvendo soma de razões convexo-côncavas. Os problemas (5.4.2), (5.4.3) e (5.4.4) são utilizados para fins da avaliação do algoritmo *branch-and-bound* PR-BB com maximizações substituídas por minimizações. Estes problemas já foram resolvidos por meio do algoritmo alternativo que utiliza enumeração de restrições.

O Algoritmo PR-BB encontra as mesmas soluções ótimas obtidas por meio do Algoritmo PFG-ENU. A Tabela 5.5 resume os resultados obtidos pelo algoritmo FMG-BB ao resolver os problemas (5.4.2)–(5.4.4). Pelos resultados encontrados, percebe-se que o algoritmo FMG-BB tem um desempenho computacional melhor quando comparado ao algoritmo FMG-ENU.

Tabela 5.5: Convergência do Algoritmo para  $(P_{FG})$  – Problemas 5.4.2–5.4.4.

Problema	Número de iterações $k$	Solução $\epsilon_{bb}$ -ótimo $x^*$	Valor $\epsilon_{bb}$ -ótimo $v^*(P_{FG})$	Número de cortes mais profundo	Número de cortes gerados na convergência	$\mu^k - \gamma^k$
(5.4.2)	4	(1,0000; 0,0000)	1,4286	10	6	1,4329-1,4286
(5.4.3)	6	(1,0000; 0,0000)	1,1500	17	13	1,1547-1,1501
(5.4.4)	8	(0,0000; 6,9282)	2,9786	43	19	2,9829-2,9785

### 5.4.3 Exemplos Gerados (Pseudo-)Aleatoriamente

Além dos exemplos ilustrativos, relatamos a seguir os resultados de experimentos mais exaustivos com os algoritmos apresentados nas seções anteriores. Consideramos duas subclasses de problemas fracionários lineares generalizados

$$P_{FLG} \left\{ \begin{array}{l} \text{minimizar} \quad \sum_{i=1}^r \frac{\langle c_i, x \rangle + \alpha_i}{\langle d_i, x \rangle + \beta_i} \\ \text{sujeito a} \quad Ax \leq b, \quad x \in \mathbb{R}_+^n, \end{array} \right.$$

e

$$Q_{FLG} \left\{ \begin{array}{l} \text{maximizar} \quad \sum_{i=1}^r \frac{\langle c_i, x \rangle + \alpha_i}{\langle d_i, x \rangle + \beta_i} \\ \text{sujeito a} \quad Ax \leq b, \quad x \in \mathbb{R}_+^n, \end{array} \right.$$

nos quais  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $c_i, d_i \in \mathbb{R}^n$ ,  $i = 1, 2, \dots, r$ , e  $\alpha_i, \beta \in \mathbb{R}$ ,  $i = 1, 2, \dots, r$  são matrizes e vetores constantes gerados pelo procedimento PT-PC.

As seguintes quantidades foram definidas (valores totais): I, número de iterações; C, números de planos de corte necessários para convergência e T, tempo de CPU (em segundos). Dez problemas com combinações selecionadas de  $m$  (número de restrições),  $n$  (número de variáveis) e  $r$  (número de razões) foram resolvidos.

Todos os algoritmos foram implementados em MATLAB versão 7.0.1/*Optimization Toolbox* versão 4, sobre uma plataforma computacional composta por um computador Intel Core Duo, 3.33 GHz com 4GB de memória RAM, 32 bits e as tolerâncias adotadas para as convergências finitas desses algoritmos foram fixadas em  $\epsilon_{en} = 0,00001$ ,  $\epsilon_{rel} = 0,00001$  e  $\epsilon_{bb} = 0,05$ . A Tabela 5.6 apresenta a média e desvio padrão (em parênteses) dos valores de I, C e T para  $r = 1$ ,  $r = 2$ ,  $r = 3$ ,  $r = 4$  e  $r = 5$ .

Pela análise da Tabela 5.6, percebe-se que o algoritmo PR-BB, apesar de gerar um número maior de cortes no espaço-imagem, apresenta tempos computacionais menores quando comparado com os do algoritmo PFG-ENU, o que pode ser explicado pela natureza enumerativa (explícita) destes últimos algoritmos. Além disso, podemos ver que em alguns casos, o desvio padrão do tempo computacional do algoritmo proposto em [KY99] ou é muito grande ou é maior do que a média, que isso significa que ou os resultados têm uma grande assimetria ou tem uma grande variabilidade de resultados para o parâmetro considerado, que nenhum dos dois casos é um comportamento desejável.

## 5.5 Resumo

Neste capítulo propusemos duas abordagens diferentes para resolver globalmente problemas de programação fracionária generalizada convexa-côncava e côncava-convexa, duas classes especiais de problemas de otimização global que envolvem a minimização e a maximização de uma soma finita de razões de funções convexas e côncavas sobre um conjunto convexo compacto e não-vazio. Resultados de análise convexa permitiram decompor o problema de programação fracionária generalizada em um problema mestre com infinitas

restrições lineares de desigualdade no espaço–imagem. De acordo com a primeira abordagem, um minimizador global do problema é obtido como o limite das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais, resolvidos por meio de um procedimento de enumeração de restrições, que explora o número relativamente pequeno de semi–espaços necessários para aproximar o problema original no espaço–imagem. Pela segunda abordagem, o problema é resolvido por um algoritmo *branch–and–bound* retangular. Os limitantes inferiores e superiores do algoritmo *branch–and–bound* são definidos por majorações dos termos fracionários dentro de regiões retangulares. As experiências computacionais atestaram a viabilidade e a eficiência das abordagens de otimização global propostas, que são, além disso, facilmente programáveis a partir de pacotes–padrões de otimização.

Tabela 5.6: Resultados dos Algoritmos PFG-ENU e PR-BB

$r$	1	1	2	3	3	3	4	4	5
$m$	50	80	80	30	80	80	10	30	10
$n$	50	100	100	50	50	100	10	50	10
ALGORITMO [KY99] – PROBLEMA ( $P_{FLG}$ )									
V	***	***	13.8	190.2	109.2	154.8	1669.4	1547.0	10063.0
	***	***	(1.8)	(49.5)	(25.8)	(75.1)	(1138.9)	(770.3)	(8307.0)
C	***	***	6.4	17.5	12.9	15.3	24.0	24.1	28.2
	***	***	(0.9)	(2.7)	(1.5)	(4.0)	(6.7)	(4.7)	(6.5)
T	***	***	175.2	381.0	206.6	1849.1	51.6	2456.1	656.2
	***	***	(0.92)	(206.9)	(65.6)	(1299.7)	(56.5)	(1926.0)	(1065.9)
$r$	1	1	2	3	3	3	4	4	5
$m$	50	80	80	30	80	80	10	30	10
$n$	50	100	100	50	50	100	10	50	10
ALGORITMO PFG-ENU – PROBLEMA ( $P_{FLG}$ )									
I	2,20	2,30	4,70	7,60	10,10	11,40	5,80	7,90	6,50
	(1,03)	(1,34)	(2,21)	(2,50)	(2,85)	(2,01)	(1,93)	(2,23)	(1,43)
C	15,20	13,50	17,40	33,40	42,10	118,40	15,00	19,30	17,60
	(12,58)	(11,28)	(7,98)	(11,92)	(21,21)	(52,07)	(5,08)	(7,95)	(5,91)
T	2,91	6,93	37,23	58,16	320,23	356,29	117,66	363,24	509,29
	(1,74)	(4,97)	(4,21)	(10,27)	(68,33)	(62,20)	(80,11)	(60,32)	(287,96)
$r$	1	1	2	3	3	3	4	4	5
$m$	50	80	80	30	80	80	10	30	10
$n$	50	100	100	50	50	100	10	50	10
ALGORITMO PR-BB – PROBLEMA ( $P_{FLG}$ )									
I	1,00	1,00	7,30	8,00	10,30	11,90	9,20	18,50	10,90
	(0,00)	(0,00)	(1,42)	(1,94)	(1,64)	(3,35)	(3,29)	(6,77)	(7,46)
C	1,70	1,30	48,60	76,70	88,20	91,80	37,90	129,40	54,70
	(0,48)	(0,47)	(15,74)	(18,71)	(17,47)	(38,68)	(15,33)	(38,23)	(44,55)
T	0,22	1,08	27,81	19,90	29,18	60,22	23,81	86,65	86,13
	(0,09)	(0,47)	(8,78)	(8,09)	(4,34)	(24,94)	(10,64)	(31,94)	(70,51)
$r$	1	1	2	3	3	3	4	4	5
$m$	50	80	80	30	80	80	10	30	10
$n$	50	100	100	50	50	100	10	50	10
ALGORITMO PR-BB – PROBLEMA ( $Q_{FLG}$ )									
I	1,00	1,00	6,23	10,08	11,25	11,50	10,23	13,20	8,30
	(0,00)	(0,00)	(1,54)	(1,48)	(1,76)	(1,83)	(1,88)	(2,16)	(2,47)
C	1,11	2,00	32,30	56,00	88,20	69,82	32,10	78,30	44,30
	(0,79)	(0,67)	(11,96)	(12,60)	(17,47)	(10,14)	(14,16)	(16,99)	(7,66)
T	0,29	0,99	19,27	23,88	25,61	48,82	34,20	67,65	74,91
	(0,13)	(0,24)	(10,10)	(4,62)	(6,08)	(17,34)	(11,47)	(20,95)	(34,24)

# 6

## Conclusões

### Sumário

---

6.1	Conclusões . . . . .	150
-----	----------------------	-----

---

## 6.1 Conclusões

Nesta tese propomos novas técnicas de otimização global com foco na resolução de duas classes importantes de problemas de programação não-convexa, problemas de programação multiplicativa e problemas de programação fracionária. Várias abordagens foram propostas e suas propriedades teóricas e computacionais foram analisadas. A essas duas classes de problemas foram associadas conceitos e formulações inspiradas em programação multiobjetiva, como o conceito Pareto-otimalidade ou eficiência e formulações no chamado espaço-imagem dos objetivos (*outcome space*). Entretanto, embora a Pareto-otimalidade ou eficiência das soluções ótimas globais encontradas esteja implícita, os algoritmos propostos não restringiram a busca a este tipo de solução.

No Capítulo 3, foi introduzida a idéia de otimização no espaço-imagem. Analizamos formulações de problemas de otimização multiobjetivos tanto no espaço das variáveis de decisão quanto no espaço-imagem, e discutimos maneiras diferentes de caracterizar e gerar soluções eficientes (ou Pareto), que constituem um conjunto de soluções candidatas a resolver problemas multiobjetivos. Um Algoritmo Básico, capaz de resolver problemas no espaço-imagem foi detalhado e serviu de ponto de partida para os algoritmos propostos nos Capítulos 4 e 5.

No Capítulo 4 foram abordadas duas variantes do problema de otimizar uma soma de produtos de funções escalares positivas sobre um conjunto viável convexo, que tem sido referido na literatura de otimização global como problema multiplicativo generalizado. Elementos de análise convexa e de programação multiobjetiva foram utilizados para mostrar como reformular esses problemas não convexos como problemas de programação quadrática indefinida com infinitas restrições lineares de desigualdade. Dois novos algoritmos com a estrutura do Algoritmo Básico foram propostos, o primeiro baseado em uma técnica de enumeração de restrições, o segundo em uma técnica *branch-and-bound*.

Uma das conclusões do Capítulo 4 foi a de que os problemas de programação multiplicativa generalizada convexa e côncava são análogos do ponto de vista do algoritmo de aproximação externa proposto neste trabalho. Observou-se que no espaço-imagem, um mesmo algoritmo (com as modificações adequadas) pode ser utilizado para resolver eficientemente problemas de ambas as classes de problemas multiplicativos generalizados.

No Capítulos 5 estendemos as abordagens do Capítulo 4 a problemas de programação fracionária generalizada nas suas formulações convexa-côncava e côncava-convexa.

Propusemos duas novas abordagens para resolver globalmente estes dois problemas de programação fracionária generalizada. De acordo com a primeira abordagem (abordagem de enumeração), os minimizadores globais do problema foram obtidos como o limite das soluções ótimas de uma sequência de problemas quadráticos indefinidos especiais, resolvidos por meio de um procedimento de enumeração de restrições, que explora um número relativamente pequeno de semi-espacos necessários para aproximar o problema original no espaço-imagem. De acordo com a segunda abordagem (abordagem *branch-and-bound*), o problema de programação fracionária é resolvido por um algoritmo de otimização que combina dois outros algoritmos: uma técnica *branch-and-bound* retangular, responsável por lidar com a não cõncavidade do problema no espaço-imagem, e uma técnica de relaxação (aproximação externa), responsável por lidar com os infinitas restrições lineares do problema.

Os resultados numéricos obtidos mostram que as abordagens propostas nesta tese possuem propriedades que as tornam competitivas em relação a outras abordagens encontradas na literatura de otimização global. Pela análise dos resultados obtidos ao resolver por meio dos métodos propostos problemas de programação multiplicativa generalizada e problemas de programação fracionária, percebe-se que os algoritmos *branch-and-bound*, apesar de gerarem um número maior de cortes no espaço-imagem, apresentam tempos computacionais menores quando comparados com os dos algoritmos de enumeração das restrições, o que pode ser explicado pela natureza enumerativa (explícita) destes últimos algoritmos. Ainda assim, pretendemos investigar técnicas alternativas de ramificação (como por exemplo, as ramificações simplicial e cônica [HPT95]) e outros refinamentos computacionais que possam conduzir a implementações mais eficientes dos nossos algoritmos de otimização global baseados em *branch-and-bound*.

Em essência, as técnicas propostas nesta tese tornam-se possíveis pela característica de monotonicidade estritamente crescente da função objetivo não-convexa em relação a cada função que a compõe. Desde que esta propriedade seja verificada e exista um envelope convexo para cada termo (bilinear, trilinear, fracionária, ...) que aparece na função objetivo, as abordagens propostas ainda se aplicam. Possíveis extensões para outras classes de problemas monotônicos mais gerais do ponto de vista teórico, porém de grande aplicação prática, estão sendo atualmente desenvolvidas pelos autor.

Outro objetivo será a busca por aplicações reais para as técnicas desenvolvidas nesta tese.



# Bibliografia

- [AAF97] C. S. Adjiman, I. P. Androulakis, and C.A. Floudas. Global optimization of MINLP problems in process synthesis and design. *Computers & Chemical Engineering*, 21:S445–S450, 1997.
- [AAF98] C. S. Adjiman, I. P. Androulakis, and C. A. Floudas. A global optimization method,  $\alpha$ BB, for general twice-differentiable constrained NLPs: II: implementation and computational results. *Computers & Chemical Engineering*, 22:1159–1179, 1998.
- [AAMF96] C. S. Adjiman, I. P. Androulakis, C. D. Maranas, and C.A. Floudas. A global optimization method,  $\alpha$ BB, for process design. *Computers & Chemical Engineering*, 20:S419–S424, 1996.
- [ADFN98] C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier. A global optimization method,  $\alpha$ BB, for general twice-differentiable constrained NLPs: I: theoretical advances. *Computers & Chemical Engineering*, 22:1137–1158, 1998.
- [ADSZ88] M. Avriel, W. E. Diewer, S. Schaible, and I. Zang. *Generalized Concavity*. Plenum Press, NY, USA, 1988.
- [AFa] A. M. Ashtiani and P. A. V. Ferreira. An image space, branch-and-bound algorithm for a class of multiplicative optimization problems. (artigo submetido ao Journal of Optimization Theory and Applications, 2012).
- [AFb] A. M. Ashtiani and P. A. V. Ferreira. Sum-of-ratios fractional programming: a decomposition approach. (artigo submetido ao European Journal of Operational Research, 2012).
- [AF10a] A. M. Ashtiani and P. A. V. Ferreira. Global maximization of a generalized concave multiplicative problem in the outcome space. In *Congresso Nacional de Matemática Aplicada e Computacional, Águas de Lindóias, SP, Brasil*, 2010.
- [AF10b] A. M. Ashtiani and P. A. V. Ferreira. Global optimization algorithms for a class of fractional programming problems. In *XLII Simpósio Brasileiro de Pesquisa Operacional, Bento Gonçalves, RS, Brasil*, 2010.

- [AF11a] A. M. Ashtiani and P. A. V. Ferreira. A branch and bound algorithm for a class of multiplicative programs. In *International Conference on Operations Research and Optimization, Teerā, Teerā, Irā*, 2011.
- [AF11b] A. M. Ashtiani and P. A. V. Ferreira. On the solution of generalized multiplicative extremum problems. *Journal of Optimization Theory and Applications*, 149:411–419, 2011.
- [AF11c] A. M. Ashtiani and P. A. V. Ferreira. Um algoritmo branch-and-bound para uma classe de problemas de programação multiplicativa. In *XLIII Simpósio Brasileiro de Pesquisa Operacional, Ubatuba, SP, Brasil*, 2011.
- [AF12] A. M. Ashtiani and P. A. V. Ferreira. Global optimization algorithm for the nonlinear sum of ratios problem. In *10th International Conference on Operations Research*, 2012.
- [AL70] Y. Almogly and O. Levin. Parametric analysis of a multi-stage stochastic shipping problem. In *J. Lawrence (Ed.), Operational Research '69*, pp. 359–370. Tavistock Publications, London, UK, 1970.
- [AMF95] I. P. Androulakis, C. D. Maranas, and C. A. Floudas.  $\alpha$ BB: A global optimization method for general constrained nonconvex problems. *Journal of Global Optimization*, 7:337–363, 1995.
- [AP96] T. W. Athan and P. Y. Papalambros. A note on weighted criteria methods for compromise solutions in multi-objective optimization. *Engineering Optimization*, 27:155–176, 1996.
- [Bar88] A. I. Barros. *Discrete and Fractional Programming Techniques for Location Models*. Kluwer Academic Publisher, The Netherlands, 1988.
- [BB97] H. P. Benson and G. M. Boger. Multiplicative programming problems: analysis and efficient point search heuristic. *Journal of Optimization Theory and Applications*, 94:487–510, 1997.
- [BBoo] H. P. Benson and G. M. Boger. Outcome-space cutting-plane algorithm for linear multiplicative programming. *Journal of Optimization Theory and Applications*, 104:301–322., 2000.
- [Ben94] K. P. Bennett. Global tree optimization: A non-greedy decision tree algorithm. *Computing Sciences and Statistics*, 25:156–160, 1994.

- [Ben99] H. P. Benson. An outcome space branch and bound-outer approximation algorithm for convex multiplicative programming. *Journal of Global Optimization*, 15:315–342, 1999.
- [Ben01] H. P. Benson. Global optimization of nonlinear sum of ratios. *Journal of Mathematical Analysis and Applications*, 263(1):301–315, November 2001.
- [Ben02a] H. P. Benson. Global optimization algorithm for the nonlinear sum of ratios problem. *Journal of Optimization Theory and Applications*, 112(1):1–29, January 2002.
- [Ben02b] H. P. Benson. Using concave envelopes to globally solve the nonlinear sum of ratios problem. *Journal of Global Optimization*, 22(1–4):343–364, January 2002.
- [Ben04] H. P. Benson. On the global optimization of sums of linear fractional functions over a convex set. *Journal of Optimization Theory and Applications*, 121(1):19–39, April 2004.
- [Ben07a] H. P. Benson. A simplicial branch and bound duality–bounds algorithm for the linear sum–of–ratios problem. *European Journal of Operational Research*, 182(16):597–611, October 2007.
- [Ben07b] H. P. Benson. Solving sum of ratios fractional programs via concave minimization. *Journal of Optimization Theory and Applications*, 135(1):1–17, October 2007.
- [Ben08] H. P. Benson. Global maximization of a generalized concave multiplicative function. *Journal of Optimization Theory and Applications*, 137:105–120, 2008.
- [Ben10] H. P. Benson. Branch–and–bound outer approximation algorithm for sum–of–ratios fractional programs. *Journal of Optimization Theory and Applications*, 146:1–8, 2010.
- [Ber65] B. Bereanu. Decisions regions and minimum risk solutions in linear programming. In A. Prekopa (Ed.), *Colloquium on Applications on Mathematics to Economics*, Budapest, 1963, pp. 37–42. Publication House of the Hungarian Academy of Sciences, Budapest, Hungary, 1965.
- [Ber95] D. P. Bertsekas. *Nonlinear Programming (2nd edn.)*. Athena Scientific, MA, USA, 1995.

- [BFSZ96] A. I. Barros, J. B. G. Frenk, S. Schaible, and S. Zhang. A new algorithm for generalized fractional programs. *Mathematical Programming*, 72:147–175, 1996.
- [BM94] K. P. Bennett and O. L. Mangasarian. Bilinear separation of two sets in  $n$ -space. *Computational Optimization and Applications*, 2:207–227, 1994.
- [BNO03] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, MA, USA, 2003.
- [BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, MA, UK, 2004.
- [CC62] A. Charnes and W. W. Cooper. Programming with linear fractional functionals quarterly. *Naval Research Logistics*, 9:181–186, 1962.
- [CC63] A. Charnes and W. W. Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11:18–39, 1963.
- [CFSYPP12] W. Chen-Feng, L. San-Yang, and S. Pei-Ping. Global minimization of a generalized linear multiplicative programming. *Applied Mathematical Modelling*, 36:2446–2451, 2012.
- [CMS89] A. Cambini, L. Martein, and S. Schaible. On maximizing a sum of ratios. *Journal of Information and Optimization Sciences*, 10:65–79, 1989.
- [CMW69] C. S. Colantoni, R. P. Manes, and A. Whinston. Programming, profit rates and pricing decision. *Accounting Review*, 44:467–481, 1969.
- [Cra88] B. D. Craven. Fractional programming. In *Sigma Series in Applied Mathematics, Vol 4*. Heldermann Verlag, Berlin, Germany, 1988.
- [CSMS00] W. Chen, A. Sahai, C. Messac, and G. Sundararaj. Exploration of the effectiveness of physical programming in robust design. *Journal of Mechanical Design*, 122:155–163, 2000.
- [Dan63] G. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, USA, 1963.
- [DHT01] M. Dür, R. Horst, and N.V. Thoai. Solving sum-of-ratios fractional programs using efficient points. *Optimization*, 49:447–466, 2001.

- [DS95] M. C. Dorneich and N. V. Sahinidis. Global optimization algorithms for chip design and compaction. *Engineering Optimization*, 25:131–154, 1995.
- [Edg81] F. Y. Edgeworth. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London C.K. Paul, UK, 1881.
- [ES89] E. Eskow and R. B. Schnabel. Mathematical modeling of a parallel global optimization algorithm. *Parallel Computing*, 12:315–325, 1989.
- [Fal69] J. E. Falk. Maximization of single-to-noise ratio in an optical filter. *SIAM Journal of Applied Mathematics*, 7:582–592, 1969.
- [Fal73] J. E. Falk. A linear max–min problem. *Mathematical Programming*, 5:159–188, 1973.
- [FDK98] J. B. G. Frenk, R. Dekker, and M. J. Kleijn. On the marginal cost approach in maintenance. *Journal of Optimization Theory and Applications*, 94:771–781, 1998.
- [Fer98] P. A. V. Ferreira. Modelagem, análise e projeto de sistemas com múltiplos objetivos (tese de livre-docência). Technical report, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil, 1998.
- [FJo1] R. W. Freund and F. Jarre. Solving the sum-of-ratios problem by an interior-point method. *Journal of Global Optimization*, 19:83–102, 2001.
- [FP92] J. E. Falk and S. W. Palocsay. Optimizing the sum of linear fractional functions. In C. A. Floudas and P.M. Pardalos (Eds.), *Recent Advances in Global Optimization*, pp. 221–258. Kluwer Academic Publishers, The Netherlands, 1992.
- [FP94] J. E. Falk and S. W. Palocsay. Image space analysis of generalized fractional programs. *Journal of Global Optimization*, 4:63–88, 1994.
- [Fre79] A. M. Freize. A bilinear programming formulation of the 3–dimensional assignment problem. *Mathematical Programming*, 7:376–379, 1979.
- [FS69] J. E. Falk and R. M. Soland. An algorithm for separable nonconvex programming problems. *Management Science*, 15:550–569, 1969.
- [FS04] J. B. G. Frenk and S. Schaible. Fractional programming. Technical report, Rotterdam School of Management and Rotterdam School of Economics, Erasmus Universiteit Rotterdam, The Netherlands, 2004.

- [FV93] C. A. Floudas and V. Visweswaran. Primal-relaxed dual global optimization approach. *Journal of Optimization Theory and Applications*, 78:187–225., 1993.
- [Ge89] R. P. Ge. A parallel global optimization algorithm for rational separable-factorable functions. *Applied Mathematics and Computation*, 32:61–72, 1989.
- [Geo68] A. M. Geoffrion. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22:618–630, 1968.
- [Geo70] A. M. Geoffrion. Elements of large-scale mathematical programming. Part I: concepts. *Management Science Series A–Theory*, 16:652–691, 1970.
- [GG63] P. C. Gilmore and R. E. Gomory. A linear programming approach to cutting stock problem (part II). *Journal of Operations Research*, 11:863–888, 1963.
- [Gia05] F. Giannessi. *Constrained Optimization and Image Space Analysis. Separation of Sets and Optimality Conditions, Vol. 1*. Springer, New York, USA, 2005.
- [HDT88] R. Horst, J. Devries, and N. V. Thoai. On finding new vertices and redundant constraints in cutting plane algorithms for global optimization. *Operations Research Letters*, 7:85–90, 1988.
- [Hir95] J. Hirsche. Optimization of sum and product of linear fractional functions under linear constraints. Technical report, Department of Computer Sciences and Scientific Computing, Martin-Luther-Universitat, 1995.
- [HJL88] P. Hansen, B. Jaumard, and S. H. Lu. Analytical approach to global optimization. *Comptes Rendus de L’Academie Des Sciences Serie I-Mathematique*, 306:29–32, 1988.
- [HK70] M. Held and R. M. Karp. The traveling-salesman problem and minimum spanning trees. *Operations Research*, 18:1138–1162, 1970.
- [HK71] M. Held and R. M. Karp. The traveling-salesman problem and minimum spanning trees: part II. *Mathematical Programming*, 1:6–25, 1971.
- [Hor88] R. Horst. Deterministic global optimization with partition sets whose feasibility is not known: Application to concave minimization, reverse convex constraints, d.c. programming and lipschitzian optimization. *Journal of Optimization Theory and Applications*, 58:11–37, 1988.

- [Hor89] R. Horst. On consistency of bounding operations in deterministic global optimization. *Journal of Optimization Theory and Applications*, 61:143–146, 1989.
- [HPRS91] W. W. Hager, P. M. Pardalos, I. M. Roussos, and H. D. Sahinoglou. Active constraints, indefinite quadratic test problems, and complexity. *Journal of Optimization Theory and Applications*, 68:499–511, 1991.
- [HPT95] R. Horst, P. M. Pardalos, and N. V. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, The Netherlands, 1995.
- [HQ71] J. M. Henderson and R. E. Quandt. *Microeconomic Theory: A Mathematical Approach*. McGraw Hill, New York, U.S., 1971.
- [HT93] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer-Verlag, Berlin, Germany, 1993.
- [IS05] A. Izmailov and M. Solodov. *Optimization, Volume 1: Optimality Conditions, Elements of Convex Analysis and Duality*. IMPA, Rio de Janeiro, Brazil, 2005.
- [JLC12] H. Jiao, S. Liu, and Y. Chen. Global optimization algorithm for a generalized linear multiplicative programming. *Journal of Applied Mathematics and Computing*, 40:551–568, 2012.
- [JMT97] B. Jaumard, C. Meyer, and H. Tuy. Generalized convex multiplicative programming via quasiconcave minimization. *Journal of Global Optimization*, 10(3):229–256, April 1997.
- [KA99] H. Konno and N. Abe. Minimization of the sum of three linear fractional functions (Dedicated to Professor Hoang Tuy’s 70th birthday). *Journal of Global Optimization*, 15:419–432, 1999.
- [KBoo] P. Kesavan and P. I. Barton. Generalized branch-and-cut framework for mixed-integer nonlinear optimization problems. *Computers & Chemical Engineering*, 24:1361–1366, 2000.
- [KFoo] H. Konno and K. Fukaiishi. A branch and bound algorithm for solving low rank linear multiplicative and fractional programming problems. *Journal of Global Optimization*, 18:283–299, 2000.

- [KG88] G. R. Kocis and I. E. Grossmann. Global optimization of nonconvex mixed–integer nonlinear programming (MINLP) problems in process synthesis. *Industrial & Engineering Chemistry Research*, 27:1407–1421, 1988.
- [KI88] H. Konno and M. Inori. Bond portfolio optimization by bilinear fractional programming. *Journal of the Operations Research Society of Japan*, 22:143–158, 1988.
- [KI89] H. Konno and M. Inori. Bond portfolio optimization by bilinear fractional programming. *Journal of the Operations Research Society of Japan*, 32:143–158, 1989.
- [KK90] H. Konno and T. Kuno. Generalized linear multiplicative and fractional programming. *Annals of Operations Research*, 25:147–162, 1990.
- [KK91] T. Kuno and H. Konno. A parametric successive underestimation method for convex multiplicative programming problems. *Journal of Global Optimization*, 1:267–285, 1991.
- [KK92] H. Konno and T. Kuno. Linear multiplicative programming. *Mathematical Programming*, 56:51–64., 1992.
- [KK95] H. Konno and T. Kuno. Multiplicative programming problems. In R. Horst and P. M. Pardalos (Eds.) *Handbook of Global Optimization*, 369–405. Kluwer Academic Publisher, The Netherlands, 1995.
- [KKY92] H. Konno, T. Kuno, and Y. Yajima. Parametric simplex algorithms for a class of np–complete problems whose average number of steps is polynomial. *Computational Optimization and Applications*, 1:227–239, 1992.
- [KKY94] H. Konno, T. Kuno, and Y. Yajima. Global minimization of a generalized convex multiplicative function. *Journal of Global Optimization*, 4(1):42–62, January 1994.
- [KR93] R. L. Keeney and H. Raiffa. *Decisions with Multiple Objective*. Cambridge University Press, Cambridge, MA, UK, 1993.
- [KSY93] H. Konno, H. Shirakawa, and H. Yamazaki. A mean–absolute deviation–skewness portfolio optimization model. *Annals of Operations Research*, 45:205–220, 1993.

- [KT97] H. Konno and P. T. Thach. *Optimization on Low Rank Nonconvex Structures*. Kluwer Academic Publisher, Dordrecht, The Netherlands, 1997.
- [KTT97] H. Konno, P. T. Thach, and H. Tuy. *Optimization on Low Rank Nonconvex Structures*. Kluwer Academic Publisher, Dordrecht, 1997.
- [Kuno02] T. Kuno. A branch-and-bound algorithm for maximizing the sum of several linear ratios. *Journal of Global Optimization*, 22:155–174., 2002.
- [KW96] H. Konno and H. Watanabe. Bond portfolio optimization problems and their applications to index tracking: a partial optimization approach. *Journal of the Operations Research Society of Japan*, 39:295–306, 1996.
- [KY] H. Konno and Y. Yajima. Solving rank two bilinear programs by parametric simplex algorithm. Institute of Human and Social Sciences Working Paper IHSS 90–17, Tokyo Institute of Technology, 1990.
- [KY92] H. Konno and Y. Yajima. Minimizing and maximizing the product of linear fractional functions. In C.A. Floudas and P.M. Pardalos (Eds.), *Recent Advances in Global Optimization*, pp. 259–273. Princeton University Press, USA, 1992.
- [KY99] H. Konno and H. Yamashita. Minimizing sums and products of linear fractional functions over a polytope. *Naval Research Logistics*, 46:583–896, 1999.
- [KYK93] T. Kuno, Y. Yajima, and H. Konno. An outer approximation method for minimizing the product of several convex functions on a convex set. *Journal of Global Optimization*, 3:325–335, 1993.
- [KYM91] H. Konno, Y. Yajima, and T. Matsui. Parametric simplex algorithms for solving a special class of nonconvex minimization problems. *Journal of Global Optimization*, 1:65–81, 1991.
- [Limo06] E. L. Lima. *Análise Real Vol.1: Funções de uma Variável*. IMPA, Rio de Janeiro, Brazil, 2006.
- [LM97] A. Lo and C. MacKinlay. Maximizing predictability in the stock and bond markets. *Macroeconomic Dynamics*, 1:102–134, 1997.
- [Lue84] D. G. Luenberger. *Linear and Nonlinear Programming*. Kluwer Academic Publishers, The Netherlands, 1984.

- [LUY99] X. J. Liu, T. Umegaki, and Y. Yamamoto. Heuristic methods for linear multiplicative programming. *Journal of Global Optimization*, 4:433–447, 1999.
- [LW66] E. L. Lawler and D. E. Wood. Branch-and-bound methods: A survey. *Operations Research*, 14:699–719, 1966.
- [MAF<sup>+</sup>97] C. D. Maranas, I. P. Androulakis, C. A. Floudas, A. J. Berger, and J. M. Mulvey. Solving long-term financial planning problems via global optimization. *Journal of Economic Dynamics & Control*, 21:1405–1425, 1997.
- [Man64] O. L. Mangasarian. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12:778–780, 1964.
- [Mar64] B. Martos. Hyperbolic programming. *Naval Research Logistics Quarterly*, 11:135–155, 1964.
- [Mar75] B. Martos. *Nonlinear Programming, Theory and Applications*. North Holland, Amsterdam, The Notherland, 1975.
- [Mat96] T. Matsui. Np-hardness of linear multiplicative programming and related problems. *Journal of Global Optimization*, 9:113–119, 1996.
- [MF95] C. D. Maranad and C. A. Floudas. Finding all solutions of nonlinearly constrained systems of equations. *Journal of Global Optimization*, 7:143–182, 1995.
- [Mie99] K. M. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publisher, Boston, MA, USA, 1999.
- [Mje83] K. M. Mjelde. *Methods of the Allocation of Limited Resources*. Wiley, NY, USA, 1983.
- [MM88] C. C. Meewella and D. Q. Mayne. An algorithm for global optimization of lipschitz continuous functions. *Journal of Optimization Theory and Applications*, 57:307–322, 1988.
- [MM89] C. C. Meewella and D. Q. Mayne. Efficient domain partitioning algorithms for global optimization of rational and lipschitz continuous functions. *Journal of Optimization Theory and Applications*, 61:247–270, 1989.
- [MMH82] K. Maling, S. H. Mueller, and W. R. Heller. On finding most optimal rectangular package plans. In *Proceedings of the 19th Design Automation Conference*, 663–667, 1982.

- [Mo091] R. E. Moore. Global optimization to prescribed accuracy. *Computers & Mathematics with Applications*, 21:25–39, 1991.
- [MR96] Cs. Mészáros and T. Rapcsák. On sensitivity analysis for a class of decision systems. *Decision Support Systems*, 16:231–240, 1996.
- [MT92] L. D. Muu and B. T. Tam. Minimizing the sum of a convex function and the product of two affine functions over a convex set. *Optimization*, 24:57–62, 1992.
- [MVZ95] J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios. Robust optimization of large-scale systems. *Operations Research*, 2:264–281, 1995.
- [NW98] G. L. Nemhauser and L. A. Wolsey. *Combinatorial Optimization*. Wiley, New York, USA, 1998.
- [NW99] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. John Wiley, New Jersey, USA, 1999.
- [OF08] R. M. Oliveira and P. A. V. Ferreira. A convex analysis approach for convex multiplicative programming. *Journal of Global Optimization*, 41:579–592, 2008.
- [OF10] R. M. Oliveira and P. A. V. Ferreira. An outcome space approach for generalized convex multiplicative programs. *Journal of Global Optimization*, 47(1):107–118, May 2010.
- [Par06] V. Pareto. *Manuale di Economica Politica, Societa Editrice Libreria*. Milan; translated into English by A. S. Schwier as *Manual of Political Economy*, edited by A. S. Schwier and A. N. Page (1971). New York, USA, 1906.
- [Par88] P. M. Pardalos. Enumerative techniques for solving some nonconvex global optimization problems. *OR Spektrum*, 10:29–35, 1988.
- [Par89] P. M. Pardalos. Parallel search algorithms in global optimization. *Applied Mathematics and Computation*, 29:219–229, 1989.
- [Par90] P. M. Pardalos. Polynomial time algorithms for some classes of constrained quadratic problems. *Optimization*, 21:843–853, 1990.
- [Par93] P. M. Pardalos. *Complexity in Numerical Optimization*. World Scientific, Singapore, 1993.

- [PAT95] T. Q. Phong, L. T. H. An, and P. D. Tao. On globally solving linearly constrained indefinite quadratic minimization problem by decomposition branch and bound method. *Operations Research Letters*, 17:215–220, 1995.
- [PR87] P. M. Pardalos and J. B. Rosen. *Constrained Global Optimization: Algorithms and Applications*. Springer–Verlag, Berlin, Germany, 1987.
- [PR02] P. M. Pardalos and M. G. C. Resende. *Handbook of Applied Optimization*. Oxford University Press, UK, 2002.
- [PT03] N. H. Phuong and H. Tuy. A unified monotonic approach to generalized linear fractional programming. *Journal of Global Optimization*, 26:229–259, 2003.
- [PY00] C. H. Papadimitriou and M. Yannakakis. On the approximability of trade-offs and optimal access of web sources. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 86–92, 2000.
- [QG96] I. Quesada and I. E. Grossmann. *Alternative bounding approximations for the global optimization of various engineering design problems*, chapter Global optimization in Engineering Design, Vol 9 Nonconvex Optimization and Its Applications, pages 309–331. Kluwer Academic Publishers, Norwell, MA, 1996.
- [Rad98] T. Radzik. Fractional combinatorial optimization. In *D. Z. Du and P. M. Pardalos (Eds.), Handbook of Combinatorial Optimization 1*, pp. 429–478. Kluwer Academic Publishers, The Netherlands, 1998.
- [Rag69] M. Raghavachari. On connections between zero–one integer programming and concave programming under linear constraints. *Operations Research*, 17:680–684, 1969.
- [Rao71] M. R. Rao. Cluster analysis and mathematical programming. *Journal of the American Statistical Association*, 66:622–626, 1971.
- [RCRL95] P. Ruiz-Canales and A. Rufian-Lizana. A characterization of weakly efficient points. *Mathematical Programming*, 68:205–212, 1995.
- [Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, USA, 1970.

- [RS95] H. S. Ryoo and N. V. Sahinidis. Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Computers & Chemical Engineering*, 19:551–566, 1995.
- [RS96] H. S. Ryoo and N. V. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8:107–138, 1996.
- [RS01] H. S. Ryoo and N. V. Sahinidis. Analysis of bounds for multilinear functions. *Journal of Global Optimization*, 19:403–424, 2001.
- [RS03] H-S. Ryoo and N. V. Sahinidis. Global optimization of multiplicative programs. *Journal of Global Optimization*, 26(4):387–418, August 2003.
- [Sch78] S. Schaible. Analyse und anwendungen von quotientenprogrammen, ein beitrage zur planung mit hilfe der nichtlinearen programmierung. In *Mathematical Systems in Economics, Vol. 42*. Hain Verlag, Meisenheim, 1978.
- [Sch81] S. Schaible. Fractional programming: applications and algorithms. *European Journal of Operational Research*, 7:111–120, 1981.
- [Sch95] S. Schaible. Fractional programming. In R. Horst and P. M. Pardalos (Eds.), *Handbook of Global Optimization*, pp. 495–608. Kluwer Academic Publishers, The Netherlands, 1995.
- [SD92] W. Stadler and J. P. Dauer. Multicriteria optimization in engineering: a tutorial and survey. In M. P. Kamat (Ed.) *Structural Optimization: Status and Promise*, pp. 211–249. Washington, DC: American Institute of Aeronautics and Astronautics, 1992.
- [SI83] S. Schaible and T. Ibaraki. Fractional programming. *European Journal of Operational Research*, 12:325–338, 1983.
- [SM97] I. M. Stancu-Minasian. *Fractional Programming*. Kluwer Academic Publishers, The Netherlands, 1997.
- [Smi96] E. M. B. Smith. *On the Optimal Design of Continuous Processes*. PhD thesis, Imperial College of Science, Technology and Medicine, University of London, UK, 1996.
- [SP97] E. M. B. Smith and C. C. Pantelides. Global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 21:S791–S796, 1997.

- [SP99] E. M. B. Smith and C. C. Pantelides. A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 23:457–478, 1999.
- [SS95] S. Schaible and C. Sadini. Finite algorithm for generalized linear multiplicative programming. *Journal of Optimization Theory and Applications*, 87:441–455, 1995.
- [SS03] S. Schaible and J. Shi. Fractional programming: the sum-of-ratios case. *Optimization Methods and Software*, 18(2):219–229, April 2003.
- [Sta79] W. Stadler. A survey of multicriteria optimization, or the vector maximum problem. *Journal of Optimization Theory and Applications*, 29:1–52, 1979.
- [Sta84] W. Stadler. Applications of multicriteria optimization in engineering and the sciences (a survey). In *Multiple Criteria Decision Making – Past Decade and Future Trends*, M. Zeleny (Ed.). JAI Press, Greenwich, Connecticut, US, 1984.
- [Sta87] W. Stadler. Initiators of multicriteria optimization. In J. Jahn and W. Krabs (Eds.) *Recent Advances and Historical Development of Vector Optimization*. Lecture Notes in Economics and Mathematical Systems, No. 294, pp. 3–25, Berlin, Springer-Verlag, 1987.
- [Sta88] W. Stadler. Fundamentals of multicriteria optimization. In W. Stadler (Ed.) *Multicriteria Optimization in Engineering and in the Sciences*, pp. 1–25. Plenum Press, New York, USA, 1988.
- [Ste86] R. E. Steuer. *Multicriteria Optimization: Theory, Computation and Application*. John Wiley & Sons, New York, USA, 1986.
- [Swa66] K. Swarup. Quadratic programming. *Cahiers du Centre d'Études de Recherche Opérationnelle*, 8:223–234, 1966.
- [SZ81] S. Schaible and W. T. Ziemba. *Generalized Concavity in Optimization and Economics*. Academic Press, New York, NY, USA, 1981.
- [TH88] H. Tuy and R. Horst. Convergence and restart in branch-and-bound algorithms for global optimization: Application to concave minimization and d.c. optimization problems. *Mathematical Programming*, 41:161–183, 1988.

- [Tho91] N. V. Thoai. A global optimization approach for solving the convex multiplicative programming problems. *Journal of Global Optimization*, 1:341–357, 1991.
- [Tuy98] H. Tuy. *Convex Analysis and Global Optimization*. Kluwer Academic Publisher, Dordrecht, The Netherlands, 1998.
- [vN37] J. von Neumann. Über ein ökonomisches gleichungssystem und eine verallgemeinerung des brouwerschen fixpunktsatzes. In *Ergebnisse eines Mathematischen Kolloquiums 8*, K. Menger (Ed.), pp. 73–83. Leipzig and Vienna, Austria, 1937.
- [vN45] J. von Neumann. A model of general economic equilibrium (English translation of "Über ein ökonomisches gleichungssystem und eine verallgemeinerung des brouwerschen fixpunktsatzes", in: *Ergebnisse Eines Mathematischen Kolloquiums*, 8:73–83, (1937). *Review of Economic Studies*, 13:73–83, 1945.
- [Wat84] H. Watanabe. *IC Layout Generation and Compaction Using Mathematical Optimization*. PhD thesis, Department of Computer Science, University of Rochester, 1984.
- [Wri97] S. J. Wright. *Primal–Dual Interior Points Methods*. SIAM, Philadelphia, 1997.
- [WS08] Ch-F. Wang and P-P. Shen. A global optimization algorithm for linear fractional programming. *Applied Mathematics and Computation*, 204:281–287, 2008.
- [Yu74] P. L. Yu. Cone convexity, cone extreme points and nondominated solutions in decision problems with multiobjectives. *Journal of Optimization Theory and Applications*, 14:319–377, 1974.
- [Yu85] P. L. Yu. *Multiple–Criteria Decision Making*. Plenum Press, New York, USA, 1985.
- [ZG99] J. M. Zamora and I. E. Grossmann. A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. *Journal of Global Optimization*, 14:217–249, 1999.

