

Algoritmos de Busca Global para Problemas de Otimização Geométricos e Multiplicativos

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, como parte dos requisitos exigidos para obtenção do título de Doutor em Engenharia Elétrica.

por

Rúbia Mara de Oliveira

Mestre em Matemática Computacional – ICMC/USP

Graduada em Matemática – IBILCE/UNESP

Orientador: **Prof. Dr. Paulo Augusto Valente Ferreira** – FEEC/UNICAMP

Banca Examinadora

Prof. Dr. Ricardo Hiroshi Caldeira Takahashi	DMA/UFGM
Prof. Dr. Aurélio Ribeiro Leite de Oliveira	DMA/IMECC/UNICAMP
Prof. Dr. Vinícius Amaral Armentano	DENSIS/FEEC/UNICAMP
Prof. Dr. Akebo Yamakami	DT/FEEC/UNICAMP
Prof. Dr. Paulo Augusto Valente Ferreira	DT/FEEC/UNICAMP (orientador)

16 de Setembro de 2005

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

OL4a Oliveira, Rúbia Mara de
Algoritmos de busca global para problemas de
otimização geométricos e multiplicativos / Rúbia Mara
de Oliveira. --Campinas, SP: [s.n.], 2005.

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

1. Programação convexa. 2. Programação
geométrica. 3. Otimização matemática. 4. Análise
numérica. I. Ferreira, Paulo Augusto Valente.
II. Universidade Estadual de Campinas. Faculdade
de Engenharia Elétrica e de Computação. III. Título.

Título em Inglês: Global search algorithms for geometric and multiplicative
optimization problems

Palavras-chave em Inglês: Global optimization, Multiplicative programming,
Geometric programming, Convex analysis, Numerical
methods, Mathematic optimization

Área de concentração: Automação

Titulação: Doutora em Engenharia Elétrica

Banca examinadora: Ricardo Hiroshi Caldeira Takahashi, Aurélio Ribeiro Leite
de Oliveira, Vinícius Amaral Armentano, Akebo Yamakami
e Paulo Augusto Valente Ferreira

Data da defesa: 16/09/2005

Dedicatória

Ao meu querido esposo, Ricardo Ribeiro dos Santos,
que esteve ao meu lado em todos os momentos,
e que através de seu amor e dedicação me deu toda
a força para concluir esta tese.

Agradecimentos

Primeiramente a Deus, por ter olhado para mim e nunca me abandonado em todos os momentos de minha vida.

Ao Prof. Dr. Paulo Augusto Valente Ferreira, pela orientação, amizade, paciência e apoio durante todos esses anos. Pela sua participação com importantes sugestões que colaboraram de forma imprescindível para o desenvolvimento deste trabalho.

Aos professores Ricardo Hiroshi Caldeira Takahashi, Aurélio Ribeiro Leite de Oliveira, Vinícius Armentano e Akebo Yamakami membros da banca examinadora pela disponibilidade e atenção dispensada ao trabalho, bem como por suas valiosas sugestões.

Ao Prof. Dr. Alfredo Del Lordelo, pela amizade, companheirismo e por toda ajuda e dedicação durante a fase de implementação do Algoritmo *Branch-and-Bound*.

Ao meu adorável Ricardo, esposo e companheiro, pelo constante apoio, motivação, compreensão e alegrias compartilhadas desde os tempos de mestrado.

A todos os amigos aqui conquistados, que fizeram e continuarão fazendo parte da minha história. Em especial aos queridos amigos do Departamento de Telemática da UNICAMP que compartilharam comigo as alegrias e dificuldades de ser um aluno de pós-graduação.

Ao amigo Alysson Mazoni, pelos bons tempos de república e pelas valiosas trocas de idéias que, sem dúvida, contribuíram muito para o meu conhecimento.

Aos meus queridos pais, Israel e Iraídes, irmãos, Érica e Misael pela constante presença em minha vida e por incentivarem todos os meus sonhos, apoiando-me sempre e ensinado-me que o amor e o respeito são os sentimentos mais plenos em nossa existência. Não poderia deixar de agradecer aos meus cunhados, Elizeu, Gisele, Rejane e o meu pequeno sobrinho Guilherme, que com seu sorriso, beleza e inocência alegrou muito a minha vida.

À FEEC, pela excelente estrutura de trabalho e pelo alto nível de seus professores e à Unicamp, pelo agradável ambiente acadêmico.

À Fapesp, pelo apoio financeiro e pela competência no acompanhamento das atividades relativas a este projeto.

Resumo

Nesta tese são propostos novos algoritmos de otimização baseados em busca global para duas importantes classes de problemas de programação não-linear: problemas geométricos, nos quais as funções envolvidas são descritas por somas de polinômios generalizados, e problemas de programação multiplicativa convexa, os quais, por sua vez, apresentam funções objetivos e/ou restrições expressas como produtos de funções convexas. Uma abordagem multiobjetivo para problemas geométricos posinomiais, que admitem reformulações convexas, é apresentada. Para problemas geométricos signomiais, que não possuem reformulações convexas conhecidas, propõe-se incorporar um procedimento de busca local a um algoritmo *branch-and-bound*, visando acelerar a convergência deste tipo de algoritmo. Elementos de análise convexa e programação multiobjetivo são usados para abordar problemas de programação multiplicativa quando estes apresentam produtos e somas de produtos de funções convexas positivas nas suas funções objetivos. Um mínimo global para o primeiro caso é obtido como o limite das soluções de uma seqüência de minimizações quase-côncavas sobre politopos, resolvidas eficientemente por meio de enumeração de vértices. Um mínimo global para o segundo caso é obtido como o limite das soluções de uma seqüência de problemas quadráticos indefinidos com características especiais, resolvidos por enumeração de restrições. O desempenho computacional dos algoritmos propostos nesta tese é avaliado por meio de problemas-testes e comparado com algoritmos alternativos existentes na literatura.

Abstract

In this thesis new optimization algorithms based on global search are proposed for two important classes of nonlinear programming problems: geometric problems, in which the functions involved are described by a sum of generalized polynomials, and convex multiplicative problems, in which, in turn, objective functions and/or constraints are expressed as a product of convex functions. A multiobjective approach for posynomial geometric problems, which admit convex reformulations, is presented. As convex reformulations for signomial geometric problems are unknown, a local search procedure with the purpose of speeding up the convergence of branch-and-bound algorithms is proposed. Elements of convex analysis and multiobjective programming are used for dealing with multiplicative programming problems presenting products and sums of products of positive convex functions in their objective functions. A global minimum in the first case is obtained as the limit of a sequence of quasi-concave minimizations on polytopes, efficiently solved by vertex enumeration. A global minimum for the second case is obtained as the limit of a sequence of special indefinite quadratic problems, solved by constraint enumeration. The computational performance of the algorithms proposed in this thesis has been evaluated by means of test problems and compared with alternate algorithms from the literature.

Notação

O símbolo “:=” significa *igual por definição*. Conjuntos são normalmente representados por letras maiúsculas no estilo caligráfico (\mathcal{Y} , \mathcal{W} , ...) ou por letras maiúsculas em estilo **bold**, no caso de conjuntos especiais (\mathbf{R} , \mathbf{C} , ...). Funções são normalmente representadas por letras minúsculas (f , g , h , ...).

\mathbf{R} , \mathbf{C} , \mathbf{Z} , \mathbf{N} , ... : Conjunto dos números reais, complexos, inteiros, naturais, ..., respectivamente. Os reais não-negativos (positivos) são representados por \mathbf{R}_+ (\mathbf{R}_{++}). Se \mathcal{C} e \mathcal{S} são conjuntos quaisquer, então

$x \in \mathcal{C}$: x pertence a \mathcal{C}

$x \notin \mathcal{C}$: x não pertence a \mathcal{C}

$\mathcal{C} \cup \mathcal{S}$: União de conjuntos

$\mathcal{C} \cap \mathcal{S}$: Interseção de conjuntos

$\mathcal{C} \subseteq \mathcal{S}$: Indica que \mathcal{C} é um subconjunto de \mathcal{S}

$conv(\mathcal{C})$: Casca convexa de \mathcal{C}

$int(\mathcal{C})$, $cls(\mathcal{C})$, $\partial(\mathcal{C})$: Interior, fecho (*closure*) e fronteira de \mathcal{C} , respectivamente

$efi(\mathcal{C})$: Subconjunto dos pontos eficientes de \mathcal{C}

$[f, \mathcal{C}]$: Epígrafo da função f sobre \mathcal{C}

$\partial f(x^0)$: Subdiferencial da função f em $x = x^0$

$\nabla f(x^0)$: Gradiente da função f em $x = x^0$

$\sup_{x \in \mathcal{C}} f(x)$: *Supremo* ou maior limitante inferior de f sobre \mathcal{C}

$\inf_{x \in \mathcal{C}} f(x)$: *Ínfimo* ou menor limitante superior de f sobre \mathcal{C}

$\lim_{k \rightarrow \infty} x^k$: Limite de x^k

$\limsup_{k \rightarrow \infty} x^k$: Limite superior de x^k

$\{x^k\}$ ou $\{x^k\}_{k=0}^{\infty}$: Seqüência infinita dos vetores x^k , $k = 0, 1, 2, \dots$

Vetores são normalmente representados por letras minúsculas (x, y, z, \dots). Um vetor com n componentes é escrito como $x = (x_1, x_2, \dots, x_n)$. Quando utilizados em operações matriciais, assumem a forma de *vetores-coluna*, isto é,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

O *vetor-linha* correspondente é

$$x^T = [x_1 \quad x_2 \quad \cdots \quad x_n],$$

e $(\cdot)^T$ indica a *transposição* de (\cdot) . Dados dois vetores x, y quaisquer,

$$x = y \quad \text{se e somente se } x_i = y_i, \quad i = 1, 2, \dots, n$$

$$x \leq y \quad \text{se e somente se } x_i \leq y_i, \quad i = 1, 2, \dots, n$$

$$x < y \quad \text{se e somente se } x_i < y_i, \quad i = 1, 2, \dots, n$$

Observe que $x \leq y$ e $x \neq y$ indica que pelo menos uma componente i é tal que $x_i < y_i$.

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i \quad (\text{Produto escalar entre os vetores } x \text{ e } y)$$

$$\|x\|_2 = \langle x, x \rangle^{\frac{1}{2}} \quad (\text{Norma Euclidiana do vetor } x)$$

Uma matriz A $m \times n$ com elementos a_{ij} (linha i , coluna j) é representada como $A = \{a_{ij}\}$.

A^T : Matriz transposta de A

$A \geq 0$: Matriz semidefinida positiva ($-A$ semidefinida negativa)

$A > 0$: Matriz definida positiva ($-A$ definida negativa)

1	Introdução	1
1.1	Proposta da Tese	3
1.2	Organização da Tese	4
2	Fundamentos de Análise Convexa e Otimização	5
2.1	Conjuntos e Funções Convexas	5
2.2	Otimização	10
2.3	Considerações Finais	15
3	Programação Multiobjetivo via Projeção	16
3.1	Formulação e Conceitos Básicos	16
3.2	Caracterização de Soluções Eficientes	19
3.3	Abordagem por Projeção	21
3.4	Algoritmo Básico	26
3.5	Considerações Finais	30
4	Programação Geométrica Generalizada	31
4.1	Programação Geométrica Posinomial	31
4.2	Programação Geométrica Multiobjetivo	37
4.3	Programação Geométrica Signomial	44
4.3.1	Transformação em Diferença de Funções Convexas	44
4.3.2	Limitante Inferior	45
4.3.3	Redução da Partição do Retângulo	47
4.3.4	Algoritmo <i>Branch-and-Bound (BB)</i>	49
4.3.5	Busca Local Via Condensação	53
4.4	Considerações Finais	62

5	Programação Convexa Multiplicativa	63
5.1	Programação Convexa Multiplicativa	63
5.1.1	Abordagem no Espaço dos Objetivos	65
5.1.2	Problema Mestre	66
5.1.3	Algoritmo Global	67
5.1.4	Enumeração de Vértices	69
5.1.5	Problema de Enumeração de Vértices	70
5.1.6	Experimentos Numéricos	73
5.2	Programação Convexa Multiplicativa Generalizada	79
5.2.1	Formulação no Espaço dos Objetivos	79
5.2.2	Solução por Enumeração	82
5.2.3	Exemplos Numéricos	83
5.3	Considerações Finais	86
6	Conclusões	87
	Referências Bibliográficas	89

Lista de Figuras

3.1	Interpretação do conceito de eficiência.	18
3.2	Interpretação do problema P_W	22
3.3	Aproximação da função ϕ	29
4.1	Aproximação da curva <i>tradeoff</i>	43
4.2	Relaxação Convexa (RC).	46
4.3	Redução dos limitantes iniciais através de relaxação convexa.	48
5.1	Interpretação da região factível.	83

Lista de Tabelas

4.1	Comparação das soluções encontradas.	37
4.2	Soluções encontradas usando PG.	43
4.3	Soluções encontradas usando o BB com a Busca Local.	62
5.1	Convergência do Algoritmo A_2	73
5.2	Resultados de [Kuno et al., 1993] para (P_{ML}) , com $m = 3$	74
5.3	Abordagem proposta para (P_{ML}) , com $m = 3$	75
5.4	Resultados de [Kuno et al., 1993] para (P_{ML}) , com $m = 4$	75
5.5	Abordagem proposta para (P_{ML}) , com $m = 4$	75
5.6	Abordagem proposta para (P_{ML}) , com $m = 5$	76
5.7	Resultados de [Kuno et al., 1993] para (P_{ML}) , com $(n, p) = (30, 20)$ e $m = 2$ a 5.	76
5.8	Abordagem proposta para (P_{ML}) , com $(n, p) = (30, 20)$ e $m = 2$ a 10.	76
5.9	Resultados de [Kuno et al., 1993] para (P_{MQ}) , com $m = 2$	76
5.10	Abordagem Proposta para (P_{MQ}) , com $m = 2$	77
5.11	Resultados de [Kuno et al., 1993] para (P_{MQ}) , com $m = 3$	77
5.12	Abordagem Proposta para (P_{MQ}) , com $m = 3$	77
5.13	Abordagem Proposta para (P_{MQ}) , com $m = 4$	78
5.14	Abordagem Proposta para (P_{MQ}) , com $m = 5$	78
5.15	Convergência do Algoritmo A_2	83
5.16	Resultados de [Konno et al., 1994].	85
5.17	Abordagem Proposta.	85
5.18	Resultados de [Konno et al., 1994].	85
5.19	Abordagem Proposta.	86

Introdução

O desenvolvimento de áreas tradicionais da Engenharia tem sido caracterizado nas últimas décadas pelo crescente emprego de modelos de otimização como paradigmas para problemas de tomada de decisões. Muitos problemas de otimização originários de aplicações industriais ou da pesquisa científica exibem não-convexidades em uma ou mais funções envolvidas. A consequência deste fato é a impossibilidade de se assegurar ótimos globais para estes problemas, uma vez que na presença de não-convexidades os métodos tradicionais de otimização são capazes apenas de fornecer ótimos locais [Bazaraa et al., 1993]. A Otimização Global é uma especialização dentro da área de otimização dedicada à caracterização e à obtenção de mínimos (ou máximos) globais de problemas não-lineares. Encontra inúmeras e importantes aplicações em problemas de economia de escala, de alocação e localização, de transporte e em diversos projetos de engenharia. Uma exposição detalhada sobre condições de otimalidade, métodos e aplicações de otimização global pode ser encontrada em [Horst & Pardalos, 1995] e [Horst et al., 1995]. Do ponto de vista computacional, problemas de otimização global pertencem à classe dos problemas NP-completos [Vavasis, 1995], que são problemas em que o tempo computacional necessário para resolvê-los cresce, exponencialmente em função do tamanho da entrada de dados para todo método conhecido. O presente trabalho propõe métodos de otimização global especializados para tratar duas importantes classes de problemas não-lineares: Problemas de Otimização Geométrica e Problemas de Otimização Multiplicativa.

A Programação Geométrica (PG) surgiu para resolver problemas algébricos de programação não-linear nos quais as funções envolvidas apresentam-se na forma posinomial [Zener, 1961]. Um posinômio é uma soma de monômios. Cada monômio é o produto de uma constante positiva pelo produto de variáveis positivas, sendo que cada variável pode ser elevada a uma

potência real. O termo Programação Geométrica foi adotado devido à relação existente entre as médias geométrica e aritmética de n números positivos. Problemas de Programação Geométrica Generalizada (PGG), também chamada de Programação Geométrica Signomial (PGS), são caracterizados por funções objetivos e restrições descritas como a diferença de dois posinômios. Os PGS's contêm um ou mais monômios com coeficientes negativos [Maranas & Floudas, 1997].

Em [Duffin et al., 1967], problemas geométricos posinomiais são resolvidos analiticamente via Dualidade. Vários métodos numéricos para obter soluções de problemas geométricos vêm sendo desenvolvidos desde a década de 1970. Grande parte desses métodos são baseados na solução de uma seqüência de problemas lineares. O primeiro método de pontos interiores para programação geométrica foi descrito em [Nesterov & Nemirovsky, 1994], no qual foi estabelecida a complexidade em tempo polinomial para o método. Trabalhos mais recentes tratando métodos de pontos interiores para PG são encontrados em [Kortanek et al., 1996], [Andersen & Ye, 1998] e [Nascimento & Andrade, 2003]. Discussões sobre os aspectos computacionais da resolução dos problemas geométricos signomiais também estão disponíveis na literatura [Dembo, 1978] e [Maranas & Floudas, 1997]. Tutoriais sobre técnicas para resolver problemas geométricos podem ser encontrados em [Phillips & Beightler, 1973], [Ecker, 1980] e [Shen & Zhang, 2004].

Uma propriedade importante da programação geométrica é que quando o problema primal tem a forma posinomial, este admite uma reformulação convexa, e a solução global pode ser encontrada resolvendo-se o dual associado, o qual exibe uma estrutura muito mais simples. Condensação é o método mais popular para resolver problemas signomiais. A idéia básica da condensação é aproximar um posinômio (soma de monômios) por um único monômio. Em [Yang & Bricker, 1996] discute-se como um problema signomial pode ser aproximado por um problema posinomial.

Programação Geométrica Signomial tem sido aplicada a diversos problemas de engenharia. Exemplos são: controle ótimo [Jefferson & Scott, 1978], teoria de controle [Yazarel & Pappas, 2004] e [Chandra et al., 2004], engenharia mecânica [Jha, 1990] e [Colleran et al., 2003], probabilidade e estatística [Bricker et al., 1997] e [Fuh & Hu, 2000], teoria da informação [Klafszky et al., 1992] e [Chiang, 2004], projeto de circuitos digitais [Boyd et al., 2005a], projeto de reatores químicos [Blau & Wilde, 1969] e de [Blau & Wilde, 1971] e trocadores de calor [Duffin & Peterson, 1966]. Em [Boyd et al., 2005b] são discutidos problemas originários de diversas áreas que podem ser resolvidos via programação geométrica.

A Programação Convexa Multiplicativa lida com problemas que envolvem um produto de funções convexas em sua função objetivo ou em suas restrições. Uma fonte importante de problemas multiplicativos são certos problemas convexas multiobjetivos nos quais o produto das funções individuais exerce o papel de uma função objetivo substituta [Yu, 1985]. Uma estratégia comumente adotada pelos algoritmos para problemas convexas multiplicativos [Konno & Kuno,

1995] e [Benson & Boger, 1997], é projetar o problema (geralmente não-convexo) sobre o espaço real m -dimensional, onde m é o número de funções convexas, de forma a coordenar sua solução global no espaço dos objetivos. Decomposição e projeção são estratégias bem estabelecidas na literatura de programação matemática [Geoffrion, 1970], e seus principais resultados têm sido progressivamente estendidos para problemas de otimização global não-convexa [Floudas & Visweswaran, 1993] e [Thoai, 2002].

Nas últimas décadas, vários algoritmos que visam resolver globalmente problemas multiplicativos têm sido propostos [Konno & Kuno, 1995]. A literatura classifica os métodos para resolver problemas multiplicativos da seguinte maneira: métodos baseados em parametrização [Konno & Kuno, 1990], métodos de aproximação externa [Thoai, 1991], [Kuno et al., 1993] e [Konno et al., 1994], métodos de enumeração de vértices [Pardalos, 1990], métodos de planos de corte no espaço dos objetivos [Benson & Boger, 2000]. Além destes, alguns métodos heurísticos têm sido desenvolvidos [Liu et al., 1999]. Veja, por exemplo, [Ryoo & Sahinidis, 2003] para maiores detalhes sobre métodos usados na resolução de problemas convexas multiplicativos.

Aplicações de Programação Multiplicativa são encontradas em áreas como microeconomia [Henderson & Quandt, 1971], otimização financeira [Maranas et al., 1997], otimização de árvores de decisões [Bennett, 1994], projeto de circuitos VLSI [Dorneich & Sahinidis, 1995], problemas de otimização multicritérios [Keeney & Raiffa, 1993] e otimização robusta [Mulvey et al., 1995].

1.1 Proposta da Tese

Esta tese tem por objetivo principal desenvolver algoritmos de otimização global para problemas de otimização geométricos e multiplicativos.

Numa primeira etapa aborda-se a implementação de métodos clássicos de Programação Geométrica por meio da Teoria da Dualidade. Num passo posterior estende-se o formalismo para Programação Geométrica Multiobjetivo. Em [Maranas & Floudas, 1997], propõe-se um algoritmo do tipo *branch-and-bound* para otimização global de problemas signomiais. Os autores propõem um excelente limitante inferior para a solução global do problema signomial. Uma das motivações desta tese é mostrar que o algoritmo discutido em [Maranas & Floudas, 1997] pode ter seu desempenho substancialmente melhorado pelo uso de um limitante superior obtido por meio da técnica de condensação. A condensação é usada como estratégia de busca local no algoritmo *branch-and-bound*, visando refinar o limitante superior, acelerando desta forma a convergência do algoritmo.

Numa segunda etapa do trabalho, aborda-se a solução de algumas classes de problemas multiplicativos. Os algoritmos propostos são baseados em elementos de Análise Convexa e Otimização Multiobjetivo, os quais combinados permitem obter aproximações externas do problema no espaço dos objetivos. A solução do problema aproximado é submetida a um sub-problema do tipo min-max, que a verifica com respeito a sua ϵ -factibilidade. Se não for ϵ -factível, a solução do problema min-max conduz a uma melhora da aproximação externa do problema original, cujo ϵ -ótimo é obtido pelo algoritmo em um número finito de iterações.

1.2 Organização da Tese

O desenvolvimento do trabalho e a apresentação dos resultados obtidos estão organizados da seguinte maneira:

- O Capítulo 2 revê conceitos e resultados básicos de Análise Convexa. O objetivo deste capítulo é fornecer um embasamento teórico para os algoritmos propostos nos capítulos seguintes;
- No Capítulo 3 alguns tópicos de Programação Multiobjetivo, e em particular, a formulação do problema no espaço dos objetivos são apresentados. Definição, formas de caracterização e propriedades básicas de problemas multiobjetivos são discutidas;
- Os principais conceitos referentes à Programação Geométrica Posinomial e Programação Geométrica Signomial são apresentados no Capítulo 4. Um algoritmo para otimização global de problemas signomiais baseado na técnica *branch-and-bound* combinada com a técnica de condensação é proposto. A implementação do algoritmo *branch-and-bound* foi realizada em colaboração com o Dr. Alfredo Del Sole Lordelo;
- O Capítulo 5 aborda problemas de Programação Convexa Multiplicativa. A abordagem adotada nesse capítulo enfatiza similaridades entre Programação Convexa Multiplicativa e Programação Convexa Multiobjetivo. Propõe-se algoritmos para duas classes de problemas multiplicativos;
- As conclusões gerais da tese e propostas para novos trabalhos são apresentadas no Capítulo 6.

Fundamentos de Análise Convexa e Otimização

O conceito de convexidade exerce um papel fundamental em otimização. O objetivo deste capítulo é apresentar os conceitos e resultados básicos sobre análise convexa e otimização que serão utilizados no decorrer desta tese.

2.1 Conjuntos e Funções Convexas

Definição 1. Diz-se que um conjunto $\mathcal{C} \subseteq \mathbf{R}^n$ é convexo se, dados quaisquer $x^1, x^2 \in \mathcal{C}$, todos os pontos da forma $\alpha x^1 + (1 - \alpha)x^2$, com $0 \leq \alpha \leq 1$, também estão contidos em \mathcal{C} .

Subespaços lineares são exemplos típicos de conjuntos convexos. Por definição, o conjunto vazio é convexo. Certas operações entre conjuntos convexos preservam convexidade.

Proposição 1. Sejam \mathcal{C} e \mathcal{S} conjuntos convexos do \mathbf{R}^n . Então, também são conjuntos convexos

1. $\mathcal{C} + \mathcal{S}$;
2. $\mathcal{C} \cap \mathcal{S}$;
3. $\alpha\mathcal{C} := \{y : y = \alpha x, x \in \mathcal{C}\}, \forall \alpha \in \mathbf{R}$.

Em certos casos torna-se necessário ou conveniente obter aproximações convexas para conjuntos originalmente não-convexos.

Definição 2. Seja \mathcal{C} um conjunto arbitrário do \mathbf{R}^n . A envoltória ou casca convexa de \mathcal{C} , representada como $\text{conv}(\mathcal{C})$ é o menor conjunto convexo que contém \mathcal{C} .

Alternativamente, pode-se definir $\text{conv}(\mathcal{C})$ como sendo a interseção de todos os conjuntos convexos que contêm \mathcal{C} . Alguns conjuntos não necessariamente convexos têm, por sua estrutura, importância destacada neste estudo.

Definição 3. Diz-se que um conjunto $\mathcal{C} \subseteq \mathbf{R}^n$ é um cone (com vértice na origem) se $x \in \mathcal{C}$ implica $\alpha x \in \mathcal{C}$, $\forall \alpha \geq 0$.

O conjunto dos vetores do \mathbf{R}^n que assumem apenas valores não-negativos ($\{x \in \mathbf{R}^n : x \geq 0\}$) é um exemplo clássico e fundamental de cone convexo. O conceito de convexidade pode ser naturalmente estendido para funções.

Definição 4. Uma função f definida sobre um subconjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$ é convexa se

$$f(\alpha x^1 + (1 - \alpha)x^2) \leq \alpha f(x^1) + (1 - \alpha)f(x^2), \quad \forall x^1, x^2 \in \mathcal{C}, \quad \forall \alpha \in [0, 1].$$

Se a desigualdade for estrita sempre que $x^1 \neq x^2$, diz-se que f é estritamente convexa. Uma função g definida sobre um subconjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$ é (estritamente) côncava se $-g$ for (estritamente) convexa.

Sob certas condições, o estudo da convexidade de funções pode ser reduzido ao da convexidade de conjuntos.

Proposição 2. Seja $\mathcal{C} \subseteq \mathbf{R}^n$ um conjunto convexo e f uma função convexa sobre \mathcal{C} . Então, o subconjunto

$$\mathcal{C}_\lambda = \{x \in \mathcal{C} : f(x) \leq \lambda\}$$

é convexo, qualquer que seja $\lambda \in \mathbf{R}$.

A convexidade de \mathcal{C}_λ para todo $\lambda \in \mathbf{R}$ não é suficiente para garantir que f seja uma função convexa. Entretanto, a relação de \mathcal{C}_λ como um conceito mais geral de convexidade pode ser perfeitamente caracterizada.

Definição 5. Uma função f definida sobre um subconjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$ é quase-convexa se

$$f(\alpha x^1 + (1 - \alpha)x^2) \leq \max \{f(x^1), f(x^2)\} \quad \forall x^1, x^2 \in \mathcal{C}, \quad \forall \alpha \in [0, 1].$$

Uma função f definida sobre um subconjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$ é quase-côncava se $(-f)$ é quase-convexa.

Teorema 1. Seja $\mathcal{C} \subseteq \mathbf{R}^n$ um conjunto convexo. Uma função f é quase-convexa sobre \mathcal{C} se e somente se \mathcal{C}_λ é um conjunto convexo para todo $\lambda \in \mathbf{R}$.

Prova: A prova deste resultado é bastante ilustrativa de vários conceitos relacionados à convexidade. Suponha que \mathcal{C}_λ é um conjunto convexo, $\forall \lambda \in \mathbf{R}$. Sejam então $x^1, x^2 \in \mathcal{C}$ e $x = \alpha x^1 + (1 - \alpha)x^2$, $\alpha \in [0, 1]$. Note que $x^1, x^2 \in \mathcal{C}_\lambda$ para

$$\lambda = \max \{f(x^1), f(x^2)\}.$$

Como \mathcal{C}_λ é convexo, $x = \alpha x^1 + (1 - \alpha)x^2 \in \mathcal{C}_\lambda$ e portanto

$$f(x) \leq \lambda = \max \{f(x^1), f(x^2)\}.$$

Uma vez que $x^1, x^2 \in \mathcal{C}$ e $\alpha \in [0, 1]$ são arbitrários, f é quase-convexa sobre \mathcal{C} . Assuma agora que f é quase-convexa sobre \mathcal{C} e sejam $x^1, x^2 \in \mathcal{C}_\lambda$. Então $x^1, x^2 \in \mathcal{C}$ e

$$\max \{f(x^1), f(x^2)\} \leq \lambda.$$

Seja $\alpha \in [0, 1]$ e $x = \alpha x^1 + (1 - \alpha)x^2 \in \mathcal{C}$. Como f é quase-convexa

$$f(x) \leq \max \{f(x^1), f(x^2)\}.$$

Logo, $x \in \mathcal{C}_\lambda$ e consequentemente \mathcal{C}_λ é convexo para todo $\lambda \in \mathbf{R}$. □

Exemplo 1: Considere o conjunto

$$\mathcal{C}_\lambda = \{x \in \mathcal{C} : x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \leq \lambda\},$$

com $\lambda \in \mathbf{R}$, $\mathcal{C} := \{x \in \mathbf{R}^n : x > 0\}$ e $\alpha_i < 0$, $i = 1, 2, \dots, n$. Se $\lambda \leq 0$, então $\mathcal{C}_\lambda = \emptyset$, e portanto convexo. Suponha então que $\lambda > 0$. Como a função logaritmo é crescente, a desigualdade em \mathcal{C}_λ é equivalente a

$$\alpha_1 \log(x_1) + \alpha_2 \log(x_2) + \dots + \alpha_n \log(x_n) \leq \log(\lambda) \tag{2.1}$$

na qual o lado esquerdo é uma soma de funções convexas, pois $\alpha_i < 0$ e $\log(x_i)$ é côncava sobre \mathcal{C} para todo $i = 1, 2, \dots, n$. Conseqüentemente, o subconjunto de \mathcal{C} que satisfaz a desigualdade (2.1) é convexo, implicando que $f(x) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ é quase-convexa sobre $\{x \in \mathbf{R}^n : x > 0\}$ com $\alpha_i < 0$, $i = 1, 2, \dots, n$. □

Uma correspondência direta entre conjuntos e funções convexas é obtida ao exigir que $\lambda \in \mathbf{R}$ e $x \in \mathcal{C}$ sejam relacionados da seguinte forma.

Definição 6. O epígrafo de uma função $f_i : \mathcal{C}_i \rightarrow \mathbf{R}$ é definido como o conjunto de pontos

$$[f, \mathcal{C}] := \{(x, \lambda) : f(x) \leq \lambda, x \in \mathcal{C}, \lambda \in \mathbf{R}\}.$$

Teorema 2. *Uma função f definida sobre um subconjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$ é convexa se e somente se $[f, \mathcal{C}]$ for um conjunto convexo.*

Geometricamente, $[f, \mathcal{C}]$ é o conjunto dos pontos que estão acima do gráfico que define a função. A demonstração do Teorema 2 segue os mesmos passos adotados na discussão do Teorema 1.

Combinações lineares da forma $\alpha x^1 + (1 - \alpha)x^2$ com $\alpha \in [0, 1]$ podem ser indutivamente estendidas para o caso da combinação envolver um número arbitrário de vetores. Resultados mais gerais relativos a funções convexas podem ser então obtidos.

Definição 7. *Sejam x^1, x^2, \dots, x^n vetores de um conjunto qualquer $\mathcal{C} \subseteq \mathbf{R}^n$. Vetores da forma*

$$x = \alpha_1 x^1 + \alpha_2 x^2 + \dots + \alpha_n x^n = \sum_{i=1}^n \alpha_i x^i,$$

onde $\alpha_i \geq 0$, $\sum_{i=1}^n \alpha_i = 1$ são chamados de combinações convexas dos vetores $x^1, x^2, \dots, x^n \in \mathcal{C}$.

Teorema 3. *Sejam $f_i : \mathcal{C}_i \rightarrow \mathbf{R}$ funções convexas sobre subconjuntos convexas $\mathcal{C}_i \subseteq \mathbf{R}^n$, $i = 1, 2, \dots, m$. Então,*

1. $\alpha f_i(x)$ é uma função convexa sobre \mathcal{C}_i , $\forall i$ e $\alpha \geq 0$;
2. $\sum_{i=1}^m \alpha_i f_i(x)$ é uma função convexa sobre $\cap_{i=1}^m \mathcal{C}_i$, com $\alpha_i \geq 0$, $i = 1, 2, \dots, m$;
3. $f_i\left(\sum_{k=1}^n \alpha_k x^k\right) \leq \sum_{k=1}^n \alpha_k f_i(x^k)$, $\forall x^1, x^2, \dots, x^n \in \mathcal{C}_i$, $\forall \alpha_k \geq 0$, $\sum_{k=1}^n \alpha_k = 1$.

Funções lineares desempenham um papel destacado em análise convexa devido a sua associação com o conceito de hiperplano.

Teorema 4. *Seja \mathcal{H} um hiperplano do \mathbf{R}^n . Então existe uma constante $\alpha \in \mathbf{R}$ e um vetor $c \in \mathbf{R}^n$, $c \neq 0$ tais que*

$$\mathcal{H} := \{x \in \mathbf{R}^n : \langle c, x \rangle = \alpha\}.$$

Para cada hiperplano do \mathbf{R}^n é sempre possível definir *semi-espacos abertos*

$$\mathcal{H}_> := \{x \in \mathbf{R}^n : f(x) > \alpha\}; \quad \mathcal{H}_< := \{x \in \mathbf{R}^n : f(x) < \alpha\},$$

e *semi-espacos fechados*

$$\mathcal{H}_\geq := \{x \in \mathbf{R}^n : f(x) \geq \alpha\}; \quad \mathcal{H}_\leq := \{x \in \mathbf{R}^n : f(x) \leq \alpha\}.$$

Definição 8. Um hiperplano \mathcal{H} é um hiperplano suporte de um conjunto convexo $\mathcal{C} \subseteq \mathbf{R}^n$, se \mathcal{C} está contido em um dos semi-espacos fechados determinados por \mathcal{H} e \mathcal{H} contém um ponto do fecho de \mathcal{C} , $cls(\mathcal{C})$.

Teorema 5. Se x^0 não é um ponto interior de um conjunto convexo \mathcal{C} , que contém pontos interiores, então existe um hiperplano \mathcal{H} contendo x^0 tal que \mathcal{C} situa-se em um dos semi-espacos determinados por \mathcal{H} .

Deste modo, para um conjunto convexo \mathcal{C} , que possui pontos interiores, é sempre possível associar um hiperplano suporte a cada ponto da fronteira de $cls(\mathcal{C})$; isto é, para todo ponto x da fronteira de \mathcal{C} , $\partial\mathcal{C}$, existe um vetor $c \neq 0$ e um escalar α tais que $\langle c, x \rangle \geq \alpha$, $\forall x \in \mathcal{C}$ se $\mathcal{C} \subset \mathcal{H}_{\geq}$ ou $\langle c, x \rangle \leq \alpha$, $\forall x \in \mathcal{C}$ se $\mathcal{C} \subset \mathcal{H}_{\leq}$.

Uma forma alternativa de se caracterizar um hiperplano suporte tendo em vista os resultados acima é como segue: $\mathcal{H} = \{x \in \mathbf{R}^n : \langle c, x \rangle = \alpha\}$ é um hiperplano suporte de \mathcal{C} em $x^0 \in \partial\mathcal{C}$ se

$$\langle c, x^0 \rangle = \inf_{x \in \mathcal{C}} \langle c, x \rangle \quad \text{ou} \quad \langle c, x^0 \rangle = \sup_{x \in \mathcal{C}} \langle c, x \rangle.$$

Observe que se $x^0 \in \mathcal{C}$, então o ínfimo ou supremo estão bem definidos; se $x^0 \in \partial\mathcal{C}$, mas $x^0 \notin \mathcal{C}$, existem pontos arbitrariamente próximos de $x^0 \in \mathcal{C}$ e invocando a continuidade das funções lineares envolvidas, o resultado se verifica.

Teorema 6. Sejam \mathcal{C}_1 e \mathcal{C}_2 conjuntos convexos do \mathbf{R}^n tais que \mathcal{C}_1 possui pontos interiores e \mathcal{C}_2 não contém nenhum ponto interior de \mathcal{C}_1 . Então existe um hiperplano \mathcal{H} separando \mathcal{C}_1 e \mathcal{C}_2 ou, de forma equivalente, existe $c \in \mathbf{R}^n$, $c \neq 0$, tal que $\sup_{x \in \mathcal{C}_1} \langle c, x \rangle \leq \inf_{x \in \mathcal{C}_2} \langle c, x \rangle$.

Prova: Seja $\mathcal{C} := \mathcal{C}_1 - \mathcal{C}_2 = \{y : y = x^1 - x^2, x^1 \in \mathcal{C}_1, x^2 \in \mathcal{C}_2\}$. Note que \mathcal{C} é um conjunto convexo e que $0 \notin int(\mathcal{C})$. Existe então $c \in \mathbf{R}^n$, $c \neq 0$, tal que $\langle c, x \rangle \leq 0$, $\forall y \in \mathcal{C}$, ou seja

$$\langle c, x^1 \rangle \leq \langle c, x^2 \rangle, \quad \forall x^1 \in \mathcal{C}_1, \forall x^2 \in \mathcal{C}_2.$$

Conseqüentemente, existe um escalar α tal que

$$\sup_{x \in \mathcal{C}_1} \langle c, x \rangle \leq \alpha \leq \inf_{x \in \mathcal{C}_2} \langle c, x \rangle$$

e $\mathcal{H} := \{x : \langle c, x \rangle = \alpha\}$ é o hiperplano procurado. \square

As próximas seções, que tratam dos aspectos de síntese via otimização convexa, discutirão a utilização prática de muitos conceitos e resultados relativos a convexidade. Dentre estes, destaca-se o de subgradiente de uma função convexa, por induzir o desenvolvimento de inúmeros algoritmos.

Definição 9. *Seja f uma função convexa definida no \mathbf{R}^n . Um vetor $\xi \in \mathbf{R}^n$ é um subgradiente de f no ponto x^0 se*

$$f(x) \geq f(x^0) + \xi^T(x - x^0), \quad \forall x \in \mathbf{R}^n.$$

O conjunto de todos os subgradientes de f no ponto x^0 , chamado de subdiferencial, é representado por $\partial f(x^0)$.

Proposição 3. *O subdiferencial $\partial f(x^0)$ é um conjunto convexo.*

Em particular, se $\partial f(x^0)$ contém um único elemento, conclui-se que $\xi = \nabla f(x^0)$ e obtém-se o resultado conhecido

$$f(x) \geq f(x^0) + \langle \nabla f(x^0), x - x^0 \rangle, \quad \forall x, x^0 \in \mathbf{R}^n,$$

para funções convexas diferenciáveis.

O conceito de subgradiente tem consequências importantes do ponto de vista de otimização. Entre outras propriedades, a função $f(x^0) + \xi(x - x^0)$ atua como *função suporte* de f no \mathbf{R}^n . No ponto $x = x^0$, os valores das funções coincidem e para $x \neq x^0$, a função suporte fornece um limitante inferior para $f(x)$, $x \in \mathbf{R}^n$. Se $\xi^T(x - x^0) \geq 0$, $\forall x \in \mathbf{R}^n$ então $f(x) \geq f(x^0)$, $\forall x \in \mathbf{R}^n$, e diz-se então que x^0 é um *mínimo global* de f sobre $x \in \mathbf{R}^n$.

2.2 Otimização

O objetivo desta seção é analisar conceitos de otimalidade associados a problemas de otimização em que uma única função objetivo deve ser minimizada sobre um subconjunto de decisões ou alternativas factíveis. Um método básico para solução de problemas convexas, mas não necessariamente diferenciáveis, é apresentado.

O problema de interesse pode ser formulado como

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{s.a } \quad x \in \Omega \subseteq \mathbf{R}^n, \end{aligned}$$

onde $f : \Omega \rightarrow \mathbf{R}$ é a função a ser minimizada e Ω é o subconjunto das decisões factíveis do problema. Observe que f mapeia vetores do \mathbf{R}^n em \mathbf{R} , um conjunto completamente ordenado de acordo com a relação de ordem “ \leq ”. Em outras palavras, dados dois elementos quaisquer $\alpha^1, \alpha^2 \in \mathbf{R}$ é sempre possível afirmar que $\alpha^1 < \alpha^2$, $\alpha^1 < \alpha^2$ ou que $\alpha^1 = \alpha^2$.

Definição 10. *Seja $f : \Omega \subseteq \mathbf{R}^n \rightarrow \mathbf{R}$. Um ponto $x^0 \in \Omega$ é chamado de mínimo local de f sobre Ω se existe uma esfera $\mathcal{S}(x^0, \epsilon) = \{x \in \mathbf{R}^n : \|x - x^0\| < \epsilon\}$ tal que*

$$f(x^0) \leq f(x), \quad \forall x \in \Omega \cap \mathcal{S}(x^0, \epsilon).$$

Diz-se que x^0 é um mínimo local estrito se a desigualdade acima sempre ocorrer de forma estrita para $x \neq x^0$. Se for possível estender a esfera $\mathcal{S}(x^0, \epsilon)$ de tal forma que $\Omega \cap \mathcal{S}(x^0, \epsilon) = \Omega$, diz-se que x^0 é um mínimo global de f em Ω , estrito caso a desigualdade seja satisfeita de modo estrito para $x \neq x^0$.

Assim, x^0 será um ponto de mínimo local de f sobre Ω se for possível definir uma vizinhança ao redor de x^0 em que a função sempre assuma valores maiores que ou iguais a $f(x^0)$. A grande maioria dos métodos de otimização é capaz de fornecer apenas mínimos locais, o que não garante que o menor valor de f (seu mínimo global) sobre Ω seja atingido. Portanto, analisar sob que condições a vizinhança de um mínimo local pode ser estendida a toda região de factibilidade é de fundamental importância.

Teorema 7. *Seja $f : \Omega \subseteq \mathbf{R}^n \rightarrow \mathbf{R}$ uma função convexa sobre um subconjunto convexo $\Omega \subseteq \mathbf{R}^n$. Se x^0 é um mínimo local de f sobre Ω , então x^0 é também um mínimo global de f sobre Ω , isto é,*

$$f(x^0) \leq f(x), \quad \forall x \in \Omega.$$

Além disso, o subconjunto $\Omega^ \subseteq \Omega$ no qual f atinge o seu ínfimo é convexo.*

Prova: Assuma que x^0 minimiza localmente f sobre $\Omega \cap \mathcal{S}(x^0, \epsilon)$, para algum $\epsilon > 0$. Então para qualquer $x^1 \in \Omega$ existe um ponto $x \in \mathcal{S}(x^0, \epsilon)$ tal que $x = \alpha x^0 + (1 - \alpha)x^1$ para algum $\alpha \in (0, 1)$. Entretanto, como f é convexa e x^0 é um mínimo local,

$$f(x^0) \leq f(x) \leq \alpha f(x^0) + (1 - \alpha)f(x^1),$$

ou $f(x^0) \leq f(x^1)$, $\forall x^1 \in \Omega$. A parte restante do teorema é demonstrada lembrando que se f é uma função convexa e $f^* = \inf_{x \in \Omega} f(x)$, então

$$\Omega^* = \{x \in \Omega : f(x) \leq f^*\}$$

é um conjunto convexo. □

Um *problema convexo* envolve a minimização de uma função convexa sobre um subconjunto convexo de alternativas factíveis. Na verdade, seria mais preciso falar em *formulação convexa*, uma vez que um mesmo problema de otimização pode admitir formulações convexas e não-convexas.

A forma padrão de um problema de *programação linear* é

$$\begin{aligned} & \text{minimizar } c^T x \\ & \text{s.a} \quad Ax = b, \\ & \quad \quad x \geq 0, \end{aligned}$$

onde $x \in \mathbf{R}^n$ e $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$ e $c \in \mathbf{R}^n$ são matrizes constantes. Problemas lineares possuem propriedades que asseguram a obtenção de mínimos globais (caso existam) em um número finito de passos através de algoritmos eficientes do tipo *Simplex*.

Um problema geral de *programação não-linear* assume a forma

$$\begin{aligned} & \text{minimizar } f(x) \\ & \text{s.a} \quad h(x) = 0, \\ & \quad \quad g(x) \leq 0, \end{aligned}$$

onde $f : \mathbf{R}^n \rightarrow \mathbf{R}$, $h : \mathbf{R}^n \rightarrow \mathbf{R}^r$ e $g : \mathbf{R}^n \rightarrow \mathbf{R}^p$. Observe que o problema possui r restrições de igualdade ($h_1(x) = 0, h_2(x) = 0, \dots, h_r(x) = 0$) e p restrições de desigualdade ($g_1(x) \leq 0, g_2(x) \leq 0, \dots, g_p(x) \leq 0$). Diz-se que num dado ponto factível x^0 (isto é, $h(x^0) = 0, g(x^0) \leq 0$), uma restrição de desigualdade $g_j(x) \leq 0$ está *ativa* se $g_j(x^0) = 0$, estando *inativa* caso contrário.

Um dos aspectos fundamentais do estudo de problemas de programação matemática é a caracterização de mínimos, locais e/ou globais, através de condições necessárias e/ou suficientes para otimalidade. No exposto a seguir, assume-se que as funções envolvidas são diferenciáveis.

Definição 11. *Seja x^* um ponto satisfazendo as restrições $h(x) = 0, g(x) \leq 0$ e \mathcal{J} o conjunto dos índices j para os quais $g_j(x^*) = 0$. Diz-se então que x^* é um ponto regular das restrições se os vetores gradientes $\nabla h_i(x^*), i = 1, 2, \dots, r$ e $\nabla g_j(x^*), j \in \mathcal{J}$ forem linearmente independentes.*

O conjunto mais geral de condições de otimalidade para problemas de programação matemática diferenciáveis é conhecido como *condições de Karush-Kuhn-Tucker*. As condições necessárias de primeira ordem são enunciadas a seguir.

Teorema 8 (Condições de Karush-Kuhn-Tucker). *Sejam $f, h_i, i = 1, 2, \dots, r$ e $g_j, j = 1, 2, \dots, p$ funções com segundas derivadas parciais contínuas. Seja x^* um mínimo local de f em $\Omega := \{x : h(x) = 0, g(x) \leq 0\}$ e assumamos que x^* é um ponto regular das restrições. Então existem multiplicadores $\lambda \in \mathbf{R}^r$ e $\mu \in \mathbf{R}_+^p$ tais que*

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^r \lambda_i \nabla h_i(x^*) + \sum_{j=1}^p \mu_j \nabla g_j(x^*) &= 0, \\ \mu_j g_j(x^*) &= 0, \quad j = 1, 2, \dots, p. \end{aligned}$$

As condições de Karush-Kuhn-Tucker adquirem carácter necessário e suficiente se o problema for convexo, o que ocorrerá quando f e g_j , $j = 1, 2, \dots, p$ forem funções convexas e h_i , $i = 1, 2, \dots, r$ forem funções lineares definidas sobre o \mathbf{R}^n .

Certas classes de problemas não-convexos podem ser abordados através de técnicas de otimização convexa quando se utiliza o conceito de envelope convexo de uma função.

Definição 12. *Seja $f : \Omega \subseteq \mathbf{R}^n \rightarrow \mathbf{R}$ uma função contínua e $\Omega \subseteq \mathbf{R}^n$, um conjunto convexo não-vazio. Então o envelope convexo de f é qualquer função F tal que:*

1. $F(x)$ é convexa em Ω ;
2. $F(x) \leq f(x)$, $\forall x \in \Omega$;
3. Se $h(x)$ é uma função convexa em Ω tal que $h(x) \leq f(x)$, $\forall x \in \Omega$, então $h(x) \leq F(x)$, $\forall x \in \Omega$.

Geometricamente, F é precisamente a função cujo epígrafo coincide com a casca convexa do epígrafo de f .

Da definição acima, o envelope convexo de uma função é o melhor subestimador convexo da função no conjunto convexo Ω . De maneira similar, o envelope côncavo de uma função sobre um conjunto convexo Ω é definido como o melhor sobre-estimador da função em Ω . A cada problema de otimização não-convexo com conjunto factível convexo pode-se associar um problema convexo cujo valor ótimo é o mesmo do problema original.

Teorema 9. *Considere o problema*

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{s.a} & x \in \Omega. \end{array}$$

no qual $\Omega \subseteq \mathbf{R}^n$ é um conjunto convexo compacto. Seja F um envelope convexo de f sobre Ω . Então,

$$f^* := \min \{f(x) : x \in \Omega\} = \min \{F(x) : x \in \Omega\}$$

e

$$\{y \in \Omega : f(y) = f^*\} \subseteq \{y \in \Omega : F(y) = f^*\}.$$

Prova: Por definição, $F(x) \leq f(x)$, $\forall x \in \Omega$. Portanto

$$\min \{F(x) : x \in \Omega\} \leq f^*.$$

A função constante igual a f^* é uma função subestimadora convexa de f . Pela definição de envelope convexo, temos que $F(x) \geq f^*$, $\forall x \in \Omega$ e assim

$$\min \{F(x) : x \in \Omega\} \geq f^*.$$

A segunda parte pode ser provada por contradição. Seja x^* um mínimo global de f em Ω , e suponha que x^* não é um ponto de mínimo global de F em Ω . Seja y^* um ponto de mínimo global de F em Ω . Então,

$$F(y^*) < F(x^*) \leq f^*.$$

Considere agora a função $H(x) = \max \{f^*, F(x)\}$. Então, H é convexa, $f(x) \leq H(x)$ e $H(x) \geq F(x)$, $\forall x \in \Omega$. Mais especificamente, $H(y^*) > F(y^*)$, o que contradiz o fato de que $F(x)$, o envelope convexo, é a maior função subestimadora convexa. Portanto $F(y^*) = F(x^*)$ e x^* minimiza globalmente $F(x)$ sobre Ω . \square

O teorema anterior sugere a possibilidade de resolver um problema não-convexo através da resolução de um problema convexo cuja função objetivo é um envelope convexo da função objetivo original. A dificuldade, no entanto, é que encontrar o envelope convexo de uma função é tão difícil quanto calcular seu mínimo global.

Existem inúmeros métodos destinados à resolução de problemas convexos. A natureza da função objetivo f e do conjunto de factibilidade Ω dita quase sempre a escolha do método mais apropriado. Se, à exceção da hipótese de convexidade, o problema não apresenta nenhuma outra propriedade importante, por exemplo, diferenciabilidade, então uma das alternativas será utilizar *métodos de planos de corte*. Um procedimento básico é descrito a seguir. Para facilitar a notação, assume-se que Ω é descrito por uma única restrição funcional, na forma,

$$\Omega := \{x \in \mathbf{R}^n : g(x) \leq 0\},$$

no qual $g : \mathbf{R}^n \rightarrow \mathbf{R}$.

Suponha que nos pontos x^1, x^2, \dots, x^k tenham sido encontrados: a) os valores $f(x^1), f(x^2), \dots, f(x^k)$ e $g(x^1), g(x^2), \dots, g(x^k)$ e b) no mínimo um subgradiente de f e g em cada ponto, $f^1 \in \partial f(x^1), f^2 \in \partial f(x^2), \dots, f^k \in \partial f(x^k), g^1 \in \partial g(x^1), g^2 \in \partial g(x^2), \dots, g^k \in \partial g(x^k)$. Observe que os pontos x^1, x^2, \dots, x^k não precisam ser factíveis.

A partir dessas informações, obtêm-se funções lineares por partes que representam os limitantes inferiores da função objetivo e da restrição na iteração k , isto é,

$$f(x) \geq f_k(x) = \max_{1 \leq i \leq k} \{f(x^i) + (f^i)^T(x - x^i)\},$$

$$g(x) \geq g_k(x) = \max_{1 \leq i \leq k} \{g(x^i) + (g^i)^T(x - x^i)\},$$

no qual a propriedade de que cada subgradiente fornece uma função suporte para f e g foi explicitamente utilizada. Uma vez que $g_k(x) \leq g(x)$, $\forall x \in \mathbf{R}^n$, a função g_k fornece uma aproximação externa do conjunto de factibilidade, ou seja,

$$\Omega \subseteq \Omega_k := \{x \in \mathbf{R}^n : g_k(x) \leq 0\}.$$

O método progride comparando os limites inferior e superior para o valor ótimo da função objetivo, f^* , os quais podem ser estimados através de

$$L_k = \inf_{x \in \Omega_k} f_k(x),$$

$$U_k = \min_i \{f(x^i)\}.$$

A idéia é que os limitantes sejam melhorados ao se incorporar as informações de um novo ponto x^{k+1} , de forma que

$$L_k \leq L_{k+1} \leq f^* \leq U_{k+1} \leq U_k,$$

fazendo com que a seqüência de pontos gerada convirja tanto do ponto de vista de otimalidade quanto de factibilidade, isto é, $\lim_{k \rightarrow \infty} f(x^k) = f^*$ e $\lim_{k \rightarrow \infty} \sup g(x^k) \leq 0$.

Uma característica importante dos métodos baseados em planos de corte é resolver problemas de otimização não-lineares através de uma seqüência de problemas lineares.

2.3 Considerações Finais

Um tratamento completo para a maior parte dos tópicos abordados neste capítulo pode ser encontrado em [Horst & Pardalos, 1995]. Veja [Rockafellar, 1970] e [Boyd & Vandenberghe, 2004] para um estudo aprofundado sobre convexidade. Condições de otimalidade e métodos de otimização para problemas de programação matemática (linear e não-linear) são estudados em [Luenberger, 1984], por exemplo.

Programação Multiobjetivo via Projeção

Problemas multiobjetivos consistem na otimização de um vetor composto por funções objetivas, escolhidas como forma de avaliar o impacto das decisões factíveis do problema de acordo com diferentes índices de desempenho. Neste capítulo as principais definições, formas de caracterização e propriedades de problemas multiobjetivos são apresentadas.

3.1 Formulação e Conceitos Básicos

A otimização multiobjetivo se distingue de todos os demais ramos da teoria de otimização quanto ao sentido que o conceito de *solução do problema* adquire. De fato, pode-se definir um problema de otimização multiobjetivo como a otimização (neste trabalho, minimização) de um vetor composto por funções escalares, escolhidas como forma de avaliar o impacto das decisões factíveis do problema sobre diferentes índices de desempenho. Entretanto, sabe-se que funções vetoriais estão definidas num conjunto parcialmente ordenado (por exemplo, o \mathbf{R}^m), e que portanto podem existir alternativas factíveis que não satisfazem nenhuma relação de ordem, do tipo “ \leq ”, inviabilizando desta forma a utilização do conceito usual de solução ótima adotado em problemas mono-objetivos (escalares). Assim sendo, a solução do problema dependerá da noção de equilíbrio utilizada para resolver os conflitos que surgem da consideração simultânea de vários objetivos. A que melhor se adapta aos problemas discutidos nesta tese é a noção de equilíbrio cooperativo de Pareto, cuja definição, formas de caracterização e propriedades básicas são discutidas neste capítulo. Dado que um mesmo problema multiobjetivo pode apresentar infinitas soluções Pareto-ótimas ou eficientes, não comparáveis entre si, torna-se evidente a necessidade de critérios adicionais para alcançar uma solução final. Esses critérios adicionais são fornecidos por

um *decisor* que seleciona uma solução que proporcione uma relação de compromisso adequada entre os objetivos do problema. Veja em [Ferreira, 1998] uma discussão mais detalhada do assunto.

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

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

no qual $f = (f_1, f_2, \dots, f_m)$ e $m \geq 2$. Observe que $f : \Omega \subseteq \mathbf{R}^n \rightarrow \mathbf{R}^m$, isto é, cada ponto $x \in \Omega$ é mapeado no espaço real m -dimensional. Uma consequência imediata deste tipo de formulação é a perda do conceito usual de otimalidade, uma vez que o contra-domínio de f é um conjunto parcialmente ordenado de acordo com a relação de ordem “ \leq ”, isto é, em geral

$$\nexists x^* \in \Omega : f(x^*) \leq f(x), \quad \forall x \in \Omega \cap \mathcal{S}(x^*, \epsilon),$$

sendo $\mathcal{S}(x^*, \epsilon) := \{x \in \Omega : \|x^* - x\| < \epsilon\}$, qualquer que seja $\epsilon > 0$. É claro que implicitamente está se excluindo a possibilidade de ocorrência de soluções triviais, como a representada pela existência de pontos factíveis que corresponderia a chamada *solução ideal ou utópica* do problema.

Definição 13. A *solução ideal* \underline{y} do problema é definida através de suas componentes como

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

sendo $x^i := \arg \min_{x \in \mathcal{X}} f_i(x)$, $i = 1, 2, \dots, m$.

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

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

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

Note que $\mathbf{R}^n = \Omega_{<}(x^0) \cup \Omega_{\geq}(x^0) \cup \Omega_{\sim}(x^0)$, $\forall x^0 \in \Omega$. Os subconjuntos $\Omega_{<}(x^0)$ e $\Omega_{\geq}(x^0)$ contêm todos os pontos de Ω 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 $\Omega \cap \Omega_{<}(x^*) = \emptyset$. Esta interpretação motiva a definição seguinte.

Definição 14. Diz-se que $x^* \in \Omega$ é uma solução eficiente se não existe qualquer outra solução $x \in \Omega$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$.

Uma solução eficiente (Pareto-ótima) não é dominada por qualquer outra solução factível 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 3.1.

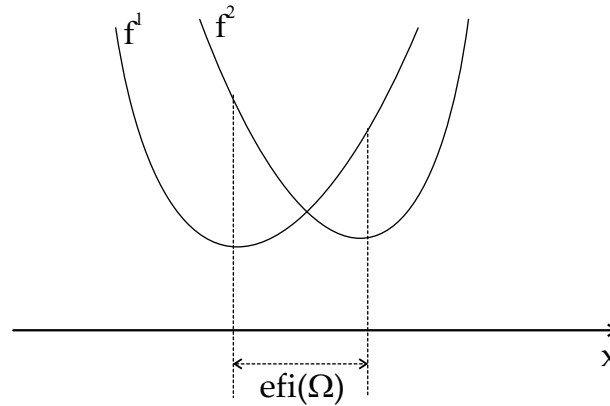


Figura 3.1: Interpretação do conceito de eficiência.

O conjunto de todas as soluções eficientes do problema no espaço das variáveis de decisão será denotado por $efi(\Omega)$. Assumindo que a solução do problema multiobjetivo deve ser eficiente, o processo de otimização fica agora restrito ao subconjunto $efi(\Omega)$. Conforme será visto, este subconjunto pode ser caracterizado através de problemas de otimização escalares. Entretanto, os algoritmos utilizados para resolver problemas deste tipo convergem para soluções ótimas locais.

Definição 15. Um vetor $x^* \in \Omega$ é uma solução localmente eficiente numa dada vizinhança $\mathcal{S}(x^*, \epsilon)$, isto é, $x^* \in efi(\Omega) \cap \mathcal{S}(x^*, \epsilon)$, se existe $\epsilon > 0$ tal que x^* é eficiente em $\Omega \cap \mathcal{S}(x^*, \epsilon)$.

Teorema 10. Sejam $f_i : \Omega \subseteq \mathbf{R}^n$, $i = 1, 2, \dots, m$ funções convexas sobre um subconjunto convexo $\Omega \subseteq \mathbf{R}^n$. Então toda solução localmente eficiente é também uma solução globalmente eficiente.

Assume-se no restante deste capítulo que f_1, f_2, \dots, f_m são funções convexas sobre um subconjunto convexo $\Omega \subseteq \mathbf{R}^n$. Quando esta suposição não é verdadeira, os resultados a seguir têm validade apenas local.

3.2 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 desses problemas. As relações entre soluções eficientes e problemas escalares passam a ser evidenciadas a partir dos teoremas seguintes.

Teorema 11. *Um vetor $x^* \in \text{efi}(\Omega)$ se e somente se x^* resolve os m problemas escalares*

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

Prova: 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. \square

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

Teorema 12. *Se para algum i , $x^* \in \Omega$ é solução única de P_i , então $x^* \in \text{efi}(\Omega)$.*

Prova: 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 $x^* \in \text{efi}(\Omega)$. \square

A caracterização de soluções eficientes através 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 13. *Se $x^* \in \Omega$ é eficiente, então existem um inteiro $k \in I := \{1, 2, \dots, m\}$ e números reais ϵ_j , $j = 1, 2, \dots, m$ ($j \neq k$) tais que x^* resolve*

$$P_k(\epsilon) : \begin{array}{l} \text{minimizar } f_k(x) \\ \text{s.a} \quad f_j(x) \leq \epsilon_j, \quad \forall j \neq k. \\ x \in \Omega \end{array}$$

onde ϵ 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

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

Prova: 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 ϵ_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 contradiz a hipótese de que x^* é eficiente. \square

Variando-se convenientemente ϵ em \mathcal{E}_k , é possível agora gerar todo o conjunto $efi(\Omega)$. 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 factível x^* seja eficiente estão resumidas no teorema seguinte.

Teorema 14. *Dado $\epsilon \in \mathcal{E}_k$, a solução x^* de $P_k(\epsilon)$ é eficiente se*

i) x^ é a solução única de $P_k(\epsilon)$ para algum k ;*

ii) x^ resolve $P_k(\epsilon)$ para todo $k \in I$.*

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

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

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

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

contradizendo $x^* \in efi(\Omega)$. \square

A caracterização de soluções eficientes através 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 15. *Assuma que f_i , $i = 1, 2, \dots, m$ são funções convexas sobre um subconjunto convexo $\Omega \subset \mathbf{R}$. Seja $x^* \in \Omega$ solução de*

$$\begin{array}{ll} P_W : & \text{minimizar } \langle w, f(x) \rangle \\ \text{s.a} & x \in \Omega \end{array}$$

para $w \in \mathbf{R}_+^m$ dado. Então x^* é uma solução eficiente do problema multiobjetivo se

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

Reciprocamente, se $x^* \in \text{efi}(\Omega)$, então existe $w \in \mathbf{R}_+^m$ tal que x^* é solução de P_W .

Prova: Considere o primeiro caso. Se x^* é a solução 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 existe $x \in \Omega$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Como $w \geq 0$, isto implicaria contradizer a unicidade de solução de x^* e portanto x^* é eficiente. A demonstração para a recíproca do Teorema pode ser encontrada em [Yu, 1985].

□

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

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

Observe que o problema P_W não altera a classe de modelos tratados, isto é, a natureza das restrições e 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 ϵ -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. Veja [Ferreira, 1998].

3.3 Abordagem por Projeção

Os conceitos e resultados sobre otimização objetivo 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 dos objetivos*. Neste sentido, seja $\mathcal{Y} = f(\Omega)$ a representação do mapeamento do subconjunto Ω no espaço dos objetivos:

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

O problema multiobjetivo pode ser formulado de forma equivalente como

$$P_y : \begin{array}{ll} \text{minimizar} & y \\ \text{s.a} & y \in \mathcal{Y}. \end{array}$$

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

Ignorando momentaneamente que este último problema é apenas uma abstração matemática, a formulação no espaço dos objetivos 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 $\Omega \subseteq \mathbf{R}^n$ e $\mathcal{Y} \subseteq \mathbf{R}^m$ e, em geral, o número de objetivos m é pequeno. Neste caso, a abordagem do problema baseada no espaço dos objetivos traria benefícios computacionais consideráveis. Além disso, o conceito de solução eficiente fica muito bem caracterizado em termos geométricos através da formulação no espaço dos objetivos.

Proposição 4. *Se $y^* \in efi(\mathcal{Y})$, então $y^* \in \partial\mathcal{Y}$.*

De fato, supor que $y^* \in int(\mathcal{Y})$ permitiria definir uma vizinhança $\mathcal{S}(y^*, \epsilon)$ entorno de y^* , e qualquer ponto $y \in \mathcal{S}(y^*, \epsilon)$ da reta que une y^* à origem do \mathbf{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 dos objetivos, conforme ilustra a Figura 3.2.

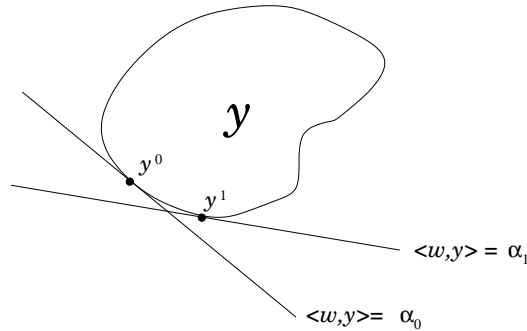


Figura 3.2: Interpretação do problema P_W .

O problema pode ser reescrito como

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

Sejam $y^0 \in \mathcal{Y}$ uma solução ótima do problema P_W e α^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-espço fechado \mathcal{H}_{\geq} determinado por \mathcal{H} . Se y^0 é a única solução de P_W e/ou $w \in \mathbf{R}_{++}^n$, então $y^0 \in \text{efi}(\mathcal{Y})$.

Se qualquer solução eficiente admitisse um hiperplano suporte, então variando-se parametricamente $w \in \mathcal{W}$, *todo* o conjunto $\text{efi}(\mathcal{Y})$ poderia ser gerado através de P_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.

O problema multiobjetivo admite uma formulação convexa equivalente no espaço dos objetivos. Nesta nova formulação, a região de factibilidade para os vetores y é redefinida como

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

no qual \mathcal{D} é o cone convexo

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

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

Teorema 16. $\text{efi}(\mathcal{Y}) = \text{efi}(\mathcal{Y} + \mathcal{D})$.

Prova: 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, assumamos 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 $\text{efi}(\mathcal{Y} + \mathcal{D}) = \text{efi}(\mathcal{Y})$. \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 17. *Sejam $f_i : \Omega \subseteq \mathcal{X}$, $i = 1, 2, \dots, m$ funções convexas sobre um subconjunto convexo $\Omega \subseteq \mathbf{R}^n$. Então o conjunto $\mathcal{Y} + \mathcal{D}$ é convexo.*

Prova: 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 $\alpha \in [0, 1]$,

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

A desigualdade surge da convexidade das funções e as igualdades subsequentes são obtidas para algum $d_\alpha \geq 0$ apropriado. Como $d_\alpha + d^1 \in \mathcal{D}$, $d_\alpha + d^2 \in \mathcal{D}$ 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 factibilidade 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.

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* $F : \mathcal{Y} \rightarrow \mathbf{R}$, 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

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

e o resultado de interesse demonstrado através do teorema seguinte.

Teorema 18. *Pelo menos uma solução de $P_{\mathcal{F}}$ é eficiente.*

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

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

Então $y^0 \in \text{efi}(\mathcal{Y})$. 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

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

dado que por definição F é uma função não-decrescente. \square

Dado que pelo menos uma solução de $P_{\mathcal{F}}$ é eficiente, pode-se aumentar a região factível de $P_{\mathcal{F}}$ de \mathcal{Y} para $\mathcal{Y} + \mathcal{D}$, obtendo-se o problema convexo equivalente

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

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

A solução 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 \mathbf{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

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

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 $y \notin \mathcal{Y} + \mathcal{D}$, obtém-se um semi-espaço adicional \mathcal{H}^{k+1} , no sentido de melhorar a aproximação anterior e a estimativa do valor ótimo da função F . Observe que $efi(\mathcal{Y}) = efi(\mathcal{Y} + \mathcal{D}) \subseteq (\mathcal{Y} + \mathcal{D})^k$, 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 espaço das decisões são inevitáveis.

Proposição 5. $\mathcal{Y} + \mathcal{D} = \mathcal{F}$, sendo

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

Prova: Note que em \mathcal{F} , os vetores y podem ser representados como $y = f(x) + d$, sendo que $f(x) \in \mathcal{Y}$ e $d \in \mathcal{D}$. \square

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

Lema 1. $y \in \mathcal{F}$ 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 direito da desigualdade, o ínfimo é obtido em Ω para cada $w \in \mathcal{W}$. As condições suficientes para que isto ocorra dependem de propriedades do \mathbf{R}^n , das funções f_1, f_2, \dots, f_m e do conjunto $\Omega \subseteq \mathbf{R}^n$. Um modo prático para se testar a pertinência de um dado $y \in \mathbf{R}^m$ 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 19. $y \in \mathcal{F}$ se e somente se

$$\Theta(y) := \sup_{w \in \mathcal{W}} \inf_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0. \quad (3.1)$$

Prova: 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 1, em particular para $w \in \mathcal{W}$ que maximiza o lado esquerdo da desigualdade, o que completa a prova. \square

Na prática, as desigualdades de que trata o Lema 1 devem ser entendidas como restrições aos valores que a variável y poderá assumir. Em outras palavras, o conjunto de soluções factíveis de $P_{\mathcal{F}}$ passa a ser representado como

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

3.4 Algoritmo Básico

Problemas de otimização que envolvam um grande número de restrições são adequadamente resolvidos através 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 factível, 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 satisfizer, uma solução ótima do problema original também terá sido encontrada. Caso contrário, determina-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: defina $\mathcal{F}^0 = \mathbf{R}^m$ e faça $k = 0$.

Passo 1. Resolva o problema multiobjetivo relaxado

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

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

Passo 2. Encontre $\Theta(y^k) = \langle w^k, f(x(w^k)) - y^k \rangle$ sendo que $(x(w^k), w^k)$ é a solução do problema representado no lado esquerdo de (3.1) para $y = y^k$. Se $\Theta(y^k) \leq \epsilon$, onde $\epsilon > 0$ é uma tolerância pré-especificada, então y^k resolve o problema no espaço dos objetivos e $x(w^k) \in \Omega$ resolve o problema no espaço das decisões. Caso contrário, defina

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

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

Na prática, é sempre conveniente considerar $\mathcal{F}^0 := \{y \in \mathbf{R}^m : \underline{y} \leq y \leq \bar{y}\}$, 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 através 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 20. *Seja y^* o ponto limite da sequência $\{y^k\}$ gerada pelo Algoritmo Básico. Então y^* resolve $P_{\mathcal{F}}$.*

Prova: Seja $\{y^{k_l}\}$, $l \in \mathbf{N}$, qualquer subsequência de $\{y^k\}$ que convirja 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,

$$\Theta(y^{k_l}) \leq \langle w^{k_l}, y^{k_{l'}} - y^{k_l} \rangle,$$

implicando que $\Theta(y^*) \leq 0$, uma vez que $\|w^{k_l}\| < 1, \forall l \in \mathbf{N}$ 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 $\mathcal{F} \subseteq \mathcal{F}^k, \forall k \in \mathbf{N}$, então $F(y^k) \leq F^*, \forall k \in \mathbf{N}$. Assim, pela continuidade de F , tem-se que $F(y^*) \leq F^*$, e portanto, y^* resolve $P_{\mathcal{F}}$. \square

Problema min-max

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

$$\begin{array}{ll} \text{maximizar} & \phi(w) \\ \text{s.a} & w \in \mathcal{W}. \end{array}$$

sendo

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

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

$$\begin{array}{ll} P_W : \text{minimizar} & \langle w, f(x) \rangle, \quad w \in \mathcal{W} \\ \text{s.a} & x \in \Omega. \end{array}$$

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

Proposição 6. *A função $\phi(w) = \inf_{x \in \Omega} \langle w, f(x) - y \rangle$ é côncava sobre o conjunto convexo \mathcal{W} .*

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

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

\square

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

Teorema 21. *Seja $\partial\phi(w^0)$ o subdiferencial de ϕ e $\Omega(w^0)$ o conjunto solução de P_W no ponto $w^0 \in \mathcal{W}$, respectivamente. Então*

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

Prova: Por definição,

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

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

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

e

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

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

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

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 ϕ na forma

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

onde k é o número de pontos considerados. A Figura 3.3 ilustra o processo de aproximação de ϕ através de funções (hiperplanos) suportes.

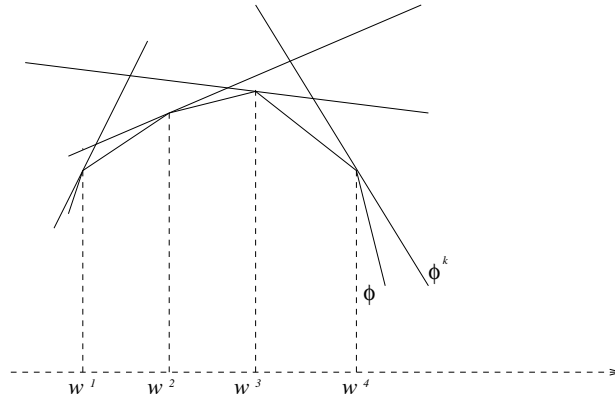


Figura 3.3: Aproximação da função ϕ .

O problema de maximização de ϕ^k sobre $w \in \mathcal{W}$ é então substituído pelo problema linear nas variáveis (w, σ) :

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

Seja (w^{k+1}, σ^{k+1}) a solução encontrada. O problema de minimização é resolvido para $w = w^{k+1}$ e a restrição linear correspondente é agregada ao problema anterior com k substituído por $k + 1$. Uma vez que σ^{k+1} é um limite superior para $\phi(w^{k+1})$, o método 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, através da abordagem por projeção, a solução do problema multiobjetivo pode ser obtida a partir de uma seqüência de problemas formulados no \mathbf{R}^m , independentemente das funções f_1, f_2, \dots, f_m e do subconjunto $\Omega \subseteq \mathbf{R}^n$ 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

$$P_W : \begin{array}{ll} \text{minimizar} & \langle w, f(x) \rangle, \quad w \in \mathcal{W} \\ \text{s.a} & x \in \Omega. \end{array}$$

uma vez que cada classe de problemas impõe condições de otimalidade específicas e fornece métodos numéricos apropriados para se obter soluções ótimas para P_W .

3.5 Considerações Finais

Este capítulo apresentou definições, formas de caracterização e propriedades básicas dos problemas multiobjetivos. Uma atenção particular foi dedicada na formulação do problema no espaço dos objetivos, uma vez que esses conceitos serão retomados no Capítulo 5. A formulação do problema no espaço dos objetivos é usual na literatura de otimização multiobjetivo e quase sempre tem fins didáticos, auxiliando em interpretações geométricas importantes. Veja, por exemplo, [Chankong & Haimes, 1983] e [Ferreira, 1998].

Programação Geométrica Generalizada

Duas classes de problemas de otimização cujas aplicações são vastas em Engenharia serão discutidas neste capítulo. A denominação Programação Geométrica é devida a relação existente entre as médias geométrica e aritmética de n números positivos. Um algoritmo de otimização global baseado na técnica de *branch-and-bound* é apresentado com vista à obtenção de uma solução mínima global para problemas geométricos signomiais.

4.1 Programação Geométrica Posinomial

Programação Geométrica é uma técnica desenvolvida para resolver problemas algébricos de programação não-linear. Os algoritmos de Programação Geométrica têm sido recentemente melhorados e atualmente são ferramentas poderosas para resolver problemas importantes em Engenharia. A Programação Geométrica surgiu com [Zener, 1961]. O matemático Richard Duffin, comprometido com o desenvolvimento da teoria de dualidade, solidificou a proposta de [Zener, 1961] aplicando-a em suas teorias. Veja [Beightler & Phillips, 1976].

O desenvolvimento matemático da Programação Geométrica está baseado na relação de desigualdade entre somatórios e produtórios de números positivos. Este desenvolvimento teórico chamado Programação Geométrica é também conhecido como Programação Posinomial. Um problema de Programação Geométrica (PG) apresenta a seguinte forma geral:

$$\begin{aligned} \text{PG:} \quad & \text{minimizar} \quad f(x) \\ & \text{s.a} \quad \quad \quad g_i(x) \leq 1, \quad i = 1, 2, \dots, p, \\ & \quad \quad \quad x_j > 0, \quad j = 1, 2, \dots, n. \end{aligned}$$

Um posinômio $g_i(x)$ é uma função composta por monômios T_k do seguinte tipo:

$$T_k = \alpha_k \prod_{j=1}^n x_j^{a_{kj}} \quad (4.1)$$

no qual $\alpha_k > 0$ e a_{kj} , $j = 1, 2, \dots, n$ são números reais. Como $x > 0$ obtém-se $T_k > 0$. Assim, a função objetivo e as restrições podem ser escritas como

$$f(x) := \sum_{k \in J_0} T_k,$$

$$g_i(x) := \sum_{k \in J_i} T_k, \quad i = 1, 2, \dots, p,$$

$$J_0 \cup J_1 \cup \dots \cup J_p := \{1, 2, \dots, P\}.$$

O conjunto J_0 descreve os termos (monômios) da função objetivo e os conjuntos J_i , com i representando uma restrição do problema, descrevem os termos de cada restrição. Esses conjuntos são conjuntos mutuamente disjuntos; P representa o número total de termos do problema, p número total de restrições e n o número de variáveis do problema. Duas trocas de variáveis baseadas nos conceitos de *dualidade lagrangeana* serão efetuadas, visando a obtenção de um problema convexo equivalente.

Aplicando-se a transformação logarítmica nas variáveis do problema original, obtém-se,

$$y_j := \ln(x_j), \quad j = 1, 2, \dots, n.$$

Logo $x_j = \exp(y_j)$ e

$$\tau_k = \alpha_k \prod_{j=1}^n \exp(y_j)^{a_{kj}} = \alpha_k \exp(a_k^T y), \quad k = 1, 2, \dots, P, \quad (4.2)$$

em que $a_k^T := (a_{k1}, a_{k2}, \dots, a_{kn})$, $k = 1, 2, \dots, P$ e $(\cdot)^T$ indica transposição.

Obtém-se a seguinte formulação do problema, denominada de Problema Geométrico Equivalente (PGE):

$$\begin{aligned} \text{PGE :} \quad & \text{minimizar} \quad \ln[F(y)] \\ & \text{s.a} \quad \quad \quad \ln[G_i(y)] \leq 0, \quad i = 1, 2, \dots, p, \\ & \quad \quad \quad y \in \mathbf{R}^n \end{aligned}$$

$$F(y) := \sum_{k \in J_0} \tau_k, \quad (4.3)$$

$$G_i(y) := \sum_{k \in J_i} \tau_k, \quad i = 1, 2, \dots, p. \quad (4.4)$$

Lema 2. *Dado um problema de Programação Geométrica Posinomial (α_k positivos), o problema equivalente obtido através da transformação logarítmica é convexo, pois $F(y)$ e $G_i(y)$, $i = 1, 2, \dots, p$ são funções convexas sobre $y \in \mathbf{R}^n$*

Pelo Teorema da Dualidade Forte, cujo enunciado e demonstração encontram-se em [Bazaraa et al., 1993], não existe *gap de dualidade* no problema geométrico e o dual lagrangeano é dado por

$$\text{DL : maximizar } \{\theta(u) \text{ s.a } u \geq 0\},$$

no qual $\theta(u) := \min_y L(y, u)$. Assim, a função Lagrangeana $L(y, u)$ é dada por

$$L(y, u) := \ln[F(y)] + \sum_{i=1}^p u_i \ln[G_i(y)]. \quad (4.5)$$

Dado $u \geq 0$, o mínimo do Lagrangeano é encontrado quando $\nabla_y L(y, u) = 0$. Logo, o dual lagrangeano é equivalente a

$$\text{maximizar } \{L(y, u) \text{ s.a } \nabla_y L(y, u) = 0, u \geq 0, y \in \mathbf{R}^n\} \quad (4.6)$$

Para simplificar a expressão (4.5), calcula-se o gradiente da função Lagrangeana:

$$\begin{aligned} \nabla_y L(y, u) &= \nabla \ln[F(y)] + \sum_{i=1}^p u_i \nabla \ln[G_i(y)], \\ &= \frac{\nabla F(y)}{F(y)} + \sum_{i=1}^p u_i \frac{\nabla G_i(y)}{G_i(y)}, \\ &= \frac{1}{F(y)} \sum_{k \in J_0} \tau_k a_k + \sum_{i=1}^p \frac{u_i}{G_i(y)} \left[\sum_{k \in J_i} \tau_k a_k \right]. \end{aligned}$$

Neste momento, aplica-se a segunda transformação de variáveis. Para isso define-se $\delta_1, \dots, \delta_p$ de acordo com,

$$\delta_k = \frac{\tau_k}{F}, \quad k \in J_0,$$

$$\delta_k = \frac{u_i \tau_k}{G_i}, \quad k \in J_i, \quad i = 1, 2, \dots, p.$$

Existe a necessidade de se tratar o argumento y como uma variável implícita, suprimindo-a na dependência de F , G_i , τ_k , δ_k , $i = 1, \dots, p$, $k = 1, \dots, P$. O problema (4.6) pode ser reescrito em termos das variáveis (δ, u) , eliminando a variável y do problema. Das expressões (4.3) e (4.4) obtém-se

$$\sum_{k \in J_0} \delta_k = 1 \quad \text{e} \quad \sum_{k \in J_i} \delta_k = u_i, \quad i = 1, 2, \dots, p. \quad (4.7)$$

Estas restrições são chamadas de *restrições de normalização*. Usando as equações (4.3), (4.4) e (4.5) pode-se concluir que

$$\begin{aligned} \nabla_y L(y, u) &= 0, \\ \frac{1}{F(y)} \sum_{k \in J_0} \tau_k a_k + \sum_{k \in J_i} \frac{u_i}{G_i(y)} \left[\sum_{k \in J_i} \tau_k a_k \right] &= 0, \\ \sum_{k \in J_0} \delta_k a_k + \sum_{k \in J_i} \delta_k a_k &= 0, \\ \sum_{k=1}^P \delta_k a_k &= 0. \end{aligned} \quad (4.8)$$

As últimas restrições de (4.8) são conhecidas como *restrições de ortogonalidade*. Para concluir a construção do problema transformado, lembre-se que

$$L(y, u) := \ln[F(y)] + \sum_{i=1}^p u_i \ln[G_i].$$

Tendo em vista as relações (4.2), (4.7) e (4.8), considere o termo $u_i \ln[G_i(y)]$ para $i \in \{1, \dots, p\}$. Assumindo que $u_i > 0$, cada um desses termos pode ser escrito como

$$\begin{aligned} u_i \ln[G_i] &= u_i \ln(u_i) + u_i \ln(G_i) - u_i \ln(u_i) \\ &= u_i \ln(u_i) + u_i \ln \left[\frac{G_i}{u_i} \right] \quad (u_i > 0) \\ &= u_i \ln(u_i) + \sum_{k \in J_i} \delta_k \ln \left[\frac{G_i}{u_i} \right] \\ &= u_i \ln(u_i) + \sum_{k \in J_i} \delta_k \ln \left[\frac{\tau_k}{\delta_k} \right] \\ &= u_i \ln(u_i) + \sum_{k \in J_i} \delta_k \ln \left[\frac{\alpha_k \exp(a_k^t y)}{\delta_k} \right] \\ &= u_i \ln(u_i) + \sum_{k \in J_i} \delta_k \ln \left[\frac{\alpha_k}{\delta_k} \right] + \sum_{k \in J_i} \delta_k a_k^t y \\ &= u_i \ln(u_i) + \sum_{k \in J_i} \delta_k \ln \left[\frac{\alpha_k}{\delta_k} \right]. \end{aligned}$$

Do mesmo modo, o termo $\ln[F(y)]$ pode ser escrito como

$$\ln[F] = \sum_{k \in J_0} \delta_k \ln \left[\frac{\alpha_k}{\delta_k} \right].$$

Obtém-se então o Problema Geométrico Dual Equivalente (PGDE):

$$\begin{aligned} \text{PGDE : } \quad & \text{maximizar} \quad \sum_{k=1}^P \delta_k \ln \left[\frac{\alpha_k}{\delta_k} \right] + \sum_{i=1}^p u_i \ln(u_i) \\ \text{s.a} \quad & \sum_{k=1}^P \delta_k a^k = A\delta = 0 \\ & \sum_{k \in J_0} \delta_k = 1 \\ & \sum_{k \in J_i} \delta_k = u_i, \quad i = 1, 2, \dots, p \\ & \delta_k \geq 0, \quad k = 1, 2, \dots, P \\ & u_i \geq 0, \quad i = 1, 2, \dots, p. \end{aligned}$$

Observe que o PGDE é um problema convexo do ponto de vista da programação não-linear: a função objetivo é côncava (e separável) e as restrições são lineares.

O PGDE envolve P número de termos e n variáveis. O grau de dificuldade desse problema é definido como $DD = P - (n + 1)$. Quando $DD = 0$, resolve-se um sistema linear $P \times P$; se $DD = 1$, resolve-se um problema de busca unidimensional. Como a função objetivo é côncava e diferenciável, a solução é simples. À medida que DD aumenta, o problema torna-se mais complicado, porém não tanto quanto o problema original.

Teorema 22. *Suponha que $(\delta^*, u^*) > 0$ resolve PGDE e que o valor ótimo associado é ν^* . Sejam $v^* := (v_1^*, v_2^*, \dots, v_n^*)$ e $w^* := (w_0^*, w_2^*, \dots, w_p^*)$ os multiplicadores ótimos associados às restrições de PGDE. Então uma solução ótima de PGE é*

$$y_j^* = v_j^*, \quad j = 1, 2, \dots, n,$$

o valor ótimo de PGE é ν^ e $u^* := (u_1^*, u_2^*, \dots, u_p^*)$ são os multiplicadores ótimos associados às restrições de PGE. Além disso,*

$$x_j^* = \exp(y_j^*), \quad j = 1, 2, \dots, n$$

resolve o problema PG.

No momento em que a solução ótima do problema PGDE é encontrada, torna-se simples encontrar a solução ótima para o problema original. Por outro lado, se o problema geométrico tem uma solução ótima e esta é ponto interior das restrições, então pode-se mostrar que o problema PGDE também possui uma solução ótima (δ^*, u^*) com o mesmo valor na função objetivo v^* , e que uma solução ótima do PGDE pode ser obtida resolvendo-se o sistema [Bazaraa et al., 1993].

$$a_k^T y = \ln \left[\frac{\delta_k^* \exp(\nu^*)}{\alpha_k} \right], \quad k \in J_0, \quad (4.9)$$

$$a_k^T y = \ln \left[\frac{\delta_k^*}{u_i^* \alpha_k} \right], \quad k \in J_i, \quad i = 1, 2, \dots, p : u_i^* > 0.$$

Os três problemas de otimização geométricos descritos a seguir e propostos em [Ruckaert & Matens, 1978] são utilizados para fins de comparação com o algoritmo implementado em Matlab[®] versão 7.1, executando sobre uma plataforma computacional composta por um computador Pentium IV, 2.4 GHz com 512MBytes de memória RAM. Todos os experimentos computacionais apresentados nesta tese foram desenvolvidos sobre esta mesma plataforma de software e hardware. A Tabela 4.1 apresenta as soluções encontradas adotando uma tolerância fixa $\epsilon = 0.001$.

Problema 1

$$\begin{array}{ll} \text{minimizar} & x_1^{-1} \\ \text{s.a} & x_1 x_2^{-1} + 0.5 x_3^{-1} \leq 1, \\ & 0.001 x_3 x_4^{-1} + 0.01 x_2 + 0.0005 x_2 x_4 \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Problema 2

$$\begin{array}{ll} \text{minimizar} & 5x_1 + 50000x_1^{-1} + 20x_2 + 72000x_2^{-1} + 10x_3 + 144000x_3^{-1} \\ \text{s.a} & 4x_1^{-1} + 32x_2^{-1} + 120x_3^{-1} \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Problema 3

$$\begin{array}{ll} \text{minimizar} & (592x_1^{0.65} + 582x_1^{0.39} + 1200x_1^{0.52} + 370x_1^{0.22}x_2^{-0.22} + 250x_1^{0.40}x_3^{-0.40} \\ & + 210x_1^{0.62}x_3^{-0.62} + 250x_1^{0.40}x_4^{-0.40} + 200x_1^{0.85}x_4^{-0.85}) \\ \text{s.a} & 500x_1^{-1} + 50x_2x_1^{-1} + 50x_3x_1^{-1} + 50x_4x_1^{-1} \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Tabela 4.1: Comparação das soluções encontradas.

Problemas	Sol. Encontrada e F.Obj	Sol. [Ruckaert & Matens, 1978] e F.Obj
1	(82.58,87.82,8.28,1.37) 0.0121	(82.847,87.924,8.293,1.364) 0.01208
2	(108.73,85.09,204.27) 6299.1	(107.4,84.9,204.5) 6300.0
3	(749.89,0.1111,1.461,3.4249) 126.310	(749.89,0.11114,1.46193,3.42481) 126.344

4.2 Programação Geométrica Multiobjetivo

Esta seção tem como objetivo estender conceitos e procedimentos da seção anterior para problemas multiobjetivos. Um problema de Programação Geométrica Multiobjetivo (PGM) correspondente à formulação P_W discutida no Capítulo 3 apresenta a seguinte estrutura geral:

$$\begin{aligned} \text{PGM: } & \text{minimizar } F(x) = w_1 f_1(x) + w_2 f_2(x) + \dots + w_m f_m(x) \\ & \text{s.a } \quad g_i(x) \leq 1 \quad i = 1, 2, \dots, p, \\ & \quad x_j > 0, \quad j = 1, 2, \dots, n. \end{aligned}$$

no qual $w \in \mathcal{W}$,

$$\begin{aligned} f_i(x) &:= \sum_{k \in J'_i} T_k, \quad i = 1, 2, \dots, m, \\ g_i(x) &:= \sum_{k \in J_i} T_k, \quad i = 1, 2, \dots, p. \end{aligned}$$

É importante ressaltar que às quantidades J_i, P, p e n definidas na seção anterior acrescenta-se a quantidade m , que define o número de funções objetivos que o problema possui e o conjunto J'_i com i representando cada função objetivo do problema. De acordo com a definição de posinômio citada na equação (4.1), as funções objetivos podem ser reescritas como:

$$\left. \begin{aligned} f_1(x) &:= \sum_{k \in J'_1} T_k \\ f_2(x) &:= \sum_{k \in J'_2} T_k \\ &\vdots \\ f_m(x) &:= \sum_{k \in J'_m} T_k \end{aligned} \right\} \Rightarrow F(x) = \sum_{i=1}^m w_i f_i(x) = \sum_{i=1}^m \sum_{k \in J'_i} w_i T_k.$$

A mesma transformação logarítmica nas variáveis do problema original será utilizada. Com a mudança de variável, cada função objetivo do problema original, assume a forma

$$\begin{aligned} f_i(x) &= \sum_{k \in J'_i} T_k \\ &= \sum_{k \in J'_i} \alpha_k \prod_{j=1}^n (e^{y_j})^{a_{kj}} = \sum_{k \in J'_i} \alpha_k e^{a_k^T y} = \sum_{k \in J'_i} \tau_k. \end{aligned}$$

O problema de Programação Geométrica Multiobjetivo Equivalente (PGME) pode ser escrito da seguinte forma:

$$\begin{aligned} \text{PGME: } & \text{minimizar } \ln[F(y)] \\ & \text{s.a } \quad \ln[G_i(y)] \leq 0, \quad i = 1, 2, \dots, p, \\ & \quad y \in \mathbf{R}^n \end{aligned}$$

no qual

$$F(y) = \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k, \quad w \in \mathcal{W} = \{w \geq 0 : \sum_{i=1}^m w_i = 1\}, \quad (4.10)$$

$$G(y) = \sum_{i=1}^p \sum_{k \in J_i} \tau_i. \quad (4.11)$$

O dual lagrangeano (DL) deste problema pode ser escrito como:

$$\text{DL : maximizar } \{ \theta(u) \text{ s.a } u \geq 0 \},$$

com $\theta(u) := \min_y L(y, u)$.

O mínimo da função Lagrangeana (4.5) é encontrado quando $\nabla_y L(y, u) = 0$. O gradiente da função Lagrangeana é dado por

$$\nabla_y L(y, u) = \nabla \ln[F(y)] + \sum_{i=1}^p u_i \nabla \ln[G_i(y)]. \quad (4.12)$$

O lado direito da equação (4.12) é composto de dois termos. O primeiro será calculado a seguir; o segundo foi calculado na Seção anterior. Dada a função,

$$F(y) = \sum_{i=1}^m \sum_{k \in J'_i} w_i \alpha_k e^{a_k^T y},$$

seu gradiente é

$$\nabla F(y) = \sum_{i=1}^m \sum_{k \in J'_i} w_i \alpha_k e^{a_k^T y} a_k = \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k a_k,$$

e o gradiente de seu logaritmo,

$$\nabla \ln F(y) = \frac{\nabla F(y)}{F(y)} = \frac{1}{F(y)} \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k a_k.$$

Assim, o gradiente da função Lagrangeana é dado por

$$\nabla_y L(y, u) = \frac{1}{F(y)} \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k a_k + \sum_{i=1}^p \sum_{k \in J_i} \frac{u_i}{G_i(y)} \left[\sum_{k \in J_i} \tau_k a_k \right].$$

Aplica-se novamente a transformação de variáveis. Para isso, definem-se $\delta_1, \delta_1, \dots, \delta_P$ de acordo com

$$\delta_k = \frac{w_i \tau_k}{F(y)}, \quad k \in J'_i, \quad i = 1, \dots, m. \quad (4.13)$$

$$\delta_k = \frac{u_i \tau_k}{G(y)}, \quad k \in J_i, \quad i = 1, \dots, p. \quad (4.14)$$

Reorganizando as dependências de $F, G_i, \tau_k, \delta_k, i = 1, \dots, p, k = 1, \dots, P$, a expressão (4.5) poderá ser reescrita em termos de (δ, u) eliminando a variável y do problema multiobjetivo. Das expressões (4.10) e (4.11), obtém-se

$$\delta_k = \sum_{i=1}^m \sum_{k \in J'_i} \frac{w_i \tau_k}{F(y)}.$$

Neste caso,

$$\begin{aligned} \sum_{i=1}^m \sum_{k \in J'_i} \delta_k &= \sum_{i=1}^m \sum_{k \in J'_i} \frac{w_i \tau_k}{F(y)} = \frac{1}{F(y)} \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k = \\ &= \frac{1}{F(y)} \sum_{i=1}^s w_i \sum_{k \in J'_i} \tau_k = \frac{1}{F(y)} F(y) = 1. \end{aligned}$$

Logo, tem-se que

$$\sum_{i=1}^m \sum_{k \in J'_i} \delta_k = 1 \quad \text{e} \quad \sum_{i=1}^p \sum_{k \in J_i} \delta_k = u_i. \quad (4.15)$$

As últimas equações são as já mencionadas *restrições de normalização*. Usando a equação (4.5) e as expressões (4.10) e (4.11), pode-se concluir que

$$\begin{aligned}\nabla_y L(y, u) &= \frac{\nabla F(y)}{F(y)} + \sum_{i=1}^p u_i \frac{\nabla G_i(y)}{G_i(y)} \\ &= \frac{1}{F(y)} \sum_{i=1}^m \sum_{k \in J'_i} w_i \tau_k a_k + \sum_{i=1}^p \sum_{k \in J_i} \frac{u_i}{G_i(y)} \left[\sum_{k \in J_i} \tau_k a_k \right] \\ &= \sum_{i=1}^m \sum_{k \in J'_i} \delta_k a_k + \sum_{i=1}^p \sum_{k \in J_i} \delta_k a_k.\end{aligned}$$

Conseqüentemente, $\nabla_y L(y, u) = 0$ implica em

$$\sum_{k=1}^P \delta_k a_k = 0. \quad (4.16)$$

sendo que P representa o número total de termos do problema.

Esta restrição é a correspondente *restrição de ortogonalidade*. Para concluir a construção do problema transformado, lembre-se que

$$L(y, u) := \ln[F(y)] + \sum_{i=1}^p u_i \ln[G_i],$$

e considere a expressão

$$u_i \ln[G_i(y)] = \sum_{i=1}^p \sum_{k \in J_i} \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^p u_i \ln(u_i), \quad (4.17)$$

obtida na Seção 4.1. Faltava calcular $\ln[F(y)]$ usando as equações (4.15) e (4.16). Como

$$\delta_k = \frac{w_i \tau_k}{F(y)}$$

obtem-se

$$\begin{aligned}\ln[F] &= \ln(w_i) + \ln(\tau_k) - \ln(\delta_k) \\ &= \ln(w_i) + \ln(\alpha_k e^{a_k^T y}) - \ln(\delta_k) \\ &= \ln(w_i) + \ln(\alpha_k) + a_k^T y - \ln(\delta_k) \\ &= \ln(w_i) + \ln\left(\frac{\alpha_k}{\delta_k}\right) + a_k^T y.\end{aligned}$$

Neste caso,

$$\sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln[F] = \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln(w_i) + \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^m \sum_{k \in J'_i} \delta_k a_k^T y,$$

implicando que

$$\ln[F] = \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln(w_i) + \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right). \quad (4.18)$$

Substituindo as equações (4.17) e (4.18) na expressão (4.5), obtém-se

$$\begin{aligned} L(y, u) &= \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln(w_i) + \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^p \sum_{k \in J_i} \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^p u_i \ln(u_i) \\ &= \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln(w_i) + \sum_{k=1}^P \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^p u_i \ln(u_i). \end{aligned}$$

O Problema Geométrico Multiobjetivo Dual equivalente assume a forma final

$$\begin{aligned} \text{maximizar} \quad & \sum_{i=1}^m \sum_{k \in J'_i} \delta_k \ln(w_i) + \sum_{k=1}^P \delta_k \ln\left(\frac{\alpha_k}{\delta_k}\right) + \sum_{i=1}^p u_i \ln(u_i) \\ \text{s.a} \quad & \sum_{k=1}^P \delta_k a_k = 0, \\ & \sum_{i=1}^m \sum_{k \in J'_i} \delta_k = 1, \\ & \sum_{i=1}^p \sum_{k \in J_i} \delta_k = u_i, \\ & \delta_k \geq 0, \\ & u_i \geq 0. \end{aligned}$$

Através de procedimentos análogos ao caso mono-objetivo, obteve-se um problema dual equivalente parametrizado por $w \in \mathcal{W}$.

O problema geométrico multiobjetivo pode ser resolvido através do seu dual, no qual as restrições são lineares e a função objetivo é côncava. A solução do dual fornece os valores ótimos das variáveis (δ, u) . O Teorema 22 é então usado para recuperar a solução ótima do problema primal resolvendo-se o sistema de equações (4.9).

A seguir são apresentados alguns exemplos adaptados de [Ruckaert & Matens, 1978], sobre os quais a técnica de programação geométrica multiobjetivo foi aplicada. Os pesos usados na resolução de cada problema foram $w_1 = 0.5$ e $w_2 = 0.5$ ($w_1 + w_2 = 1$). As funções objetivo e as restrições são detalhadas em cada exemplo e o problema geométrico multiobjetivo assume a

seguinte forma geral:

$$\begin{array}{ll} \text{minimizar} & w_1 f_1(x) + w_2 f_2(x) \\ \text{s.a} & g_1(x) \leq 1, \\ & g_2(x) \leq 1, \\ & x > 0. \end{array}$$

Problema 1

$$\begin{array}{ll} \text{Objetivos:} & f_1(x) = x_1^{-1}, \\ & f_2(x) = 5x_1 + 50000x_1^{-1} + 20x_2 + 72000x_2^{-1} + 10x_3 + 144000x_3^{-1}, \\ \text{Restrições:} & g_1(x) = x_1x_2^{-1} + 0.5x_3^{-1} \leq 1, \\ & g_2(x) = 0.001x_3x_4^{-1} + 0.01x_2 + 0.0005x_2x_4 \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Problema 2

$$\begin{array}{ll} \text{Objetivos:} & f_1(x) = 5x_1 + 50000x_1^{-1} + 20x_2 + 72000x_2^{-1} + 10x_3 + 144000x_3^{-1}, \\ & f_2(x) = x_1^{-1}, \\ \text{Restrições:} & g_1(x) = 4x_1^{-1} + 32x_2^{-1} + 120x_3^{-1} \leq 1, \\ & x_1, x_2, x_3 > 0. \end{array}$$

Problema 3

$$\begin{array}{ll} \text{Objetivos:} & f_1(x) = 168x_1x_2 + 3651.2x_1x_2x_3^{-1} + 3651.2x_1 + 40.000x_4^{-1}, \\ & f_2(x) = 5x_1 + 50000x_1^{-1} + 20x_2 + 72000x_2^{-1} + 10x_3 + 144000x_3^{-1}, \\ \text{Restrições:} & g_1(x) = 1.0425x_1x_2^{-1} \leq 1, \\ & g_2(x) = 0.00035x_1x_3 \leq 1, \\ & g_3(x) = 1.25x_1^{-1}x_4 + 41.63x_1^{-1} \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Problema 4

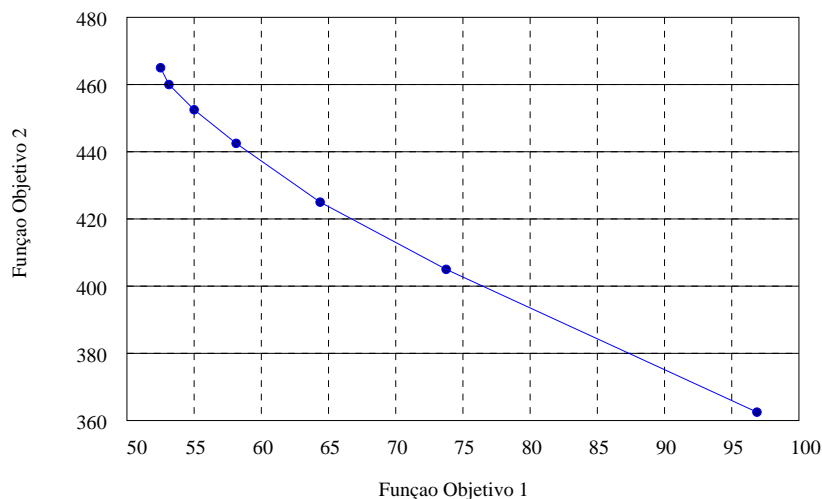
$$\begin{array}{ll} \text{Objetivos:} & f_1(x) = (592x_1^{0.65} + 582x_1^{0.39} + 1200x_1^{0.52} + 370x_1^{0.22}x_2^{-0.22} \\ & + 250x_1^{0.40}x_3^{-0.40} + 210x_1^{0.62}x_3^{-0.62} + 250x_1^{0.40}x_4^{-0.40} + 200x_1^{0.85}x_4^{-0.85}), \\ & f_2(x) = x_1^{-1}, \\ \text{Restrições:} & g_1(x) = x_1x_2^{-1} + 0.5x_3^{-1} \leq 1, \\ & g_2(x) = 0.001x_3x_4^{-1} + 0.01x_2 + 0.0005x_2x_4 \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

A Tabela 4.2 apresenta os resultados obtidos pela técnica de PG ao se resolver os problemas 1 a 4 com uma tolerância fixa $\epsilon = 0.001$. As soluções encontradas, o valor das funções $(f_1(x), f_2(x))$ e os tempos computacionais requeridos são apresentados.

Tabela 4.2: Soluções encontradas usando PG.

Problemas	Solução encontrada	Valor da função	Tempo (Seg)
1	(15.62,47.06,90.98,6.21)	(0.06,8241.61)	0.312
2	(25.42,85.10,204.27)	(9432.65,0.03)	0.0393
3	(40.96,35.08,74.31,1.52)	(487842.66,453.76)	0.344
4	(0.018,3.45,62.94,44.21)	(55.05,453.05)	0.281

Técnicas de escalarização como as implementadas neste trabalho são bastante utilizadas para gerar aproximações da curva de *trade-off* (curva de pontos eficientes) de problemas multiobjetivos, para posterior escolha de uma solução eficiente por parte de um tomador de decisões. A principal desvantagem deste método é o alto custo computacional envolvido. No caso de problemas multiobjetivos posinomiais, esse custo é substancialmente reduzido ao se adotar a formulação dual. A Figura 4.1 apresenta uma aproximação da curva de *trade-off* do Problema 4, obtida resolvendo-se o dual multiobjetivo para diferentes valores de w_1 e w_2 ($w_1 + w_2 = 1$) [Oliveira & Ferreira, 2003b].

Figura 4.1: Aproximação da curva *tradeoff*.

4.3 Programação Geométrica Signomial

O problema de Programação Geométrica Signomial (PGS) é caracterizado por uma função objetivo e/ou restrições que são diferenças de dois posinômios; o problema geométrico signomial contém um ou mais termos com coeficientes negativos. Cada monômio do problema é um produto de variáveis positivas, sendo que cada uma delas pode estar elevada a uma potência real, multiplicado por uma constante α_k real:

$$T_k = \alpha_k \prod_{j=1}^n x_j^{a_{jk}} .$$

Assim, um posinômio $g_i(x)$ é um somatório de monômios $T_k(x)$. Veja [Maranas & Floudas, 1997]. Agrupando-se os monômios de mesmo sinal, o problema (PGS) pode ser formulado como o seguinte problema de otimização não-linear:

$$\begin{aligned} \text{PGS} \quad & \text{minimizar} \quad g_0(x) = g_0^+(x) - g_0^-(x) \\ \text{s.a} \quad & g_i(x) = g_i^+(x) - g_i^-(x) \leq 0, \quad i = 1, 2, \dots, p. \\ & x_j^L \leq x_j \leq x_j^U, \quad j = 1, 2, \dots, n. \end{aligned} \quad (4.19)$$

sendo

$$\begin{aligned} g_i^+(x) &= \sum_{k \in k_i^+} \alpha_{ik} \prod_{j=1}^n x_j^{a_{jki}}, \\ g_i^-(x) &= \sum_{k \in k_i^-} \alpha_{ik} \prod_{j=1}^n x_j^{a_{jki}}, \quad i = 1, 2, \dots, p. \end{aligned}$$

Em (4.19), $x = (x_1, \dots, x_n)$ é o vetor de variáveis positivas; g_i^+ e g_i^- , para $i = 1, \dots, p$ são funções posinomiais em x ; a_{jki} são constantes reais arbitrárias e α_k são coeficientes positivos. Os conjuntos k_i^+ e k_i^- indicam quais monômios positivos e negativos formam os posinômios g_i^+ e g_i^- , respectivamente.

4.3.1 Transformação em Diferença de Funções Convexas

A função objetivo e as restrições da formulação original (PGS) são geralmente funções não-convexas. Aplicando-se em (4.19) a transformação $x_j = \exp(y_j)$, para $j = 1, \dots, n$, obtém-se o seguinte problema, denotado por (DC), cuja função objetivo e restrições são a diferenças de funções convexas:

$$\begin{aligned}
 \text{DC: } & \text{minimizar } G_0(y) = G_0^+(y) - G_0^-(y) \\
 & \text{s.a } G_i(y) = G_i^+(y) - G_i^-(y) \leq 0, \quad i = 1, 2, \dots, p, \\
 & y_j^L \leq y_j \leq y_j^U, \quad j = 1, 2, \dots, n,
 \end{aligned}$$

sendo

$$\begin{aligned}
 G_i^+(y) &= \sum_{k \in k_i^+} \alpha_{ik} \exp \left\{ \sum_{j=1}^n a_{jki} y_j \right\}, \\
 G_i^-(y) &= \sum_{k \in k_i^-} \alpha_{ik} \exp \left\{ \sum_{j=1}^n a_{jki} y_j \right\}, \quad i = 1, 2, \dots, p.
 \end{aligned}$$

Dado que $y_j^L = \ln(x_j^L)$, é necessário que $x_j^L > 0$, para $j = 1, 2, \dots, n$.

4.3.2 Limitante Inferior

Um limitante inferior para o problema (DC) pode ser obtido através da minimização de uma relaxação convexa. Tal relaxação convexa é conseguida subestimando cada função côncava $-G_i^-$ por uma função linear $-L_i^-$, para $i = 1, \dots, p$. A Figura 4.2 permite visualizar essa situação. Assim, obtém-se um problema Relaxado Convexo, denotado por (RC), que proporciona um limitante inferior para o problema (DC).

$$\begin{aligned}
 \text{RC: } & \text{minimizar } G_0^{\text{Conv}}(y) = G_0^+(y) - L_0^-(y) \\
 & \text{s.a } G_i^{\text{Conv}}(y) = G_i^+(y) - L_i^-(y) \leq 0, \quad i = 1, 2, \dots, p, \\
 & y_j^L \leq y_j \leq y_j^U, \quad j = 1, 2, \dots, n,
 \end{aligned}$$

sendo

$$G_i^+(y) = \sum_{k \in k_i^+} \alpha_{ik} \exp \left\{ \sum_{j=1}^n a_{jki} y_j \right\},$$

$$L_i^-(y) = \sum_{k \in k_i^-} \alpha_{ik} \left\{ A_{jk} + B_{jk} \left\{ \sum_{i=1}^N a_{jki} y_i \right\} \right\}, \quad j = 1, 2, \dots, P,$$

$$A_{jk} = \frac{Y_{jk}^U \exp(Y_{jk}^L) - Y_{jk}^L \exp(Y_{jk}^U)}{Y_{jk}^U - Y_{jk}^L},$$

$$B_{jk} = \frac{\exp(Y_{jk}^U) - \exp(Y_{jk}^L)}{Y_{jk}^U - Y_{jk}^L},$$

$$Y_{jk}^L = \sum_{j=1}^n \min(a_{jki} y_j^L, a_{jki} y_j^U),$$

$$Y_{jk}^U = \sum_{j=1}^n \max(a_{jki} y_j^L, a_{jki} y_j^U).$$

Veja [Maranas & Floudas, 1997] para detalhes sobre a obtenção das funções lineares $L_i^-(y)$. Observe que o limitante inferior de $-G_i^-$, $-L_i^-(y)$, é composto da soma de funções lineares. Claramente, quanto menor a diferença entre a função original $G_i^-(y)$ e a linearização $L_i^-(y)$, mais próxima a solução do problema (RC) estará da solução do problema (DC). A qualidade desses limitantes inferiores está relacionada à forma com que termos do tipo $-\exp(y)$ são aproximadas por funções lineares dentro do mesmo intervalo $[Y^L, Y^U]$.

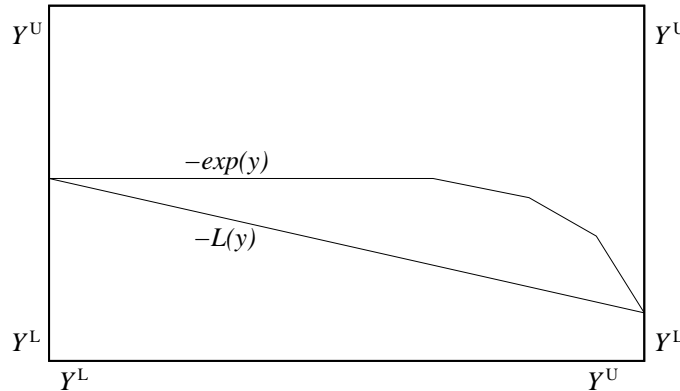


Figura 4.2: Relaxação Convexa (RC).

4.3.3 Redução da Partição do Retângulo

Um limitante inferior para (PGS) pode ser obtido resolvendo-se o problema (RC). A qualidade do limitante inferior é melhorada pelo particionamento do hiper-retângulo inicial ($y_j^L \leq y_j \leq y_j^U, j = 1, 2, \dots, n.$) e, em seguida, resolvendo o problema associado a cada um deles. Para alcançar este objetivo, pode-se tentar localizar os valores máximo e mínimo da variável y_j para os quais a restrição não-linear do problema (DC) permanece factível e o valor da função objetivo é menor que ou igual ao valor atual do limitante superior, U . Para cada $j = 1, 2, \dots, n$ resolve-se então

$$\begin{aligned} & \text{minimizar/maximizar} && y_j \\ & \text{s.a} && G_i(y) \leq 0, i = 1, 2, \dots, p, \\ & && G_0(y) \leq U, \\ & && y_j^L \leq y_j \leq y_j^U, j = 1, 2, \dots, n. \end{aligned}$$

No entanto, este é um problema não-convexo de dificuldade equivalente ao problema inicial (DC). Pode-se entretanto resolver o problema relaxado, substituindo todas as funções não-lineares por suas respectivas relaxações convexas.

$$\begin{aligned} & \text{minimizar/maximizar} && y_j, \\ & \text{s.a} && G_i^{conv}(y) \leq 0, \\ & && G_0^{conv}(y) \leq U, \\ & && y_j^L \leq y_j \leq y_j^U, j = 1, 2, \dots, n. \end{aligned}$$

A Figura 4.3 ilustra a minimização/maximização das variáveis y_j na região factível do problema relaxado, o que leva a novos limitantes para y (limites aprimorados). A relaxação convexa (área tracejada maior) é obtida através da substituição de toda função côncava por uma função linear. Enquanto a restrição convexa (área tracejada menor) é criada sempre que um ponto corrente for factível. Conceitos relacionados à restrição convexa será discutido com maiores detalhes na seção 4.3.5.

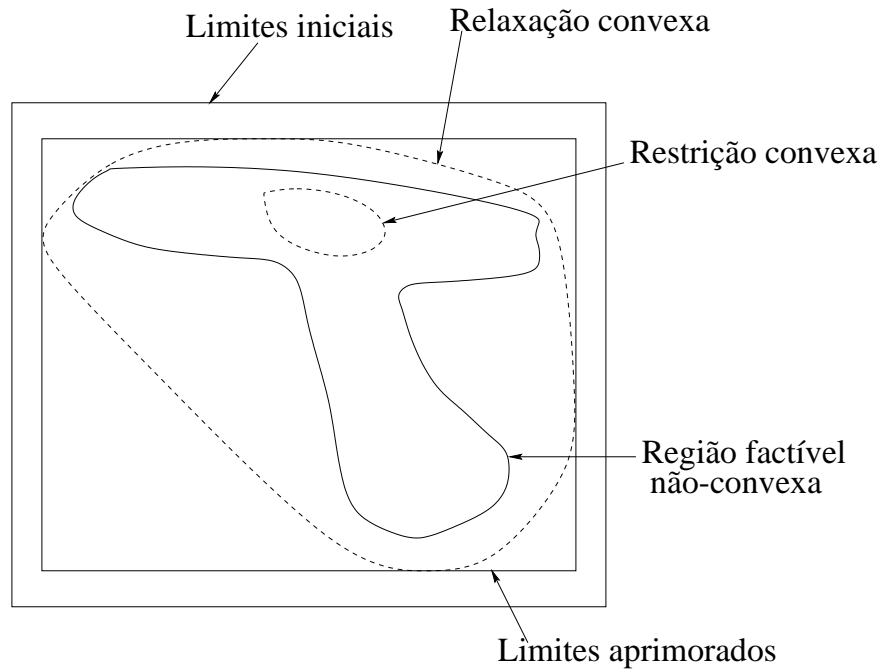


Figura 4.3: Redução dos limitantes iniciais através de relaxação convexa.

Apresenta-se a seguir os passos de um algoritmo que visa reduzir os limitantes iniciais do problema.

Algoritmo para Redução de Limitantes

Passo 0. Inicialização $j' \leftarrow 1$

$$y_{j'}^L = y_j^L$$

$$y_{j'}^U = y_j^U$$

Passo 1. Resolva

minimizar $y_{j'}$

s.a

$$G_i^{conv}(y) \leq 0, \quad i = 1, 2, \dots, p,$$

$$G_0^{conv}(y) \leq U,$$

$$y_j^L \leq y_j \leq y_j^U, \quad j = 1, 2, \dots, n,$$

obtendo $y_{j'}^*$;

Passo 2. Atribua

$$y_{j'}^{L,old} \leftarrow y_{j'}^L,$$

$$y_{j'}^L \leftarrow y_{j'}^*,$$

Passo 3. Resolvamaximizar $y_{j'}$

s.a

$$G_i^{conv}(y) \leq 0, \quad i = 1, 2, \dots, p,$$

$$G_0^{conv}(y) \leq U,$$

$$y_j^L \leq y_j \leq y_j^U, \quad j = 1, 2, \dots, n,$$

obtendo $y_{j'}^*$;**Passo 4. Atribua**

$$y_{j'}^{U,old} \leftarrow y_{j'}^U,$$

$$y_{j'}^U \leftarrow y_{j'}^*,$$

Passo 5. TesteSe $j' < n$ então $j' \leftarrow j' + 1$; volte para o **Passo 1**.

O algoritmo de otimização global *branch-and-bound* descrito em [Maranas & Floudas, 1997] é usado para localizar a solução mínima global do problema (PGS). Esse algoritmo é baseado na redução da diferença entre os limitantes superior e inferior do problema através da solução de sub-problemas de programação convexa.

4.3.4 Algoritmo *Branch-and-Bound* (BB)

O algoritmo *branch-and-bound* é uma técnica de otimização global que pode ser aplicada a problemas NP-Completo, encontrando o mínimo global de uma função não-convexa f sobre um hiper-retângulo $Q_{ini} \subset \mathbf{R}^n$. O hiper-retângulo é iterativamente subdividido até que o ótimo global seja obtido com uma dada precisão $\epsilon > 0$. Apesar de algumas vezes demandar um elevado esforço computacional, o algoritmo encontra uma solução ϵ -ótima global em tempo finito.

O método *branch-and-bound* pode ser interpretado como uma maneira conveniente de gerar subdivisões do espaço de busca, utilizando limites superiores e inferiores para refinar progressivamente as áreas de interesse, evitando que todo o espaço de busca seja investigado. Este procedimento termina quando a diferença entre o limite superior e o inferior é menor que a tolerância ϵ .

No algoritmo BB aplica-se uma regra para o particionamento do retângulo inicial Q_{ini} que contribui para a redução do número de buscas necessárias para encontrar a solução do problema. Para um retângulo $Q \subset Q_{ini}$, define-se $\Phi_{min}(Q) = \min_{x \in Q} f(x)$. O algoritmo encontra $\Phi_{min}(Q_{ini})$

com precisão $\epsilon > 0$, usando duas funções $\Phi_{lb}(Q)$ e $\Phi_{ub}(Q)$, definidas em $\{Q : Q \subset Q_{ini}\}$, tais que $\Phi_{lb}(Q) \leq \Phi_{min}(Q) \leq \Phi_{ub}(Q)$.

É apresentada a seguir uma descrição básica do algoritmo BB, cujos parâmetros são: k , contador de iteração; \mathcal{L}_k denota a lista de retângulos; L_k é o limitante inferior e U_k o limitante superior para $\Phi_{min}(Q_i)$ no final de cada iteração k .

Algoritmo *Branch-and-Bound*

Passo 1. Inicialização

$$k = 0;$$

$$\mathcal{L}_0 = \{Q_{init}\};$$

Passo 2. Processo de Divisão

Enquanto $U_k - L_k > \epsilon$

escolha $Q \in \mathcal{L}_k$ tal que $\Phi_{lb}(Q) = L_k$;

divida Q ao longo do maior lado em Q_1 e Q_2 ;

$$\mathcal{L}_{k+1} = (\mathcal{L}_k - Q) \cup \{Q_1, Q_2\};$$

$$L_{k+1} = \min_{Q \in \mathcal{L}_{k+1}} \Phi_{lb}(Q);$$

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

$$k = k + 1;$$

fim enquanto.

No processo de *ramificação* (particionamento de retângulos) descrito acima, o número de retângulos é igual ao número de iterações. Retângulos podem ser eliminados caso o ponto de mínimo não pertença ao retângulo considerado. Os retângulos $Q \in \mathcal{L}_k$ tais que $\Phi_{lb}(Q) > U_k$ são eliminados da lista \mathcal{L}_k .

O algoritmo de otimização global *BB* descrito em [Maranas & Floudas, 1997] é usado para localizar a solução mínima global do problema (PGS).

Algoritmo *Branch-and-Bound* [Maranas & Floudas, 1997]

Passo 1. Inicialização

Uma tolerância de convergência $\epsilon_c > 0$ e uma tolerância factibilidade $\epsilon_f > 0$ são definidas, um contador *Iter* é inicializado. Limitantes globais apropriados y_j^{LBD} e y_j^{UBD} para as variáveis y_j são calculados através dos procedimentos da Seção 4.3.3. Os limitantes locais $y_j^{L,Iter}$, $y_j^{U,Iter}$ na primeira iteração são considerados iguais aos globais. Finalmente, os limitantes inferior e superior G_0^{LBD} e G_0^{UBD} do valor mínimo global G_0^* são inicializados e o ponto inicial corrente

$y^{c,Iter}$, factível para o problema relaxado, é determinado através de um algoritmo baseado no método dos elipsóides [Kachian, 1979].

Passo 2. Checar a factibilidade e atualizar o limitante superior G_0^{UBD}

Se o máximo sobre todas as restrições G_i calculado no ponto corrente $y_j^{c,Iter}$ for menor que ϵ_f , isto é, se

$$\max_{i=1,2,\dots,p} G_i(y^{c,Iter}) \leq \epsilon_f,$$

então as restrições são ϵ_f -factíveis no ponto corrente. O valor da função objetivo G_0 é calculado no ponto corrente $y^{c,Iter}$ e o limitante superior G_0^{UBD} é avaliado como segue:

$$G_0^{UBD} = \min (G_0^{UBD}, G_0(y^{c,Iter}))$$

Passo 3. Partição do retângulo corrente

O retângulo corrente $[y_j^{L,Iter}, y_j^{U,Iter}]$ $j = 1, 2, \dots, n$ é particionado em dois sub-retângulos ($r = 1, 2$) na forma

$$\left[\begin{array}{cc} y_1^{L,Iter} & y_1^{U,Iter} \\ \vdots & \vdots \\ y_{l^{Iter}}^{L,Iter} & \frac{(y_{l^{Iter}}^{L,Iter} + y_{l^{Iter}}^{U,Iter})}{2} \\ \vdots & \vdots \\ y_n^{L,Iter} & y_n^{U,Iter} \end{array} \right], \left[\begin{array}{cc} y_1^{L,Iter} & y_1^{U,Iter} \\ \vdots & \vdots \\ \frac{(y_{l^{Iter}}^{L,Iter} + y_{l^{Iter}}^{U,Iter})}{2} & y_{l^{Iter}}^{U,Iter} \\ \vdots & \vdots \\ y_n^{L,Iter} & y_n^{U,Iter} \end{array} \right],$$

nos quais l^{Iter} corresponde à variável com o maior lado no retângulo inicial,

$$l^{Iter} = \arg \max_j (y_j^{U,Iter} - y_j^{L,Iter}).$$

Passo 4. Checar a factibilidade da relaxação convexa (RC)

Atualize em ambos os sub-retângulos ($r = 1, 2$) os parâmetros Y_{jk}^L , Y_{jk}^U , A_{jk} , B_{jk} , como segue:

$$\begin{aligned} Y_{jk}^L &= \sum_{j=1}^n \min (\alpha_{jki} y_j^L, \alpha_{ijk} y_j^U), \\ Y_{jk}^U &= \sum_{j=1}^n \max (\alpha_{jki} y_j^L, \alpha_{ijk} y_j^U), \\ A_{jk} &= \frac{Y_{jk}^U \exp(Y_{jk}^L) - Y_{jk}^L \exp(Y_{jk}^U)}{Y_{jk}^U - Y_{jk}^L}, \\ B_{jk} &= \frac{\exp(Y_{jk}^U) - \exp(Y_{jk}^L)}{Y_{jk}^U - Y_{jk}^L}. \end{aligned}$$

Se existir $i = 0, 1, \dots, p$ tal que um dos limitantes inferiores

$$G_i^{conv,L} = \sum_{k \in k_i^+} \alpha_{ik} \exp \left(\sum_{j=1}^n Y_{jk}^L \right) - \sum_{k \in k_i^-} \alpha_{ik} \left[A_{jk} + B_{jk} \left(\sum_{j=1}^n Y_{jk}^U \right) \right], \quad i = 0, 1, \dots, p,$$

satisfaça

$$G_0^{conv,L} \geq G_0^{UBD} \text{ ou } G_i^{conv,L} \geq 0 \text{ para algum } i = 1, 2, \dots, p,$$

então, elimine o sub-retângulo correspondente e vá para o Passo 6.

Passo 5. Solução do problema convexo dentro dos sub-retângulos

Resolva o problema de otimização convexa (RC) em ambos os sub-retângulos ($r = 1, 2$). Se a solução $G_0^{r,Iter}$ é menor que o limitante superior, então G_0^{UBD} é armazenado junto com a solução $y_{j,sol}^{r,Iter}$ encontrada.

Passo 6. Atualização do contador $Iter$ e do limitante inferior G_0^{LBD}

O contador é incrementado,

$$Iter \leftarrow Iter + 1,$$

e o novo limitante inferior é definido como o menor valor encontrado para os retângulos das iterações anteriores:

$$G_0^{LBD} = G_{0,sol}^{r',Iter'},$$

sendo $G_{0,sol}^{r',Iter'} = \min_{r,i} G_{0,sol}^{r,I}$, $r = 1, 2$, $I = 1, \dots, Iter - 1$. ($Iter'$ armazena a iteração na qual ocorre o mínimo e r' indica o retângulo associado.)

Passo 7. Atualizar o ponto corrente $y^{c,Iter}$ e os limitantes do retângulo $y^{L,Iter}$, $y^{U,Iter}$.

O novo ponto corrente é selecionado como aquele que leva ao mínimo do Passo 6:

$$y^{c,Iter} = y_{j,sol}^{r',Iter'}.$$

Particiona-se o retângulo correspondente ($r' = 1$ ou $r' = 2$) para obter os novos limitantes.

Passo 8. Checar a convergência

Se $(G_0^{UBD} - G_0^{LBD}) > \epsilon_c$, vá para o Passo 2. Caso contrário, o ponto atual satisfaz a condição ϵ_c -convergência e o ponto corrente é um mínimo global.

$$G_0^* = G_0(y^{c, Iter''}) \text{ no qual } Iter'' = \arg \{G_0(y^{c, I}) = G_0^{UBD}, I = 1, \dots, Iter\},$$

$$y^* = y^{c, Iter''}.$$

4.3.5 Busca Local Via Condensação

Um tipo particular de transformação baseado na desigualdade geométrica-aritmética é chamada de *Condensação*. Esta técnica foi proposta inicialmente por [Duffin et al., 1967]. O princípio básico da condensação é aproximar uma função posinomial multitermos por uma função de um único termo, isto é, por uma função *monomial*. Define-se o problema geométrico condensado como um problema geométrico na qual todas as funções posinomiais tenham sido reduzidas a funções de um único termo. Veja [Beightler & Phillips, 1976] para maiores detalhes.

Considere o problema de programação geométrico generalizado apresentado a seguir.

$$P_1 : \text{ minimizar } g_0(x)$$

$$\text{s.a } g_i(x) \leq \sigma_i, \quad i = 1, 2, \dots, p.$$

sendo que $g_i(x) = \sum_{k \in J_i} \alpha_k \prod_{j=1}^n x_j^{a_{kj}}$, $i = 0, 1, 2, \dots, p$. Os conjuntos J_i , com cada i representando a função objetivo e todas as restrições do problema, descrevem os termos da função objetivo e de cada restrição respectivamente. Através de uma substituição direta de variáveis, o problema P_1 pode ser escrito de forma equivalente como

$$P_2 : \text{ minimizar } x_0$$

$$\text{s.a } x_0^{-1} g_0(x) \leq 1.$$

$$g_i(x) \leq \sigma_i, \quad i = 1, 2, \dots, p,$$

$$x_j > 0, \quad j = 0, 1, 2, \dots, n.$$

No sentido de otimalidade, o problema P_2 é completamente equivalente ao problema P_1 , pois ambos os problemas possuem os mesmos pontos estacionários. Logo, o problema P_2 pode ser escrito na seguinte forma padrão:

$$P_3 : \text{ minimizar } x_0$$

$$\text{s.a } g_i(x) \leq \sigma_i, \quad i = 1, 2, \dots, p,$$

sendo que $g_i(x) = \sum_{k \in J_i} \alpha_k \prod_{j=1}^n x_j^{a_{kj}}$ com σ_i podendo assumir os valores 1 ou -1 .

Sem perda de generalidade, considere uma restrição i do conjunto de restrições do problema P_3 e defina $g_i(x) = g_i^+ - g_i^-(x) \leq \sigma_i$. Nota-se que $g_i^+(x)$ e $g_i^-(x)$ são, respectivamente, o conjunto de todos os termos positivos e negativos da restrição i . Consequentemente, $g_i^+(x)$ e $g_i^-(x)$ são funções posinomiais multitermos. A seguir analisa-se os dois possíveis casos para σ_i .

1º Caso: $\sigma_i = +1$

Obtém-se $g_i^+(x) - g_i^-(x) \leq 1$. Define-se $G_i^+(x) = g_i^+(x)$, e $G_i^-(x) = 1 + g_i^-(x)$.

Portanto, $\frac{G_i^+(x)}{G_i^-(x)} \leq 1$.

2º Caso: $\sigma_i = -1$

Obtém-se $g_i^+(x) - g_i^-(x) \leq -1$. Define-se $G_i^+(x) = 1 + g_i^+(x)$, e $G_i^-(x) = g_i^-(x)$.

Portanto, $\frac{G_i^+(x)}{G_i^-(x)} \leq 1$.

Novamente, $G_i^+(x)$ e $G_i^-(x)$ são funções posinomiais; o valor de σ_i adotado nesta tese foi $\sigma_i = 1$. Desta maneira, o problema P_3 pode ser reescrito como a seguir:

$$P_4 : \begin{array}{ll} \text{minimizar} & x_0 \\ \text{s.a} & \frac{G_i^+(x)}{G_i^-(x)} \leq 1, \quad i = 1, 2, \dots, p, \end{array}$$

sendo que $G_i^+(x) = g_i^+(x)$ e $G_i^-(x) = 1 + g_i^-(x)$.

O passo seguinte é condensar o posinômio $G_i^-(x)$ num monômio $G_i^-(x, \bar{x})$ no ponto $x = \bar{x}$, a partir da desigualdade média aritmética-geométrica.

$$\sum_{j=1}^n w_j x_j \geq \prod_{j=1}^n x_j^{w_j} \quad (4.20)$$

A desigualdade (4.20) é verdadeira sempre que $x_j > 0$ e $w_j > 0$ tal que $\sum_{j=1}^n w_j = 1$.

Considere a função posinomial $G_i^-(x) = \sum_{k \in J_i} \alpha_k \prod_{j=1}^n x_j^{a_{kj}}$ avaliada em um ponto de condensação \bar{x} e defina o peso w_k em $x = \bar{x}$ como

$$w_k = \frac{\alpha_k \prod_{j=1}^n \bar{x}_j^{a_{kj}}}{\sum_{k \in J_i} \alpha_k \prod_{j=1}^n \bar{x}_j^{a_{kj}}}.$$

A aplicação direta da desigualdade da média aritmética-geométrica (4.20) resulta no seguinte posinômio condensado:

$$G_i^-(x) \geq G_i^-(x, \bar{x}) = \prod_{k \in J_i} \left(\frac{\alpha_k \prod_{j=1}^n x_j^{a_{kj}}}{w_k} \right)^{w_k} = \prod_{k \in J_i} \left(\frac{\alpha_k \prod_{j=1}^n x^{a_{kj}} G_i^-(\bar{x})}{\alpha_k \prod_{j=1}^n \bar{x}^{a_{kj}}} \right) \left(\frac{\alpha_k \prod_{j=1}^n \bar{x}^{a_{kj}}}{G_i^-(\bar{x})} \right).$$

O qual se reduz a

$$G_i^-(x) \geq G_i^-(x, \bar{x}) = G_i^-(\bar{x}) \prod_{j=1}^n \left(\frac{x_j}{\bar{x}_j} \right) \left(\frac{\sum_{k \in J_i} a_{kj} \alpha_k \prod_{j=1}^n \bar{x}^{a_{kj}}}{G_i^-(\bar{x})} \right).$$

O monômio $G_i^-(x, \bar{x})$ condensa $G_i^-(x)$ no ponto \bar{x} . Se o ponto de condensação \bar{x} for factível para o conjunto de restrições originais, então

$$\frac{G_i^+(x)}{G_i^-(x, \bar{x})} \leq 1, \text{ uma vez } G_i^-(x) \geq G_i^-(x, \bar{x}).$$

Portanto, qualquer solução factível para $\frac{G_i^+(x)}{G_i^-(x)} \leq 1$ será factível para $\frac{G_i^+(x)}{G_i^-(x, \bar{x})} \leq 1$. Em geral, o inverso não será verdadeiro. Por fim, o problema condensado assume a seguinte forma padrão:

$$P_5 : \begin{array}{l} \text{minimizar } x_0 \\ \text{s.a} \quad \frac{G_i^+(x)}{G_i^-(x, \bar{x})} \leq 1, \quad i = 1, 2, \dots, p. \end{array}$$

O problema P_5 é um problema posinomial, pois a divisão do posinômio $G_i^+(x)$ pelo monômio $G_i^-(x, \bar{x})$ gera posinômios. Se x é factível para o problema condensado P_5 , então satisfaz o problema de programação geométrico original: $G_i^+(x) \leq G_i^-(x, \bar{x}) \leq G_i^-(x)$. Logo, a região factível do problema condensado está inteiramente contida no espaço de soluções do problema original. Desta forma, uma solução ótima para o problema condensado é subótima para o problema original. Um procedimento subótimo para resolução de problemas signomiais via condensação é resumido a seguir.

Algoritmo de Condensação

Passo 1. Reescrever o problema de programação geométrica signomial na forma de função objetivo linear, na forma de uma variável adicional x_0 , e restrições não-lineares como razões de

funções posinomiais:

$$\frac{G_i^+(x)}{G_i^-(x)} \leq 1, \quad i = 1, 2, \dots, p;$$

Passo 2. Condensar cada função posinomial $G_i^-(x), i = 1, 2, \dots, m$ num monômio $G_i^-(x, \bar{x})$ (função condensada), num dado ponto factível \bar{x} ;

Passo 3. Resolver o problema posinomial resultante através de técnica adequada (Seção 4.1);

Passo 4. O algoritmo termina quando a solução corrente for igual ao ponto de condensação. Não havendo convergência, retorna-se ao **Passo 2**, com a solução atual como o novo ponto da condensação.

Como exemplo, considere o problema de programação geométrica a seguir.

$$\begin{aligned} \text{minimizar} \quad & 2x_1x_2^{\frac{1}{2}} + x_3^{-1}x_2x_4^2 + x_1^{-2}x_2^{-1}x_3^2 \\ \text{s.a} \quad & x_1x_2^{\frac{1}{2}}x_3 - x_2^{-1}x_4^2 \leq 1, \\ & x_1x_2x_3 \geq 5, \\ & x_1, x_2, x_3, x_4 > 0. \end{aligned}$$

Este problema pode ser reescrito na seguinte forma equivalente:

$$\begin{aligned} \text{minimizar} \quad & x_0 \\ \text{s.a} \quad & 2x_0^{-1}x_1x_2^{\frac{1}{2}} + x_0^{-1}x_2x_3^{-1}x_4^2 + x_0^{-1}x_1^{-2}x_2^{-1}x_3^2 \leq 1, \\ & \frac{x_1x_2^{\frac{1}{2}}x_3}{1 + x_2^{-1}x_4^2} \leq 1, \\ & 5x_1^{-1}x_2^{-1}x_3^{-1} \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{aligned}$$

Assim, a primeira e a terceira restrições já estão na forma posinomial. É necessário condensar apenas a segunda restrição.

$$\begin{aligned} G_1(x) &= G_1^+(x) = 2x_0^{-1}x_1x_2^{\frac{1}{2}} + x_0^{-1}x_2x_3^{-1}x_4^2 + x_0^{-1}x_1^{-2}x_2^{-1}x_3^2; \\ G_2(x) &= G_2^+(x) - G_2^-(x), \quad G_2^+(x) = x_1x_2^{\frac{1}{2}}x_3, \quad G_2^-(x) = 1 + x_2^{-1}x_4^2; \\ G_3(x) &= G_3^+(x) = 5x_1^{-1}x_2^{-1}x_3^{-1}. \end{aligned}$$

A condensação de $G_2^-(x)$ em $x = \bar{x}$ é igual a [Beightler & Phillips, 1976]

$$G_2^-(x, \bar{x}) = \left[1 + \bar{x}_2^{-1}\bar{x}_4^2\right] \left[\frac{x_2}{\bar{x}_2}\right]^{\frac{-\bar{x}_2^{-1}\bar{x}_4^2}{1+\bar{x}_2^{-1}\bar{x}_4^2}} \left[\frac{x_4}{\bar{x}_4}\right]^{\frac{2\bar{x}_2^{-1}\bar{x}_4^2}{1+\bar{x}_2^{-1}\bar{x}_4^2}}.$$

Torna-se necessário escolher um ponto factível \bar{x} do problema original no qual a condensação será executada. O ponto escolhido foi $\bar{x}_1 = 10.0, \bar{x}_2 = 0.10, \bar{x}_3 = 5.0, \bar{x}_4 = 2.0$. Assim, função condensada é dada por

$$G_2^-(x, \bar{x}) = 1.22x_2^{-0.975}x_4^{1.95}.$$

O problema geométrico condensado é agora descrito da seguinte forma:

$$\begin{aligned} &\text{minimizar } x_0 \\ &\text{s.a} \quad 2x_0^{-1}x_1x_2^{\frac{1}{2}} + x_0^{-1}x_2x_3^{-1}x_4^2 + x_0^{-1}x_1^{-2}x_2^{-1}x_3^2 \leq 1, \\ &\quad 0.94x_1x_2^{1.475}x_3x_4^{-1.95} \leq 1, \\ &\quad 5x_1^{-1}x_2^{-1}x_3^{-1} \leq 1, \\ &\quad x_1, x_2, x_3, x_4 > 0. \end{aligned}$$

Como este problema posinomial tem grau de dificuldade igual a zero, a solução é facilmente obtida. A solução ótima do problema condensado (sub-ótima do problema original) é $x_1^* = 3.27, x_2^* = 0.54, x_3^* = 2.82, x_4^* = 1.90$.

A técnica da condensação foi usada como técnica de busca local no algoritmo *BB* com a finalidade de refinar o limitante superior. O Passo 2 do algoritmo *BB* foi adaptado para explorar esta técnica. A idéia fundamental consiste em criar uma *restrição convexa* do problema original sempre que o ponto corrente for factível. Com isso, resolvendo-se o problema convexo implicado, a melhoria do limitante superior seria acelerada. Os detalhes envolvidos na adaptação da técnica da condensação no Passo 2 do algoritmo *BB* serão apresentados a seguir.

Passo 2. Se $\max G_i(y^{c,Iter}) \leq \epsilon_f$ para $i = 1, 2, \dots, p$, então, o conjunto de restrições é ϵ_f -factível em $y^{c,Iter}$. Formula-se o problema condensado nas variáveis x_j , através da transformação $y_j = \ln(x_j)$ no ponto factível $y_j^{c,Iter} = \ln(x_j^{c,Iter})$ para $j = 1, 2, \dots, n$:

$$\begin{aligned} &\text{minimizar } x_{n+1} \\ &\text{s.a} \quad \frac{G_i^+(x)}{G_i^-(x, x^{c,Iter})} \leq 1, \quad i = 1, 2, \dots, p, \end{aligned}$$

sendo que

$$\begin{aligned}
 G_i^+(x) &= \sum_{k \in k_i^+} \alpha_{ik} \prod_{j=1}^{n+1} x_j^{a_{jki}}, \\
 G_i^-(x^{c,Iter}) &= \sum_{k \in k_i^-} \alpha_{ik} \prod_{j=1}^{n+1} (x_j^{c,Iter})^{a_{jki}}, \\
 H_{ji}(x^{c,Iter}) &= \sum_{k \in k_i^-} \alpha_{ik} \prod_{j=1}^{n+1} (x_j^{c,Iter})^{a_{jki}}, \\
 G_i^-(x, x^{c,Iter}) &= G_i^-(x^{c,Iter}) \prod_{j=1}^{n+1} \left(\frac{x_j}{x_j^{c,Iter}} \right) \left[\frac{H_{ji}(x^{c,Iter})}{G_i^-(x^{c,Iter})} \right].
 \end{aligned}$$

Se

$j = 1$, então $a_{jk(N+1)} = -1$;

$j > 1$, então

$a_{jk(N+1)} = 1$, se $\alpha_{ik} = 0$;

$a_{jk(N+1)} = 0$, caso contrário.

Definindo-se,

$$d_i = G_i^-(x^{c,Iter}) \quad \text{e} \quad \beta_{ji} = \frac{H_{ji}(x^{c,Iter})}{G_i^-(x^{c,Iter})},$$

tem-se que

$$\begin{aligned}
 G_i^-(x, x^{c,Iter}) &= d_i \prod_{j=1}^{n+1} \left(\frac{x_j}{x_j^{c,Iter}} \right)^{\beta_{ji}}, \\
 &= d_i \prod_{j=1}^{n+1} \left(\frac{1}{x_j^{c,Iter}} \right)^{\beta_{ji}} \prod_{j=1}^{n+1} x_j^{\beta_{ji}}, \\
 &= d_i \prod_{j=1}^{n+1} (x_j^{c,Iter})^{-\beta_{ji}} \prod_{j=1}^{n+1} x_j^{\beta_{ji}}.
 \end{aligned}$$

Logo,

$$\begin{aligned}
 \frac{G_i^+(x)}{G_i^-(x, x^{c,Iter})} &= \frac{\sum_{k \in k_i^+} \alpha_{ik} \prod_{j=1}^{n+1} x_j^{a_{jki}}}{d_i \prod_{j=1}^{n+1} (x_j^{c,Iter})^{-\beta_{ji}} \prod_{j=1}^{n+1} x_j^{\beta_{ji}}}, \\
 &= \sum_{k \in k_i^+} \frac{\alpha_{ik}}{d_i \prod_{j=1}^{n+1} (x_j^{c,Iter})^{-\beta_{ji}}} \prod_{j=1}^{n+1} x_j^{(a_{jki} - \beta_{ji})}, \\
 &= \sum_{k \in k_i^+} \frac{\alpha_{ik}}{d_i} \prod_{j=1}^{n+1} (x_j^{c,Iter})^{\beta_{ji}} \prod_{j=1}^{n+1} x_j^{(a_{jki} - \beta_{ji})}. \tag{4.21}
 \end{aligned}$$

Aplica-se em (4.21) a transformação $x_j = \exp(y_j)$ e $x_j^{c,Iter} = \exp(y_j^{c,Iter})$, para $j = 1, 2, \dots, n$ e resolve-se o problema condensado convexo:

$$\begin{aligned}
 &\text{minimizar } y_{N+1} \\
 &\text{s.a } \frac{G_i^+(y)}{G_i^-(y, y^{c,Iter})} \leq 0, \quad i = 1, 2, \dots, p.
 \end{aligned}$$

com

$$\frac{G_i^+(y)}{G_i^-(y, y^{c,Iter})} = \sum_{k \in k_i^+} \frac{\alpha_{ik}}{d_i} \exp \left\{ \sum_{j=1}^{n+1} \beta_{ji} y_j^{c,Iter} \right\} \exp \left\{ \sum_{j=1}^{n+1} (a_{jki} - \beta_{ji}) y_j \right\}.$$

Com a solução y_j^* , $j = 1, 2, \dots, n$, atualiza-se o limitante superior:

$$G_0^{UBD} = \min (G_0^{UBD}, G_0(y^*)).$$

O algoritmo BB com a busca local via condensação foi testado em diversos exemplos da literatura. Uma característica fundamental para o melhor desempenho do método proposto é a quantidade de pontos factíveis encontrados durante a execução do algoritmo. Nos problemas apresentados a seguir, esta característica não foi encontrada com frequência, de forma que a busca local não pôde ser muito efetiva. As soluções dos problemas são apresentadas na Tabela 4.3.

O Problema 1 encontra um ponto factível na primeira iteração. O limitante superior é atualizado para 18.32314. Depois de 8 iterações, o algoritmo encontra novamente outro ponto factível (ótimo global) e o critério de parada é satisfeito. Esse mesmo exemplo foi resolvido usando o algoritmo BB sem a busca local, o limitante superior inicial foi 49.88694. Sem a adoção da busca local, a região de busca torna-se maior, aumentando as possibilidades de pesquisa por

parte do algoritmo e, eventualmente, o esforço computacional necessário para a obtenção de ótimos globais.

Problema 1

$$\begin{array}{ll} \text{minimizar} & 3.7x_1^{0.85} + 1.985x_1 + 700.3x_2^{-0.75} \\ \text{s.a} & 0.7673x_2^{0.05} - 0.05x_1 \leq 1, \\ & x_1, x_2 > 0. \end{array}$$

Problema 2

$$\begin{array}{ll} \text{minimizar} & 0.5x_1x_2 - x_1 - 5x_2^{-1} \\ \text{s.a} & 0.01x_2x_3^{-1} + 0.01x_1 + 0.0005x_1^{-1}x_3 \leq 1, \\ & x_1, x_2, x_3 > 0. \end{array}$$

Problema 3

$$\begin{array}{ll} \text{minimizar} & -x_1 + 0.4x_1^{0.67}x_3^{-0.67} \\ \text{s.a} & 0.05882x_3x_4 + 0.1x_1 \leq 1, \\ & 4x_2x_4^{-1} + 2x_2^{-0.71}x_4^{-1} + 0.05882x_2^{-1.3}x_3 \leq 1, \\ & x_1, x_2, x_3, x_4 > 0. \end{array}$$

Problema 4

$$\begin{array}{ll} \text{minimizar} & -x_1 - x_5 + 0.4x_1^{0.67}x_3^{-0.67} + 0.4x_5^{0.67}x_7^{-0.67} \\ \text{s.a} & 0.05882x_3x_4 + 0.1x_1 \leq 1, \\ & 0.05882x_7x_8 + 0.1x_1 + 0.1x_5 \leq 1, \\ & 4x_2x_4^{-1} + 2x_2^{-0.71}x_4^{-1} + 0.05882x_2^{-1.3}x_3 \leq 1, \\ & 4x_6x_8^{-1} + 2x_6^{-0.71}x_8^{-1} + 0.05882x_6^{-1.3}x_7 \leq 1, \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8 > 0. \end{array}$$

Problema 5

$$\begin{array}{ll}
\text{minimizar} & x_6 + 0.4x_4^{0.67} + 0.4x_9^{0.67} \\
\text{s.a} & x_1^{-1}x_2^{-1.5}x_3x_4^{-1}x_5^{-1} + 5x_1^{-1}x_2^{-1}x_3x_5^{1.2} \leq 1, \\
& 0.05x_3 + 0.05x_2 \leq 1, \\
& 10x_3^{-1} - x_1x_3^{-1} \leq 1, \\
& x_6^{-1}x_7^{-1.5}x_8x_9^{-1}x_{10}^{-1} + 5x_6^{-1}x_7^{-1}x_8x_{10}^{1.2} \leq 1, \\
& x_2^{-1}x_7 + x_2^{-1}x_8 \leq 1, \\
& x_1x_8^{-1} - x_6x_8^{-1} \leq 1, \\
& 10x_{10} \leq 1, \\
& x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10} > 0.
\end{array}$$

Problema 6

$$\begin{array}{ll}
\text{minimizar} & 5.3578x_3^2 + 0.8357x_1x_5 + 37.2392x_1 \\
\text{s.a} & 0.00002584x_3x_5 + 0.00006663x_2x_5 + 0.0000734x_1x_4 \leq 1, \\
& 0.000853007x_2x_5 + 0.00009395x_1x_4 - 0.00033085x_3x_5 \leq 1, \\
& 1330.3294x_2^{-1}x_5^{-1} - 0.42x_1x_5^{-1} - 0.30586x_2^{-1}x_3x_5^{-1} \leq 1, \\
& 0.00024186x_2x_5 + 0.00010159x_1x_2 + 0.00007379x_3^2 \leq 1, \\
& 2275.1327x_3^{-1}x_5^{-1} - 0.2668x_1x_5^{-1} - 0.40584x_4x_5^{-1} \leq 1, \\
& 0.00029955x_3x_5 + 0.00007992x_1x_3 + 0.00012157x_3x_4 \leq 1, \\
& 78 \leq x_1 \leq 102, \quad 33 \leq x_2 \leq 45, \quad 27 \leq x_3 \leq 45, \\
& 27 \leq x_4 \leq 45, \quad 27 \leq x_5 \leq 45.
\end{array}$$

Problema 7

$$\begin{array}{ll}
\text{minimizar} & x_1 \\
\text{s.a} & \frac{x_1}{4} + \frac{x_2}{2} - \frac{x_1^2}{16} - \frac{x_2^2}{16} \leq 1, \\
& \frac{x_1^2x_2^{-1}}{6} + \frac{x_1}{6} - \frac{7x_1^{-1}}{3} - x_1x_2^{-1} \leq 1, \\
& x_1, x_2 > 0.
\end{array}$$

A Tabela 4.3 exibe uma comparação entre duas soluções (Solução 1 e Solução 2). A tolerância fixa considerada para esses problemas foi $\epsilon = 0.001$. A Solução 1 adota o algoritmo BB com a busca local via condensação; a Solução 2 é a fornecida em [Ruckaert & Matens, 1978].

O número de iterações necessárias para encontrar uma solução ótima global para cada problema também é apresentado. Não houve alteração no número de iterações quando a busca local foi removida. Analisando a coluna FunObj1 (FunObj2) pode-se observar que os valores ótimos globais são praticamente os mesmos nas duas abordagens.

Tabela 4.3: Soluções encontradas usando o BB com a Busca Local.

Prob	Iters	Solução 1	Solução 2	FunObj1 (FunObj 2)
1	8	(0.746,420.792)	(0.819,446)	11.9046 (11.91)
2	2	(87.97,8.21,1.36)	(88.31,7.45,1.31)	-83.23 (-83.21)
3	2	(7.42,0.63,0.75,5.76)	(8.13,0.61,0.56,5.63)	-5.5811 (-5.7398)
4	2	(5.87,0.68,1.16,6.03, 2.52,0.60,0.48,5.58)	(6.46,0.66,1.01,5.93, 2.23,0.59,0.40,5.52)	-6.0104 (-6.0482)
5	116	(1.73,10.86,8.25,0.91,0.29, 0.57,8.34,1.16,0.94,0.09)	(2.09,12.09,7.90,0.45,0.35, 0.45,10.45,1.64,1.19)	1.3354 (1.1436)
6	37	(78,33.32,31.07,45,34.13)	(78 33, 29.99,45, 36.7673)	10302.0 (10127.13)
7	10	(1.1790,2.1662)	(1.178, 2.178)	1.1790 (1.178)

4.4 Considerações Finais

Este capítulo abordou problemas de Programação Geométrica. Através de procedimentos análogos aos adotados no caso mono-objetivo, obteve-se um problema dual equivalente para problemas multiobjetivos representados como soma ponderada de funções. O problema dual encontrado representa uma pequena contribuição visando o futuro desenvolvimento de algoritmos especializados para Programação Geométrica Multiobjetivo.

Outro ponto abordado foi a implementação de um algoritmo *branch-and-bound* com busca local via condensação voltado para problemas de Programação Geométrica Signomial. Embora o algoritmo tenha apresentado bom desempenho na busca de soluções globais, experiências realizadas mostraram que a busca local foi pouco efetiva no sentido de reduzir o número total de iterações necessárias para convergência. Acredita-se que os procedimentos usados para geração de soluções iniciais factíveis, necessárias a várias rotinas, respondem em parte pelo baixo rendimento da busca local. Além disso, testes exaustivos não puderam ser realizados com o protótipo Matlab implementado. Uma implementação do tipo C++ combinada a pacotes de otimização especializados para programação não-linear permitiria tirar conclusões mais precisas sobre a busca local.

Programação Convexa Multiplicativa

A abordagem adotada neste capítulo para tratar problemas de Programação Convexa Multiplicativa explora conceitos e métodos de Programação Multiobjetivo. Um aspecto central do algoritmo proposto é a redução do problema à minimização de uma função quase-côncava sobre um politopo e sua resolução por meio de um procedimento de enumeração de vértices.

5.1 Programação Convexa Multiplicativa

A Programação Convexa Multiplicativa é uma classe de problemas de minimização envolvendo um produto de funções convexas em sua função objetivo ou em suas restrições. Pode-se interpretar os problemas convexas multiplicativos do ponto de vista da Programação Multiobjetivo, abordada no Capítulo 3. Cada fator da função objetivo do problema multiplicativo é visto como uma função objetivo individual e o produto de todos os fatores representa uma forma possível da função de preferências, a qual agrega os objetivos. Nesse contexto, as ferramentas matemáticas e algoritmos para tratar os problemas estudados neste capítulo guardam estreita relação com os do capítulo anterior, conforme será verificado ao longo do texto.

Considere o problema convexo multiplicativo

$$P_M : \begin{array}{l} \text{minimizar } F(f(x)) = \prod_{i=1}^m f_i(x) \\ \text{s.a} \quad g_i(x) \leq 0, \quad i = 1, 2, \dots, p, \end{array}$$

no qual $f_i : \mathbf{R}^n \rightarrow \mathbf{R}$ ($i = 1, 2, \dots, m$) e $g_j : \mathbf{R}^n \rightarrow \mathbf{R}$ ($j = 1, 2, \dots, p$) são funções convexas contínuas. Como é usual para a classe de problemas considerada, assume-se que

$$\Omega := \{x \in \mathbf{R}^n : g_j(x) \leq 0, j = 1, 2, \dots, p\} \quad (5.1)$$

é um conjunto convexo, compacto e não-vazio, e que cada f_i é positiva em Ω . Associamos a P_M a minimização da função objetivo vetorial $f := (f_1, f_2, \dots, f_m)$ sobre Ω , com $F : \mathbf{R}^n \rightarrow \mathbf{R}$ exercendo o papel de uma função de preferências especial [Yu, 1985] que agrega os objetivos f_1, f_2, \dots, f_m . A função F é geralmente não-convexa em Ω , mas quase-côncava sobre $\{f(x) : x \in \Omega\}$ [Horst et al., 1995].

Uma solução $x^* \in \Omega$ é uma solução eficiente do problema multiplicativo (multiobjetivo) P_M se não existir outro $x \in \Omega$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Como no Capítulo 3, o conjunto de todas as soluções eficientes de P_M é denotado como $efi(\Omega)$. Dado que

$$\frac{\partial F(f(x))}{\partial f_i(x)} = \prod_{j \neq i}^m f_j(x) > 0, \quad i = 1, 2, \dots, m$$

para todo $x \in \Omega$, segue que $F(f(x))$ é crescente em relação a cada $f_i(x)$, o que possibilita garantir a validade de uma propriedade fundamental demonstrada em [Geoffrion, 1967].

Teorema 23. *Seja $x^* \in \Omega$ uma solução ótima do problema multiplicativo (multiobjetivo) P_M . Então $x^* \in efi(\Omega)$.*

Da literatura de programação multiobjetivo convexa [Yu, 1985], sabe-se que se $x \in \Omega$ é uma solução eficiente de P_M então existe $w \in \mathbf{R}_+^m$ tal que x é também uma solução ótima do problema de programação convexa

$$P_W : \begin{array}{ll} \text{minimizar} & \langle w, f(x) \rangle \\ \text{s.a} & g_i(x) \leq 0, \quad i = 1, 2, \dots, p. \end{array}$$

Por outro lado, considere $x(w)$ uma solução ótima de P_W . Então $x(w)$ é eficiente se $w \in \mathbf{R}_+^m$. Lembrando que

$$\mathcal{W} := \{w \in \mathbf{R}_+^m : \sum_{i=1}^m w_i = 1\}, \quad (5.2)$$

o conjunto completo $efi(\Omega)$ pode ser gerado resolvendo-se P_W sobre $w \in \mathcal{W}$. A seguinte caracterização da solução ótima de P_M em termos do problema ponderado P_W foi introduzida por [Kathoh & Ibaraki, 1987].

Teorema 24. *Seja x^* uma solução ótima de P_M . Então qualquer solução ótima de P_W para $w = w^*$, com*

$$w_i^* = \prod_{j \neq i} f_j(x^*) > 0, \quad i = 1, 2, \dots, m, \quad (5.3)$$

é também ótima para (P_M) .

O Teorema 24 é um teorema de existência: o vetor de ponderação ótimo w^* depende da solução ótima (desconhecida) de P_M . Um método para obter w^* como o limite da seqüência gerada por um algoritmo é proposto neste capítulo.

5.1.1 Abordagem no Espaço dos Objetivos

A formulação no espaço dos objetivos do problema P_M é

$$P_{\mathcal{Y}} : \begin{array}{l} \text{minimizar } F(y) = \prod_{i=1}^m y_i \\ \text{s.a} \quad y \in \mathcal{Y}, \end{array}$$

no qual

$$\mathcal{Y} := \{y \in \mathbf{R}^m : y = f(x), \quad x \in \Omega\} \quad (5.4)$$

é o espaço dos objetivos. Formulações no espaço dos objetivos têm sido utilizadas com sucesso em Programação Convexa Multiplicativa [Kuno et al., 1993] e [Benson, 1999]. Nesta seção é proposta uma manipulação diferente de $P_{\mathcal{Y}}$, a qual combina elementos de análise convexa e otimização multiobjetivo no espaço dos objetivos [Ferreira & Machado, 1996].

Note que a continuidade de f e a compacidade de Ω implicam na compacidade de \mathcal{Y} . O conjunto de todas as soluções eficientes no espaço dos objetivos é dado por $\text{efi}(\mathcal{Y}) = f(\text{efi}(\Omega))$. Sabe-se que se $y \in \text{efi}(\mathcal{Y})$ então $y \in \partial\mathcal{Y}$, sendo $\partial\mathcal{Y}$ a fronteira de \mathcal{Y} . Além disso, \mathcal{Y} possui um hiperplano suporte em cada $y \in \text{efi}(\mathcal{Y})$.

Como discutido no Capítulo 3, o problema ($P_{\mathcal{Y}}$) admite uma formulação equivalente no espaço dos objetivos com uma região factível convexa:

$$P_{\mathcal{F}} : \begin{array}{l} \text{minimizar } F(y) \\ \text{s.a} \quad y \in \mathcal{F}, \end{array}$$

com $\mathcal{F} := \mathcal{Y} + \mathcal{D}$ e $\mathcal{D} := \{z \in \mathbf{R}^m : z = y + d, \quad d \in \mathbf{R}_+^m, \quad y \in \mathcal{Y}\}$. Além disso, $y \in \mathcal{F}$ se e somente se y satisfaz o sistema de infinitas desigualdades lineares

$$\min_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0 \quad \text{para todo } w \in \mathcal{W}. \quad (5.5)$$

Em termos práticos, $y \in \mathcal{F}$ se e somente se $\theta(y) \leq 0$, com

$$\theta(y) := \max_{w \in \mathcal{W}} \phi(w) \quad (5.6)$$

e

$$\phi(w) := \min_{x \in \Omega} \langle w, f(x) - y \rangle. \quad (5.7)$$

O algoritmo utilizado para obter $\theta(y)$ foi discutido no Capítulo 3, e é reintroduzido a seguir por conveniência.

Algoritmo A_1

Passo 0. Escolha $w^0 \in \mathcal{W}$ e faça $l \leftarrow 0$;

Passo 1. Resolva o problema de programação convexa

$$P_W : \begin{array}{ll} \text{minimizar} & \langle w^l, f(x) \rangle \\ \text{s.a} & x \in \Omega, \end{array}$$

obtendo $x(w^l)$;

Passo 2. Resolva o problema de programação linear

$$P_L : \begin{array}{ll} \text{maximizar} & \sigma \\ \text{s.a} & \sigma \leq \langle w, f(x(w^i)) - y \rangle, \quad i = 0, 1, \dots, l \\ & w \in \mathcal{W}, \sigma \in \mathbf{R}. \end{array}$$

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

5.1.2 Problema Mestre

Seja $\mathcal{F}^0 := \{y \in \mathbf{R}^m : 0 < \underline{y} \leq y \leq \bar{y}\}$ uma aproximação externa inicial de \mathcal{F} . Com as hipóteses iniciais sobre P_M , o mínimo de F sobre \mathcal{F}^0 é atingido em $y^0 = \underline{y}$. Pelo Algoritmo A_1 deve-se concluir que $\theta(y^0) > 0$, isto é, $y^0 \notin \mathcal{F}$. Ao convergir, o Algoritmo A_1 fornece o hiperplano

$$\mathcal{H}^0 := \{y \in \mathbf{R}^m : \langle w^0, y \rangle = \langle w^0, f(x(w^0)) \rangle\}, \quad (5.8)$$

que suporta \mathcal{F} em $y = f(x(w^0))$. O semi-espaço

$$\mathcal{H}_+^0 := \{y \in \mathbf{R}^m : \langle w^0, y \rangle \geq \langle w^0, f(x(w^0)) \rangle\} \quad (5.9)$$

não contém y^0 ; sua intersecção com \mathcal{F}^0 fornece uma aproximação externa melhorada de \mathcal{F} pois, ao remover y^0 , uma parte significativa de \mathcal{F}^0 é também removida pelo *plano de corte* \mathcal{H}^0 . Imaginando \mathcal{F} como um conjunto infinitamente restrito, a restrição mais violada por y^0 é \mathcal{H}_+^0 , no sentido que w^0 maximiza o lado esquerdo de (5.5). Com base neste argumento, pode-se dizer que \mathcal{H}^0 produz um *corte máximo* em \mathcal{F} .

O problema a ser resolvido em uma iteração arbitrária k é formulado como

$$P_{\mathcal{F}^k} : \begin{array}{ll} \text{minimizar} & F(y) \\ \text{s.a} & y \in \mathcal{F}^k. \end{array}$$

Dado que F é uma função contínua quase-côncava e \mathcal{F}^k é um politopo, um mínimo global de $(P_{\mathcal{F}^k})$ é atingido em um vértice de \mathcal{F}^k [Horst et al., 1995]. O número de vértices de \mathcal{F}^k está associado a m . Quando m é pequeno, o que é frequentemente verdade na prática [Kuno et al., 1993], o problema $(P_{\mathcal{F}^k})$ pode ser resolvido por um procedimento de enumeração de vértices. O procedimento aqui adotado é o seguinte: inicialmente, \mathcal{F}^0 tem 2^m vértices e a solução de $(P_{\mathcal{F}^0})$ é, obviamente, o vértice $y^0 = \underline{y}$. Em uma iteração arbitrária k , o conjunto dos vértices de

$$\mathcal{F}^k \leftarrow \mathcal{F}^k \cap \mathcal{H}_+^k, \quad (5.10)$$

denotado como $\mathcal{V}(\mathcal{F}^k)$, é determinado por um algoritmo denominado de Algoritmo de Lista de Adjacências [Chen et al., 1991], a ser detalhado na Seção 5.1.5. A implementação proposta deste algoritmo pode resolver problemas degenerados, discutidos em detalhes em [Horst et al., 1995]. Então

$$y^k = \arg \min_{y \in \mathcal{V}(\mathcal{F}^k)} F(y) \quad (5.11)$$

resolve globalmente $(P_{\mathcal{F}^k})$.

5.1.3 Algoritmo Global

Nesta Seção será formalizado um algoritmo de otimização global para resolver problemas de Programação Convexa Multiplicativa por aproximações externas; algumas de suas características serão discutidas.

Algoritmo A_2

Passo 0. Encontre \mathcal{F}^0 e faça $k \leftarrow 0$;

Passo 1. Resolva o problema multiplicativo

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

obtendo y^k ;

Passo 2. Encontre $\theta(y^k) = \langle w^k, f(x(w^k)) - y^k \rangle$ através do Algoritmo A_1 . Se $\theta(y^k) < \epsilon_2$, onde $\epsilon_2 > 0$ é uma tolerância pequena, pare: y^k e $x(w^k)$ são soluções ϵ_2 -ótimas de $(P_{\mathcal{F}})$ e (P_M) , respectivamente. Caso contrário, defina

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

faça $k \leftarrow k + 1$ e retorne ao Passo 1.

Teste de convergência: Assumindo que um mínimo global é sempre determinado no Passo 1 do Algoritmo A_2 , a convergência infinita do Algoritmo A_2 para um mínimo global de $(P_{\mathcal{F}})$ e de (P_M) decorre do Teorema 21 do Capítulo 3.

Politopo Inicial: A aproximação externa inicial de \mathcal{F} assume a forma $\mathcal{F}^0 = \{y \in \mathbf{R}^m : 0 < \underline{y} \leq y \leq \bar{y}\}$, com o limitante inferior \underline{y} dado por

$$\underline{y}_i := \min_{x \in \Omega} f_i(x), \quad i = 1, 2, \dots, m. \quad (5.12)$$

O limitante superior \bar{y} pode ser definido como os máximos individuais das funções convexas f_1, f_2, \dots, f_m sobre Ω , o que exige a solução de m problemas de maximização de funções convexas. As características particulares dos problemas multiplicativos considerados ou as abordagens analíticas propostas por [Kuno et al., 1993] e [Benson, 1999], podem ser usadas para encontrar \bar{y} .

Critério de Convergência: Suponha que $y^k \notin \mathcal{F}$ e considere o problema de programação convexa

$$\begin{aligned} P : \quad & \text{minimizar} \quad \eta \\ & \text{s.a} \quad \quad \quad y \leq \eta e + y^k, \\ & \quad \quad \quad y \in \mathcal{F}, \eta \in \mathbf{R}, \end{aligned}$$

sendo que e denota o vetor m -dimensional cujos elementos possuem valores iguais a um. Pelos resultados de [Geoffrion, 1970], o problema dual Lagrangeano associado a (P) é

$$D : \quad \text{maximizar} \min_{y \in \mathcal{F}} \langle w, y - y^k \rangle,$$

no qual w é o multiplicador de Lagrange associado às restrições lineares de desigualdade de (P) e \mathcal{W} é o domínio da função dual

$$\min_{y \in \mathcal{F}} \langle w, y - y^k \rangle.$$

Seja (w^*, y^*) o par ótimo associado a (D) e suponha que $w^* \in \mathbf{R}_{++}^m$. Se $y^* \in \mathcal{F}$ mas $y^* \notin \mathcal{Y}$, a otimalidade de y^* com $w^* \in \mathbf{R}_{++}^m$ não seria válida. Portanto, $y^* \in \mathcal{Y}$ e o valor ótimo dual é

$$\theta(y^k) =: \langle w^*, y^* - y^k \rangle.$$

Uma vez que os valores ótimos de (P) e (D) devem ser iguais, $\eta^* = \theta(y^k)$. Além disso, como $w^* \in \mathbf{R}_{++}^m$, as restrições de desigualdade de P devem estar ativas, implicando que $y^* = \theta(y^k)e + y^k$. Então, uma interessante interpretação geométrica é estabelecida: $\theta(y^k)$ é a norma infinito entre y^k e a região factível de \mathcal{F} . É importante notar que $w^* \in \mathbf{R}_{++}^m$ para um mínimo global de P_M (Teorema 24). Assim, ao fixar $\epsilon_2 > 0$, estabelece-se uma distância máxima para \mathcal{F} ao término do Algoritmo A_2 .

Esforço Computacional: A maior parte do esforço computacional requerido pelo Algoritmo A_2 está concentrado no Passo 2, em que $\theta(y^k)$ é calculado pelo Algoritmo A_1 . Enquanto o esforço computacional empregado nas minimizações por programação linear (Passo 2 do A_1) é relativamente pequeno, para as não-lineares (Passo 1 do A_1) um esforço maior é exigido, apesar da convexidade das funções permitir o uso de métodos eficientes de programação convexa. Os esforços de codificação e preparação relacionados à abordagem proposta (Algoritmos A_1 e A_2) parecem ser pequenos comparados com outras abordagens disponíveis na literatura [Oliveira & Ferreira, 2003a].

5.1.4 Enumeração de Vértices

Um mínimo global de uma função f côncava ou quase-côncava (como no problema $\mathcal{P}_{\mathcal{F}}$) sobre um polítopo $\mathcal{P} := \{x \in \mathbf{R}^n : Ax \leq b\}$ é obtido em um dos seus vértices. Podemos então resolver um problema de minimização quase-côncava sobre polítopos determinando todos os vértices de \mathcal{P} e comparando os valores correspondentes da função objetivo. Assim, está-se diante do problema clássico de encontrar a representação de um polítopo descrito por um sistema de desigualdades lineares através de seus vértices.

Freqüentemente é preferível construir uma aproximação externa de \mathcal{P}_s por um polítopo simples $\mathcal{P}_s \supset \mathcal{P}$, com um conjunto de vértices $\mathcal{V}(\mathcal{P}_s)$ conhecido. Então, um vértice $v^* \in \mathcal{V}(\mathcal{P}_s)$ satisfazendo

$$f(v^*) = \min\{f(v) : v \in \mathcal{V}(\mathcal{P}_s)\} = \min\{f(x) : x \in \mathcal{P}_s\} \quad (5.13)$$

fornece o limitante inferior $f(v^*) \leq \min\{f(x) : x \in \mathcal{P}\} = f(x^*)$ ($\mathcal{P}_s \supset \mathcal{P}$).

Se acontecer de v^* pertencer a \mathcal{P} , então, obviamente, v^* deve minimizar f sobre \mathcal{P} , e o problema está resolvido. Quando v^* não pertence a \mathcal{P} e um ponto z interior a \mathcal{P} é conhecido, a interseção y do segmento $[z, v^*]$ com a fronteira de \mathcal{P} produz uma solução factível aproximada.

Para melhorar a aproximação externa inicial, define-se o conjunto

$$\mathcal{P}_s \leftarrow \mathcal{P}_s \cap \{x : (a^i)^T x \leq b_i\}, \quad (5.14)$$

no qual $(a^i)^T x \leq b_i$ é uma das restrições que define \mathcal{P} , a qual é violada por v^* ($(a^i)^T v^* > b_i$) e determina-se um novo v^* via (5.13). Claramente, a desigualdade $(a^i)^T x \leq b_i$ elimina uma parte não-vazia do polítopo inicial \mathcal{P} (porque $a_i^T v^* > b_i$). Por esta razão, o hiperplano correspondente $\mathcal{H} = \{x : (a^i)^T x = b_i\}$ é frequentemente chamado de plano de corte.

Repetindo este procedimento obtém-se um método de aproximações externas por (hi-per)planos de corte, que termina depois de um número finito de iterações com $v^* \in \mathcal{P}$, ou quando a aproximação y correspondente é suficientemente boa devido a $f(y) - f(v^*) \leq c$ para uma dada tolerância $c > 0$. Quando $\mathcal{P} = \{x : (a^k)^T x \leq b_k, k = 1, 2, \dots, m\}$, pode acontecer

que, no pior caso, o método leve m iterações para reconstruir \mathcal{P} através de cortes sucessivos. Neste caso, certamente o método não será melhor do que uma enumeração direta dos vértices de \mathcal{P} .

A seguir, discute-se o sub-problema de determinar novos vértices de $\mathcal{P}_s \leftarrow \mathcal{P}_s \cap \{x : (a^i)^T x \leq b_i\}$, que surge de (5.13) quando \mathcal{P} é substituído por (5.14).

5.1.5 Problema de Enumeração de Vértices

Seja $\mathcal{P}_s := \{x \in \mathbf{R}^n : g_i(x) = (a^i)^T x - b_i \leq 0, i = 1, 2, \dots, m\}$, com $a^i \in \mathbf{R}^n \setminus \{0\}, b^i \in \mathbf{R}, i = 1, 2, \dots, m$, um politopo com um conjunto de vértices conhecido, $\mathcal{V}(\mathcal{P})$, e considere $\mathcal{H} := \{x \in \mathbf{R}^n : g_{m+1}(x) = (a^{m+1})^T x - b_{m+1} = 0\}$ com $a^{m+1} \in \mathbf{R}^n \setminus \{0\}, b_{m+1} \in \mathbf{R}$, um hiperplano tal que

$$\overline{\mathcal{P}_s} = \mathcal{P}_s \cap \mathcal{H}$$

não é vazio, nem uma face de \mathcal{P} . Determine o conjunto de vértices $\mathcal{V}(\overline{\mathcal{P}_s})$ de $\overline{\mathcal{P}_s}$.

Este problema difere do problema geral de encontrar vértices de um politopo com representação na forma de desigualdades porque dispomos do conhecimento adicional do conjunto de vértices $\mathcal{V}(\mathcal{P}_s)$. Considere

$$\begin{aligned} \mathcal{V}^+ &= \mathcal{V}^+(\mathcal{P}_s) := \{v \in \mathcal{V}(\mathcal{P}_s) : g_{m+1}(v) > 0\}, \\ \mathcal{V}^- &= \mathcal{V}^-(\mathcal{P}_s) := \{v \in \mathcal{V}(\mathcal{P}_s) : g_{m+1}(v) < 0\}, \end{aligned}$$

e defina $\nu := |\mathcal{V}(\mathcal{P}_s)|, \nu^+ := |\mathcal{V}^+(\mathcal{P}_s)|, \nu^- := |\mathcal{V}^-(\mathcal{P}_s)|$ e $\bar{\nu} := |\mathcal{V}(\overline{\mathcal{P}_s})|$ como sendo o número de vértices de $\mathcal{P}_s, \mathcal{P}_s \cap \{x : g_{m+1}(x) > 0\}, \mathcal{P}_s \cap \{x : g_{m+1} < 0\}$ e $\overline{\mathcal{P}_s}$, respectivamente.

Todos os algoritmos conhecidos para resolver o problema considerado são baseados na seguinte caracterização [Horst & Pardalos, 1995].

Teorema 25. $w \in \mathcal{V}(\overline{\mathcal{P}_s})$ se e somente se w é um vértice de \mathcal{P}_s em \mathcal{H} ou um ponto onde uma aresta $[u, v]$ de $\mathcal{P}_s, u \in \mathcal{V}^-(\mathcal{P}_s), v \in \mathcal{V}^+(\mathcal{P}_s)$, intercepta \mathcal{H} .

Uma aplicação direta do Teorema 25 leva ao *problema da vizinhança*. A suposição de que $\overline{\mathcal{P}_s}$ não é vazio, nem face de \mathcal{P}_s , que surge do método de aproximação externa subjacente, implica que $\nu^- \neq 0, \nu^+ \neq 0$.

Sem perda de generalidade, supõe-se que $\nu^- \leq \nu^+$. Caso contrário, os papéis \mathcal{V}^- e \mathcal{V}^+ terão de ser trocados na discussão abaixo. Um vértice v de um politopo \mathcal{P}_s é um vizinho do vértice u de \mathcal{P}_s se o segmento $[u, v]$ é uma aresta de \mathcal{P}_s . Então, pelo Teorema 25, o problema é equivalente a encontrar, para cada vértice $u \in \mathcal{V}^-$, seus vizinhos em \mathcal{V}^+ . Seja $\mathcal{N}(u)$ o conjunto

de vizinhos (lista de adjacência) de $u \in \mathcal{V}^-$ em \mathcal{P}_s . Dado $u \in \mathcal{V}^-$ e $v \in \mathcal{N}(u) \cap \mathcal{V}^+$, o ponto w onde a aresta $[u, v]$ intercepta o hiperplano \mathcal{H} , é,

$$w = \alpha u + (1 - \alpha)v,$$

com

$$\alpha = g_{m+1}(v)/(g_{m+1}(v) - g_{m+1}(u)),$$

é um novo vértice, e todos os novos vértices em $\mathcal{V}(\overline{\mathcal{P}}_s)$ podem ser determinados dessa forma. Assuma primeiramente que todos os vértices de \mathcal{P}_s e $\overline{\mathcal{P}}_s$ são *não-degenerados*. Um vértice não-degenerado é um vértice de \mathcal{P}_s que tem exatamente n vértices adjacentes a ele, onde n é a dimensão do politopo. A degenerescência pode ocorrer quando num processo de sucessivos planos de cortes acontecer de um ou mais vértices do politopo \mathcal{P}_s estar sobre o hiperplano \mathcal{H} . Esta situação pode ser tratada de várias maneiras. A abordagem adotada neste trabalho, consiste em checar, no Passo 2 do algoritmo de lista de Adjacências, se algum dos vértices do politopo está sobre o hiperplano, e então introduzir uma pequena perturbação em b_{m+1} , evitando assim a degenerescência [Oliveira & Ferreira, 2004b].

O algoritmo a seguir assume ainda que para cada vértice $u \in \mathcal{P}_s$ a lista de adjacência $\mathcal{N}(u)$ de seus vizinhos em \mathcal{P}_s é conhecida. Seja $\mathcal{J}(u) := \{i \in \{1, 2, \dots, m\} : g_i(u) = 0\}$, o conjunto de índices das restrições ativas no vértice u , $\nu^- \leq \nu^+$ e $\varepsilon > 0$ um escalar muito pequeno.

Algoritmo de Lista de Adjacências

Passo 1. Inicialização

Determinar $\mathcal{V}^+, \mathcal{V}^-, \nu^+, \nu^-$;

Se $\nu^- = 0$ **então** Pare ($\overline{\mathcal{P}}_s$ é vazio);

Se $\nu^+ = 0$ **então** Pare (restrição redundante);

Passo 2. Evitar degenerescência

enquanto $|\nu^+ \cup \nu^-| < |\overline{\mathcal{P}}_s|$

$$b_{m+1} = b_{m+1} + \varepsilon;$$

fim enquanto

Passo 3. Novos vértices

$\mathcal{V}(\overline{\mathcal{P}}_s) \leftarrow \emptyset$;

para todo $u \in \mathcal{V}^-$ **faça**

para todo $v \in \mathcal{N}(u) \cap \mathcal{V}^+$ **faça**
determinar $w = [u, v] \cap \mathcal{H}$;
 $(\overline{\mathcal{P}}_s) \leftarrow \mathcal{V}(\overline{\mathcal{P}}_s) \cup \{w\}$;
 $\mathcal{N}(u) = (\mathcal{N}(u) \setminus \{v\}) \cup \{w\}$;
 $\mathcal{N}(w) \leftarrow \{u\}$;
 $\mathcal{J}(w) \leftarrow (\mathcal{J}(u) \cap \mathcal{J}(v)) \cup \{m + 1\}$;

fim para

fim para

Passo 4. Vizinhos dos novos vértices

para todo $u \in \mathcal{V}(\overline{\mathcal{P}}_s), v \in \mathcal{V}(\overline{\mathcal{P}}_s)$ **faça**

se $|\mathcal{J}(u) \cap \mathcal{J}(v)| = n - 1$ **então**

$\mathcal{N}(u) \leftarrow \mathcal{N}(u) \cup \{v\}$;

$\mathcal{N}(v) \leftarrow \mathcal{N}(v) \cup \{u\}$;

fim se

fim para

Com respeito à complexidade do algoritmo, assumindo não-degenerescência verifica-se que, na implementação direta do algoritmo, a complexidade do Passo 1 é $O(n(v^+ + v^-))$, a complexidade do Passo 3 é $O(n(v^- + v^+))$ e a complexidade do Passo 4 é $O(n\bar{v})$ no pior caso. Se ocorrer degenerescência, então ao invés de simplesmente checar se $|\mathcal{J}(u) \cap \mathcal{J}(v)| = n - 1$, deve-se checar se o posto da matriz dos coeficientes das restrições em $\mathcal{J}(u) \cap \mathcal{J}(v)$ é igual a $n - 1$. No Passo 2, esse procedimento requer $O(mn^2v^-v^+)$ operações, no pior caso.

Alguns comentários sobre a implementação do Algoritmo de Lista de Adjacências.

Planos de Corte: O principal objetivo do Passo 1 é evitar planos de corte e execuções desnecessárias do algoritmo. Ao separar os vértices em dois conjuntos, acima (\mathcal{V}^+) e abaixo (\mathcal{V}^-) do plano de corte, garante-se que, se não houver vértices em \mathcal{V}^+ , a restrição não elimina vértice algum do politopo; se não houver vértices em \mathcal{V}^- , o plano de corte elimina todo o politopo.

Degenerescência: O Passo 2 tem como objetivo a eliminação da degenerescência. Para eliminá-la, uma pequena perturbação (ϵ) é aplicada sobre o termo b_{m+1} , sucessivamente até que a alteração seja suficiente para que nenhum vértice esteja mais sobre o plano de corte, dentro da precisão estipulada.

5.1.6 Experimentos Numéricos

Considere inicialmente o exemplo ilustrativo discutido em [Benson, 1999], onde um algoritmo alternativo para problemas convexos multiplicativos que combina técnicas do tipo *branch-and-bound* com técnicas baseadas em aproximações externas são propostas. Os dados envolvidos são: $n = m = p = 2$,

$$\begin{aligned} f_1(x) &= (x_1 - 2)^2 + 1, & f_2(x) &= (x_2 - 4)^2 + 1, \\ g_1(x) &= 25x_1^2 + 4x_2^2 - 100, & g_2(x) &= x_1 + 2x_2 - 4. \end{aligned}$$

Considerando $\underline{y} = (1, 1)$, $\bar{y} = (18, 38)$ (como em [Benson, 1999]), $\epsilon_1 = 0.001$ e $\epsilon_2 = 0.01$, obtém-se os resultados apresentados na Tabela 5.1.

Tabela 5.1: Convergência do Algoritmo A_2 .

k	y^k	w^k	$x(w^k)$	$\Theta(y^k)$
0	(1.0000,1.0000)	(0.4074,0.5926)	(0.0000,2.0000)	4.0000
1	(1.0000,7.7500)	(0.6585,0.3415)	(1.3547,1.3226)	0.4170
2	(1.0000,8.9711)	(0.8129,0.1871)	(1.7014,1.1493)	0.1016
3	(1.0000,9.5139)	(0.8907,0.1093)	(1.8509,1.0745)	0.0247
4	(1.0000,9.7394)	(0.9451,0.0549)	(1.9009,1.0495)	0.0074

O algoritmo A_2 convergiu após 5 iterações para a solução ϵ_2 -global $x^4 = (1.9009, 1.0495)$. O valor ótimo da função multiplicativa foi $f_1(x^4)f_2(x^4) = 9.8008$. Como esperado, x^4 é uma solução eficiente do problema convexo bi-objetivo associado, visto que ambas as componentes de w^4 são positivas. De fato, todas as soluções intermediárias geradas pelo algoritmo A_2 são eficientes. O algoritmo proposto em [Benson, 1999], com um critério de convergência equivalente a $\epsilon_2 = 0.025$, convergiu após 8 iterações.

Além desse exemplo ilustrativo, a abordagem proposta foi testada tendo como base duas classes de problemas de programação multiplicativa, ambas com restrições lineares:

$$\begin{aligned} P_{ML} : \quad & \text{minimizar} \quad \prod_{i=1}^m \langle c^i, x \rangle \\ & \text{s.a} \quad Ax \leq b, \quad x \geq 0, \end{aligned}$$

e

$$\begin{aligned} P_{MQ} : \quad & \text{minimizar} \quad (\langle c^0, x \rangle + d^0) \prod_{i=1}^{m-1} [\langle c^i, x \rangle + x^T \text{diag}(d_1^i, d_2^i, \dots, d_n^i)x] \\ & \text{s.a} \quad Ax \leq b, \quad x \geq 0, \end{aligned}$$

nos quais $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$, $c^i \in \mathbf{R}^n$ e $d^i \in \mathbf{R}^n$ representam matrizes constantes com elementos gerados aleatoriamente com distribuição uniforme no intervalo $[0, 100]$.

As seguintes quantidades foram definidas (valores totais): W , número de problemas P_W resolvidos; C , números de planos de corte necessários para convergência; V , número de vértices gerados no processo e T , tempo de CPU (em segundos). A quantidade W foi introduzida visando estabelecer comparações aproximadas com os resultados de [Kuno et al., 1993]. Dez problemas com combinações selecionadas de n (número de variáveis), p (número de restrições) e m (número de funções) foram resolvidos. A tolerância fixa adotada para resolver esses problemas foi $\epsilon = 0.001$. As tabelas a seguir apresentam a média e desvio padrão (em parênteses) dos valores de C , W , V e T .

Primeiramente, considere os problemas na forma P_{ML} , envolvendo produtos de funções lineares. As Tabelas 5.2, 5.4 e 5.7 reproduzem os resultados obtidos por [Kuno et al., 1993] para $m = 3$, $m = 4$ e $m = 2$ a 5 respectivamente, com um método de aproximação externa alternativo. Os resultados obtidos com o Algoritmo A_2 são apresentados nas Tabelas 5.3, 5.5, 5.6 e 5.8. Os valores médios de W nas Tabelas 5.4 e 5.5 são comparáveis, embora sejam menores e relativamente constantes na Tabela 5.5. Por outro lado, os valores médios de V na Tabela 5.5 são significativamente menores que os da Tabela 5.4. Uma razão para isso é que, enquanto em [Kuno et al., 1993], V é uma função crescente de W , no método proposto neste capítulo, V cresce com C , o número de cortes (máximos) gerados pelo Algoritmo A_2 , que é significativamente menor que W . Outra razão para isso é o uso do Algoritmo de Lista de Adjacências para enumeração de vértices [Oliveira & Ferreira, 2005b].

Quando o método proposto neste capítulo é adotado, o crescimento de V é muito mais lento o que leva [Kuno et al., 1993], a limitar o número máximo de funções a $m = 5$. Na abordagem proposta, problemas relativamente maiores podem ser resolvidos. A Tabela 5.8 apresenta os resultados obtidos com o Algoritmo A_2 para $m = 2, 3, \dots, 10$ ($n = 30$ e $p = 20$).

Tabela 5.2: Resultados de [Kuno et al., 1993] para (P_{ML}) , com $m = 3$.

n	100	100	120	120	140	140	160	180
p	80	100	100	120	120	150	150	200
W	41.8 (5.98)	38.3 (4.67)	46.9 (12.51)	47.5 (7.47)	45.5 (7.23)	46.1 (9.98)	46.1 (9.67)	42.5 (3.75)
V	176.5 (30.91)	159.0 (23.14)	200.0 (61.42)	204.9 (37.07)	197.1 (36.99)	200.4 (46.86)	200.0 (50.55)	180.3 (19.29)

Considere agora o problema da forma (P_{MQ}) , produtos de uma função linear e $m - 1$ funções quadráticas convexas [Oliveira & Ferreira, 2004a]. As Tabelas 5.9 e 5.11 apresentam os resultados obtidos por [Kuno et al., 1993], e as Tabelas 5.10 e 5.12 apresentam os resultados

Tabela 5.3: Abordagem proposta para (P_{ML}) , com $m = 3$.

n	100	100	120	120	140	140	160	180
p	80	100	100	120	120	150	150	200
W	45.40 (11.13)	40.20 (10.86)	42.30 (9.03)	45.80 (18.45)	48.09 (9.20)	42.39 (9.46)	47.10 (9.82)	41.7 (10.70)
C	9.60 (5.81)	9.33 (6.76)	7.40 (3.06)	10.00 (5.45)	9.20 (3.85)	7.50 (4.00)	9.00 (3.62)	6.60 (3.71)
V	33.20 (18.06)	33.70 (26.70)	28.80 (13.71)	43.00 (29.83)	34.70 (19.52)	27.20 (13.46)	35.60 (15.83)	25.40 (16.48)
T	27.83 (5.96)	26.46 (7.83)	42.87 (10.22)	47.80 (9.84)	75.38 (13.68)	68.94 (16.24)	89.63 (50.60)	122.94 (52.56)

Tabela 5.4: Resultados de [Kuno et al., 1993] para (P_{ML}) , com $m = 4$.

n	40	60	80	100	100	120	120
p	50	50	60	80	100	100	120
W	77.9 (21.6)	81.9 (11.41)	86.8 (15.09)	100.1 (17.84)	101.5 (24.62)	98.5 (13.68)	99.8 (18.65)
V	983.7 (365.13)	1060.6 (199.37)	1153.8 (258.10)	1386.0 (311.34)	1414.7 (422.3)	1370.6 (251.71)	1385.3 (327.6)

Tabela 5.5: Abordagem proposta para (P_{ML}) , com $m = 4$.

n	40	60	80	100	100	120	120
p	50	50	60	80	100	100	120
W	52.87 (12.51)	51.87 (10.60)	52.66 (6.81)	58.06 (14.32)	56.25 (5.20)	56.27 (9.03)	57.56 (8.65)
C	11.40 (4.37)	11.10 (4.55)	10.20 (2.29)	11.50 (5.06)	10.20 (2.20)	10.60 (3.40)	10.10 (2.84)
V	48.80 (26.85)	47.90 (28.60)	52.20 (17.16)	51.80 (16.95)	55.00 (15.24)	48.70 (16.40)	52.20 (26.98)
T	5.52 (1.36)	11.08 (2.82)	19.91 (3.04)	36.41 (9.15)	38.10 (4.22)	56.72 (9.91)	59.66 (8.26)

obtidos pelo algoritmo A_2 , respectivamente, para o caso $m = 2$ e $m = 3$. Comparando os dados dessas tabelas, pode-se observar que quando o Algoritmo A_2 é utilizado, o número de vértices gerados é muito menor, permitindo assim que problemas maiores venham a ser resolvidos. As Tabelas 5.13 e 5.14 apresentam os resultados do Algoritmo A_2 para $m = 4$ e $m = 5$ respectivamente.

Tabela 5.6: Abordagem proposta para (P_{ML}) , com $m = 5$.

n	40	60	80	100	100	120	120
p	50	50	60	80	100	100	120
W	61.04 (11.44)	62.44 (12.84)	70.94 (7.18)	76.43 (9.26)	78.91 (9.80)	79.04 (10.66)	84.91 (12.75)
C	14.10 (4.01)	14.20 (5.73)	14.10 (2.92)	15.60 (3.40)	15.00 (4.00)	13.40 (3.47)	18.40 (8.36)
V	81.80 (18.73)	85.40 (39.91)	108.20 (45.88)	100.80 (37.74)	100.70 (36.16)	76.90 (23.88)	102.80 (47.36)
T	6.24 (1.55)	11.59 (3.68)	23.90 (3.61)	40.96 (6.41)	44.15 (5.40)	65.24 (7.98)	72.17 (11.99)

Tabela 5.7: Resultados de [Kuno et al., 1993] para (P_{ML}) , com $(n, p) = (30, 20)$ e $m = 2$ a 5.

m	2	3	4	5
W	10.1 (3.78)	37.4 (4.2)	62.8 (12.66)	118.3 (19.3)
V	19.2 (7.56)	152.5 (20.42)	733.2 (207.12)	5406.2 (1585.29)

Tabela 5.8: Abordagem proposta para (P_{ML}) , com $(n, p) = (30, 20)$ e $m = 2$ a 10.

m	2	3	4	5	6	7	8	9	10
W	17.19 (3.04)	30.47 (5.66)	43.28 (6.25)	63.89 (19.74)	64.73 (9.96)	73.63 (9.08)	98.13 (23.48)	92.72 (23.77)	105.01 (11.13)
C	2.20 (0.63)	7.60 (2.79)	10.50 (2.32)	15.20 (4.13)	16.10 (1.96)	19.60 (1.89)	26.30 (8.51)	26.30 (6.43)	28.30 (0.67)
V	7.00 (1.69)	28.20 (10.20)	51.70 (26.63)	113.50 (48.40)	144.30 (40.28)	274 (52.29)	423.40 (88.50)	730.70 (154.49)	1686 (433.79)
T	0.91 (0.19)	1.65 (0.37)	2.67 (0.55)	4.79 (1.29)	7.79 (1.99)	22.66 (5.51)	89.50 (43.19)	521.21 (349.27)	2692 (999.48)

Tabela 5.9: Resultados de [Kuno et al., 1993] para (P_{MQ}) , com $m = 2$.

n	40	60	80	100	100	120	120
p	50	50	60	80	100	100	120
W	34.6 (8.62)	45.5 (19.41)	43.1 (12.51)	43.7 (10.63)	43.0 (14.72)	52.7 (10.74)	51.4 (17.60)
V	140.7 (40.71)	192.9 (99.68)	181.9 (63.14)	185.3 (50.54)	181.3 (70.93)	226.9 (51.36)	222.6 (87.26)

Tabela 5.10: Abordagem Proposta para (P_{MQ}) , com $m = 2$.

n	40	60	80	100	100	120	120
p	50	50	60	80	100	100	120
W	47.48 (6.52)	45.48 (6.39)	42.09 (4.33)	46.18 (4.54)	48.58 (7.87)	43.98 (10.36)	45.98 (5.86)
C	6.40 (1.34)	6.50 (1.08)	5.40 (0.51)	6.90 (1.28)	6.20 (1.03)	6.00 (1.15)	5.90 (0.87)
V	25.10 (7.04)	26.00 (5.47)	20.30 (2.05)	26.80 (5.49)	24.00 (4.24)	23.00 (4.32)	23.20 (3.35)
T	5.00 (0.69)	10.77 (1.36)	18.24 (2.37)	34.78 (3.25)	38.50 (5.92)	55.10 (14.64)	58.27 (8.78)

Tabela 5.11: Resultados de [Kuno et al., 1993] para (P_{MQ}) , com $m = 3$.

n	20	40	40	60	80
p	30	30	50	50	60
W	69.3 (12.84)	93.6 (27.52)	91.1 (13.63)	109.1 (26.09)	101.3 (23.39)
V	829.1 (199.55)	1234.2 (473.77)	1194.6 (236.4)	1505.6 (439.97)	1352.6 (378.39)

Tabela 5.12: Abordagem Proposta para (P_{MQ}) , com $m = 3$.

n	20	40	40	60	80
p	30	30	50	50	60
W	76.34 (13.05)	61.11 (6.28)	63.89 (22.27)	69.95 (6.96)	76.34 (13.05)
C	12.90 (2.80)	9.80 (1.31)	9.50 (1.50)	10.10 (1.52)	11.50 (3.97)
V	97.4 (30.63)	64.2 (12.76)	59.8 (16.26)	70.4 (21.03)	86.40 (55.85)
T	3.16 (0.74)	5.9 (0.69)	7.43 (1.22)	16.10 (1.56)	35.73 (6.90)

Como consideração final, é importante mencionar que devido ao fato da função multiplicativa ser um produto de funções quadráticas convexas ou lineares, o problema P_W (Algoritmo A_1) é um problema de programação convexa quadrática ou linear, respectivamente, para os quais algoritmos eficientes de solução estão disponíveis. Essa característica foi levada em consideração ao se implementar o Algoritmo A_1 [Oliveira & Ferreira, 2005a].

Tabela 5.13: Abordagem Proposta para (P_{MQ}) , com $m = 4$.

n	20	40	40	60	80
p	30	30	50	50	60
W	89.04 (15.45)	95.44 (7.71)	93.73 (13.17)	85.06 (8.53)	93.66 (9.15)
C	15.4 (1.17)	13.8 (1.47)	14.30 (2.71)	11.70 (0.82)	15.20 (3.42)
V	208.80 (25.65)	176.70 (48.04)	179.30 (75.83)	118.50 (18.13)	198.80 (82.89)
T	6.60 (0.78)	11.60 (2.18)	12.76 (2.77)	21.00 (3.54)	47.22 (7.05)

Tabela 5.14: Abordagem Proposta para (P_{MQ}) , com $m = 5$.

n	20	40	40	60	80
p	30	30	50	50	60
W	111.75 (21.77)	114.91 (10.85)	121.52 (8.15)	124.74 (9.29)	126.40 (14.96)
C	19 (2.98)	16.4 (2.45)	20 (3.16)	17.4 (2.63)	17.30 (3.46)
V	423 (105.05)	379.70 (163.04)	554.40 (198.16)	408.40 (135.82)	419.60 (234.14)
T	21 (7.58)	25.40 (10.38)	42.60 (16.43)	43.60 (9.73)	73.70 (22.21)

5.2 Programação Convexa Multiplicativa Generalizada

Os problemas multiplicativos generalizados [Konno et al., 1994] são formulados como

$$P_{MG} : \text{minimizar } F(f(x)) = f_1(x) + \sum_{i=1}^r f_{2i}(x)f_{2i+1}(x) \quad (r \geq 1)$$

$$\text{s.a} \quad g_j(x) \leq 0, \quad j = 1, 2, \dots, p,$$

no qual $f_i : \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, 2, \dots, m = 2r + 1$, e $g_j : \mathbf{R}^n \rightarrow \mathbf{R}$, $j = 1, 2, \dots, p$, são funções convexas contínuas. Como no caso multiplicativo clássico, assume-se que

$$\Omega := \{x \in \mathbf{R}^n : g_j(x) \leq 0, \quad j = 1, 2, \dots, p\} \quad (5.15)$$

é um conjunto compacto não-vazio e que f_1, f_2, \dots, f_m são positivas em Ω . Ao problema (P_{MG}) também pode-se associar o problema de minimizar a função objetivo vetorial $f := (f_1, f_2, \dots, f_m)$ em Ω , com $F : \mathbf{R}^n \rightarrow \mathbf{R}$ fazendo o papel de uma função de agregação, contínua e crescente em relação às funções individuais f_1, f_2, \dots, f_m [Yu, 1985].

Observa-se que F é geralmente não-convexa em Ω , e que diferentemente do caso multiplicativo, não se pode nem mesmo afirmar que F é quase-côncava em $\{f(x) : x \in \Omega\}$, uma vez que F envolve agora uma soma de funções quase-côncavas em $\{f(x) : x \in \Omega\}$, e a soma de funções quase-côncavas não é geralmente quase-côncava [Bazaraa et al., 1993]. Embora não seja possível explorar propriedades como quase-concavidade, que permite resolver problemas multiplicativos através de enumeração de vértices, quando o problema multiplicativo generalizado (P_{MG}) é reformulado no espaço dos objetivos, algoritmos eficientes ainda podem ser desenvolvidos.

5.2.1 Formulação no Espaço dos Objetivos

A projeção do problema (P_{MG}) no espaço dos objetivos assume a forma

$$P_{\mathcal{Y}} : \text{minimizar } F(y) = y_1 + \sum_{i=1}^r y_{2i}y_{2i+1}$$

$$\text{s.a} \quad y \in \mathcal{Y},$$

sendo

$$\mathcal{Y} := \{y \in \mathbf{R}^m : y = f(x), \quad x \in \Omega\} \quad (5.16)$$

o espaço dos objetivos. A reformulação do problema ($P_{\mathcal{Y}}$) como o problema de minimizar F sobre um conjunto convexo \mathcal{F} e a caracterização de \mathcal{F} como um conjunto infinitamente restrito seguem as mesmas linhas adotadas para o caso multiplicativo clássico. Em particular, como F é crescente em relação aos objetivos f_1, f_2, \dots, f_m , as soluções ótimas globais de P_{MG} serão

eficientes, o que implica que se x^* é uma solução ótima de P_{MG} , então existe $w \in \mathcal{W}$ tal que x^* também é uma solução ótima de

$$P_W : \begin{array}{ll} \text{minimizar} & \langle w, f(x) \rangle \\ \text{s.a} & g_i(x) \leq 0, \quad i = 1, 2, \dots, p. \end{array}$$

O algoritmo de otimização global (Algoritmos A_1 e A_2) e todas as propriedades derivadas anteriormente, como convergência para uma solução ótima global, permanecem válidas. Com isso, o problema se resume na obtenção da solução ótima global do problema mestre

$$P_{\mathcal{F}}^k : \begin{array}{ll} \text{minimizar} & F(y) = y_1 + \sum_{i=1}^r y_{2i} y_{2i+1} \\ \text{s.a} & y \in \mathcal{F}^k. \end{array}$$

numa iteração genérica k , isto é, após a geração de k hiperplanos de corte, responsáveis pela aproximação corrente de \mathcal{F} . O conjunto \mathcal{F}^k é descrito por k restrições de desigualdade da forma

$$\sum_{j=1}^m w_{ij} y_j \geq v_i, \quad i = 1, 2, \dots, k,$$

nas quais w_{ij} e v_i são quantidades calculadas pelo Algoritmo A_1 , e pelas restrições de limitação $\underline{y}_i \leq y_i \leq \bar{y}_i$, $i = 1, 2, \dots, m$. Na forma matricial,

$$\begin{array}{l} Wy \geq v, \quad W \in \mathbf{R}^{k \times m}, \quad v \in \mathbf{R}^k, \\ \underline{y} \leq y \leq \bar{y}, \end{array}$$

sendo $W := \{w_{ij}\}$ e v o vetor-coluna formado por v_1, v_2, \dots, v_k .

Proposição 7. *O problema $(P_{\mathcal{F}^k})$ é equivalente ao problema de programação quadrática*

$$P_{Q^k} : \begin{array}{ll} \text{minimizar} & F(y) = \frac{1}{2} y^T Q y + c^T y \\ \text{s.a} & Wy \geq v, \\ & \underline{y} \leq y \leq \bar{y}, \end{array}$$

no qual

$$Q := \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad Q \in \mathbf{R}^{m \times m},$$

e

$$c^T := [1 \ 0 \ 0 \ \cdots \ 0 \ 0], \quad c \in \mathbf{R}^m$$

Proposição 8. *A matriz Q é indefinida.*

Prova: Dada a estrutura bloco-diagonal de Q , seu polinômio característico pode ser escrito como

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

Logo, os autovalores de Q são 0, 1 e -1 , e portanto Q é indefinida. \square

Uma conseqüência importante da Proposição 8 é que (P_y) representa um problema de programação quadrática indefinida, uma vez que a matriz Hessiana de F , igual a Q , é indefinida. Embora seja necessário abordar o problema $(P_{\mathcal{F}^k})$ por meio de algoritmos mais gerais, a reformulação (P_{Q^k}) exhibe como possíveis atrativos sua forma altamente estruturada, pequeno número de variáveis, m , o número de objetivos, geralmente muito menor do que n , o número de variáveis de (P_{MG}) , e, em função de experiências numéricas realizadas, um número máximo de restrições, isto é, um número máximo para k , relativamente pequeno.

Por conveniência, define-se

$$A := \begin{bmatrix} -W \\ \cdots \\ I \\ \cdots \\ -I \end{bmatrix}, \quad A \in \mathbf{R}^{(k+2m) \times m}, \quad \text{e} \quad b := \begin{bmatrix} -v \\ \cdots \\ \bar{y} \\ \cdots \\ -\underline{y} \end{bmatrix}, \quad b \in \mathbf{R}^{k+2m},$$

de forma a representar as restrições de (P_{Q^k}) como $Ay \leq b$. Sabe-se que qualquer solução ótima global de (P_{Q^k}) deve satisfazer as condições de Karush-Kuhn-Tucker do problema:

$$\begin{aligned} Ay &\leq b, \\ Qy + A^T \lambda &= -c, \\ \lambda^T (Ay - b) &= 0, \\ \lambda &\geq 0, \end{aligned}$$

nas quais $\lambda \in \mathbf{R}^{k+2m}$ é o vetor de multiplicadores de Lagrange associados às restrições de desigualdade.

Uma propriedade fundamental para a validade do algoritmo a ser proposto, demonstrável a partir de resultados gerais para a classe de problemas quadráticos indefinidos [Horst et al., 1995], é enunciada a seguir.

Teorema 26. *Uma solução ótima global de (P_{Q^k}) encontra-se na fronteira do polítopo*

$$\mathcal{P} := \{y \in \mathbf{R}^m : Ay \leq b\}.$$

Do Teorema 26 conclui-se que para uma solução ótima global de (P_{Q^k}) , existe um subconjunto de índices $\mathcal{I} \subseteq \{1, 2, \dots, k + 2m\}$, tal que

$$\begin{aligned} A_{\mathcal{I}}y &= b_{\mathcal{I}}, \\ Qy + A^T\lambda &= -c, \\ \lambda_{\bar{\mathcal{I}}} &= 0, \end{aligned}$$

sendo $A_{\mathcal{I}}y = b_{\mathcal{I}}$ o subconjunto de restrições ativas determinado por \mathcal{I} , e $\bar{\mathcal{I}}$ o conjunto complementar de \mathcal{I} . Às restrições inativas são associados multiplicadores de Lagrange nulos, de forma a satisfazer as restrições de complementariedade $\lambda^T(Ay - b) = 0$.

5.2.2 Solução por Enumeração

Um método simples para resolver problemas quadráticos indefinidos, como (P_{Q^k}) , consiste em enumerar todos os possíveis conjuntos \mathcal{I} , resolver os sistemas lineares associados, checar se $y \in \mathcal{P}$ e $\lambda \geq 0$, e então, dentre as soluções factíveis encontradas, determinar aquela que fornece o menor valor para F [Horst et al., 1995].

Na presente aplicação do método enumerativo, $O(2^{k+2m}(k+3m)^3)$ operações aritméticas são necessárias: existem 2^{k+2m} maneiras de se selecionar subconjuntos de $\{1, 2, \dots, k + 2m\}$ e cada sistema de equações é resolvido em no máximo $O((k+3m)^3)$ operações aritméticas. Entretanto, como as restrições de limitação $\underline{y}_i \leq y_i$ e $y_i \leq \bar{y}_i$ não podem estar simultaneamente ativas se os limites forem diferentes, o número de subconjuntos a serem explorados é significativamente menor.

Para $k = 0$, a solução ótima global de (P_{Q^0}) é $y^0 = \underline{y}$, devido ao fato de F ser crescente em relação a y . Suponha que numa iteração qualquer k o problema (P_{Q^k}) é resolvido por enumeração, obtendo-se a solução ótima global y^k . O Algoritmo A_1 verifica y^k , retornando com uma nova restrição de desigualdade, caso y^k não seja factível ($y \notin \mathcal{F}$). Neste caso 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 y^k anterior, que se revelou infactível; é suficiente explorar as combinações de $\{1, 2, \dots, k + 2m\}$ como a nova restrição. Se o algoritmo global convergir após l iterações, este procedimento terá evitado 2^l combinações desnecessárias.

5.2.3 Exemplos Numéricos

O algoritmo desenvolvido para problemas de programação multiplicativa generalizada foi avaliado com base em problemas clássicos da literatura. O problema a seguir foi utilizado para ilustrar o algoritmo alternativo proposto em [Konno et al., 1994].

$$\begin{aligned} &\text{minimizar } F(x) = (3x_1 - 4x_2 + 15) + (x_1 + 2x_2 - 1.5)(2x_1 - x_2 + 4) + \\ &\quad (x_1 - 2x_2 + 8.5)(2x_1 + x_2 - 1) \\ \text{s.a} \quad &5x_1 - 8x_2 \geq -24, \\ &5x_1 + 8x_2 \leq 44, \\ &6x_1 - 3x_2 \leq 15, \\ &4x_1 + 5x_2 \geq 10, \\ &x_1 \geq 0. \end{aligned}$$

A Figura 5.1 mostra a região factível Ω formada pela intersecção das restrições do problema. O algoritmo baseado na projeção do problema no espaço dos objetivos forneceu os resultados apresentados na Tabela 5.15.

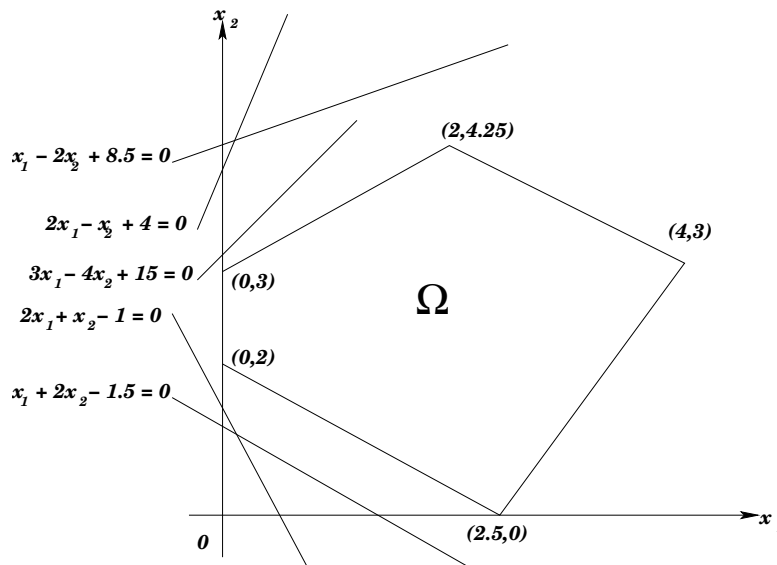


Figura 5.1: Interpretação da região factível.

Tabela 5.15: Convergência do Algoritmo A_2 .

k	y^k	w^k	$x(w^k)$	$\Theta(y^k)$
0	(3,1,1,2,2)	(0.3333,0.6667,0,0,0)	(0,2)	2.3333
1	(3,4.5,1,9,2)	(1,0,0,0,0)	(0,3)	0

O algoritmo A_2 convergiu depois de apenas 2 iterações para a solução ϵ_2 -global $x^* = (0, 3)$ com $\epsilon_2 = 0.01$. O valor ótimo da função multiplicativa foi $f_1(x^*) + f_2(x^*)f_3(x^*) + f_4(x^*)f_4(x^*) =$

12.50. Foram gerados 2 planos de cortes, o problema (P_W) foi resolvido 6 vezes e o tempo computacional requerido foi de 0.20 segundos. O algoritmo proposto em [Konno et al., 1994], com um critério de convergência equivalente, convergiu para a mesma solução, após um número não indicado de iterações.

Em seguida, o algoritmo proposto foi avaliado de forma mais geral por meio de problemas-teste aleatórios envolvendo a seguinte classe de problemas multiplicativos:

$$P_{MG} : \begin{array}{l} \text{minimizar} \quad (\langle c^1, x \rangle + \frac{1}{2}x^T C^T C x) + \sum_{i=1}^r \langle c^{2i}, x \rangle \langle c^{2i+1}, x \rangle \\ \text{s.a} \quad \quad \quad Ax \leq b, \quad x \geq 0, \end{array}$$

com $A \in \mathbf{R}^{p \times n}$, $C \in \mathbf{R}^{n \times n}$, $b \in \mathbf{R}^p$, $c^i \in \mathbf{R}^n$, $i = 1, 2, \dots, m = 2r + 1$ representando matrizes e vetores constantes com elementos gerados aleatoriamente com distribuição uniforme no intervalo $[0, 100]$. As condições impostas para os testes e as convenções adotadas para utilizar as quantidades presentes nas tabelas a seguir são as mesmas do caso multiplicativo clássico, ou seja, W , número de problemas P_W resolvidos; C , números de planos de corte necessários para convergência; V , número de vértices gerados no processo 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 funções quadráticas) foram resolvidos. A tolerância fixa adotada foi $\epsilon = 0.001$. As tabelas a seguir apresentam a média e desvio padrão (em parênteses) dos valores de C , W , V e T .

Considere o problema na forma (P_{MG}), envolvendo a soma de uma função quadrática com a soma de produtos de funções lineares. As Tabelas 5.16 e 5.18 reproduzem os resultados obtidos por [Konno et al., 1994] para $r = 1$, $r = 2$ e $r = 3$, respectivamente, com um método de aproximação externa alternativo baseado em enumeração de vértices. Os resultados obtidos com o uso do Algoritmo A_2 em conjunto com o algoritmo enumerativo são apresentados nas Tabelas 5.17 e 5.19. Observa-se que os valores médios de W , apresentados nas Tabelas 5.16 e 5.18 são ligeiramente menores que os valores médios de W apresentados nas Tabelas 5.17 e 5.19. Os números médios de vértices gerados pelo método de [Konno et al., 1994] são progressivamente maiores, o que dificulta a aplicação do algoritmo para $r > 3$. Por outro lado, a abordagem proposta gera um número pequeno de cortes no espaço dos objetivos. Com isso, o algoritmo enumerativo proposto, mesmo sendo de natureza exponencial consegue excelente desempenho nos casos testados.

Tabela 5.16: Resultados de [Konno et al., 1994].

r	1	1	1	1	1	1	1
n	20	20	50	50	100	100	150
p	10	30	30	70	70	130	130
W	8.9 (3.5)	10.6 (3.0)	10.5 (2.7)	10.2 (4.3)	8.1 (3.4)	11.7 (4.2)	11.5 (3.5)
V	15.8 (7.0)	19.2 (6.0)	19.0 (5.3)	18.4 (8.7)	14.2 (6.8)	21.4 (8.5)	21.0 (7.1)
T	0.5 (0.1)	1.9 (1.0)	8.7 (4.0)	27.9 (13.6)	83.5 (28.7)	288.2 (99.9)	482.2 (131.7)

Tabela 5.17: Abordagem Proposta.

r	1	1	1	1	1	1	1
n	20	20	50	50	100	100	150
p	10	30	30	70	70	130	130
C	2.5 (0.7)	2.3 (0.48)	3 (0.47)	2.7 (0.48)	2.8 (0.48)	2.9 (0.73)	2.6 (0.52)
W	16.6 (4.25)	15.2 (3.64)	21.3 (3.47)	21.5 (5.46)	21.3 (3.92)	22 (5.31)	19.6 (4.38)
T	0.55 (0.15)	0.53 (0.13)	3.14 (0.61)	3.61 (0.98)	16.80 (4.25)	23 (16.88)	45.34 (10.9)

Tabela 5.18: Resultados de [Konno et al., 1994].

r	2	2	2	3	3	3
n	20	50	50	20	50	50
p	30	30	70	30	30	70
W	19.6 (3.0)	22.2 (3.8)	23.9 (3.0)	28.1 (2.9)	33.0 (4.1)	35.1 (5.2)
V	212.8 (65.2)	270.8 (96.9)	308.8 (80.0)	2048.0 (625.0)	3193.0 (1089.4)	3896.6 (1816.5)
T	7.0 (4.1)	56.2 (43.0)	183.8 (57.4)	421.4 (1096.5)	536.0 (193.1)	2101.4 (881.4)

Tabela 5.19: Abordagem Proposta.

r	2	2	2	3	3	3
n	20	50	50	20	50	50
p	30	30	70	30	30	70
C	4.3 (1.16)	4.1 (0.32)	4.2 (0.42)	3.66 (0.5)	4.55 (1.33)	6.11 (0.33)
W	34.9 (5.04)	41.22 (15.82)	48.11 (2.93)	31 (5.05)	41.22 (15.82)	48.11 (2.93)
T	2.35 (1.66)	5.24 (0.59)	6.65 (1.25)	12.85 (4.04)	35.63 (22.28)	70.34 (17.80)

5.3 Considerações Finais

Um algoritmo para problemas multiplicativos convexos baseado em elementos de análise convexa foi proposto neste capítulo. O uso de análise convexa permitiu decompor o problema multiplicativo em subproblemas mestre, quase-côncavo no espaço dos objetivos, solucionado por enumeração de vértices, e min-max, responsável pelo teste de factibilidade das soluções encontradas pelo subproblema mestre. O uso de cortes máximos no espaço dos objetivos, aliado a uma implementação eficiente de um procedimento de enumeração de vértices resulta num desempenho computacional significativamente melhor do que os apresentados por um algoritmo alternativo bastante referenciado na literatura de programação convexa multiplicativa. A extensão da abordagem proposta para problemas convexos multiplicativos generalizados implicou no uso de uma técnica mais geral para resolver o subproblema mestre, representado por um problema quadrático indefinido de pequena dimensão e estrutura particular. O algoritmo global apresentou bom desempenho nos testes realizados. A abordagem proposta também evidencia conexões entre programação multiplicativa e programação multiobjetivo convexas.

Conclusões

O desenvolvimento teórico e a implementação computacional de métodos de otimização global especializados para duas classes de problemas de programação não-linear – problemas geométricos e problemas multiplicativos – foram as principais metas perseguidas nesta tese.

No Capítulo 4, após uma extensão relativamente simples da abordagem dual para problemas geométricos posinomiais a problemas multiobjetivos representados pela soma ponderada de funções objetivos, investigou-se a possibilidade de melhoria de um algoritmo do tipo *branch-and-bound* voltado para problemas geométricos signomiais. A tentativa de melhoria consistiu na incorporação ao algoritmo *branch-and-bound* de um procedimento de busca local baseado no conceito de condensação, bastante utilizado na literatura no desenvolvimento de métodos subótimos para problemas signomiais.

Embora o algoritmo tenha apresentado bom desempenho no sentido de obter soluções ótimas globais nos problemas testados, a adoção da busca local como estratégia de refinamento do limitante superior do algoritmo não se mostrou particularmente vantajosa, no sentido de acelerar sua convergência.

No Capítulo 5, elementos de análise convexa e de programação multiobjetivo foram utilizados no desenvolvimento de algoritmos para os chamados problemas convexos multiplicativos. A estratégia comum aos algoritmos propostos foi a projeção e resolução do problema no espaço dos objetivos por meio da técnica de relaxação. Ao ser projetado, cada problema considerado passou a exibir propriedades especiais que tornaram sua solução mais simples. No caso de problemas multiplicativos clássicos, por enumeração dos vértices gerados por planos de corte;

no caso de problemas multiplicativos generalizados, por enumeração dos próprios planos de corte introduzidos pelo algoritmo.

Testes realizados vêm demonstrando maior eficiência dos algoritmos propostos em relação aos algoritmos alternativos existentes na literatura de programação multiplicativa.

Trabalhos futuros na linha de programação geométrica signomial deverão necessariamente envolver uma implementação mais eficiente do algoritmo *branch-and-bound*, provavelmente numa linguagem de programação como C++, que faça uso de pacotes especializados em programação não-linear. (O protótipo atual encontra-se implementado em MATLAB/*Optimization Toolbox*.)

Dois aspectos poderão contribuir para que a busca local via condensação possa se mostrar eficaz. Em [Maranas & Floudas, 1997] é proposto um escalamento especial de variáveis e outros refinamentos voltados para problemas signomiais, os quais chegaram a ser testados, mas não incorporados ao protótipo MATLAB. Um tratamento mais completo para o problema de inicializar várias rotinas de otimização restrita com soluções iniciais factíveis seria outro aspecto a considerar. Aplicações da versão aprimorada do algoritmo a problemas importantes nas áreas de Pesquisa Operacional e Teoria de Controle poderiam então ser desenvolvidas.

Trabalhos futuros na linha de programação multiplicativa deverão contemplar funções multiplicativas descritas como produtos de razões de funções lineares, ou outras classes de funções multiplicativas que favoreçam a aplicação da abordagem proposta nesta tese. Problemas que apresentem restrições convexas multiplicativas também poderão ser abordados.

Durante o desenvolvimento dos algoritmos propostos para programação multiplicativa, alguns algoritmos alternativos baseados na idéia de se condensar funções multiplicativas no espaço dos objetivos convergiram, infelizmente sem prova teórica, para soluções globais em todos os problemas testados. Estudos mais aprofundados sobre condensação de posinômios poderão eventualmente levar a propriedades interessantes desses algoritmos alternativos.

Referências Bibliográficas

- [Andersen & Ye, 1998] ANDERSEN, E., YE, Y. (1998). A Computational Study of the Homogeneous Algorithm for Large-Scale Convex Optimization. *Computational Optimizations and Applications*, 10(3):243–269.
- [Bazaraa et al., 1993] BAZARAA, M. S., SHERALI, H. D., SHETTY, C. M. (1993). *Nonlinear Programming - Theory and Algorithms*. 2nd Edition, John Wiley & Sons, Inc.
- [Beightler & Phillips, 1976] BEIGHTLER, C., PHILLIPS, D. T. (1976). *Applied Geometric Programming*. John Wiley & Sons.
- [Bennett, 1994] BENNETT, K. P. (1994). Global Tree Optimization: A Non-Greedy Decision Tree Algorithm. *Computing Sciences and Statistics*, 26:156–160.
- [Benson, 1999] BENSON, H. P. (1999). An Outcome Space Branch and Bound-Outer Approximation Algorithm for Convex Multiplicative Programming. *Journal of Global Optimization*, 15:315–342.
- [Benson & Boger, 1997] BENSON, H. P., BOGER, G. M. (1997). Multiplicative Programming Problems: Analysis and Efficient Point Search Heuristic. *Journal of Optimization Theory and Applications*, 94:487–510.
- [Benson & Boger, 2000] BENSON, H. P., BOGER, G. M. (2000). Outcome Space Cutting-Plane Algorithm for Linear Multiplicative Programming. *Journal of Optimization Theory and Applications*, 104(2):301–322.
- [Blau & Wilde, 1969] BLAU, G. E., WILDE, D. J. (1969). Generalized Polynomial Programming. *The Canadian Journal of Chemical Engineering*, 47.
- [Blau & Wilde, 1971] BLAU, G. E., WILDE, D. J. (1971). A Lagrangian Algorithm for Equality Constrained Generalized Polynomial Optimization. *The Canadian Journal of Chemical Engineering*, 17.

- [Boyd et al., 2005a] BOYD, S., KIM, S. J., PATIL, D., HOROWITZ, M. (2005a). Digital Circuit Optimization Via Geometric Programming. *Control and Dynamic Systems*. URL: http://www.stanford.edu/~boyd/gp_tutorial.html.
- [Boyd & Vandenberghe, 2004] BOYD, S., VANDENBERGHE, L. (2004). Convex Optimization. *Cambridge University Press*, Cambridge.
- [Boyd et al., 2005b] BOYD, S. P., KIM, S. J., VANDENBERGHE, L., HASSIBI, A. (2005b). A Tutorial on Geometric Programming. Relatório técnico, Stanford University. URL: http://www.stanford.edu/~boyd/gp_tutorial.html.
- [Bricker et al., 1997] BRICKER, D., KORTANEK, K., XUI, L. (1997). Maximum Likelihood Estimates with Order Restrictions on Probabilities and Odds Ratios: A Geometric Programming Approach. *Journal of Applied Mathematics and Decision Sciences*, 1:53–65.
- [Chandra et al., 2004] CHANDRA, D., SINGH, V., KAR, H. (2004). Improved Routh-Padé Approximants: A Computer-Aided Approach. *IEEE Transactions on Automatic Control*, 49(2):292–296.
- [Chankong & Haimes, 1983] CHANKONG, V., HAIMES, Y. Y. (1983). Multiobjective Decision Making-Theory and Methodology. *Elsevier Science Publishing* - New York.
- [Chen et al., 1991] CHEN, P. C., HANSEN, P., JAUMARD, B. (1991). On-line and Vertex Enumeration by Adjacency Lists. *Operations Research Letters*, 10:403–409.
- [Chiang, 2004] CHIANG, M. (2004). Geometric Programming Duals of Channel Capacity and Rate Distortion. *IEEE Transactions on Information Theory*, 50(2):245–258.
- [Colleran et al., 2003] COLLERAN, D., PORTMANN, C., HASSIBI, A., CRUSIUS, C., MOHAN, S., BOYD, S., LEE, T., HERSHENSON, M. (2003). Optimization of Phase-locked Loop Circuits Via Geometric Programming. In *Proceedings of the Custom Integrated Circuits Conference (CICC)*, San Jose-CA, 326–328.
- [Dembo, 1978] DEMBO, R. S. (1978). Current State of the Art of Algorithms and Computer Software for Geometric Programming. *Journal of Optimization Theory and Applications*, 26:149–184.
- [Dorneich & Sahinidis, 1995] DORNEICH, M. C., SAHINIDIS, N. V. (1995). Global Optimization Algorithms for Chip Design and Compaction. *Engineering Optimization*, 25(2):131–154.
- [Duffin & Peterson, 1966] DUFFIN, R. J., PETERSON, E. L. (1966). Duality Theory for Geometric Programming. *SIAM Journal of Applied Mathematics*, 14.

- [Duffin et al., 1967] DUFFIN, R. J., PETERSON, E. L., ZENER, C. (1967). Geometric Programming - Theory and Application. *Wiley* - New York.
- [Ecker, 1980] ECKER, J. G. (1980). Geometric Programming: Methods Computations and Applications. *SIAM Review*, 22(3):338–362.
- [Ferreira, 1998] FERREIRA, P. A. V. (1998). Modelagem, Análise e Projeto de Sistemas com Múltiplos Objetivos. Tese de Livre-Docência, FEEC/UNICAMP.
- [Ferreira & Machado, 1996] FERREIRA, P. A. V., MACHADO, M. E. S. (1996). Solving Multiple Objective Problems in the Objective Space. *Journal of Optimization Theory and Applications*, 89:659–680.
- [Floudas & Visweswaran, 1993] FLOUDAS, C. A., VISWESWARAN, V. (1993). Primal-Relaxed Dual Global Optimization Approach. *Journal of Optimization Theory and Applications*, 78:187–225.
- [Fuh & Hu, 2000] FUH, C. D., HU, I. (2000). Asymptotically Efficient Strategies for a Stochastic Scheduling Problem with Order Constraints. *The Annals of Statistics*, 28(6):1670–1695.
- [Geoffrion, 1967] GEOFFRION, A. M. (1967). Solving Bicriterion Mathematical Programs. *Operations Research*, 15:39–54.
- [Geoffrion, 1970] GEOFFRION, A. M. (1970). Elements of Large-Scale Mathematical Programming. *Management Science*, 16:652–691.
- [Henderson & Quandt, 1971] HENDERSON, J. M., QUANDT, R. E. (1971). *Microeconomic Theory*. 2nd Edition, JMcGraw-Hill.
- [Horst & Pardalos, 1995] HORST, R., PARDALOS, P. M. (1995). *Handbook of Global Optimization*. Kluwer Academic Publishers, Netherlands.
- [Horst et al., 1995] HORST, R., PARDALOS, P. M., THOAI, N. V. (1995). *Introduction to Global Optimization*. Kluwer Academic Publishers, Netherlands.
- [Jefferson & Scott, 1978] JEFFERSON, T. R., SCOTT, C. H. (1978). Generalized Geometric Programming Applied to Problems of Optimal Control: I Theory. *Journal of Optimization Theory and Applications*, 26:117.
- [Jha, 1990] JHA, N. (1990). A Discrete Data-Base Multiple Objective Optimization of Milling Operation Through Geometric Programming. *Journal of Engineering for Industry*, 112:368–374.

- [Kachian, 1979] KACHIAN, L. G. (1979). A Polynomial Algorithm in Linear Programming. *Soviet Mathematica Doklady*, 20:191–194.
- [Kathoh & Ibaraki, 1987] KATHOH, N., IBARAKI, T. (1987). A Parametric Characterization and an ϵ -Approximation Scheme for the Minimization of a Quasiconcave Program. *Discrete Applied Mathematics*, 17:39–66.
- [Keeney & Raiffa, 1993] KEENEY, R. L., RAIFFA, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University, Cambridge.
- [Klafszky et al., 1992] KLAFSZKY, E., MAYER, J., TERLAKY, T. (1992). A Geometric Programming Approach to the Channel Capacity Problem. *Engineering Mathematics*, 19:115–130.
- [Konno & Kuno, 1990] KONNO, H., KUNO, T. (1990). Generalized Linear Multiplicative and Fractional Programming. *Annals of Operations Research*, 25:147–162.
- [Konno & Kuno, 1995] KONNO, H., KUNO, T. (1995). *Multiplicative Programming Problems*, Handbook of Global Optimization Problems, 369–405. Kluwer Academic Publishers, Netherlands.
- [Konno et al., 1994] KONNO, H., KUNO, T., YAJIMA, Y. (1994). Global Minimization of a Generalized Convex Multiplicative Function. *Journal of Global Optimization*, 4:47–62.
- [Kortanek et al., 1996] KORTANEK, K., XU, X., YE, Y. (1996). An Infeasible Interior-Point Algorithm for Solving Primal and Dual Geometric Programs. *Mathematical Programming*, 76:155–181.
- [Kuno et al., 1993] KUNO, T., YAJIMA, Y., KONNO, H. (1993). An Outer Approximation Method for Minimizing the Product of Several Convex Functions on a Convex Set. *Journal of Global Optimization*, 3:325–335.
- [Liu et al., 1999] LIU, X. J., UMEGAKI, T., YMAMOTO, Y. (1999). Heuristic Methods for Linear Multiplicative Programming. *Journal of Global Optimization*, 4(15):433–447.
- [Luenberger, 1984] LUENBERGER, D. G. (1984). *Linear and Nonlinear Programming*. 2nd Edition, Addison Wesley, California.
- [Maranas & Floudas, 1997] MARANAS, C. A., FLOUDAS, C. A. (1997). Global Optimization in Generalized Geometric Programming. *Computers & Chemical Engineering*, 21:351–367.

- [Maranas et al., 1997] MARANAS, C. D., ANDROULAKIS, I. P., FLOUDAS, C. A., BERGER, A. J., MULVEY, J. M. (1997). Solving Long-Term Financial Planning Problems Via Global Optimization. *Journal of Economic Dynamics Control*, 21:1405–1425.
- [Mulvey et al., 1995] MULVEY, J. M., J. VANDERBEI, R., ZENIOS, S. A. (1995). Robust Optimization of Large-Scale Systems. *Operations Research*, 43:264–281.
- [Nascimento & Andrade, 2003] NASCIMENTO, R. Q., ANDRADE, M. G. (2003). Otimização de um Sistema Hidrotérmico Via Programação Geométrica. *In Anais do XXXV Simpósio Brasileiro de Pesquisa Operacional*, Natal-Brasil.
- [Nesterov & Nemirovsky, 1994] NESTEROV, Y., NEMIROVSKY, A. (1994). Interior-Point Polynomial Methods in Convex Programming. *Studies in Applied Mathematics. SIAM*, 13.
- [Oliveira & Ferreira, 2003a] OLIVEIRA, R. M., FERREIRA, P. A. V. (2003a). Global Optimization of Convex Multiplicative Programs by Duality Theory. *Notes of the 2nd International Workshop on Global Constrained Optimization and Constraint Satisfaction*, Lausanne - Switzerland, 21–28.
- [Oliveira & Ferreira, 2003b] OLIVEIRA, R. M., FERREIRA, P. A. V. (2003b). Programação Geométrica Multiobjetivo: Uma Implementação Computacional Via Matlab. *In Anais do XXXV Simpósio Brasileiro de Pesquisa Operacional*, Natal-Brasil.
- [Oliveira & Ferreira, 2004a] OLIVEIRA, R. M., FERREIRA, P. A. V. (2004a). Otimização Global de Problemas de Programação Multiplicativa Via Decomposição. *Notes of the XII Congreso Latino Iberoamericano de Investigación de operaciones y Sistemas*, La Ciudad de La Habana-Cuba.
- [Oliveira & Ferreira, 2004b] OLIVEIRA, R. M., FERREIRA, P. A. V. (2004b). Um Algoritmo de Otimização Global para Problemas de Programação Multiplicativa. *In Anais do XV Congresso Brasileiro de Automática*, Gramado-Brasil.
- [Oliveira & Ferreira, 2005a] OLIVEIRA, R. M., FERREIRA, P. A. V. (2005a). A Convex Analysis Approach for Convex Multiplicative Programming. *Annals of Operations Research*, submetido.
- [Oliveira & Ferreira, 2005b] OLIVEIRA, R. M., FERREIRA, P. A. V. (2005b). Global Optimization of Convex Multiplicative Programs by Duality Theory. *Lecture Notes in Computer Science*, 3478:101–111.
- [Pardalos, 1990] PARDALOS, P. M. (1990). Polynomial Time Algorithms for Some Classes of Constrained Quadratic Problems. *Optimization*, 21(6):843–853.

- [Phillips & Beightler, 1973] PHILLIPS, D. T., BEIGHTLER, C. S. (1973). Geometric Programming: A Technical State-of-the-Art Survey. *AIIE Transactions*, 97:97–112.
- [Rockafellar, 1970] ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton University, New Jersey.
- [Ruckaert & Matens, 1978] RUCKAERT, M. J., MATENS, X. M. (1978). Comparison of Generalized Programming Algorithms. *Journal of Optimization Theory and Applications*, 26(2):205–242.
- [Ryoo & Sahinids, 2003] RYOO, H. S., SAHINIDS, N. V. (2003). Global Optimization of Multiplicative Programs. *Journal of Global Optimization*, 26:387–418.
- [Shen & Zhang, 2004] SHEN, P., ZHANG, K. (2004). Global Optimization of Signomial Geometric Programming Using Linear Relaxation. *Elsevier Applied Mathematics and Computation*, 150:99–114.
- [Thoai, 1991] THOAI, N. V. (1991). A Global Optimization Approach for Solving the Convex Multiplicative Programming Problems. *Journal of Global Optimization*, 1(4):341–357.
- [Thoai, 2002] THOAI, N. V. (2002). Convergence and Application of a Decomposition Method Using Duality Bounds for Nonconvex Global Optimization. *Journal of Optimization Theory and Applications*, 133:165–193.
- [Vavasis, 1995] VAVASIS, S. A. (1995). *Complexity Issues in Global Optimization: A Survey*, Handbook of Global Optimization Problems, 27–40. Kluwer Academic Publishers, Netherlands.
- [Yang & Bricker, 1996] YANG, H. H., BRICKER, D. L. (1996). Investigation of Path-Following Algorithms for Signomial Geometric Programming Problems. *European Journal of Operational Research*, 103:230–241.
- [Yazarel & Pappas, 2004] YAZAREL, H., PAPPAS, G. (2004). Geometric Programming Relaxations for Linear System Reachability. *In Proceedings of the 2004 American Control Conference (ACC)*, Boston-MA, 553–559.
- [Yu, 1985] YU, P. L. (1985). Multiple-Criteria Decision Making. *Plenum Press*, 10.
- [Zener, 1961] ZENER, C. (1961). A Mathematical Aid in Optimizing Engineering Designs. *Proceedings of the National Academy of Sciences*, 47(4):537–539.