UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM
INSTITUTO DE CIÊNCIAS EXATAS - ICE
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

ALGUMAS CONTRIBUIÇÕES PARA A OTIMIZAÇÃO MULTIOBJETIVO VIA TEORIA DOS CONES

RAPHAEL RIBEIRO COSTA

MANAUS - 2013

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM
INSTITUTO DE CIÊNCIAS EXATAS - ICE
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Raphael Ribeiro Costa

ALGUMAS CONTRIBUIÇÕES PARA A OTIMIZAÇÃO MULTIOBJETIVO VIA TEORIA DOS CONES

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática, na área de concentração em Otimização.

Orientador: Prof. Dr. Sandro Dimy Barbosa Bitar

MANAUS - 2013

FICHA CATALOGRÁFICA

COSTA, R. R.,

Algumas contribuições para a otimização multiobjetivo via teoria dos cones, Manaus, 2013.

X, 83 p. 21,0 cm x 29,7 cm (UFAM, M.Sc., Matemática, 2013)
Dissertação de Mestrado - Universidade Federal do Amazonas, UFAM

1 - Otimização multiobjetivo.
2 - Teoria dos cones.
3 - Pareto-ótimo.
4 - Condições de otimalidade.

I. PPGM/UFAM II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

COSTA, R. R. *Algumas contribuições para a otimização multiobjetivo via teoria dos cones*. Manaus, 2013. Dissertação de Mestrado. PPGM - Universidade Federal do Amazonas.

CESSÃO DE DIREITOS

AUTOR: Raphael Ribeiro Costa.
TÍTULO DA DISSERTAÇÃO: Algumas contribuições para a otimização multiobjetivo via teoria dos cones
GRAU/ANO: Mestre em Matemática/2013

É concedida à Universidade Federal do Amazonas a permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação, e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem sua autorização por escrito.

Raphael Ribeiro Costa
Rua 248, Nº29, QUADRA 435
Cidade Nova 5
CEP: 69097-500 - Manaus/AM
e-mail: *raphael.ufam@gmail.com*

Raphael Ribeiro Costa

ALGUMAS CONTRIBUIÇÕES PARA A OTIMIZAÇÃO MULTIOBJETIVO VIA TEORIA DOS CONES

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática, na área de concentração em Otimização.

Manaus, 11 de Janeiro de 2013.

BANCA EXAMINADORA

Prof. Dr. Sandro Dimy Barbosa Bitar - Presidente
Universidade Federal do Amazonas

Prof. Dr. João Xavier Da Cruz Neto - Membro
Universidade Federal do Piauí

Prof. Dra. Flávia Morgana de Oliveira Jacinto - Membro
Universidade Federal do Amazonas

"*À minha mãe*
Maria do Socorro
Dedico"

AGRADECIMENTOS

Agradeço primeiro a Deus, por permitir a minha existência.

Aos meus pais Raimundo Pereira Costa e Maria do Socorro Costa Ribeiro, pelo carinho, força e apoio incondicional.

Ao meu orientador e amigo, Prof. Sandro Bitar pelas excelentes orientações e confiança depositadas em mim para realização deste trabalho.

Ao meu amigo, Prof. Mário Salvatierra Júnior pelas ótimas contribuições e orientações para realização da dissertação.

Ao programa de Pós Graduação em Matemática (PPGM) da Universidade Federal do Amazonas, por possibilitar o desenvolvimento do trabalho.

À Fundação de Amparo à Pesquisa do Estado do Amazonas - FAPEAM, pelo suporte financeiro para realização deste trabalho.

A todos os meus amigos do curso em mestrado em Matemática da UFAM, pelos excelentes momentos que passamos.

A todos que contribuiram direta ou indiretamente para realização deste trabalho.

"*A confiança em si próprio*
é o primeiro segredo do êxito"

Ralph Waldo Emerson

RESUMO

ALGUMAS CONTRIBUIÇÕES PARA A OTIMIZAÇÃO MULTIOBJETIVO VIA TEORIA DOS CONES

Neste trabalho apresenta-se uma caracterização das soluções (eficientes ou pareto-ótimo) para problemas de otimização multiobjetivo baseado no cálculo de conjuntos tangentes. Os fundamentos teóricos discorrem sobre alguns elementos de análise convexa, teoria dos cones bem como elementos de otimização multiobjetivo necessários para formulação do modelo. Além disso, apresentam-se alguns métodos multiobjetivos clássicos que são classificados em três tipos: à priori, à posteriori e interativos. A partir destes elementos, formula-se as condições de otimalidade para problemas multiobjetivos gerais usando conjuntos de aproximação do conjunto viável do respectivo problema. Ao final, o conjunto viável do problema é definido por restrições de igualdade e desigualdade e, com condições de qualificação e regularidade, obtém-se os multiplicadores de Lagrange.
Palavras-chave: Otimização Multiobjetivo, Teoria dos cones, Condições de Otimalidade, Pareto-ótimo.

ABSTRACT

SOME CONTRIBUTIONS TO MULTIOBJECTIVE OPTIMIZATION THEORY VIA CONES THEORY

This paper presents a characterization of solutions (efficient or pareto-optimal) for multiobjective optimization problems based on the calculation of sets tangent the theoretical discuss some elements of convex analysis, cones theory as well as multiobjective optimization elements needed for the formulation of the model. Furthermore, some classic multiobjectives methods are presented and classified into three types: priori, posteriori and interactive. From these elements, makes up the optimality conditions for general multiobjective problems, using sets of approximation of all of their issue viable set of the respective problem. At the end, the feasible set of the problem is defined by equality and inequality constraints and, with skill and regularity conditions, we obtain the Lagrange multipliers.
Keywords: Multiobjective Optimization, Cones Theory, Optimality Conditions, Pareto-optimal.

Sumário

Introdução **2**
Objetivos . . . 2
Objetivo geral . . . 2
Objetivos específicos . . . 2
Descrição da dissertação . . . 2

1 Fundamentos de Otimização Multiobjetivo **3**
1.1 Definições e Formulação do Problema . . . 3
1.2 Relação de ordem parcial . . . 4
1.3 Pareto Otimalidade . . . 6

2 Elementos da Teoria dos Cones e Análise Convexa **10**
2.1 Conjuntos e Cones Tangentes . . . 11
2.2 Eficiência . . . 18

3 Alguns Métodos de Programação Multiobjetivo Matemática **22**
3.1 Métodos à Priori . . . 23
3.1.1 Métodos baseados em Limitantes . . . 23
3.1.2 Método Lexicográfico . . . 23
3.1.3 Programação por Metas . . . 24
3.2 Métodos à posteriori . . . 26
3.2.1 Método das ponderações . . . 26
3.2.2 Método das ϵ - restrições . . . 27
3.3 Métodos Interativos . . . 28
3.3.1 Geoffrion, Dyer e Feinberg . . . 28
3.4 Algoritmos Genéticos . . . 29
3.4.1 Algoritmos Genéticos aplicados a Otimização Multiobjetivo . . . 31
3.5 Métricas de Desempenho . . . 33
3.5.1 Métricas de convergência . . . 33
3.5.2 Métricas de Diversidade . . . 34

4 Condições de Otimalidade em Otimização Multiobjetivo via Teoria dos Cones **36**
4.1 Condições Necessárias no caso geral . . . 36
4.2 Condição Suficiente no caso geral . . . 38
4.3 Condições necessárias no caso diferenciável . . . 39
4.4 Condição suficiente no caso diferenciável . . . 43
4.5 Condições de Otimalidade de John e Kuhn-Tucker . . . 45
4.5.1 Condições de Otimalidade Necessárias do Tipo John . . . 47
4.5.2 Condições de Otimalidade Necessárias do Tipo 1 de Kuhn-Tucker . . . 50

4.5.3 Condições de Otimalidade Necessárias do Tipo 2 de Kuhn-Tucker . . 51

Considerações Finais **55**
Sugestões para trabalhos futuros 55

Referências Bibliográficas **56**

Introdução

A Otimização Multiobjetivo (ou Otimização Vetorial), embora tenha suas origens no século XIX, tornou-se uma ferramenta prática de amplo alcance apenas nos últimos anos do século XX, em parte devido ao aumento da capacidade computacional disponível. Ao definir de forma precisa o que são as soluções de um problema de otimização com várias funções-objetivo (as soluções Pareto-ótimas), a otimização multiobjetivo fornece o conceito básico sobre o qual é possível estruturar a noção de trade-off (ou compromisso), ou seja, quais escolhas de fato estão disponíveis em um problema determinado. Recentemente, vários campos do conhecimento têm incorporado a noção de conjunto Pareto-ótimo para refinar tanto a qualidade de soluções de problemas identificados em projeto quanto o "insight"obtido em problemas de análise.

Para a otimização vetorial, análises de muitos resultados teóricos para condições de otimalidade de primeira ordem são reveladas pelo teorema fundamental de Lin (teorema 5.1 em LIN (1987)). Há um crescente interesse na generalização de condições de otimalidade de segunda ordem em problemas de otimização multiobjetivo e diferentes abordagens tem sido sugeridas (veja AGHEZZZAF (1999), PENOT (1998)). Nesse contexto o cone tangente desempenha um papél fundamental para estabelecer essas condições. Estendendo tal cone tangente, é possível construir outros conjuntos tangentes de segunda ordem. Em Aghezzaf, Hachimi (1999), os autores obtiveram condições necessárias de segunda ordem por meio de um conjunto tangente de segunda ordem que pode ser considerado uma extensão do cone tangente. Cambini (1997) e Penot (1998) formularam um novo conjunto tangente de segunda ordem que é chamado de cone assintótico de segunda ordem com a finalidade de diminuir a diferença a um nível aceitável as condições necessárias e suficientes de otimalidade para problemas de otimização multiobjetivo.

Nesta dissertação aborda-se condições de otimalidade para problemas de otimização multiobjetivo, também condições do tipo Fritz John e de Kuhn Tucker utilizando a teoria dos cones e algumas condições obtidas recentemente em Aghezzaf, Hachimi (2007) para a otimização escalar.

Objetivo

Objetivo geral

Apresentar condições de otimalidade para problemas de otimização multiobjetivo via teoria dos cones.

Objetivos Específicos

Esta dissertação têm alguns objetivos específicos, tais como, destacar elementos da otimização multiobjetivo e da análise convexa, bem como a teoria dos cones e conjuntos tangentes, que são necessários para a descrição do modelo matemático investigado. Um outro objetivo é apresentar resultados sobre condições de otimalidade para problemas de otimização multiobjetivo por meio da teoria dos cones.

Além disso, para problemas de otimização multiobjetivo com restrição de igualdade e desigualdade apresenta-se a versão para os Multiplicadores de Lagrange associados às funções objetivo, utilizando-se condições de qualificação e regularidade sobre a solução.

Descrição da dissertação

Este trabalho está organizado em quatro capítulos. O Capítulo 1 trata de alguns elementos matemáticos utilizados com a otimização multiobjetivo, assim como a formulação geral de um problema vetorial. No Capítulo 2, apresenta-se fundamentos matemáticos sobre análise convexa bem como teoria dos cones e conjuntos tangentes. No capítulo 3, apresenta-se os métodos clássicos de otimização multiobjetivo assim como são descritas algumas técnicas que utilizam algoritmos genéticos. No capítulo 4, descreve-se as condições de otimalidade para problemas de otimização multiobjetivo (POM), condições de otimalidade para os casos geral e diferenciável via teoria dos cones e também os multiplicadores de Lagrange e, algumas condições de qualificação e regularidade em programação não-linear para tal problema. Por último, as perspectivas para trabalhos futuros são apresentadas na Conclusão deste Trabalho.

Capítulo 1

Fundamentos de Otimização Multiobjetivo

A teoria de Otimização Multiobjetivo teve início pelas contribuições do economista britânico Francis Ysidro Edgeworth($1845 - 1926$) que primeiro introduziu o conceito de otimalidade em 1881 (EDGEWORTH (1881)) e que depois foi generalizado pelo engenheiro italiano Vilfredo Pareto ($1848 - 1923$) em *Cours d'économie politique* ($1896 - 97$). Nessa obra além de outras teorias, estudava-se uma condição que otimiza o bem-estar de uma sociedade. Pareto dizia que a sociedade estava em estado ótimo se a melhora de um indivíduo na população significasse a piora de outro. Esse conceito dá-se o nome ótimo de Pareto sendo esse relevante para o estudo da programação multiobjetivo.

O estudo de problemas de otimização tornou-se aplicável a várias áreas do conhecimento e auxilia na solução de vários problemas práticos, como os processos de Tomada de Decisão que envolvem a consideração de determinando número de objetivos conflitantes entre si (SILVA NETO (2011)). A solução desses problemas tem sido motivo da investigação pela comunidade científica.

Devido a sua proeminência para o estudo, neste capítulo apresenta-se aspectos gerais relacionados aos modelos matemáticos que envolvem Otimização Multiobjeivo. Defini-se ordem parcial no espaço $\mathbb{R}^n$ e o conceito de soluções Pareto-ótimo, que são utilizadas no decorrer dessa dissertação.

1.1 Definições e Formulação do Problema

Um Problema de Otimização Multiobjetivo(também chamada de Otimização Vetorial) tem como objetivo encontrar um vetor de variáveis de decisão que satisfaça a um conjunto de restrições e otimize um vetor de funções cujas componentes representam as funções objetivos.

O modelo geral de otimização multiobjetivo pode ser formulado da seguinte forma:

$$\begin{aligned} \min \quad & f(x) \\ s.a. \quad & x \in X \end{aligned} \tag{1.1}$$

Em (1.1), $f : R^n \to R^l$ onde $X \subset R^n$ é o conjunto viável ou espaço de decisões e $Y = f(X) \subset \mathbb{R}^l$ denota o espaço de objetivos. Nesse contexto x é o vetor de decisão e $y = f(x)$ é chamado vetor de objetivos.

Um problema de programação linear multiobjetivo é um caso particular ao Problema (1.1) sendo que as funções coordenadas de f são lineares (ou seja, $f_k(x) = \sum_{i=1}^{n} c_{ki}x_i,\ k = 1, 2, ..., l$) e o conjunto viável X é definido por restrições lineares.

Vale ressaltar que as funções objetivos muitas vezes são conflitantes entre si.

1.2 Relação de ordem parcial

Devido à necessidade de um ordenamento para definição de eficiência em problemas de otimização multiobjetivo(ou seja, definição de solução Pareto-ótimo), apresenta-se nessa seção a definição de ordem parcial em um subconjunto X do espaço vetorial $\mathbb{R}^n$.

Uma relação binária A em $\mathbb{R}^n$ é um subconjunto do produto cartesiano $\mathbb{R}^n \times \mathbb{R}^n$. Assim, diz-se que um elemento $x \in \mathbb{R}^n$ está relacionado com $y \in \mathbb{R}^n$ por A se $(x, y) \in A$.

Definição 1.1. *Seja A uma relação binária em $\mathbb{R}^n$. Dizemos que ela é*

(1) *reflexiva, se $(x, x) \in A$ para todo $x \in \mathbb{R}^n$;*

(2) *anti-simétrica, se $(x, y) \in A$ e $(y, x) \in A$ implica que $x = y$ para $x, y \in \mathbb{R}^n$;*

(3) *transitiva, se $(x, y) \in A$ e $(y, z) \in A$ implica que $(x, z) \in A$ para $x, y, z \in \mathbb{R}^n$;*

(4) *completa ou conectada, se $(x, y) \in A$ ou $(y, x) \in A$ para todo $x, y \in \mathbb{R}^n$ com $x \neq y$;*

(5) *linear, se $(x_1, y_1),\ (x_2, y_2) \in A$ e $t \geq 0$ então $(tx_1, ty_1) \in A$ e $(x_1 + x_2, y_1 + y_2) \in A$.*

Definição 1.2. ***(Conjunto parcialmente ordenado)*** *Sejam $X \subset \mathbb{R}^n$. Dizemos que um conjunto X é parcialmente ordenado se $\forall\ x, y \in X$ valem as propriedades* (1)*,* (2) *e* (3) *da definição* (1.1)*.*

Vale destacar que uma ordem parcial é uma relação binária que possui as propriedades $(1), (2), (3)$. Se além disso, satisfaz a propriedade (5), dizemos que a ordem parcial é linear. Quando uma ordem parcial satisfaz a propriedade (4), ela é chamada de ordem total. Por exemplo, o conjunto dos números reais é totalmente ordenado.

A otimização mono-objetivo é fundamentalmente diferente do caso da otimização multiobjetivo devido a propriedade de ordenamento das soluções, pois na otimização mono-objetivo o espaço dos objetivos é totalmente ordenado e no multiobjetivo parcialmente ordenado. A seguir, apresenta-se para o $\mathbb{R}^n$, ordens parciais usuais.

Definição 1.3. ***(Comparação de vetores)*** *Sejam $x, y \in \mathbb{R}^n$. As relações "$\leqq$", "$\leq$", "$<$"e "$=$", para dados dois vetores $x, y \in \mathbb{R}^n$, serão definidas como segue:*

$$x \leqq y \iff x_i \leqq y_i, \forall i = 1, \ldots, n; \tag{1.2}$$

$$x \leq y \iff x \leqq y \text{ e } x \neq y; \tag{1.3}$$

$$x < y \iff x_i < y_i, \forall i = 1, \ldots, n; \tag{1.4}$$

$$x = y \iff x_i = y_i, \forall i = 1, \ldots, n; \tag{1.5}$$

as relações "$\geqslant$"e "$>$"são definidos de forma análoga. Para relação "$\neq$"tem-se:

$$x \neq y \iff \exists\, i \in \{1, ..., n\}\ ;\ x_i \neq y_i. \tag{1.6}$$

Quando entre vetores x e y não for possível relacioná-los através de "$\leq$"(isto é, não for possível escrever $x \leq y$ ou $y \leq x$), será utilizada a notação $x \nleq y$.

Observação 1.1. *Na definição* (1.3) *para não gerar novos símbolos a desigualdade "menor do que ou igual a "($\leqq$)"simboliza a desigualdade usual dos números reais($x_i \leqq y_i,\ i = 1, ..., n$) e também a desigualdade vetorial($x \leqq y$), ou seja, nessa definição símbolos iguais operam conjuntos diferentes.*

As Figuras 1.1(a) e 1.1(b) ilustram os casos (1.3) e (1.4), respectivamente.

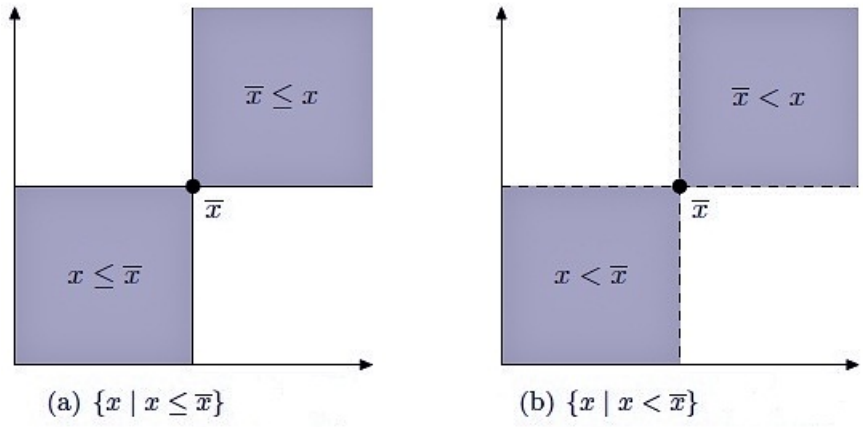


Figura 1.1: *Relações $\leq$ e $<$ no $\mathbb{R}^2$. Fonte Sampaio (2011).*

O conjunto $\mathbb{R}^n_+$ representa o ortante não-negativo dos vetores reais n-dimensionais e o conjunto $\mathbb{R}^n_{++}$ representa o ortante estritamente positivo do mesmo.

Exemplo 1.1. *Considere os vetores $u, v, w \in \mathbb{R}^3$ tais que:*

$$u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, w = \begin{pmatrix} 7 \\ 4 \\ 8 \end{pmatrix}$$

Então verificam-se as seguintes relações:

$$u \leq v,\ u \leq w,\ v \nleq w \, e \, w \nleq v$$

A seguir define-se a ordem lexicográfica que será utilizada no desenvolvimento matemático das condições de otimalidade apresentadas neste trabalho.

Definição 1.4. ***(Ordem lexicográfica)*** *Sejam $x, y \in \mathbb{R}^n \times \mathbb{R}^n$ com $x = (x_1, x_2),\ x_i \in \mathbb{R}^n$ e $y = (y_1, y_2),\ y_i \in \mathbb{R}^n$, definimos $x \leqq_{lex} y$ se:*

$$x_1 < y_1 \ ou \ (x_1 = y_1 \ e \ x_2 \leqq y_2).$$

Similarmente, $x <_{lex} y$:

$$x_1 < y_1 \ ou \ (x_1 = y_1 \ e \ x_2 < y_2).$$

1.3 Pareto Otimalidade

Em problemas de otimização multiobjetivo, devido a conflitos entre as funções objetivos ou até mesmo indefinições em algumas delas, não é possível encontrar uma solução (sentido usual) para o Problema (1.1) (solução viável que satisfaça todas as funções objetivo simultaneamente)(SILVA NETO (2011)).

O problema da solução usual não passa pelos conflitos simplesmente. Os conflitos são modelados com essa técnica(programação multiobjetivo), porém não são eles que causam esse problema. A questão é estrutural, o fato de não se conseguir definir uma solução x tal que $f(x) \leqq f(y)\ \forall y \in X$(sentido usual), torna o estudo de otimalidade, nesse ambiente, "refém"dessa patologia.

Para contornar essa dificuldade é necessário dar um novo sentido ao que vem a ser solução para um problema multiobjetivo(ótimo de Pareto), e isso define o conceito de Pareto Dominância(ou somente dominância)(SILVA NETO (2011)).

Definição 1.5. ***(Pareto Dominância no espaço dos objetivos)*** *Sejam $y_1, y_2 \in Y$, dois vetores no espaço dos objetivos. Dizemos que y_1 domina y_2 se $y_1 \leq y_2$. Em outras palavras, $y_{1i} \leq y_{2i}$, $\forall i \in \{1, ..., l\}$ e $y_{1i} < y_{2i}$ para pelo menos um $i \in \{1, ..., l\}$. Caso ocorra que $y_1 \nleq y_2$ e $y_2 \nleq y_1$ (y_1 não domina y_2 e y_2 não domina y_1), então y_1, y_2 são chamados indiferentes(não comparáveis).*

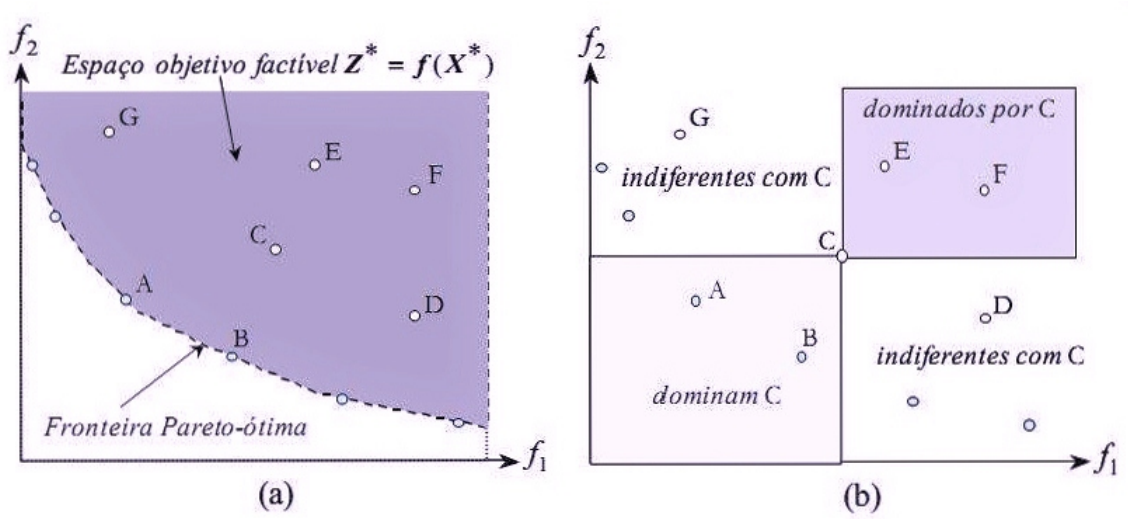


Figura 1.2: *Pareto Dominância no espaço de objetivos. Fonte: Arroyo(2002) apud Silva Neto(2011).*

Na figura 1.2(b), no espaço dos objetivos os pontos G, C, D são não comparáveis, o ponto C domina os pontos pertencentes ao retângulo superior direito e os pontos A, B dominam o ponto C.

Um objeto fundamental da otimização multiobjetivo é o conjunto de soluções, denominado conjunto Pareto-ótimo. As soluções Pareto-ótimo são também conhecidas como soluções não-dominadas ou soluções eficientes. Para uma melhor compreensão do que venha ser um Pareto-ótimo e consequentemente o conjunto Pareto-ótimo, defini-se o conceito de Pareto Dominância. Estabelecido esse conceito, fica mais simples o entendimento de um conjunto formado pelas soluções do Problema (1.1), chamado de ótimo de Pareto ou conjunto Pareto-ótimo.

Definição 1.6. ***(Pareto Dominância no espaço de decisão)*** *Sejam $x_1, x_2 \in X$. Dizemos que x_1 domina x_2 se $f(x_1)$ dominar $f(x_2)$, ou seja, $f(x_1) \leq f(x_2)$. Caso aconteça $f(x_1) \nleq f(x_2)$ e $f(x_2) \nleq f(x_1)$, então x_1, x_2 são chamados não comparáveis ou indiferentes.*

Definição 1.7. ***(Conjunto Pareto-ótimo)*** *Um vetor de decisão $\overline{x} \in X$ é uma solução Pareto-ótima para o Problema* (1.1) *se nenhum $x \in X$ dominar $\overline{x} \in X$, isto é, se não existe outro vetor de decisão $x \in X$ tal que $f(x) \leq f(\overline{x})$. O conjunto de todas as soluções Pareto-ótimas é denominado Conjunto Pareto-ótimo.*

Vale destacar que um vetor $y = f(x) \in Y$ é Pareto-ótimo se seu vetor de decisão correspondente x é Pareto-ótimo. A imagem do conjunto Pareto-ótimo em $Y = f(X)$ é chamada fronteira Pareto-ótima, essa fronteira Pareto-ótima é formada por pontos não comparáveis dois a dois.

Na figura 1.2(a), temos um exemplo da fronteira Pareto-ótima cujos pontos pertencentes a essa fronteira são os pontos Pareto-ótimos.

Vale ressaltar que o interesse inicial em problemas de otimização mono-objetivo é sobre o vetor de decisão aquele que orientará o decisor com respeito a Tomada de Decisão. Contudo,

a estratégia para o resgate de um vetor eficiente para problemas de otimização multiobjetivo passa pela análise do espaço dos objetivos, independente da dimensão.

A definição de Pareto-ótimo pode ser entendida como aquela em que, não conseguiremos outra solução, se melhorarmos um dos objetivos sem piorar qualquer outro objetivo. Porém é muito difícil de obtermos soluções Pareto-ótimo e de trabalharmos condições de otimalidade para pontos eficientes e, por esse motivo, apresenta-se uma definição um pouco mais fraca que é conhecida como solução fracamente Pareto-ótimo.

Definição 1.8. ***(Solução fracamente Pareto-Ótimo(eficiente))*** *Um vetor de decisão $\overline{x} \in X$ é uma solução fracamente Pareto-ótima para o Problema* (1.1) *se não existe outra solução viável tal que $f(x) < f(\overline{x})$. O conjunto de todas as soluções fracamente Pareto-ótimas é denominado Conjunto Fracamente Pareto-ótimo.*

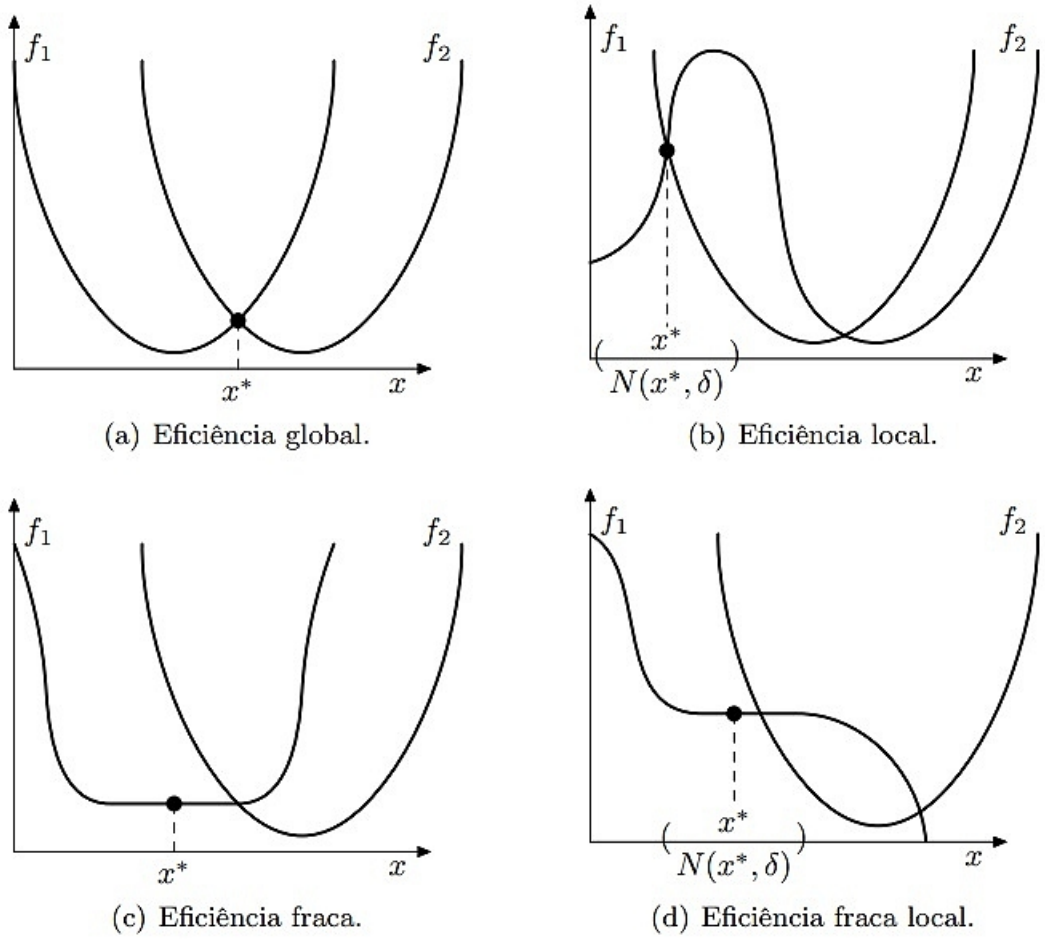
Uma solução fracamente eficiente é obtida se não existe outra solução que possa reduzir todos os objetivos, que é diferente da solução Pareto-ótima para a qual não é possível melhorar algum objetivo sem piorar qualquer outro objetivo. Logo um vetor eficiente(Pareto-ótimo) é fracamente eficiente, mas não vale a recíproca. No caso geral, é utilizado um sentido global para as soluções Pareto-ótimas, mas podemos caracterizar as Pareto-otimalidade e Fracamente Pareto-otimalidade locais para essas soluções.

Definição 1.9. ***(Solução localmente Pareto-ótimo)*** *Um vetor de decisão $\overline{x} \in X$ é localmente Pareto-ótimo se existe $\delta > 0$ tal que $\overline{x}$ é Pareto-ótimo em $X \cap B(\overline{x}, \delta)$. Um vetor de objetivos $y* \in Y$ é localmente Pareto-ótimo se seu vetor de decisão correspondente é localmente Pareto-ótimo.*

Definição 1.10. ***(Solução fracamente Pareto-ótimo local)*** *Um vetor de decisão $\overline{x} \in X$ é fracamente Pareto-ótimo local se existe $\delta > 0$ tal que $\overline{x}$ é fracamente Pareto-ótimo em $X \cap B(\overline{x}, \delta)$. Um vetor de objetivos $y* \in Y$ é fracamente Pareto-ótimo local se seu vetor de decisão correspondente é fracamente Pareto-ótimo local.*

Quando estivermos nos referindo tanto a soluções eficientes quanto a soluções eficientes locais em algum momento, denotaremos as primeiras por soluções eficientes globais. A figura 1.3 ilustra os quatro tipos de eficiência definidos anteriormente.

Na figura 1.3 observe que, nos exemplos, as soluções eficientes locais, fracamente eficientes e fracamente eficientes locais não são soluções eficientes globais. Além disso, a solução fracamente eficiente não é uma solução eficiente local e a solução fracamente eficiente local não é uma solução fracamente eficiente.



(a) Eficiência global. (b) Eficiência local.

(c) Eficiência fraca. (d) Eficiência fraca local.

Figura 1.3: *Exemplos de Eficiência. Fonte Sampaio (2011).*

Capítulo 2

Elementos da Teoria dos Cones e Análise Convexa

Neste capítulo, apresenta-se as definições, exemplos e teoremas básicos necessários para o desenvolvimento deste trabalho.

Para a Otimização Multiobjetivo, toda solução globalmente Pareto-ótima será localmente Pareto-ótima, analogamente ao caso da Otimização Mono-objetivo. A recíproca só será verdadeira adicionando algumas outras hipóteses dependentes de alguns resultados de Análise Convexa, apresentados a seguir.

Definição 2.1. ***(Conjuntos Convexos)*** *Dizemos que um conjunto $X \subset \mathbb{R}^n$ é convexo se, para todo $x, y \in X$ e $\alpha \in [0,1]$*

$$\alpha x + (1-\alpha)y \in X \tag{2.1}$$

Exemplo 2.1. *O conjunto vazio, o espaço $\mathbb{R}^n$, um conjunto que contém só um ponto, qualquer subespaço $\mathbb{R}^n$, qualquer semi-espaço em $\mathbb{R}^n$, sendo $A \in \mathbb{R}^{mxn}$ e $b \in \mathbb{R}^m$ o conjunto dado por $X = \{x \in R^n | Ax = b, x \geq 0\}$ são conjuntos convexos. Qualquer conjunto não conexo é não convexo.*

Definição 2.2. ***(Combinação convexa)*** *Sejam $x_1, x_2, ..., x_n \in \mathbb{R}^n$. A combinação convexa desses vetores é dada por:*

$$\sum_{i=1}^{n} \alpha_i x_i = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n, \ com \ \sum_{i=1}^{n} \alpha_i = 1, \alpha_i \geq 0 \tag{2.2}$$

Definição 2.3. ***(Fecho convexo)*** *Seja $X \subset \mathbb{R}^n$. O fecho convexo de X é a interseção de todos os conjuntos convexos em $\mathbb{R}^n$ que contêm X e denota-se por $convX$.*

Um exemplo que mostra que o fecho convexo de um conjunto fechado pode não ser fechado.

Exemplo 2.2. *Seja $X = \{x = (x_1, x_2) \in \mathbb{R}^2 \ ; \ x_2 = 0\} \cup \{(0,1)\}$. Claramente, X é um conjunto fechado. Mas como pode ser visto, $convX = \{x = (x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 \leq x_2 < 1\} \cup \{(0,1)\}$ que não é fechado.*

Definição 2.4. ***(Ponto extremo)*** *Seja $X = \{x \in R^n | Ax = b, x \geq 0\}$. Um ponto $x \in X$ é um ponto extremo de X se não existirem $y, z \in X$ distintos de x e $\alpha \in [0,1]$ tais que*

$$x = \alpha y + (1 - \alpha)z \tag{2.3}$$

Definição 2.5. ***(Função convexa)*** *Seja $X \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f_i : X \longrightarrow \mathbb{R}$ é convexa se, para quaisquer $x, y \in X$,*

$$f_i(\alpha x + (1 - \alpha)y) \leq \alpha f_i(x) + (1 - \alpha) f_i(y) \tag{2.4}$$

para todo $\alpha \in [0,1]$. Se, para quaisquer $x, y \in X$, sendo $x \neq y$ e $0 < \alpha < 1$, a desigualdade for estrita, então f_i é estritamente convexa.

Definição 2.6. ***(Programação Multiobjetivo Convexa)*** *O problema de Otimização Multiobjetivo definido em (1.1) é convexo se todas as funções coordenadas da função vetorial f e o conjunto $X \subset \mathbb{R}^n$ são convexos.*

A partir destas definições, o próximo teorema fornece condições para que uma solução localmente Pareto-ótima seja globalmente Pareto-ótima.

Teorema 2.1. *Se o problema de otimização definido em (1.1) for convexo, então toda solução localmente Pareto-ótima é também globalmente Pareto-ótima.*

Demonstração. Ver por exemplo em Miettinem (1999). □

2.1 Conjuntos e Cones Tangentes

Para desenvolver as condições de otimalidade no caso de problemas com restrições, precisamos estudar não só o comportamento da função vetorial numa vizinhança da solução, mas também a estrutura do conjunto viável nessa vizinhança. As noções de direções viáveis e direções tangentes são fundamentais neste contexto e são muito importantes para um entendimento correto das condições de otimalidade (IZMAILOV E SOLODOV (2009)).

Definição 2.7. ***(Direção viável)*** *Seja $X \subset \mathbb{R}^n$ um conjunto não-vazio. Dizemos que $d \in \mathbb{R}^n$ é uma direção viável em relação ao conjunto X no ponto $\overline{x} \in X$, se existe $\epsilon > 0$ tal que*

$$\overline{x} + td \in X \ \forall \ t \in [0, \epsilon] \tag{2.5}$$

Para tornar mais natural o entendimento sobre conjuntos de direções, a seguinte definição é útil.

Definição 2.8. ***(Cone)*** *Um conjunto $D \subset \mathbb{R}^n$ é um cone, se para todo $x \in D$, $\lambda \geq 0, \lambda x \in D$. Se D é convexo, então dizemos que D é um cone convexo.*

Exemplo 2.3. *O espaço $\mathbb{R}^n$, qualquer subespaço de $\mathbb{R}^n$, o ortante não-negativo $\mathbb{R}^n_+$ são exemplos de cones.*

Proposição 2.1. ***(Cone Convexo)*** *Seja D um cone. Então, D é convexo se, e somente se, $D + D \subseteq D$.*

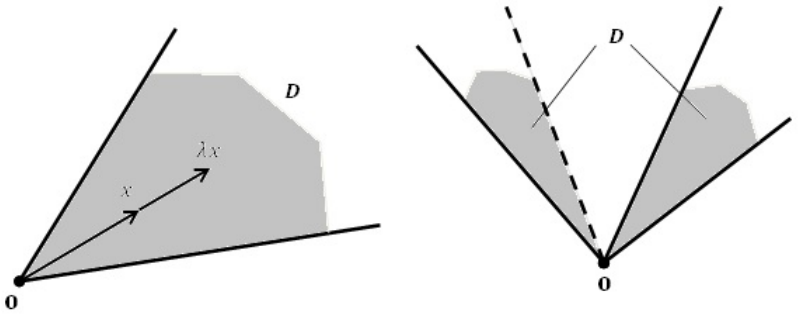


Figura 2.1: *Ilustração do cone de um conjunto D.*

Definição 2.9. ***(Cone Pontudo)***. *Dizemos que um cone D é pontudo se $D \cap -D = \{0\}$.*

Informalmente, um cone é um conjunto de direções. Um cone não vazio sempre é ilimitado, com exceção do cone trivial $\{0\}$.

Definição 2.10. *Um ponto $a \in \mathbb{R}^n$ diz-se aderente a um conjunto $X \subset \mathbb{R}^n$ quando é limite de uma sequência de pontos desse conjunto.*

Definição 2.11. *O fecho (ou aderência) de um conjunto X no espaço $\mathbb{R}^n$ é o conjunto $\overline{X}$ dos pontos de $\mathbb{R}^n$ que são aderentes a X.*

A próxima definição matemática abordada nesse trabalho é encontrado em Penot (1994).

Definição 2.12. ***(Cone Tangente)*** *Seja X um conjunto não-vazio de $\mathbb{R}^n$. O cone tangente de X em $\overline{x} \in \overline{X}$ (fecho de X) é o conjunto definido por*

$$T(X, \overline{x}) \;=\; \{d \in \mathbb{R}^n |\; \exists\; t_n \to 0^+, \exists\; d_n \to d \text{ tal que } x_n = \overline{x} + t_n d_n \in X\} \qquad (2.6)$$

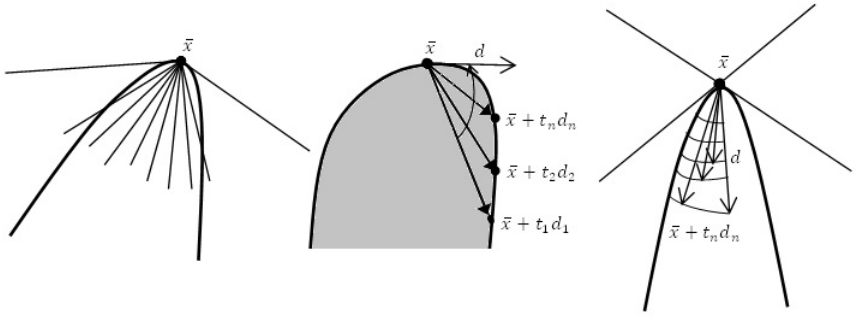


Figura 2.2: *Ilustrações do cone tangente*

Exemplo 2.4. *Seja $S^{n-1} \subset \mathbb{R}^n$ a esfera unitária. Então $T(S^{n-1}, \overline{x}) = \overline{x}^{\perp}$.*

Vamos mostrar que $T(S^{n-1}, \overline{x}) \subset \overline{x}^{\perp}$. Dado $d \in T(S^{n-1}, \overline{x})$, $\exists\ t_n \to 0^+$ e $d_n \to d$ tal que $x_n = \overline{x} + t_n d_n$. Como $x_n \in S^{n-1}$ tem-se $||x_n|| = 1$. Logo

$$||\overline{x} + t_n d_n|| = 1 \Rightarrow ||\overline{x} + t_n d_n||^2 = 1$$

$$\Rightarrow < \overline{x} + t_n d_n, \overline{x} + t_n d_n >= 1$$

$$\Rightarrow < \overline{x}, \overline{x} > + < t_n d_n, \overline{x} > + < \overline{x}, t_n d_n > + < t_n d_n, t_n d_n >= 1$$

$$\Rightarrow 2 < d_n, \overline{x} > + t_n < d_n, d_n >= 0$$

$$\lim_{n \to +\infty} (2 < d_n, \overline{x} > + t_n < d_n, d_n >) = 0 \Rightarrow < d, \overline{x} >= 0.$$

Portanto $T(S^{n-1}, \overline{x}) \subset \overline{x}^{\perp}$.

Vamos mostrar que $T(S^{n-1}, \overline{x}) \supset \overline{x}^{\perp}$. Dado $d \in \overline{x}^{\perp}$, ou seja, $< d, \overline{x} >= 0$. Considere $d_n = \alpha_n d - \beta_n \overline{x}$, $< d_n, \overline{x} >= -\beta_n$ em que $\alpha_n \to 1^+$ e $\beta_n \to 0^+$. Definindo

$$t_n = -\frac{< d_n, \overline{x} >}{< d_n, d_n >} \ tem - se\ t_n = \frac{2\beta_n}{< d_n, d_n >} \ com\ d_n \to d\ e\ t_n \to 0^+$$

Seja $d \in \overline{x}^{\perp}$, então $< d, \overline{x} >= 0$. Sendo $x_n = \overline{x} + t_n d_n$ temos que

$$x_n = \overline{x} + \frac{2\beta_n}{< d_n, d_n >} d_n$$

Vamos provar que $||x_n|| = 1$. Então

$$||x_n||^2 = ||\overline{x} + t_n d_n||^2$$

$$\Rightarrow < \overline{x}, \overline{x} > + 2t_n < d_n, \overline{x} > + t_n^2 < d_n, d_n > .$$

Como $\overline{x} \in S^{n-1}$ tem-se $< \overline{x}, \overline{x} >= 1$. Agora vamos mostrar que

$$2t_n < d_n, \overline{x} > + t_n^2 < d_n, d_n >= 0.$$

De fato,

$$2t_n < d_n, \overline{x} > + t_n^2 < d_n, d_n >= \frac{-4\beta_n^2}{< d_n, d_n >} + \frac{4\beta_n^2}{< d_n, d_n >} = 0.$$

Logo

$$T(S^{n-1}, \overline{x}) \supset \overline{x}^{\perp}.$$

Portanto $T(S^{n-1}, \overline{x}) = \overline{x}^{\perp}$.

Estendendo a definição de cone tangente, pode-se construir outros conjuntos tangentes de segunda ordem. Essas aproximações da região viável X, em $\overline{x} \in \overline{X}$ ($\overline{X}$ representa nessa dissertação o fecho do conjunto X), podem ser consideradas como duas diferentes extensões do cone tangente $T(X, \overline{x})$. Essas aproximações serão usadas para a formulação das condições de otimalidade de segunda ordem.

Definição 2.13. ***(Conjunto tangente de segunda ordem)*** *Seja X um conjunto não-vazio de $\mathbb{R}^n$. O conjunto tangente de segunda ordem de X em $\overline{x} \in \overline{X}$ é o conjunto definido por*

$$\begin{aligned} T^2(X, \overline{x}) &= \{(d, z) \in \mathbb{R}^n \times \mathbb{R}^n | \ \exists\ t_n \to 0^+, \exists\ z_n \to z \\ &\quad tal\ que\ x_n = \overline{x} + t_n d + (1/2) t_n^2 z_n \in X\} \end{aligned} \tag{2.7}$$

Cambini e Penot (2007) definem um novo conjunto tangente de segunda ordem chamado de cone assintótico de segunda-ordem.

Definição 2.14. ***(Cone tangente assintótico de segunda ordem)*** *Seja X um conjunto não-vazio de $\mathbb{R}^n$. O cone assintótico de segunda ordem de X em $\overline{x} \in \overline{X}$ é o conjunto definido por*

$$\begin{aligned} T''(X, \overline{x}) &= \{(d, z) \in \mathbb{R}^n \times \mathbb{R}^n | \ \exists\ t_n \to 0^+, \exists\ r_n \to 0^+, \exists\ z_n \to z \\ &\quad tal\ que\ \frac{t_n}{r_n} \to 0, x_n = \overline{x} + t_n d + (1/2) r_n t_n z_n \in X\} \end{aligned} \tag{2.8}$$

Outras aproximações da região viável de X em $\overline{x} \in \overline{X}$ podem ser inspiradas pela noção dada no trabalho de Cambini (1997).

Definição 2.15. ***(Conjunto tangente de segunda ordem com índice k)*** *Seja $k \geqq 0$ um número real, define-se o conjunto tangente de segunda ordem com índice k de X em $\overline{x} \in \overline{X}$ por*

$$\begin{aligned} T_k''(X, \overline{x}) &= \{(d, z) \in \mathbb{R}^n \times \mathbb{R}^n | \ \exists t_n \to 0^+, \exists r_n \to 0^+, \exists\ z_n \to z \\ &\quad tal\ que\ \frac{t_n}{r_n} \to k, x_n = \overline{x} + t_n d + (1/2) r_n t_n z_n \in X\} \end{aligned} \tag{2.9}$$

As d-seções de $T_k''(X, \overline{x})$ são denotadas por

$$T_k''(X, \overline{x})(d) = \{z \in \mathbb{R}^n | (d, z) \in T_k''(X, \overline{x})\} \tag{2.10}$$

A seguinte proposição mostra a relação entre o conjunto tangente de segunda ordem de índice $k > 0$ e o clássico conjunto tangente de segunda ordem.

Proposição 2.2. *Seja X um subconjunto não-vazio de R^n e seja $\overline{x} \in \overline{X}$. Então, para todo $k > 0$, $T_k''(X, \overline{x}) = (1/k) T^2(X, \overline{x})$. Em particular, $T_1''(X, \overline{x}) = T^2(X, \overline{x})$.*

Demonstração. Segue-se imediatamente da definição de $T_k''(X,\overline{x})$. □

Observação 2.1. *Observe que* $T_0''(X,\overline{x}) = T''(X,\overline{x})$ *e* $T_1''(X,\overline{x}) = T^2(X,\overline{x})$. *Mais ainda, para todo* $k \geqq 0$, *tem-se* $T_k''(X,\overline{x})(0) = T(X,\overline{x})$ *e se* $d \notin T(X,\overline{x})$ *então* $T_k''(X,\overline{x})(d) = \emptyset$.

Exemplo 2.5. *Seja o conjunto* Z *usado por Cambini* (1997) :
$Z = \{(x,y) \in \mathbb{R}^2 : y = x\sqrt{x}, x \geq 0\}$, $\overline{y} = (0,0)$.

Então

(i) $T(Z,\overline{y}) = \{(d_1,d_2) \in \mathbb{R}^2 : d_2 = 0\}$;

(ii) $T_1''(Z,\overline{y}) = T^2(Z,\overline{y}) = \{(0,0)\} \times \{(z_1,z_2) \in \mathbb{R}^2 : z_2 = 0\}$;

(iii) $T_0''(Z,\overline{y}) = T''(Z,\overline{y}) = \{(d,z) \in \mathbb{R}^4 : \ d_1 > 0, \ d_2 = 0 \ e \ d_1.z_2 \geqq 0\} \cup T^2(Z,\overline{y})$.

Verificação:

(i) Seja $d = (d_1,d_2) \in T(Z,\overline{y})$. *Tome* $d_n = (d_1^n, d_2^n) \to d$, $t_n \to 0^+$, $y_n = (y_1^n, y_2^n)$ *e* $y_0 = (0,0)$.

Sendo $y_n = \overline{y} + t_n d_n = (0,0) + t_n(d_1^n, d_2^n) = (t_n d_1^n, t_n d_2^n)$. *Logo*

$(y_1^n, y_2^n) = (t_n d_1^n, t_n d_2^n) \in Z \Rightarrow t_n d_2^n = t_n d_1^n \sqrt{(t_n d_1^n)} \Rightarrow \ t_n d_2^n = \sqrt{(t_n d_1^n)^3}$. *Logo*

$$d_2^n = \sqrt{t_n (d_1^n)^3} \tag{2.11}$$

Para $n \to +\infty$ *em* (2.11) *temos que* $d_2 = 0$ *e portanto* $T(Z,\overline{y}) = \{(d_1,d_2) \in \mathbb{R}^2 : d_2 = 0\}$.

(ii) Seja $(d,z) \in T_1''(Z,\overline{y})$, *com* $d = (d_1,d_2)$ *e* $z = (z_1,z_2)$. *Considere* $z_n = (z_1^n, z_2^n) \to (z_1,z_2) = z$, $t_n \to 0^+$, $y_n = (y_1^n, y_2^n) \to (y_1,y_2) = y$ *e* $\overline{y} = (0,0)$.

Sendo $y_n = \overline{y} + t_n d + \frac{1}{2}t_n^2 z_n = (t_n d_1, t_n d_2) + (\frac{1}{2}t_n^2 z_1^n, \frac{1}{2}t_n^2 z_2^n) = (t_n d_1 + \frac{1}{2}t_n^2 z_1^n, t_n d_2 + \frac{1}{2}t_n^2 z_2^n) \in Z$. *Logo*

$$t_n(d_2 + \frac{1}{2}t_n z_2^n) = \sqrt{(t_n(d_1 + \frac{1}{2}t_n z_1^n))^3} \ \Rightarrow$$

$$d_2 + \frac{1}{2}t_n z_2^n = \sqrt{t_n}\sqrt{(d_1 + \frac{1}{2}t_n z_1^n)^3} \ com \ d_1 \geqq 0. \tag{2.12}$$

Aplicando o limite em (2.12) *quando* $n \to +\infty$ *temos que* $d_2 = 0$.

Substituindo $d_2 = 0$ *na equação* (2.12) *temos*

$$\frac{1}{2}t_n z_2^n = \sqrt{t_n}\sqrt{(d_1 + \frac{1}{2}t_n z_1^n)^3}$$

Então

$$\frac{1}{2}\sqrt{t_n} z_2^n = \sqrt{(d_1 + \frac{1}{2}t_n z_1^n)^3} \tag{2.13}$$

Aplicando o limite em (2.13) *quando* $n \to +\infty$ *temos que* $d_1 = 0$.

Substituindo $d_1 = 0$ *e* $d_2 = 0$ *na equação* (2.12) *temos*

$$\frac{1}{2}t_n z_2^n = \sqrt{t_n}\sqrt{(\frac{1}{2}t_n z_1^n)^3}$$

Logo

$$z_2^n = 2.t_n\sqrt{(\frac{1}{2}z_1^n)^3} \tag{2.14}$$

Aplicando o limite em (2.14) *quando* $n \to +\infty$ *obtemos* $z_2 = 0$.

Nessas observações $T^2(Z, \overline{y}) = \{(0,0)\} \times \{(z_1, z_2) \in \mathbb{R}^2 : z_2 = 0\}$.

(iii) Sejam $z_n = (z_1^n, z_2^n) \to (z_1, z_2) = z$, $d = (d_1, d_2)$, $y_n = (y_1^n, y_2^n) \to (y_1, y_2) = y$, $\overline{y} = (0,0)$, $(t_n, r_n) \to (0^+, 0^+)$ *tal que* $\frac{t_n}{r_n} \to 0$.

Como $y_n = \overline{y} + t_n d + \frac{1}{2}r_n t_n z_n = (t_n d_1, t_n d_2) + (\frac{1}{2}r_n t_n z_1^n,\ \frac{1}{2}r_n t_n z_2^n) = (t_n(d_1 + \frac{1}{2}r_n z_1^n,\ t_n(d_2 + \frac{1}{2}r_n z_2^n) \in E$. *Logo*

$$t_n(d_2 + \frac{1}{2}r_n z_2^n) = \sqrt{(t_n(d_1 + \frac{1}{2}r_n z_1^n))^3} \ \Rightarrow$$

$$d_2 + \frac{1}{2}r_n z_2^n = \sqrt{t_n}\sqrt{(d_1 + \frac{1}{2}r_n t_n z_1^n)^3} \text{ com } d_1 \geqq 0. \tag{2.15}$$

Aplicando o limite quando $n \to +\infty$ *em* (2.15) *temos que* $d_2 = 0$.

Substituindo $d_2 = 0$ *na equação* (2.15) *obtemos*

$$\frac{1}{2}r_n z_2^n = \sqrt{t_n}\sqrt{(d_1 + \frac{1}{2}r_n t_n z_1^n)^3}$$

Se $d_1 = 0$ *na equação* (2.15) *obtemos* $z_2 = 0$.

Se $d_1 > 0$ *na equação* (2.15) *tem-se* $z_2^n = \frac{\sqrt{t_n}}{r_n}\sqrt{(d_1 + \frac{1}{2}r_n t_n z_1^n)^3}$. *Para* $n \in \mathbb{N}$ *suficiente-*

mente grande temos que

$$z_2^n > \sqrt{t_n}\sqrt{(d_1 + \frac{1}{2}r_n t_n z_1^n)^3} \tag{2.16}$$

Aplicando o limite quando $n \to +\infty$ *em* (2.16) *tem-se* $z_2 \geqq 0$.

Portanto $T''(Z, \overline{y}) = \{(d, z) \in \mathbb{R}^4 : \ d_1 > 0, \ d_2 = 0 \ e \ d_1.z_2 \geqq 0\} \cup T^2(Z, \overline{y})$.

Exemplo 2.6. *Seja o conjunto E usado por Jiménez e Novo* (2004) :
$E = \{(y_1, y_2) \in \mathbb{R}^2 : y_2 = -|y_1|^{\frac{3}{2}}\}$, $y_0 = (0, 0)$ *e* $d = (-1, 0)$.

Então
(i) $T(E, y_0) = \{(d_1, d_2) : d_2 = 0\}$;

(ii) $T^2(E, y_0)(d) = \emptyset$;

(iii) $T''(E, y_0)(d) = \{(z_1, z_2) : z_2 \leqq 0\}$.

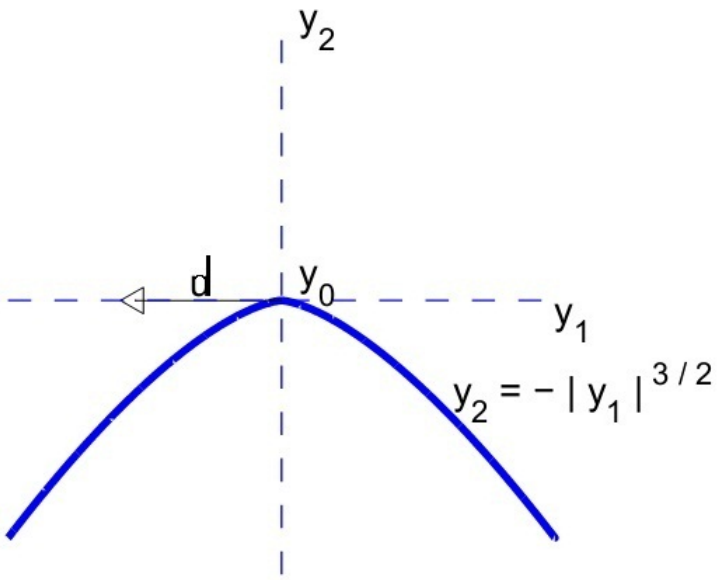


Figura 2.3: *Ilustração do conjunto E*

Verificação:

(i) Sejam $d_n = (d_1^n, d_2^n) \to d = (-1, 0)$, $t_n \to 0^+$, $y_n = (y_1^n, y_2^n)$ *e* $y_0 = (0, 0)$.

Sendo $y_n = y_0 + t_n d_n = (0, 0) + t_n(d_1^n, d_2^n) = (t_n d_1^n, t_n d_2^n)$. *Logo*

$(y_1^n, y_2^n) = (t_n d_1^n, t_n d_2^n) \in E \Rightarrow t_n d_2^n = -|t_n d_1^n|^{\frac{3}{2}} \Rightarrow t_n d_2^n = -\sqrt{|t_n d_1^n|^3} \Rightarrow$

$$d_2^n = -\sqrt{t_n |d_1^n|^3} \tag{2.17}$$

Para $n \to +\infty$ *em* (2.17) *temos que* $d_2 = 0$ *e portanto* $T(E, y_0) = \{(d_1, d_2) : d_2 = 0\}$.

(ii) Sejam $z_n = (z_1^n, z_2^n) \to (z_1, z_2) = z$, $(t_n, r_n) \to (0^+, 0^+)$ $y_n = (y_1^n, y_2^n) \to (y_1, y_2) = y$, $d = (-1, 0)$ *e* $y_0 = (0, 0)$.

Sendo $y_n = y_0 + t_n d + \frac{1}{2} t_n^2 z_n = (-t_n, 0) + (\frac{1}{2} t_n^2 z_1^n, \frac{1}{2} t_n^2 z_2^n) = (-t_n + \frac{1}{2} t_n^2 z_1^n, \frac{1}{2} t_n^2 z_2^n) \in E$. *Logo*

$$\frac{1}{2} t_n^2 z_2^n = -\sqrt{|t_n + \frac{1}{2} t_n^2 z_1^n|^3} \Rightarrow$$

$$z_2^n = -2\sqrt{\frac{1}{t_n} |(\frac{t_n}{2} - 1) z_1^n|^3}. \tag{2.18}$$

Aplicando o limite quando $n \to +\infty$ *em* (2.18) *temos que* $z_2 \to +\infty$.

Portanto $T^2(E, y_0)(d) = \emptyset$.

(iii) Sejam $z_n = (z_1^n, z_2^n) \to (z_1, z_2) = z$, $y_n = (y_1^n, y_2^n) \to (y_1, y_2) = y$, $d = (-1, 0)$, $y_0 = (0, 0)$ $(t_n, r_n) \to (0^+, 0^+)$ *tal que* $\frac{t_n}{r_n} \to 0$.

Como $y_n = y_0 + t_n d + \frac{1}{2} r_n t_n z_n = (-t_n, 0) + (\frac{1}{2} r_n t_n z_1^n,\ \frac{1}{2} r_n t_n z_2^n) = (t_n(-1 + \frac{1}{2} r_n z_1^n,\ \frac{1}{2} r_n t_n z_2^n) \in E$. *Logo*

$\frac{1}{2} r_n t_n z_2^n = -\sqrt{|t_n(-1 + \frac{1}{2} r_n z_1^n|^3} \Rightarrow -z_2^n = 2\frac{\sqrt{t_n}}{r_n}\sqrt{|-1 + \frac{1}{2} r_n z_1^n|^3}$. *Para* $n \in \mathbb{N}$ *suficientemente grande tem-se*

$$-z_2^n > 2\sqrt{t_n}\sqrt{|-1 + \frac{1}{2} r_n z_1^n|^3} \tag{2.19}$$

Aplicando o limite quando $n \to +\infty$ *em* (2.19) *temos que* $z_2 \leqq 0$

Portanto $T''(E, y_0)(d) = \{(z_1, z_2) : z_2 \leqq 0\}$.

2.2 Eficiência

Miettinen (1999) afirma que é possível definir otimalidade em contextos mais gerais para problemas de otimização multiobjetivo como uma opção à otimalidade de Pareto e fracamente Pareto-ótimo. Seja D um cone convexo pontudo em $\mathbb{R}^l$. Tal cone D pode ser usado para induzir uma ordem parcial em R^l da seguinte maneira:

Proposição 2.3. *Seja A uma relação binária em $\mathbb{R}^l$. A é uma ordem parcial linear se, e somente se, existe um cone convexo pontudo $D \subseteq \mathbb{R}^l$ tal que*

$$(y_1, y_2) \in A \Longleftrightarrow y_2 - y_1 \in D. \tag{2.20}$$

Demonstração. ($\Longrightarrow$) Sejam A uma ordem parcial linear e $D = \{y_1 \in \mathbb{R}^l \mid (0, y_1) \in A\}$. Pela definição 1.1 a propriedade (5) garante que D é um cone. Dados quaisquer $y_1, y_2 \in D$, temos que $(0, y_1) \in A$ e $(0, y_2) \in A$. Pela propriedade (5) de linearidade, temos $(-y_2, 0) \in A$ e, pela transitividade, obtemos $(-y_2, y_1) \in A$. Pela linearidade de A, temos $(0, y_1 + y_2) \in A$. Logo $y_1 + y_2 \in D$. Portanto, pela Proposição 2.1, D é convexo. Se $y_1 \in D \cap -D$, então $(0, y_1) \in A$ e $(0, -y_1) \in A$. Como A é linear, temos $(y_1, 0) \in A$. Como A é anti-simétrica, segue que $y_1 = 0$. Portanto, D é pontudo.
($\Longleftarrow$) Dado um cone convexo pontudo $D \in \mathbb{R}^l$, mostraremos que a relação binária A estabelecida na definição 1.1 é uma ordem parcial linear. A propriedade (5) segue de forma imediata da definição de D. Como $0 \in D$, A é reflexiva. Dados $y_1,\ y_2,\ y_3 \in D$, suponha que $(y_1, y_2) \in A$ e $(y_2, y_3) \in A$. Por (2.20), temos $y_2 - y_1 \in D$ e $y_3 - y_2 \in D$. Pela convexidade de D e pela Proposição 2.1, $y_3 - y_1 \in D$ e, portanto, $(y_1, y_3) \in A$. Logo, A é transitiva. Como $(y_1, y_2) \in A$ e $(y_2, y_1) \in A$ implica que $y_2 - y_1 \in D \cap -D$, podemos concluir que A é anti-simétrica pelo fato de que D é pontudo. □

Observação 2.2. *Denota-se $(y_1, y_2) \in A$ por $y_1 \leq_D y_2$, assim pode-se reescrever* (2.20) *por*

$$y_1 \leq_D y_2 \Leftrightarrow y_2 - y_1 \in D \tag{2.21}$$

Sejam y_1 *e* $y_2 \in Y$ dois vetores no espaço dos objetivos. Diz-se que y_1 domina y_2, denotado por $y_1 \leq_D y_2$, se $y_2 - y_1 \in D$ e $y_1 \neq y_2$, ou seja, $y_2 - y_1 \in D \backslash \{0\}$. Escreve-se também como $y_2 \in y_1 + D$ e $y_1 \neq y_2$, isto é, $y_2 \in y_1 + D \backslash \{0\}$.

Uma alternativa para a definição de otimalidade é feita considerando um cone convexo pontudo (MIETTINEN (1999)). Para essa definição utiliza-se o termo eficiência.

Definição 2.16. ***(Eficiência)*** *Sejam $D \subset \mathbb{R}^l$ um cone convexo pontudo e $Y \subset \mathbb{R}^l$ o espaço dos objetivos. Um vetor de decisão $\overline{x} \in X$ é eficiente (com respeito a D) para o problema* (1.1) *se não existe outro vetor de decisão $x \in X$ tal que $f(x) \leq_D f(\overline{x})$.*

Sejam $y, \overline{y} \in Y$ dois vetores no espaço dos objetivos. Uma vetor $\overline{y} \in Y$ é eficiente se não existe outro vetor $y \in Y$ tal que $y \leq_D \overline{y}$.

A definição 2.14 significa que um vetor é eficiente se ele não é dominado por qualquer outro vetor viável. Tal definição pode ser reformulada de várias formas equivalentes, por exemplo:

1. $\overline{y} \in Y$ é eficiente se $(Y - \overline{y}) \cap (-D) = \{0\}$ (PASCOLETTI E SERAFINI (1984) e WEIDNER (1988));

2. $(\overline{y} - D\backslash\{0\}) \cap Y = \emptyset$ (TAPIA e MURTAGH (1989));

3. $(Y - \overline{y}) \cap Y = \overline{y}$ (CHEN (1984) e JAHN (1987)).

Reformulando a definição 2.14 com uma representação equivalente obtém-se à:

Definição 2.17. *Seja D um cone convexo pontudo. Um vetor de decisão $\overline{x} \in X$ é eficiente (com relação a D) para o problema* (1.1) *se não existe outro vetor de decisão $x \in X$ tal que $f(\overline{x}) \in f(x) + D - \{0\}$, isto é, $(f(\overline{x}) - D\backslash\{0\}) \cap Y = \emptyset$.*

Definição 2.18. *Sejam y, $\overline{y} \in Y$ dois vetores no espaço de objetivos e $Y \subset \mathbb{R}^l$. Diz-que um vetor $\overline{y} \in Y$ é eficiente se não existe outro vetor $y \in Y$ tal que $\overline{y} \in y + D\backslash\{0\}$, ou seja, $(\overline{y} - D\backslash\{0\}) \cap Y = \emptyset$.*

Observação 2.3. *Essas definições são equivalentes a Pareto Otimalidade quando $D = R^l_+$.*

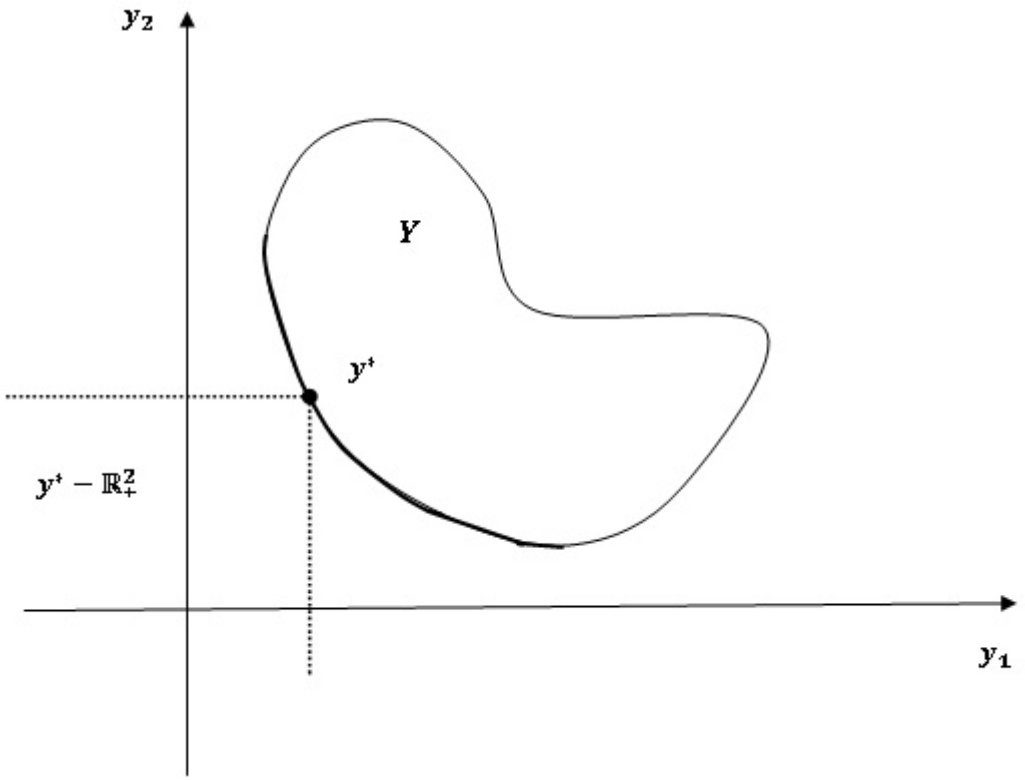


Figura 2.4: *Representação da fronteira eficiente com respeito a ordem parcial do cone R^2_+. Fonte: Miettinen*(1999).

Define-se as soluções fracamente eficientes utilizando a ordem parcial induzida pelo cone convexo pontudo D (JAHN (1987)) do seguinte modo:

Definição 2.19. ***(Solução fracamente eficiente induzida por um cone convexo pontudo)*** *Seja D um cone convexo pontudo. Um vetor de decisão $\overline{x} \in X$ é fracamente eficiente (com relação a D) para o problema* (1.1) *se não existe qualquer outro vetor de decisão $x \in X$ tal que $f(\overline{x}) \in f(x) + intD$($intD$ representa o interior do conjunto D), isto é, $(f(\overline{x}) - intD) \cap Y = \emptyset$.*

Definição 2.20. *Sejam* $y,\ \overline{y} \in Y$ *dois vetores no espaço dos objetivos e* $Y \in \mathbb{R}^l$*. Um vetor* $\overline{y} \in Y$ *é fracamente eficiente se não existe outro vetor* $y \in Y$ *tal que* $\overline{y} \in y + intD$*, ou seja,* $(\overline{y} - intD) \cap Y = \emptyset$.

Uma formulação alternativa segundo Sawaragi (1985) e Jiménez B, Novo V (2004) é que um vetor objetivo $\overline{y} \in Y$ é fracamente eficiente se $(Y - \overline{y}) \cap (-intD) = \emptyset$.

Capítulo 3

Alguns Métodos de Programação Multiobjetivo Matemática

De acordo com o estágio de decisão em que o decisor envolve-se no processo de Tomada de Decisão expressando sua preferência, Hwang e Masud (1979) apud Mavrotas (2006) afirmam que os métodos para resolver problemas de Programação Matemática Multiobjetivo (PMM) podem ser classificados em três categorias : *os métodos à priori, à posteriori e os interativos.* Embora os métodos *à priori* são os mais populares, os *métodos interativos e à posteriori* levam mais informações ao tomador de decisão, veja Mavrotas (2006). Especialmente, os métodos *à posteriori*(ou de geração) que tem como exemplo os algoritmos genéticos (que apresenta-se na seção 3.4) aos quais informam ao tomador de decisão sobre todo o contexto das alternativas antes da sua determinação final.

Segundo Miettinen (1999) apud Silva Neto (2011), matematicamente, toda solução eficiente (Pareto-ótima) é igualmente aceitável como solução do problema de programação multiobjetivo, lembrando que a fronteira Pareto ótima é formada por pontos indiferentes entre si. Na ausência de qualquer informação, nenhuma dessas soluções podem ser ditas melhores do que as outras. Então se faz necessário algum critério ou alguém que possa classificá-lo e/ou escolher alguma ou algumas soluções aceitáveis. Nesse momento insere-se a figura do tomador de decisão que é uma pessoa ou um grupo de pessoas que supostamente tem experiência a respeito do processo ao qual o (PMM) está inserido. Para o decisor é interessante selecionar somente uma solução, ou no máximo um número finito(pequeno) de soluções.

Os métodos de otimização multiobjetivo usualmente estão interessados na busca de soluções no espaço dos objetivos ao invés do espaço de decisões. Segundo Miettinen (1999), uma das razões se deve ao fato que a dimensão do espaço dos objetivos é geralmente menor do que o espaço de decisão. Outra razão, conforme Silva Neto (2011), seria que os decisores tem maior interesse nos valores da função objetivo.

3.1 Métodos à Priori

Os métodos pertencentes a essa classe são caracterizados pela aplicação de informações sobre prioridades, ordenação ou intervalos aceitáveis para os objetivos, fornecidos pelo decisor, como parâmetros do problema a ser resolvido. As técnicas mais comuns transformam o problema original multiobjetivo em problemas com formatos mais simples (problemas escalares), através da inserção das preferências do decisor na definição matemática do problema. Tem como vantagem um baixo custo computacional. No entanto, caso haja mudança de ideia do decisor, o método tem que ser executado novamente.

3.1.1 Métodos baseados em Limitantes

O decisor deve selecionar um objetivo de referência f_i, $i \in \{1, ..., l\}$, e indicar limitantes inferior e superior (l_j e L_j, respectivamente) para os demais objetivos.

Deve-se então resolver o problema modificado, para cada i :

$$\begin{aligned} \min \quad & f_i(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \\ & l_j(x) \leq f_j(x) \leq L_j, \ \forall \ j \neq i \end{aligned} \tag{3.1}$$

Sua desvantagem se deve à dificuldade de especificar à priori os limitantes para que tenha um problema viável e se tenha uma solução satisfatória do problema para o decisor.

3.1.2 Método Lexicográfico

O decisor deve listar os objetivos em ordem decrescente de preferência ou importância. Assim as funções objetivos são classificada nessa ordem de preferência. Ou seja, $f_1(x)$ denota a função de maior importância e $f_l(x)$ denota a função de menor importância. Para obter a solução Pareto ótima $\overline{x}$ realiza-se minimizações sucessivas como segue:

Considera-se primeiramente a função de maior importância somente, minimiza-se sobre o conjunto viável X do problema multiobjetivo definido em (1.1) sem considerar os demais objetivos. Isto é,

$$\begin{aligned} \min \quad & f_1(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \end{aligned} \tag{3.2}$$

encontrando o valor ótimo f_1^*. Em seguida, minimiza-se a função $f_2(x)$ sobre o mesmo conjunto viável X adicionando uma nova restrição $f_1(x) = f_1^*$, tornando-se o seguinte problema:

$$\begin{aligned} \min \quad & f_2(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \\ & f_1(x) = f_1^* \end{aligned} \tag{3.3}$$

obtendo o valor ótimo f_2^*. Depois, é feito semelhante para $f_3(x)$ adicionando as restrições $f_1(x) = f_1^*$ e $f_2(x) = f_2^*$, isto é:

$$\begin{aligned} \min \quad & f_2(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \\ & f_1(x) = f_1^* \\ & f_2(x) = f_2^* \end{aligned} \tag{3.4}$$

Para o i-ésimo objetivo, a expressão anterior ficará:

$$\begin{aligned} \min \quad & f_i(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \\ & f_j(x) = f_j^*,\ j = 1, 2, ..., i-1 \end{aligned} \tag{3.5}$$

Continua-se o processo até a função de menor importância. Depois desse processo, o $\overline{x}$ tal que $f_l(\overline{x}) = f_l^*$ é a solução do problema lexicográfico. O próximo teorema encontrado em Miettinen (1999) mostra que toda solução do problema lexicográfico é Pareto-ótimo.

Teorema 3.1. *A solução encontrada no método lexicográfico é Pareto-ótimo.*

Demonstração. Ver Miettinen (1999) □

3.1.3 Programação por Metas

Em problemas práticos, muitas vezes é difícil e até impossível atingir os objetivos e ainda satisfazer uma série de restrições. Nesses casos, é razoável que os gestores procurem satisfazer ou aproximar-se dos objetivos e restrições estabelecidos, ao invés de considerá-los rígidos.

A programação por metas é uma técnica que permite a modelagem e a busca de soluções para problemas com múltiplos objetivos ou metas a serem otimizadas (DEB (2001)). Esse método foi introduzido por Charnes e Cooper em 1961, a idéia desse método é determinar níveis de aspiração pelo tomador de decisão, representados por M_i, para as funções objetivos através de desvios que serão minimizados. As metas são definidas pela junção das funções

objetivos com essas aspirações. Miettinen (1999) afirma que as restrições, que geralmente são funcionais, podem ser observadas como subconjuntos das metas.

Com os valores fornecidos para as metas as variáveis de desvio d_i podem ser formuladas da seguinte forma:

$$d_i = M_i - f_i(x) \tag{3.6}$$

em que $d_i = d_i^- - d_i^+$ e $d_i^-, d_i^+ > 0$. A expressão anterior resultará em $f_i(x) + d_i^- - d_i^+ = M_i,\ \forall i = 1, ..., l$.

A partir das definições das metas, o problema multiobjetivo torna-se um problema monobjetivo onde se deseja minimizar os desvios d_i. É comum formular este método, chamado programação por metas ponderada, como:

$$\begin{aligned} \min \quad & \left(\sum_{i=1}^{l} w_i (d_i^+ + d_i^-) \right) \\ s.a. \quad & f_i(x) + d_i^- - d_i^+ = M_i \\ & d_i^-, d_i^+ \geq 0 \\ & d_i^+ d_i^- = 0 \ \forall i = 1, ..., l \\ & x \in X \end{aligned} \tag{3.7}$$

Note que w_i e M_i é o ponderador e o nível de aspiração (ou valor desejado) para o objetivo de índice i, respectivamente. O d_i^+ indica quanto o i-ésimo objetivo excedeu M_i e d_i^- indica quanto o i-ésimo objetivo ficou abaixo de M_i.

Vale ressaltar que não pode ocorrer simultaneamente $d_i^+ > 0$ e $d_i^- > 0$. Por isso, a restrição $d_i^+ d_i^- = 0$ sempre é satisfeita, podendo (3.7) ser reescrito como:

$$\begin{aligned} \min \quad & \left(\sum_{i=1}^{l} w_i (d_i^+ + d_i^-) \right) \\ s.a. \quad & f_i(x) + d_i^- - d_i^+ = M_i \\ & d_i^-, d_i^+ \geq 0 \forall i = 1, ..., l \\ & x \in X \end{aligned} \tag{3.8}$$

O problema (3.8) é um caso particular com $p = 1$ da formulação feita através da métrica $L_p.$, conforme visto em Miettinen (1999). Com isso a programação por metas ponderada é relacionada com o método da métrica ponderada ou programação por compromisso.

$$\begin{aligned} \min \quad & \left(\sum_{i=1}^{l} w_i (d_i^+ + d_i^-)^p \right)^{\frac{1}{p}} ; 1 \leq p \leq \infty \\ s.a. \quad & f_i(x) + d_i^- - d_i^+ = M_i \\ & d_i^-, d_i^+ \geq 0 \forall i = 1, ..., l \\ & x \in X \end{aligned} \tag{3.9}$$

A métrica L_1 é a mais utilizada segundo Miettinen (1999) pelo fato da origem da programação por metas vir do estudo da programação linear.

Em Romero (2004), são descritas outras formas tradicionais de programação por metas: programação por metas lexicográfica e programação por metas minmax, também conhecida como programação por metas nebulosa.

Teorema 3.2. *A solução de um problema de programação por metas ponderadas é Pareto-ótimo para Problema* (1.1) *se as variáveis de desvio* d_i^+, d_i^- *são soluções ótimas.*

Demonstração. Ver Miettinen (1999) □

3.2 Métodos à posteriori

Nos métodos a posteriori (ou métodos de geração) as soluções eficientes do problema (todas ou uma representação suficiente delas) são geradas e então o tomador de decisão envolve-se, para selecionar dentre elas, a solução de maior preferência.

3.2.1 Método das ponderações

Esse é um dos métodos mais conhecidos e foi apresentado por Zadeh (1963), a ideia é associar a cada função objetivo um coeficiente de pesos e minimizar a soma ponderada de todas as funções coordenadas f_i. Esse método é formulado da seguinte maneira:

$$\begin{aligned} \min \quad & \sum_{i=1}^{l} w_i f_i(x) \\ s.a. \quad & x \in X \end{aligned} \tag{3.10}$$

onde $w_i \geq 0, \forall i = 1, ..., l$. O peso w_i representa a importância da função objetivo i comparado com as demais. Geralmente, é feita a normalização desses pesos, ou seja,

$$\sum_{i=1}^{l} w_i = 1 \tag{3.11}$$

tornando a função objetivo de (3.11) uma combinação convexa das funções objetivos do problema multiobjetivo.

O teorema a seguir encontrado em Takahashi (2007) apresenta condições para obtenção de soluções Pareto ótimas através do problema ponderado.

Teorema 3.3. *Seja $\overline{x}$ a solução do problema* (3.11) *para $w \geq 0$. Então $\overline{x}$ é uma solução Pareto ótima se:*

(i) $\overline{x}$ é solução única de (3.11) *ou*

(ii) $w_i > 0$, $\forall i = 1, ...l$

Segundo Takahashi (2007) a vantagem desse método é a simplicidade computacional. Equivale a soluções de problemas mono-objetivos, um para cada ponto gerado e cada problema monobjetivo possui a mesma estrutura de restrições do problema original. A desvantagem desse método é que ele não se aplica a problemas não convexos, nos quais talvez não seja possível gerar várias soluções Pareto-ótimas e várias ponderações podem estar associadas à mesma solução Pareto-ótima. Sendo este método recomendado, por exemplo, para problemas de programação linear multiobjetivo.

3.2.2 Método das ϵ - restrições

Introduzido por Haimes (1971) apud Miettinen (1999). Nesse método otimiza-se uma das funções objetivos f_i, julgada a mais relevante, e todas as funções restantes são convertidas em restrições através de definição de limites superiores para cada uma delas (ϵ_j). O método é formulado pelo problema:

$$\begin{aligned} \min \quad & f_i(x) \\ s.a. \quad & x \in X \subset \mathbb{R}^n \\ & f_j(x) \leq \epsilon_j, \ \forall \ j = 1, ..., l; \ j \neq i \end{aligned} \tag{3.12}$$

em que $i \in \{1, ..., l\}$.

Arroyo (2002) apud Silva Neto (2011) comenta que variando convenientemente os limitantes ϵ_j, torna-se possível gerar o conjunto Pareto- ótimo, mesmo quando o espaço de objetivos é não convexo. Segundo Takahashi (2007) a desvantagem desse método é justamente a determinação dos valores de ϵ_j, pois o ponto obtido pode não ser Pareto-ótimo e os valores de ϵ_j podem tornar o problema (3.3) infactível. Como implicação prática do próximo resultado conclui-se que, variando parametricamente ϵ_j, $\forall i$ e $\forall j \neq i$, é possível gerar completamente o conjunto Pareto-ótimo.

Teorema 3.4. *Se $\overline{x}$ é Pareto-ótimo então existem um inteiro $i \in \{1, ..., l\}$ e números reais ϵ_j, $j = 1, ..., l$ $(j \neq i)$ tais que $\overline{x}$ resolve o problema* (3.13).

Demonstração. Ver Takahashi (2007). □

3.3 Métodos Interativos

Nesse método o decisor é consultado enquanto é executado o algoritmo de busca de soluções do problema multiobjetivo e suas preferências são utilizadas na busca da solução (SILVA NETO (2011)). A busca só termina quando o tomador de decisão se satisfaz com a solução.

3.3.1 Geoffrion, Dyer e Feinberg

O método GDF(Geoffrion, Dyer e Feinberg) baseia-se no Algoritmo de Frank-Wolfe, que foi desenvolvido para resolver problemas de programação quadrática mono-objetivo com restrições lineares, ou seja, problema de otimização convexa. Segundo Ferreira (1999), recorre-se a uma função $u = U(f(x)),\ u : \mathbb{R}^n \to \mathbb{R}$ que não precisa ser explicitamente conhecida, apenas localmente. O problema a ser resolvido é

$$\begin{aligned} \min \quad & u = U(f(x)) \\ s.a. \quad & x \in X \end{aligned} \tag{3.13}$$

em que u é uma função convexa crescente diferenciável que reflete as preferências gerais do decisor.

O problema (3.13) é convexo para a utilização do método GDF porque vem do uso do Algoritmo de Frank-Wolfe. Miettinem (1999) apud Silva Neto (2011) afirma que o método de Geoffrion, Dyer e Feinberg também é chamado de método de Frank-Wolfe interativo.

Conforme encontrado em Ferreira (1999) apud Silva Neto (2011), o algoritmo GDF é similar ao algoritmo de Frank-Wolfe sendo que os passos 2, 3 *e* 4 são feitos interativamente consultando o decisor:

- 1- Determinar uma solução $x^0 \in X$; fazer $k = 0$;
- 2- Obter direção de busca d^k resolvendo o seguinte problema:

$$\begin{aligned} \min \quad & \nabla u(x^k)^t \xi \\ s.a. \quad & \xi \in X \end{aligned} \tag{3.14}$$

  Seja ξ^k uma solução do problema; fazer $d^k := \xi^k - x^k$;

- 3- Determinar passo ótimo α_k através de busca unidimensional

$$\min_{\alpha \in [0,1]} \quad u(x^k + \alpha d^k) \tag{3.15}$$

- 4- Calcular $x^{k+1} = x^k + \alpha_k d^k$; se $||x^{k+1} - x^k|| < \epsilon,\ \epsilon > 0$ suficientemente pequeno, então parar. A solução é x^{k+1}. Caso contrário, $k = k + 1$ e retornar ao passo 2.

Para maiores informações sobre o método Geoffrion, Dyer e Feinberg podem ser encontrados em Miettinen (1999), Ferreira (1999).

3.4 Algoritmos Genéticos

O algoritmo genético (*Genetic Algorithm*-GA) foi descrito, originalmente, por Holland (1975) com o intuito de reproduzir e estudar o processo adaptativo dos sistemas naturais. Os algoritmos genéticos são métodos heurísticos, usualmente aplicados em problemas complexos, esses métodos imitam a evolução natural, ou seja, o GA trabalha em uma população de soluções potenciais aplicando o princípio da sobrevivência dos mais adaptados a produzir melhores aproximações para uma solução.

Por ser um algoritmo evolutivo, sua estrutura é baseada nos conceitos de reprodução celular e evolução natural, podendo ser descrito por quatro elementos básicos: uma *população de soluções*, geralmente viáveis, representada através de uma estrutura computacional adequada; uma função de *fitness* que associa a cada indivíduo da população um valor da medida de desempenho na resolução do problema; *operadores genéticos (crossover e mutação),* que realizam a troca de informações entre as soluções, introduzem ou modificam indivíduos e mantêm a diversidade da população; e *processo de seleção* de indivíduos que, em analogia à seleção natural, promove a sobrevivência e reprodução das melhores soluções, acarretando um aumento do *fitness* médio da população ao longo das gerações. A estrutura básica desse tipo de algoritmo pode ser resumida da seguinte maneira.

- Passo 1: Gerar população inicial;
- Passo 2: Avaliar população;
- Passo 3: Aplicar crossover e mutação;
- Passo 4: Selecionar indivíduos para nova geração;
- Retornar ao Passo 2 até que uma condição de parada seja alcançada.

Cada indivíduo da população (cromossomo), codificado através de estruturas computacionais (grafos, vetores, etc) representa uma possível solução para o problema tratado. Na versão clássica do algoritmo em Holland (1975), os indivíduos são representados através de vetores binários de tamanho fixo.

A escolha da codificação dos cromossomos a ser adotada é uma etapa bastante importante na construção do algoritmo genético, já que deve permitir a representação da solução completa do problema, sem dificultar a medida do *fitness* ou a aplicação dos *operadores genéticos.*

De modo geral, inicia-se a população de forma aleatória mantendo tamanho fixo. Informações e restrições do problema podem ser incorporados na geração da população inicial de forma a viabilizar as soluções ou melhorar a qualidade da população.

A função de *fitness* deve ser escolhida de forma a medir o desempenho de cada indivíduo como solução do problema. Na maioria dos casos, essa medida não é absoluta e sim, relativa aos outros indivíduos da geração atual.

Conforme o autor Michalewicz (1995) nos problemas restritos, a função de fitness pode incorporar a medida de inviabilidade do indivíduo, na intenção de que os indivíduos inviáveis convirjam para a viabilidade e indivíduos viáveis convirjam para soluções com melhores *fitness.*

O *crossover* e a *mutação* são os operadores genéticos mais utilizados. O *crossover* é o responsável pela geração de novos indivíduos através da troca de informações entre duas ou mais soluções da população. A *mutação* é responsável por modificações aleatórias em um cromossomo, a fim de promover a variabilidade da população, sem comprometer a convergência já alcançada. Os operadores *crossover* e *mutação* são aplicados com probabilidades determinadas na definição do algoritmo.

Os mais comuns operadores *crossover* são o algoritmo genético clássico *crossover simples*, em que dois indivíduos pais selecionados geram dois novos indivíduos através da troca do segmento de cromossomo iniciado em um ponto de corte escolhido aleatoriamente, o *crossover uniforme* (SYSWERDA (1989)). Nesse uma quantidade determinada de genes aleatórios (unidade de informação do cromossomo - um bit na codificação binária) é trocada entre dois indivíduos pais e o *crossover aritmético* (MICHALEWICZ (1995)) utiliza-se em populações com codificação em ponto flutuante, que gera dois indivíduos a partir de combinações convexas das soluções pais.

Para problemas com codificação binária, a *mutação simples* troca o valor de um ou mais genes em um cromossomo selecionado. Em algoritmos com codificação real pode-se utilizar por exemplo: *mutação uniforme*, em que o valor do gene a ser modificado é selecionado aleatoriamente ou de acordo com uma distribuição de probabilidade adequada, e o operador de *mutação indutiva* o valor selecionado é somado ao valor atual do gene.

No algoritmo genético clássico, as soluções são selecionadas para a próxima geração através de uma técnica denominada *roulette wheel,* que atribui a cada indivíduo da população uma probabilidade de seleção proporcional a seu valor de *fitness.* A grande desvantagem desse esquema é a possibilidade de perda melhor indivíduo da população atual na montagem da nova geração. Essa dificuldade pode ser superada com uma estratégia etilista, que consiste em manter o melhor indivíduo ou um grupo dos melhores da geração atual na seguinte.

Tem-se também as técnicas de *seleção por diversidade* onde os indivíduos mais diversos são selecionados e *seleção por torneio*, em que duas soluções são indicadas aleatoriamente e a melhor é selecionada.

Essas técnicas de seleções apresentadas também podem ser utilizadas na identificação de indivíduos para aplicação dos operadores genéticos.

3.4.1 Algoritmos Genéticos aplicados a Otimização Multiobjetivo

A solução de um problema de otimização multiobjetivo é constituída por um conjunto de pontos, a utilização de algoritmos genéticos parece imediata, dado que a cada geração no algoritmo considera-se simultaneamente uma população de soluções (MOURA (2002)).

Em problemas de otimização multiobjetivo, os algoritmos genéticos podem ser aplicados como métodos *à posteriori*, ou seja, direcionados à geração de pontos que caracterizam o conjunto Pareto-ótimo, para possibilitar ao decisor a escolha de uma solução. Essa caracterização deve ser tal que os pontos gerados apresentem distribuição aproximadamente uniforme no conjunto.

Autores como Deb (2001), Coello (2006) destacam o emprego dos algoritmos genéticos em problemas de otimização multiobjetivo.

- Apresentam facilidade e flexibilidade da modelagem;
- Exigem um menor esforço computacional para encontrar cada solução do conjunto Pareto-ótimo se comparado ao esforço de encontrar uma solução para uma formulação mono-objetivo;
- São menos susceptíveis as características da fronteira de Pareto (não convexa, descontínua);

Os algoritmos genéticos de otimização multiobjetivo mais importantes, desenvolvidos entre 1989 e 1998 (AZUMA (2011)) são:

- NSGA (*Nondominated Sorting Genetic Algorithm*): proposto em Srinivas & Deb (1994), é baseado em vários níveis de classificação dos indivíduos. Todos os indivíduos não-dominados são classificados em uma categoria. Esse grupo de indivíduos classificados é removido e outro nível de indivíduos não-dominados é considerado. O processo continua até que todos os indivíduos da população sejam classificados. Esse algoritmo não é muito eficiente porque requer um número elevado de repetições de classificação.
- NPGA (*Niched-Pareto Genetic Algorithm*): proposto em Horn et al. (1994), usa seleção por torneio baseado na relação de dominância. Dois indivíduos são escolhidos aleatoriamente e comparados com um subconjunto da população (tipicamente 10%). Se um deles é dominado por esse subconjunto e outro não, então o outro não dominado ganha. No caso dos dois serem dominados ou dois dominarem, o resultado do torneio é decidido através de compartilhamento de *fitness.* O objetivo de compartilhamento de *fitness* é distribuir a população ao longo dos picos do espaço de decisão. Cada pico recebe uma parte proporcional a altura do pico.
- MOGA (*Multi-Objective Genetic Algorithm*): proposto em Fonseca & Fleming (1993), a classificação de um indivíduo corresponde ao número de indivíduos na população

corrente que o dominam. A todos os indivíduos não dominados é atribuído o maior valor possível de *fitness*, enquanto que os dominados são penalizados de acordo com a densidade da população na região em que se encontram.

Esses algoritmos representam a primeira geração de algoritmos genéticos para otimização multiobjetivo. A segunda geração começou quando o elitismo passou a ser considerado. Tal elitismo geralmente refere-se ao uso de uma população externa para manter os indivíduos não dominados, encontrados ao longo do processo evolutivo. A motivação surgiu do fato de uma solução que não é dominada por soluções candidatas de uma população corrente não necessariamente ser não dominada em relação a todas as soluções candidatas já produzidas pelo algoritmo (AZUMA (2011)). Os principais são:

- SPEA (*Streng Pareto Evolutionary Algorithm*): proposto em Zitzler & Thiele (1998) faz uso de uma população externa, atualizada a cada geração, de forma a conter as soluções não dominadas em relação aos indivíduos gerados até o momento. Como o algoritmo faz uso de uma estratégia de agrupamento, os pontos da população externa apresentam distribuição aproximadamente uniforme. Ao final, a população externa é utilizada como conjunto Pareto-ótimo. A função de *fitness* é associada à medida de dominação de indivíduos da população externa em relação a cada solução da geração atual. Na intenção de manter a diversidade, o algoritmo SPEA busca distribuir uniformemente a população convencional em torno dos pontos da população externa, de forma que cada solução da segunda domine aproximadamente o mesmo número de indivíduos da primeira.

- PAES (*Pareto Archived Evolution Strategy*): proposto em Horn et al. (1994), consiste em uma estratégia evolutiva $(1 + 1)$ (BEYER (2001)), ou seja, um pai gera um filho, que emprega busca local e usa um arquivo externo que armazena as soluções não dominadas encontradas até o momento. Esse arquivo é usado como referência para obter uma classificação de dominância aproximada para os candidatos que sofreram mutação.

- NSGA-II (*Nondominated Sorting Genetic Algorithm II*): proposto em Deb et al. (2002) como uma melhoria do NSGA. Ele estima a densidade das soluções vizinhas a uma solução da população através do cálculo da distância de agrupamento(*crowding distance*). A distância de agrupamento de uma determinada solução i é uma estimativa do tamanho do maior cubóide que engloba a solução i sem incluir uma ou mais dentre as soluções da população. Essa estimativa de tamanho usa a média de compromisso dos lados. Esse método não usa uma população ou arquivo externo, sendo que o mecanismo de elitismo consiste em combinar os melhores pais com os melhores filhos obtidos. Esse algoritmo é muito mais eficiente do que o seu predecessor e é considerado um algoritmo de referência.

Para boas referências de utilização de algoritmos genéticos na resolução de problemas multiobjetivo veja Coello (1999), Deb (2001), Fonseca & Fleming (1993).

3.5 Métricas de Desempenho

A comparação de desempenho de um ou vários métodos de otimização multiobjetivo não é uma tarefa fácil. Duas metas da otimização multiobjetivo são a convergência e a diversidade das soluções encontradas (AZUMA (2011)). Portanto, são necessárias pelo menos duas métricas para avaliar ambos os algoritmos, uma métrica para avaliar a convergência e outra para avaliar a diversidade das soluções encontradas. Nessa seção discuti-se algumas dessas metodologias adotadas para comparar o desempenho entre os métodos de otimização multiobjetivo.

3.5.1 Métricas de convergência

As métricas de convergência podem comparar soluções P encontradas com o conjunto de soluções Pareto-ótimo, denominado P'. É possível também utilizar essas métricas para comparar (de forma relativa) soluções P_1 produzidas por um algoritmo com soluções P_2 produzidas por outro algoritmo (AZUMA (2011)).

1. ***Taxa de Erro***.
   Calcula o número de soluções P que não estão em P':

$$ER(Error\ Ratio) \ = \ \frac{|\{i \in P \cap P'\}|}{|P|} \tag{3.16}$$

   onde $|.|$ indica a cardinalidade de um conjunto de elementos. Quanto menor for o valor de ER, melhor será a convergência. Se ER=0, significa que $P \subseteq P'$. Aplica-se somente em problemas de otimização discreta.

2. ***Distância Geracional***.
   Representa a distância euclidiana média (no espaço dos objetivos) entre as soluções P e P':

$$GD(Generational\ Distance) \ = \ \frac{\sum_{i \in P}(mindist_i)^{1/2}}{|P|} \tag{3.17}$$

   onde $mindist_i$ representa a distância euclidiana no espaço dos objetivos entre a solução i e a solução mais próxima dela no conjunto P':

$$mindist_i = min_{j \in P'} \sqrt{\sum_{n=1}^{l_{obj}} [f_l(i) - f_l(j)]^2} \tag{3.18}$$

   Quanto mais próximo de zero for o valor de GD, melhor será a convergência de P. Pode ser aplicado a problemas de otimização discreta e contínua.

3. ***Métrica de Cobertura***.
   Dado dois conjuntos de soluções P e Q, essa métrica calcula a proporção de soluções de Q que são fracamente dominadas pelas soluções de P:

$$SC(P,Q)(Set\ Coverage) \ = \ \frac{|\{i \in P \mid \exists j \in Q \ e \ i \leq j\}|}{|Q|} \tag{3.19}$$

   Quando SC(P,Q) = 1 todas as soluções de Q são dominadas por soluções de P. Se SC(P,Q)=0, então nenhuma solução de Q é fracamente dominada pelas soluções de P.

3.5.2 Métricas de Diversidade

As métricas de diversidade calculam a distribuição das soluções de um conjunto P.

1. ***Espaçamento***.
   Calcula o desvio padrão entre as distâncias de soluções consecutivas (no espaço dos objetivos) do conjunto P:

$$SP(Spacing) \ = \ \sqrt{\frac{1}{|P-1|}\sum_{1 \in P}[neardist_i - \overline{neardist}]^2} \tag{3.20}$$

   onde $neardist_i$ representa a distância euclidiana (no espaço dos objetivos) entre a solução i e a solução mais próxima dela no conjunto P, sendo expresso como:

$$neardist_i \ = \ \min_{j \in P, j \neq i} \sqrt{\sum_{n=1}^{l_{obj}}[f_l(i) - f_l(j)]^2} \tag{3.21}$$

   O valor $\overline{neardist}$ é a média dos valores $neardist_i$:

$$\overline{neardist} \ = \ \frac{\sum_{i \in P} neardist_i}{|P|} \tag{3.22}$$

   Quanto menor for o valor da métrica SP, melhor distribuídas estão as soluções do conjunto P.

2. ***Número de Nichos***.
   Essa métrica calcula o número de nichos dentro de um conjunto de soluções P:

$$NC(Niche\ Count) \ = \ \frac{1}{|P|-1}\sum_{i,j \in P} |\{j \ ; \ dist_{i,j} > \sigma\}| \tag{3.23}$$

   onde $dist_{i,j}$ é a distância entre as soluções i e j do conjunto P. O valor NC representa o número de soluções cuja distância entre elas é maior que o parâmetro σ. O valor de σ é atualizado dinamicamente. Quando $dist_{i,j} < \sigma$, as soluções i e j estão no mesmo

nicho. Quanto maior a quantidade de nichos formados em P, melhor distribuídas estão as soluções.

3. ***Espalhamento***.
   Avalia a dispersão das soluções no conjunto P ao longo da fronteira de Pareto, assim como a distribuição entre soluções contíguas (no espaço dos objetivos) de P:

$$SPREAD \;=\; \frac{\sum_{n=1}^{l_{obj}} extdist_l + \sum_{i \in P} |neardist_i - \overline{neardist}|}{\sum_{n=1}^{l_{obj}} extdist_l + |P| \times \overline{neardist}} \tag{3.24}$$

   onde $extdist_l$ representa a distância euclidiana entre as soluções extremas na l-ésima função objetivo dos conjuntos P e P':

$$extdist_l \;=\; dist_{i,j} \; tal \; que \; i = \min_{k \in P} f_l(k) \; e \; j = \min_{k \in P'} f_l(k) \tag{3.25}$$

   no caso de minimização da função f_l. O valor ideal para SPREAD é 0.

Atualmente, uma métrica bastante usada na avaliação de algoritmos multiobjetivos é o indicador de hiper volume (AZUMA (2011)). Calcula-se o volume da região coberta entre os pontos das soluções do conjunto P encontradas e um ponto de referência. Para cada solução i pertencente a P, é construído um hipercubo v_i com referência a um ponto W. O ponto de referência pode ser encontrado construindo-se um vetor com os piores valores da função objetivo. O resultado da métrica é a união de todos os hipercubos encontrados. Quanto maior o valor do hiper volume melhor porque um alto valor de hiper volume indica que houve um elevado espalhamento entre as soluções de P e indica também que houve uma melhor convergência. Essa métrica é calculada através de:

$$HV(Hipervolume) \;=\; \sum_{i \in P} v_i \tag{3.26}$$

onde v_i representa o hipercubo da solução i.

Capítulo 4

Condições de Otimalidade em Otimização Multiobjetivo via Teoria dos Cones

O desenvolvimento de condições de otimalidade via teoria dos cones é, de certo modo, inovador. Este desenvolvimento é baseado nas noções de direções tangentes e de cone tangente do conjunto viável do problema. Esta abordagem torna-se mais natural e permite ao leitor uma melhor intuição geométrica, facilitando assim um entendimento correto do assunto. Além do mais, as condições de otimalidade baseadas no cone tangente permitem um desenvolvimento unificado e natural para vários formatos de conjunto viável do problema a resolver(descrição abstrata, restrições de igualdade, de desigualdade, restrições mistas) e para várias condições de qualificação e de regularidade das restrições (IZMAILOV E SOLODOV (2009)).

Nesse capítulo, estabelece-se condições de otimalidade utilizando cones tangentes, conjuntos tangentes de segunda ordem e o cone assintótico de segunda ordem, ambos no caso geral e também no caso diferenciável. Depois o conjunto viável do problema (1.1) é definido por restrições mistas e com condições de qualificação e regularidade das restrições obtém-se os multiplicadores de Lagrange. Primeiramente, fornece-se condições necessárias e suficientes no espaço dos objetivos(caso geral) que nos serve para deduzir todo o resto.

4.1 Condições Necessárias no caso geral

Teorema 4.1. *Se $\overline{y}$ é um vetor Pareto-ótimo local(fraco) de $Y \subset \mathbb{R}^l$, então para todo $k \in \{0,1\}$, a seguinte condição vale:*

$$T_k''(Y, \overline{y}) \cap \Omega = \emptyset, \tag{4.1}$$

onde

$$\Omega = \{(d, z) \in \mathbb{R}^l \times \mathbb{R}^l \mid (d_i, z_i)^t <_{lex} (0,0)^t, \forall i\} \tag{4.2}$$

Demonstração. Suponha por absurdo que (4.1) não vale para $k \in \{0, 1\}$, isto é, existem $(t_n, r_n) \to (0^+, 0^+)$ e $z_n \to z$ tal que $\frac{t_n}{r_n} \to k$ e

$$y_n = \overline{y} + t_n d + \frac{1}{2} r_n t_n z_n \in Y \tag{4.3}$$

Se $r_n t_n (z_n)_i = o(t_n d_i)$ e $d_i < 0$ então $(y_n)_i - (\overline{y})_i < 0$ para $n \in \mathbb{N}$ suficientemente grande. Se $d_i = 0$, temos que $(z_n)_i < 0$. Então, para $n \in \mathbb{N}$ suficientemente grande,

$$(y_n)_i - (\overline{y})_i = \frac{1}{2} r_n t_n (z_n)_i < 0.$$

Consequentemente $y_n < \overline{y}$ para $n \in \mathbb{N}$ suficientemente grande, o que contradiz a hipótese do teorema. □

Observação 4.1. *Como foi mostrado na prova do Teorema* 4.1*, a condição* (4.1) *é válida para todo número real k. Entretanto, nada de especial é perdido quando k está restrito pela condição* $k \in \{0, 1\}$*. De fato, se* $k < 0$*, então* $T_k''(Z, \overline{y}) = \emptyset$ *e* (4.1) *vale trivialmente; se* $k > 0$*, temos a seguinte equivalência:*

$$T_1''(Z, \overline{y}) \cap \Omega = \emptyset \iff (1/k) T_1''(Z, \overline{y}) \cap (1/k)\Omega = \emptyset \iff T_k''(Z, \overline{y}) \cap \Omega = \emptyset,$$

em que Ω*(conjunto definido em* (4.2)*) é um cone e* $(1/k) T_1''(Z, \overline{y}) = T_k''(Z, \overline{y})$ *pela proposição* 2.2*.*

Exemplo 4.1. *Consideremos o conjunto Z usado por Cambini* (1997)*:*

$$Z = \{(x, y) \in \mathbb{R}^2 \mid y = x\sqrt{x}, x \geqq 0\}, \text{ onde } \overline{y} = (0, 0).$$

Como foi visto no exemplo 2.4 :

$$T_1''(Z, \overline{y}) = T^2(Z, \overline{y}) = \{(0, 0)\} \times \{(z_1, z_2) \in \mathbb{R}^2 | z_2 = 0\}$$

e

$$T_0''(Z, \overline{y}) = T''(Z, \overline{y}) = \{(d, z) \in \mathbb{R}^4 | d_1 > 0,\ d_2 = 0 \text{ e } d_1.z_2 \geqq 0\} \cup T_1''(Z, \overline{y})$$

Portanto, a condição (4.1) *é satisfeta para* $k \in \{0, 1\}$*, o que implica* $\overline{y}$ *é um vetor Pareto-ótimo de Z.*

Corolário 4.1. *Se* $\overline{y}$ *é um vetor Pareto-ótimo local(fraco) de* $Y \subset \mathbb{R}^l$*, então para toda*

direção $d \in T(Y,\overline{y}) \cap Fr(\mathbb{R}^l_-)$*(fronteira de* $\mathbb{R}^l_-$*) a seguinte condição vale:*

$$T''_k(Y,\overline{y})(d) \cap \Omega(d) = \emptyset, \quad k \in \{0,1\} \tag{4.4}$$

onde

$$\Omega(d) = \{z \in \mathbb{R}^l \mid (d,z) \in \Omega\} \tag{4.5}$$

em que Ω *é o conjunto definido em* (4.2)

Demonstração. Segue imediatamente do Teorema 4.1 □

4.2 Condição Suficiente no caso geral

Teorema 4.2. *Uma condição suficiente para que um ponto* $\overline{y} \in Y \subset \mathbb{R}^l$ *seja um ponto eficiente local de* Y *é que:*

$$T(Y,\overline{y}) \cap (\mathbb{R}^l_-) \subseteq Fr(\mathbb{R}^l_-) \tag{4.6}$$

onde $Fr(\mathbb{R}^l_-$ *denota a fronteira de* $\mathbb{R}^l_-$

e para qualquer $d \in H = T(Y,\overline{y}) \cap Fr(\mathbb{R}^l_-) - \{0\}$*, se* $H \neq \emptyset$*, obtemos*

$$T''_0(Y,\overline{y})(d) \cap d^{\perp} \cap \overline{\Omega(d)} = \{0\}, \tag{4.7}$$
$$T''_1(Y,\overline{y})(d) \cap d^{\perp} \cap \overline{\Omega(d)} = \emptyset, \tag{4.8}$$

onde $d^{\perp}$ *denota o subespaço ortogonal de* d *e* $\overline{\Omega(d)}$ *é o fecho de* $\Omega(d)$ (4.3)*.*

Demonstração. Seja $T(Y,\overline{y}) \cap Fr(\mathbb{R}^l_-) = \{0\}$. Então $\overline{y}$ é um vetor Pareto-ótimo de Y pelo teorema 5.2 em Lin (1976). Agora, assumi-se que $T(Z,\overline{y}) \cap (\mathbb{R}^l_-)$ possui mais elementos do que $\{0\}$ e que as condições (4.7) e (4.8) valem, mas $\overline{y}$ não é um vetor Pareto-ótimo de Y. Pela definição 1.5, existe uma sequência $\{y_n\} \subset Y$ tal que

$$y_n \leq \overline{y}, \quad y_n \to \overline{y}$$

Podemos assumir que $y_n = \overline{y} + t_n d_n$, onde $||d_n|| = 1$, $d_n \to d$ e $t_n = ||y_n - \overline{y}||$. Assim, $d \in T(Y,\overline{y}) \cap S^{n-1}$ onde $S^{n-1} = \{v \in \mathbb{R}^l; ||v|| = 1\}$. Denotando $z_n = \frac{2(d_n - d)}{t_n}$ temos que

$$d_n = d + \frac{1}{2}t_n z_n \text{ e } y_n = \overline{y} + t_n d + \frac{1}{2}t_n^2 z_n \tag{4.9}$$

Para essa sequência temos duas possibilidades: (a) ela é limitada ou (b) ela não é limitada.

(a) Se z_n é limitada, então(escolhendo uma subsequência adequada) temos $z_n \to z$ para

algum $z \in T^2 = T_1''(Y, \overline{y})(d)$. Então, de (4.9), temos

$$z \in T_1''(Y, \overline{y})(d) \cap T(S^{n-1}, d)$$

Sendo $T(S^{n-1}, d) = d^{\perp}$ obtemos $z \in T_1''(Y, \overline{y})(d) \cap d^{\perp}$ e também $z \in \overline{\Omega(d)}$, o que contradiz a condição (4.8).

(b) Se z_n não é limitada. Então, tomando uma subsequência se necessário, podemos supor que $||z_n|| \to +\infty$ e que existe $z' \neq 0$ tal que $z'_n = ||z_n||^{-1} z_n \to z'$. Por (4.9) obtemos

$$d_n = d + \frac{1}{2} r_n z'_n \ e \ y_n = \overline{y} + t_n d + \frac{1}{2} r^n t_n^n z'_n. \tag{4.10}$$

onde

$$r_n = ||z_n|| t_n = 2||d_n - d|| \to 0^+ \ e \ \frac{t_n}{r_n} \to 0.$$

Isso implica que $z' \in T_0''(Z, \overline{y})(d) \cap d^{\perp}$ e também que $z' \in \overline{\Omega(d)}$, o que contradiz a condição (4.7). □

Exemplo 4.2. *Seja Z o conjunto definido por*

$$Z = \{(x, y) \in \mathbb{R}^2 \mid y = x\sqrt{|x|}, \ com \ \overline{y} = (0, 0)\}$$

. Então

$$T(Y, \overline{y}) = \{(d_1, d_2) \in \mathbb{R}^2 \mid d_2 = 0\}$$

e que, para cada $d \in T(Y, \overline{y})$ e $d \neq 0$, temos que

$$T_1''(Y, \overline{y})(d) = \emptyset, \quad T_0''(Y, \overline{y})(d) = \{z \in \mathbb{R}^2 \mid z_2.d_1 \geqq 0\}.$$

Note que $T(Y, \overline{y}) \cap \mathbb{R}^2_- \subseteq Fr(\mathbb{R}^2_-)$ onde $Fr(\mathbb{R}^2_-) = \{d_2 = 0, \ d_1 < 0\} \cup \{d_1 = 0, \ d_2 < 0\}$. Observe também que as condições (4.6), (4.7) *e* (4.8) *são satisfeitas, portanto $\overline{y}$ é um ponto eficiente de Z.*

A partir de agora fornece-se as condições de otimalidade no espaço de decisão supondo que a função objetivo f é duas vezes continuamente diferenciável no problema (1.1).

4.3 Condições necessárias no caso diferenciável

Teorema 4.3. *Se $\overline{x}$ é uma solução Pareto-ótima local(fraco) do problema* (1.1)*, então o sistema*

$$(\nabla f_j(\overline{x})d, \nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d, d)) <_{lex} (0, 0), \ \forall j = 1, .., l$$

não tem solução $(d, z) \in T_k''(X, \overline{x})$ para $k \in \{0, 1\}$.

Demonstração. Sendo $\overline{x}$ pareto-ótimo local (fraco) do problema (1.1) então existe $B(\overline{x},\delta)$ com $\delta > 0$ tal que $f(\overline{x})$ é um ponto Pareto-ótimo local de $f(X \cap B(\overline{x},\delta),\overline{x})$.

Por outro lado, para qualquer k fixado, seja (d,z) um elemento qualquer de $T_k''(X,\overline{x})$. Então existem $(t_n,r_n) \to (0^+,0^+)$ e $z_n \to z$ tal que $\frac{t_n}{r_n} \to k$ e $x_n = \overline{x} + t_n d + \frac{1}{2} r_n t_n z_n \in X$. Para $n \in \mathbb{N}$ suficientemente grande $T_k''(X,\overline{x}) = T_k''(X \cap B(\overline{x},\delta))$. Logo $x_n \in X \cap B(\overline{x},\delta)$.

Pela expansão de Taylor, para cada j existe um $\epsilon_n \to 0$ tal que

$$\begin{aligned} f_j(x_n) &= f_j(\overline{x}) + t_n \nabla f_j(\overline{x})(d + (1/2) r_n z_n) + \\ &\quad + (1/2) t_n^2 \nabla^2 f_j(\overline{x})(d + (1/2) r_n z_n, d + (1/2) r_n z_n) + t_n^2 \epsilon_n \end{aligned}$$

Assim, definindo $k_n = \frac{t_n}{r_n} \to k$ temos

$$\begin{aligned} f_j(x_n) &= f_j(\overline{x}) + t_n \nabla f_j(\overline{x}) d + \\ &\quad (1/2) t_n r_n [\nabla f_j(\overline{x})(z_n) + k_n \nabla^2 f_j(\overline{x})(d + (1/2) r_n z_n, d + (1/2) r_n z_n) + 2 k_n \epsilon_n] \end{aligned}$$

com $k_n \to k$. Como f é duas vezes diferenciável, então

$$\nabla f_j(\overline{x})(z_n) + k_n \nabla^2 f_j(\overline{x})(d + (1/2) r_n z_n, d + (1/2) r_n z_n) \xrightarrow[n \to +\infty]{} \nabla f_j(\overline{x}) z + k \nabla^2 f_j(\overline{x})(d,d)$$

logo

$$(\nabla f_j(\overline{x}) d, \nabla f_j(\overline{x}) z + k \nabla^2 f_j(\overline{x})(d,d)) \in T_k''(f_j(X \cap B(\overline{x},\delta)), f_j(\overline{x}))$$

Sendo $f(\overline{x})$ Pareto-ótimo local(fraco) de $f(X \cap B(\overline{x},\delta))$, pelo Teorema 4.1 e condição (4.2)

$$T_k''(f_j(X \cap B(\overline{x},\delta)), f_j(\overline{x})) \cap \Omega = \emptyset$$

Portanto,

$$(\nabla f_j(\overline{x}) d, \nabla f_j(\overline{x}) z + k \nabla^2 f_j(\overline{x})(d,d)) \notin \Omega.$$

□

Observação 4.2. *O conjunto* Λ *definido por*

$$\Lambda = \{(d,z) \in R^{2n} \mid (\nabla f_l(\overline{x}) d, \nabla f_l(\overline{x}) z + k \nabla^2 f_l(\overline{x})(d,d)) <_{lex} (0,0), \forall l\} \tag{4.11}$$

é um cone. Daí, para cada $k > 0$,

$$\Lambda \cap T_1''(X,\overline{x}) = \emptyset \Leftrightarrow (1/k)\Lambda \cap (1/k) T_1''(X,\overline{x}) = \emptyset \Leftrightarrow \Lambda \cap T_k''(X,\overline{x}) = \emptyset$$

Definição 4.1. *O conjunto das direções críticas para o problema* (1.1) *em* $\overline{x} \in \overline{X}$ *é definido por*

$$\mathbb{K} = \{d \in \mathbb{R}^n \mid d \in T(X,\overline{x}),\ \nabla f(\overline{x}) d \leqq 0,\ \nabla f_i(\overline{x}) d = 0,\ \textit{em pelo menos um } i\} \tag{4.12}$$

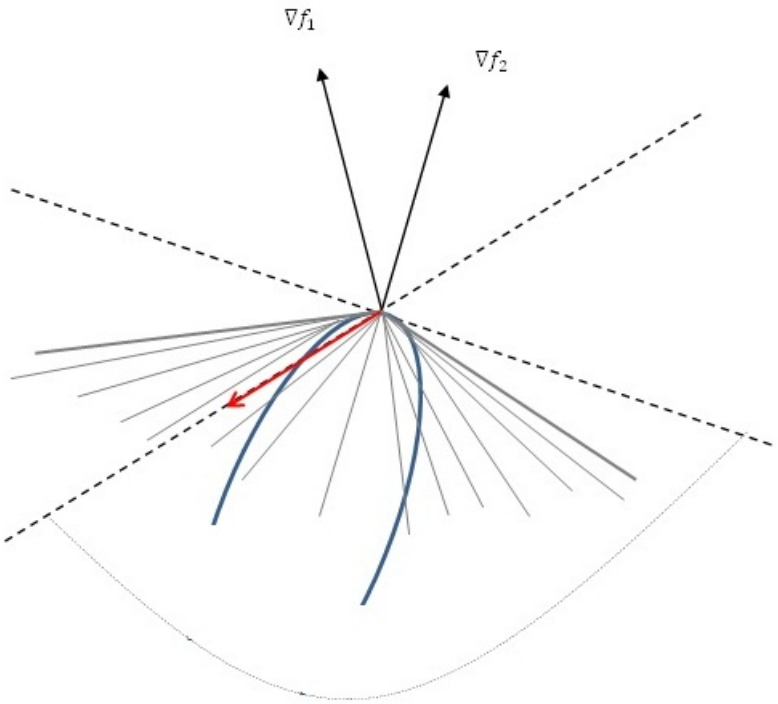


Figura 4.1: *Ilustração da definição 4.1*

Para qualquer vetor $\overline{x} \in X$ e $d \in \mathbb{R}^n$, defini-se

$$I_f = \{1, ..., l\}, \quad I_f(\overline{x}, d) = \{j \in I_f \mid \nabla f_j(\overline{x})d = 0\}$$

Corolário 4.2. *Se $\overline{x}$ é uma solução Pareto-ótima local(fraco) do problema* (1.1)*, então para cada direção crítica $d \in \mathbb{K}$, o sistema*

$$\nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d,d) < 0, \quad j \in I_f(\overline{x}, d) \tag{4.13}$$

não tem solução z em $T_k''(X, \overline{x})(d)$ para $k \in \{0, 1\}$.

Demonstração. Segue imediatamente do Teorema 4.3. □

Exemplo 4.3. *Sejam $\overline{x} = (0, 0)$ e $d = (1, 0)$. Definindo o conjunto $X = \{x \in \mathbb{R}^2 \ ; \ x_2 \geq 0\}$ e $f(x_1, x_2) = x_2 - x_1^2$ para o problema* (1.1) *com $T^2(M, \overline{x})(d) = T''(M, \overline{x})(d) = \{w \ ; \ w_2 \geq 0\}$ (JIMÉNEZ E NOVO* (2004)*). Vamos verificar que $\overline{x}$ não é uma solução Pareto-ótima local.*

Verificação:

$$\nabla f(x_1, x_2) = \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix} \ e \ \nabla^2 f(x_1, x_2) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$$

Então

$$\nabla f_l(\overline{x})z = < \nabla f(0, 0), (z_1, z_2) > = z_2 \ e \ \nabla^2 f(x_1, x_2)(d, d) = -2 \tag{4.14}$$

Substituindo (4.15) *em* (4.14) *do Corolário* 4.2 *tem-se* $z_2 - 2k < 0$.

Para $k = 0$ *obtemos* $z_2 < 0$, *logo o sistema* (4.14) *não tem solução em* $T''(X, \overline{x})(d)$. *Para* $k = 1$ *temos* $z_2 < 2$, *logo o sistema* (4.14) *tem solução em* $T^2(X, \overline{x})(d)$.

Portanto $\overline{x}$ *não é uma solução Pareto-ótima local.*

Exemplo 4.4. *Seja o problema* (1.1) *definido por* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *dada por* $f(x_1, x_2) = (-x_2 - x_1x_2 - x_1^2,\ x_1)$ *em que* $X = \{x\ ;\ x_2^2 + x_1^3 = 0\}$, $\overline{x} = (0, 0)$ *e* $d = (-1, 0)$ *com* $T''(X, \overline{x})(d) = \mathbb{R}^2$ *e* $T^2(X, \overline{x})(d) = \emptyset$. *Verifica-se que* $\overline{x}$ *não é uma solução Pareto-ótimo local.*

De fato,

$$\nabla f_1(x_1, x_2) = \begin{bmatrix} -2x_1 - x_2 \\ -x_1 - 1 \end{bmatrix},\ \nabla f_2(x_1, x_2) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$
$$\nabla^2 f_1(x_1, x_2) = \begin{bmatrix} -2 & -1 \\ -1 & 0 \end{bmatrix} \text{ e } \nabla^2 f_2(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Então

$$\nabla f_1(\overline{x})z = < \nabla f_1(0, 0), (z_1, z_2) > = -z_2,$$
$$\nabla f_2(\overline{x})z = < \nabla f_2(0, 0), (z_1, z_2) > = z_1,$$
$$\nabla^2 f_1(\overline{x})(d, d) = -2 \text{ e } \nabla^2 f_2(\overline{x})(d, d) = 0$$

(i) Para $j = 1$ *em* (4.14) *obtemos* $-z_2 - 2k < 0$.

(ii) Para $j = 2$ *em* (4.14) *obtemos* $z_1 < 0$.

Se $k = 1$ *os sistemas* (i) *e* (ii) *não tem solução pelo Corolário* 4.2 *em* $T^2(X, \overline{x})$.

Se $k = 0$ *os sistemas* (i) *e* (ii) *tem solução pelo Corolário* 4.2 *em* $T''(X, \overline{x})$.
Portanto $\overline{x}$ *não é solução Pareto-ótimo local.*

Estendendo o corolário 4.2 pode-se obter condições de otimalidade para problemas de otimização mono-objetivo(escalar). Para tal o próximo lema será útil.

Lema 4.1. *Sejam* Γ *e* T *subconjuntos em* $\mathbb{R}^n$. *Suponha que* Γ *é um semi espaço aberto. Então*

$$\Gamma \cap T = \emptyset \Rightarrow \overline{conv}T \cap \Gamma = \emptyset$$

em que $\overline{conv}T$ *representa o conjunto fechado do fecho convexo de* T.

Demonstração. Veja Aghezzaf, B., Hachimi (2001). □

Teorema 4.4. *Seja $l = 1$ no problema* (1.1)*. Se $\overline{x}$ é um minimizador (local) para tal problema, então para cada direção crítica $d \in \mathbb{K}$, o sistema*

$$\nabla f(\overline{x})z + k\nabla^2 f(\overline{x})(d,d) < 0 \tag{4.15}$$

não tem solução em $\overline{conv}[T_k''(X,\overline{x})(d)]$, para $k \in \{0,1\}$.

Demonstração. Para $k \in \{0,1\}$, seja $\Gamma = \{z \in \mathbb{R}^n \mid \nabla f(\overline{x})z + k\nabla^2 f(\overline{x})(d,d) < 0\}$. O conjunto Γ pode ser o conjunto vazio, ou o conjunto $\mathbb{R}^n$ ou um semi espaço aberto. Em cada caso, o Lema 4.1 é válido. Consequentemente o resultado segue do Corolário 4.2 e Lema 4.1. □

Segundo Sawaragi, Y., Nakayama (1985) o próximo lema fornece uma relação entre problemas de otimização multiobjetivo e o seu correspondente problema mono-objetivo. Para isso defini-se o i-ésimo problema objetivo restrito (P^i)

$$\begin{aligned} (P^i) \qquad & \min \ f_i(x) \\ & s.a. \ x \in Q^i \end{aligned} \tag{4.16}$$

onde $Q^i = \{x \in \mathbb{R}^n \mid x \in X, \ f_j(x) \leqq f_j(\overline{x}), \ j = 1,...,l \ e \ j \neq i\}$ $i = 1,...,l$.

Lema 4.2. *Um vetor $\overline{x}$ é uma solução Pareto-ótima (local) do problema* (1.1) *se, e somente se, $\overline{x}$ é um minimizador (local) do problema (P^i), para todo $i = 1,...,l$.*

Demonstração. Segue da prova do Teorema 3.4.11 em Sawaragi, Y., Nakayama (1985). □

Teorema 4.5. *Se $\overline{x}$ é uma solução Pareto-ótima (local) para o problema* (1.1)*, então para cada direção crítica $d \in \mathbb{K}$, o sistema*

$$\nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d,d) < 0, \ \textit{para pelo menos um } j \in I_f(\overline{x},d) \tag{4.17}$$

não tem solução z em $\bigcap_{i=1}^{l} \overline{conv}[T_k''(Q^i,\overline{x})(d)]$ para $k \in \{0,1\}$.

Demonstração. Segue do Teorema 4.4 e Lema 4.2. □

4.4 Condição suficiente no caso diferenciável

Teorema 4.6. *Uma condição suficiente para um ponto $\overline{x} \in X$ ser uma solução Pareto-ótimo local do problema* (1.1) *é que*

$$T(X,\overline{x}) \cap \{d \in \mathbb{R}^n | \nabla f(\overline{x})d \leqq 0\} \subseteq \mathbb{K} \tag{4.18}$$

e para qualquer direção crítica $d \in \mathbb{K} - \{0\}$, *se* $\mathbb{K} - \{0\} \neq \emptyset$, *os dois sistemas a seguir não tem solução* $z \in \mathbb{R}^n$:

$$\nabla f_j(\overline{x})z + \nabla^2 f_j(\overline{x})(d,d) \leqq 0,\ j \in I_f(\overline{x},d),$$

$$z \in T_1''(X,\overline{x})(d) \cap d^{\perp},\ z \neq 0.$$

e

$$\nabla f_j(\overline{x})z \leqq 0,\ j \in I_f(\overline{x},d),$$

$$z \in T_0''(X,\overline{x})(d) \cap d^{\perp},\ z \neq 0,$$

Demonstração. Suponha que $\overline{x}$ não é uma solução Pareto-ótimo local do problema (1.1). Então existe uma sequência $f(x_n) \in Y$ tal que $f(x_n) \leqq f(\overline{x}),\ f(x_n) \to f(\overline{x})$. Assumimos que $x_n = \overline{x} + t_n d_n$, onde $||d_n|| = 1$, $d_n \to d$ e $t_n = ||x_n - \overline{x}||$. Assim $d \in T(X,\overline{x})$ com $||d|| = 1$.

Pela expansão de Taylor

$$f(x_n) = f(\overline{x}) + t_n \nabla f(\overline{x}) d_n + r(t_n d_n),\ com \lim_{n\to+\infty} \frac{r(t_n d_n)}{||t_n d_n||} = 0$$

Para $n \to +\infty$ tem-se $\nabla f(\overline{x})d_n = \frac{f(x_n)-f(\overline{x})}{t_n} \leqq 0$. Logo $\nabla f(\overline{x})d \leqq 0$.
Definindo $z_n = \frac{2(d_n - d)}{t_n}$ temos que

$$d_n = d + (1/2)t_n z_n \ e\ x_n = \overline{x} + t_n d + (1/2)t_n^2 z_n. \tag{4.19}$$

Vamos considerar os casos quando $\{z_n\}$ é limitado e quando $\{z_n\}$ é ilimitado.

(i) Se $\{z_n\}$ é limitado. Tomando uma subsequência se necessário, existe z tal que $z_n \to z$. Lembrando que $T(S^{n-1},\overline{x}) = \overline{x}^{\perp}$ exemplo 2.3 e definindo $s_n = (1/2)t_n$, $h_n = z_n$ e $d_n = d + s_n h_n$ com $d_n \in T(X,\overline{x})$. De (4.19) temos que $z \in T_1''(X,\overline{x}) \cap d^{\perp}$.

Sendo f duas vezes diferenciável temos que

$$\begin{aligned} f(x_n) &= f(\overline{x}) + t_n \nabla f(\overline{x})(d + (1/2)t_n z_n) + (1/2)t_n^2 \nabla^2 f(\overline{x})(d + (1/2)t_n z_n, \\ & \quad d + (1/2)t_n z_n) + \rho_n \ \ com \lim_{n\to+\infty} \rho_n = 0 \end{aligned}$$

Dado $j \in I_f(\overline{x},d)$ com $\nabla f_j(\overline{x})d = 0$ temos que

$$\begin{aligned} & f_j(x_n) - f_j(\overline{x}) = (1/2)\nabla f_j(\overline{x}) z_n + (1/2)\nabla^2 f_j(\overline{x})(d + (1/2)t_n z_n, \\ & d + (1/2)t_n z_n) + \frac{\rho_n}{t_n^2} \end{aligned} \tag{4.20}$$

Aplicando o limite quando $n \to +\infty$ em (4.20) tem-se

$$\nabla f_j(\overline{x})z + \nabla^2 f_j(\overline{x})(d,d) \leqq 0,\ \forall\ j \in I_f(\overline{x},d)$$

o que contraria a hipótese.

(ii) Se z_n não é limitada. Então, tomando uma subsequência se necessário, podemos supor que $||z_n|| \to +\infty$ e que, existe $z' \neq 0$ tal que $z'_n = ||z_n||^{-1} z_n \to z'$. De (4.19) obtemos $d_n = d + (1/2) r_n z'_n$ e $x_n = \overline{x} + t_n d + (1/2) r_n t_n z'_n$ onde $r_n = ||z_n|| t_n = 2||d_n - d|| \to 0^+$ e $\frac{t_n}{r_n} \to 0$. Isto implica que $z' \in T''_0(X, \overline{x})(d) \cap d^{\perp}$ com $z' \neq 0$.

Pela expansão de Taylor

$$f(x_n) = f(\overline{x}) + t_n \nabla f(\overline{x})(d + (1/2) r_n z'_n) \Rightarrow$$

$$\frac{f(x_n) - f(\overline{x})}{t_n} = \nabla f(\overline{x})(d + (1/2) r_n z'_n) \leqq 0 \Rightarrow$$

$$\nabla f(\overline{x})(d) + (1/2) \nabla f(\overline{x}) r_n z'_n \leqq 0$$

Dado $j \in I_f(\overline{x}, d)$ temos que

$$\nabla f_j(\overline{x})(d) + (1/2) \nabla f_j(\overline{x}) r_n z'_n \leqq 0 \tag{4.21}$$

Dividindo por r_n em (4.21) tem-se $\nabla f_j(\overline{x}) z'_n \leqq 0$. Para $n \to +\infty$ obtemos $\nabla f_j(\overline{x}) z' \leqq 0$ o que é um absurdo.

□

4.5 Condições de Otimalidade de John e Kuhn-Tucker

Na área de Otimização Multiobjetivo, muitos artigos tem mostrado condições necessárias de primeira ordem e segunda ordem ao longo dos anos e geralmente essas condições são fundamentadas pelo Teorema Fundamental de Lin's (Teorema 5.1 em LIN, J.G. (1987)). Entre estas, condições do tipo Fritz John tem sido analisadas para estes problemas. Diferentemente do caso de problemas com um único objetivo, há várias possibilidades para os valores dos multiplicadores de Lagrange associados à função objetivo. O vetor de multiplicadores pode ser inteiramente igual ao vetor zero ou não, mas isso não garante que todos os multiplicadores serão diferentes de zero. Quando um multiplicador associado a alguma função objetivo é igual a zero, essa função não desempenha papel algum na condição de otimalidade, o que não é desejado. As condições de qualificação em programação não-linear, envolvendo apenas as funções objetivo, permitem que o vetor de multiplicadores associados as funções objetivo sejam diferente de zero. Entretanto, tais condições não propiciam a estrita positividade de cada multiplicador. Para tal efeito, é necessário usar condições que envolvam as funções objetivo. Estas condições são chamadas de condições de regularidade e são um tipo de extenção das condições de qualificação usuais.

Aborda-se nesta seção sobre os multiplicadores de Lagrange para o problema (1.1), incluindo as condições do tipo Fritz John e Kuhn-Tucker. Para isso o conjunto viável X desse

problema é definido por restrições de igualdade e desigualdade.

O modelo de otimização multiobjetivo utilizado agora é formulado da seguinte forma:

$$\begin{array}{ll} \min & f(x) \\ s.a. & x \in X \end{array} \tag{4.22}$$

Em (4.22), $X = \{x \in \mathbb{R}^n \mid g(x) \leqq 0, h(x) = 0\}$ em que $f : R^n \to R^l$, $g : \mathbb{R}^n \to \mathbb{R}^p$ *e* $h : \mathbb{R}^n \to \mathbb{R}^q$ são funções duas vezes continuamente diferenciável e $I_g = \{1, ..., p\}$ *e* $I_h = \{1, ..., q\}$ os conjuntos de índices de g e h respectivamente.

Para fornecer as condições de otimalidade necessárias (os multiplicadores de Lagrange) para o problema (4.22), as próximas definições e lemas serão úteis.

Para qualquer ponto viável $\overline{x} \in X$ e $d \in \mathbb{R}^n$, defini-se

$$I_g(\overline{x}) = \{i \in I_g \mid g_i(\overline{x}) = 0\},\ I_g(\overline{x}, d) = \{i \in I_g(\overline{x}) \mid \nabla g_i(\overline{x})d = 0\}$$

Definição 4.2. *O k-ésimo conjunto linearizado de segunda-ordem X em $\overline{x}$ é*

$$\begin{array}{rcl} L''_k(X, \overline{x}) & = & \{(d, z) \in \mathbb{R}^{2n} \mid \\ & & (\nabla g_i(\overline{x})d, \nabla g_i(\overline{x})z + k\nabla^2 g_i(\overline{x})(d, d)) \leqq_{lex} (0, 0),\ i \in I_g(\overline{x}), \\ & & (\nabla h_r(\overline{x})d, \nabla h_r(\overline{x})z + k\nabla^2 h_r(\overline{x})(d, d)) = (0, 0),\ r \in I_h\} \end{array}$$

Definição 4.3. *O k-ésimo conjunto fracamente linearizado de segunda-ordem X em $\overline{x}$ é*

$$\begin{array}{rcl} WL''_k(X, \overline{x}) & = & \{(d, z) \in \mathbb{R}^{2n} \mid \\ & & (\nabla g_i(\overline{x})d, \nabla g_i(\overline{x})z + k\nabla^2 g_i(\overline{x})(d, d)) <_{lex} (0, 0),\ i \in I_g(\overline{x}), \\ & & (\nabla h_r(\overline{x})d, \nabla h_r(\overline{x})z + k\nabla^2 h_r(\overline{x})(d, d)) = (0, 0),\ r \in I_h\} \end{array}$$

O lema a seguir relaciona os conjuntos tangentes de segunda ordem com o conjunto linearizado de segunda ordem.

Lema 4.3. *Seja $\overline{x} \in X$ uma solução eficiente qualquer para o problema* (4.22)*. Então,*

$$T''_k(X, \overline{x}) \subseteq L''_k(X, \overline{x}),\ k \in \{0, 1\} \tag{4.23}$$

Mais ainda, se os vetores $\{\nabla h_r(\overline{x}),\ r \in I_h\}$ são linearmente independentes, então

$$WL''_k(X, \overline{x}) \subseteq T''_k(X, \overline{x}),\ k \in \{0, 1\} \tag{4.24}$$

Demonstração. Veja o Lema 4.1 de Bigi, G., Castellani, M. (2000) e também Hachimi, M. (2000). □

Segue imediatamente de (4.23) que, para toda direção $d \in \mathbb{R}^n$,

$$T_k''(X, \overline{x})(d) \subseteq L_k''(X, \overline{x})(d), \ k \in \{0, 1\} \tag{4.25}$$

onde

$$L_k''(X, \overline{x})(d) = \{z \in \mathbb{R}^n \mid (d, z) \in L_k''(X, \overline{x})\}$$

Sendo $L_k''(X, \overline{x})(d)$ um conjunto convexo fechado então para cada $d \in \mathbb{R}^n$, de (4.25) obtemos

$$\overline{conv}[T_k''(X, \overline{x})(d)] \subseteq L_k''(X, \overline{x})(d), \ k \in \{0, 1\} \tag{4.26}$$

Lema 4.4. *Seja $\overline{x}$ uma solução eficiente qualquer para o problema* (4.22)*. Então,*

$$\bigcap_{i=1}^{l} \overline{conv}[T_k''(Q^i, \overline{x})(d)] \subseteq L_k''(Q, \overline{x})(d), \ k \in \{0, 1\} \tag{4.27}$$

Demonstração. Como $\cap_{i=1}^{l} L_k''(Q^i, \overline{x})(d)] \subseteq L_k''(Q, \overline{x})(d)$, a prova segue imediatamente da condição (4.26). □

4.5.1 Condições de Otimalidade Necessárias do Tipo John

Agora, as condições necessárias no caso diferenciável do Corolário 4.2 são reduzidas em termos dos gradientes e hessianas das funções objetivos e das restrições.

Teorema 4.7. *Seja $k \in \{0, 1\}$. Suponha que os vetores $\{\nabla h_r(\overline{x}), \ r \in I_h\}$ são linearmente independentes. Se $\overline{x}$ é uma solução eficiente local(fraco) para o problema* (4.22)*, então para cada direção crítica $d \in \mathbb{K}$, o sistema*

$$\nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d, d) < 0, \ j \in I_f(\overline{x}, d), \tag{4.28}$$

$$\nabla g_i(\overline{x})z + k\nabla^2 g_i(\overline{x})(d, d) < 0, \ i \in I_g(\overline{x}, d), \tag{4.29}$$

$$\nabla h_r(\overline{x})z + k\nabla^2 h_r(\overline{x})(d, d) = 0, \ r \in I_h, \tag{4.30}$$

não tem solução z em $\mathbb{R}^n$.

Demonstração. O resultado segue do Corolário 4.2 e da condição(4.24) do Lema 4.3. □

Para obtermos a regra dos multiplicadores, o próximo lema será provado.

Lema 4.5. *Sejam A, B, C matrizes reais dadas, com $A \neq 0$. Sejam a, b, c vetores reais. Se o sistema*

$$(\gamma) \ Ax + a < 0, \ Bx + b \leqq 0, \ Cx + c = 0$$

não tem solução x em $\mathbb{R}^n$, então exatamente um dos seguintes dois sistemas (I) e (II) tem uma solução:

$$\begin{aligned}(I) \qquad & \lambda^t A + \mu^t B + \nu^t C = 0,\\ & \lambda^t a + \mu^t b + \nu^t c \geqq 0,\\ & \lambda \geq 0,\ \mu \geqq 0,\\ (II) \quad & Az < 0,\ Bz \leqq 0,\ Cz = 0.\end{aligned}$$

Demonstração. Se (I) e (II) valem, então teríamos z, λ, μ, ν tal que

$$zA^t\lambda + zB^t\mu + zC^t\nu < 0$$

pois

$$zC^tv = 0,\ zC^t\nu \leqq 0 \ e\ zA^t\lambda < 0$$

Contradição pela primeira igualdade de (I).
Agora, assuma que

$$\begin{aligned}(I) \quad & \lambda^t A + \mu^t B + \nu^t C = 0,\\ & \lambda^t a + \mu^t b + \nu^t c \geqq 0,\\ & \lambda \geq 0,\ \mu \geqq 0,\end{aligned}$$

não tem solução. O sistema (I) é equivalente com a incosistência do seguinte sistema

$$\begin{aligned}(I) \quad & \lambda^t A + \mu^t B + \nu^t C + \xi.0 = 0,\\ & \lambda^t a + \mu^t b + \nu^t c - \xi = 0,\\ & \lambda \geq 0,\ \mu \geqq 0 \ e\ \xi \geqq 0.\end{aligned}$$

Pelo Teorema de alternativa de Motzkin's (MANGASARIAN (1969)), existem z e $t \leqq 0$ satisfazendo

$$Az + at < 0,\ Bz + bt \leqq 0,\ Cz + ct = 0$$

Como o sistema (γ) não tem solução, tomando t=0 tem-se que o sistema (II) tem uma solução z, o que causa um absurdo. □

Teorema 4.8. *Seja $k \in \{0, 1\}$. Se $\overline{x}$ é uma solução eficiente local (fraco) para o problema* (4.22)*, então para cada direção crítica $d \in \mathbb{K}$, existem $\lambda \in \mathbb{R}^l_+, \mu \in \mathbb{R}^p_+, \nu \in \mathbb{R}^q$ tais que*

$$\sum_{j=1}^{l} \lambda_j \nabla f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla h_r(\overline{x}) = 0, \tag{4.31}$$

$$k\left(\sum_{j=1}^{l} \lambda_j \nabla^2 f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla^2 g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla^2 h_r(\overline{x})\right)(d,d) \geqq 0, \tag{4.32}$$

$$\mu_r g_r(\overline{x}) = 0, \ \textit{para cada } j \in I_g, \tag{4.33}$$

$$\lambda_j \nabla f_j(\overline{x}) d = 0, \ \textit{para cada } r \in I_f, \tag{4.34}$$

$$\mu_i \nabla g_i(\overline{x}) d = 0, \ \textit{para cada } i \in I_g. \tag{4.35}$$

Demonstração. Se os vetores $\{\nabla h_r(\overline{x}),\ r \in I_h\}$ são linearmente dependentes, então pode-se encontrar vetores $\lambda \geqq 0,\ \mu \geqq 0,\ \nu$ não nulos tais que as condições (4.31) à (4.35) sejam verdadeiras, ou seja, se $\{\nabla h_r(\overline{x}),\ r \in I_h\}$ é linearmente dependente , escolha

$$\lambda_j = \mu_i = 0 \ e \ \nu_r \neq 0 \ tais \ que \ \sum_{r=1}^{q} \nu_r \nabla h_r(\overline{x}) = 0$$

As condições $(4.31), (4.33), (4.34)$ *e* (4.35) são satisfeitas trivialmente, se (4.32) não vale, é suficiente trocar ν_r por $-v_r$.

Chamando de (I) o sistema (4.31) à (4.35) e (II) o sistema (4.28) à (4.30) e sejam $A = \nabla f_j(\overline{x}),\ B = \nabla g_i(\overline{x}),\ C = \nabla h_r(\overline{x})$ e $a = \nabla^2 f_j(\overline{x}),\ b = \nabla^2 g_i(\overline{x}),\ c = \nabla^2 h_r(\overline{x})$. Suponha que os vetores $\{\nabla h_r(\overline{x}), r \in I_h\}$ são linearmente independentes. Então para $d \in \mathbb{K}$ o sistema (II) não tem solução z. Vamos assumir que o sistema (I) não tem solução. Pelo Lema 4.5 existe $z \in \mathbb{R}^n$ tal que

$$\nabla f_j(\overline{x}) z < 0,\ j \in I_f(\overline{x}, d),\ \nabla g_i(\overline{x}) z < 0,\ i \in I_g(\overline{x}, d),\ \nabla h(\overline{x}) z = 0$$

Por outro lado,

$$\nabla f_j(\overline{x}) d = 0,\ j \in I_f(\overline{x}, d),\ \nabla f_j(\overline{x}) d < 0,\ j \notin I_f(\overline{x}, d)$$

$$\nabla g_i(\overline{x}) d = 0,\ i \in I_g(\overline{x}, d),\ \nabla g_i(\overline{x}) d < 0,\ i \notin I_g(\overline{x}, d),\ \nabla h(\overline{x}) d = 0$$

Pois d é uma direção crítica. Assim, segue para $t > 0$ suficientemente pequeno que

$$\nabla f(\overline{x})(d + tz) < 0,\ \nabla g_i(d + tz) < 0,\ i \in I_g(\overline{x}),\ \nabla h(\overline{x})(d + tz) = 0$$

Para $t > 0$ suficientemente pequeno o sistema (II) do Teorema 4.7 teria solução o que é um absurdo.

□

4.5.2 Condições de Otimalidade Necessárias do Tipo 1 de Kuhn-Tucker

Não há garantias que no teorema 4.8 os multiplicadores $\lambda_j \neq 0$ para pelo menos um $j \in I_f$. Nos casos onde $\lambda = 0$, a função objetivo não desempenha qualquer papel nas condições necessárias de otimalidade. Para evitar esse tipo de situação desagradável, algumas hipóteses na região viável serão acrescentadas.

Definição 4.4. *A k-ésima condição de qualificação de restrição de segunda ordem de Abadie ($(ASOCQ)_k$ - kth Abadie second-order constraint qualification) em $\overline{x} \in X$ na direção $d \in \mathbb{R}^n$ é tal que*

$$T_k''(X, \overline{x})(d) = L_k''(X, \overline{x})(d) \tag{4.36}$$

Se $(ASOCQ)_k$ vale em $\overline{x} \in X$ para toda direção $d \in \mathbb{R}^n$, diz-se que $(ASOCQ)_k$ vale em $\overline{x} \in X$.

Definição 4.5. *A k-ésima condição de qualificação de restrição de segunda ordem de Guignard ($(GSOCQ)_k$ - kth Guignard second-order constraint qualification) em $\overline{x} \in X$ na direção $d \in \mathbb{R}^n$ é tal que*

$$\overline{conv}[T_k''(X, \overline{x})(d)] = L_k''(X, \overline{x})(d) \tag{4.37}$$

Observação 4.3. *Observe que no caso $d = 0$, as condições (4.36) e (4.37) são reduzidas para qualificação de restrições de primeira-ordem.*

Teorema 4.9. *Seja $k \in \{0, 1\}$. Suponha que $(ASOCQ)_k$ vale em $\overline{x} \in X$ na direção crítica $d \in \mathbb{K}$. Se $\overline{x}$ é uma solução eficiente local(fraco) para o problema (4.22), então para cada direção crítica $d \in \mathbb{K}$, o sistema*

$$\nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d, d) < 0, \ j \in I_f(\overline{x}, d), \tag{4.38}$$

$$\nabla g_i(\overline{x})z + k\nabla^2 g_i(\overline{x})(d, d) \leqq 0, \ i \in I_g(\overline{x}, d), \tag{4.39}$$

$$\nabla h_r(\overline{x})z + k\nabla^2 h_r(\overline{x})(d, d) = 0, \ r \in I_h, \tag{4.40}$$

não tem solução z em $\mathbb{R}^n$.

Demonstração. Segue-se do Corolário 4.2 e da condição (4.36). □

Teorema 4.10. *Seja $k \in \{0,1\}$. Suponha que $(ASOCQ)_k$ vale em $\overline{x} \in X$ na direção crítica $d \in \mathbb{K}$. Se $\overline{x}$ é uma solução eficiente local (fraco) para o problema* (4.22)*, então existem $\lambda \in \mathbb{R}^l_+, \mu \in \mathbb{R}^p_+, \nu \in \mathbb{R}^q$, com $\lambda \neq 0$ tais que*

$$\sum_{j=1}^{l} \lambda_j \nabla f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla h_r(\overline{x}) = 0, \tag{4.41}$$

$$k \left(\sum_{j=1}^{l} \lambda_j \nabla^2 f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla^2 g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla^2 h_r(\overline{x}) \right) (d,d) \geqq 0, \tag{4.42}$$

$$\mu_r g_r(\overline{x}) = 0, \ \textit{para cada } j \in I_g, \tag{4.43}$$

$$\lambda_j \nabla f_j(\overline{x}) d = 0, \ \textit{para cada } r \in I_f, \tag{4.44}$$

$$\mu_i \nabla g_i(\overline{x}) d = 0, \ \textit{para cada } i \in I_g. \tag{4.45}$$

Demonstração. A prova segue-se do Lema 4.5 aplicado no sistema (4.38) à (4.40). □

Para os problemas de otimização mono-objetivo refina-se os teoremas 4.9 e 4.10 para tais problemas da seguinte maneira:

Teorema 4.11. *Para $l = 1$, os teoremas* 4.9 *e* 4.10 *ainda valem se a condição de qualificação de segunda-ordem $(ASOCQ)_k$ é substituído pela condição de qualificação de segunda-ordem $(GSOCQ)_k$.*

Demonstração. Segue-se dos Teoremas 4.9 e 4.10, levando em conta o teorema 4.4 em vez do corolário 4.2. □

4.5.3 Condições de Otimalidade Necessárias do Tipo 2 de Kuhn-Tucker

Observe que no teorema 4.10 não se tem garantia que $\lambda > 0$ mesmo assumindo como hipótese as condições de qualificação. Nesses casos onde $\lambda_j = 0$ para algum $j \in I_f$, a função objetivo correspondente não participa das condições necessárias de otimalidade. Para evitar essas situações, algumas hipóteses são construídas no problema.

Definição 4.6. *A k-ésima condição de regularidade de restrição de segunda ordem de Guignard $(GSORQ)_k$ - kth Guignard second-order regularity condition) em $\overline{x} \in X$ na direção $d \in \mathbb{R}^n$ é tal que*

$$\bigcap_{i=1}^{l} \overline{conv}[T''_k(Q^i, \overline{x})(d)] = L''_k(Q, \overline{x})(d) \tag{4.46}$$

O termo "condição de regularidade"é usado em vez de "qualificação de restrição" porque ambas as funções objetivos e as funções de restrição são envolvidas.

Teorema 4.12. *Seja $k \in \{0,1\}$. Suponha que $(GSORQ)_k$ vale em $\overline{x} \in X$ na direção crítica $d \in \mathbb{K}$. Se $\overline{x}$ é uma solução eficiente local(fraco) para o problema* (4.21)*, então o sistema*

$$\nabla f_j(\overline{x})z + k\nabla^2 f_j(\overline{x})(d,d) \leqq 0,\ j \in I_f(\overline{x},d), \tag{4.47}$$

$$\nabla f_s(\overline{x})z + k\nabla^2 f_s(\overline{x})(d,d) < 0,\ \textit{para pelo menos um } s, \tag{4.48}$$

$$\nabla g_i(\overline{x})z + k\nabla^2 g_i(\overline{x})(d,d) \leqq 0,\ i \in I_g(\overline{x},d), \tag{4.49}$$

$$\nabla h_r(\overline{x})z + k\nabla^2 h_r(\overline{x})(d,d) = 0,\ r \in I_h, \tag{4.50}$$

não tem solução z em $\mathbb{R}^n$.

Demonstração. Segue-se do Teorema 4.5 e condição (4.46). □

Para obter a regra dos multiplicadores, o seguinte lema vai ser utilizado.

Lema 4.6. *Sejam A, B, C matrizes reais dadas, com $A \neq 0$. Sejam a, b, c vetores reais. Se o sistema*

$$(\Theta)\ Ax + a < 0,\ Bx + b \leqq 0,\ Cx + c = 0$$

não tem solução x em $\mathbb{R}^n$, então exatamente um dos seguintes dois sistemas (I) e (II) tem uma solução:

$$\begin{aligned}
(I) \qquad & \lambda^t A + \mu^t B + \nu^t C = 0,\\
& \lambda^t a + \mu^t b + \nu^t c \geqq 0,\\
& \lambda > 0,\ \mu \geqq 0,\\
(II) \quad & Az \leq 0,\ Bz \leqq 0,\ Cz = 0.
\end{aligned}$$

Demonstração. A prova é similar a prova do Lema 4.5 com suas devidas alterações. □

Teorema 4.13. *Seja $k \in \{0,1\}$. Suponha que $(GSORQ)_k$ vale em $\overline{x} \in X$ na direção crítica $d \in \mathbb{K}$. Se $\overline{x}$ é uma solução eficiente local (fraco) para o problema* (4.22)*, então existem $\lambda \in \mathbb{R}^l_+, \mu \in \mathbb{R}^p_+, \nu \in \mathbb{R}^q$, com $\lambda_j > 0$ para $j \in I_f(\overline{x},d)$ tais que*

$$\sum_{j=1}^{l} \lambda_j \nabla f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla h_r(\overline{x}) = 0, \tag{4.51}$$

$$k\left(\sum_{j=1}^{l} \lambda_j \nabla^2 f_j(\overline{x}) + \sum_{i=1}^{p} \mu_i \nabla^2 g_i(\overline{x}) + \sum_{r=1}^{q} \nu_r \nabla^2 h_r(\overline{x})\right)(d,d) \geqq 0, \tag{4.52}$$

$$\mu_r g_r(\overline{x}) = 0,\ \textit{para cada } j \in I_g, \tag{4.53}$$

$$\lambda_j \nabla f_j(\overline{x})d = 0,\ \textit{para cada } r \in I_f, \tag{4.54}$$

$$\mu_i \nabla g_i(\overline{x})d = 0,\ \textit{para cada } i \in I_g. \tag{4.55}$$

Demonstração. A prova segue-se do Lema 4.6 aplicado no sistema (4.47) à (4.50). □

Observação 4.4. *Observe que a diferença entre as condições de Fritz John e KKT é exatamente sobre esse particular ($\lambda \in \mathbb{R}^l_+$). Contudo quando se impõe a restrição de qualificação assumindo a regularidade na solução então as duas condições coincidem (KKT e Fritz).*

Observação 4.5. *No Teorema* 4.10 *o $\nabla g_i(\overline{x})$ está definido para $i \in I_g(\overline{x}, d)$ logo a combinação linear dos vetores $\nabla g_i(\overline{x})$ e $\nabla h_r(\overline{x})$ não são L.D (linearmente dependentes), pois se fossem, haveria uma contraposição na hipótese de Abadie, e com isso garante-se que pelo menos um $\lambda \neq 0$.*

Observação 4.6. *As condições de otimalidade de primeira ordem são obtidas tomando $d = 0$ nos teoremas tratados neste trabalho veja Marusiac* (1982) *e Singh* (1987).

Considerações Finais

Atualmente a otimização multiobjetivo tornou-se pesquisa de ponta porque várias empresas necessitam de especialistas nesse segmento. Um problema que atinge esses profissionais é garantir que a solução obtida é "boa"ou não no sentido de solução eficiente (fracamente eficiente). Por isso é importante desenvolver pesquisas no campo da teoria de otimização multiobjetivo pois na prática as condições de otimalidade garantem se essas soluções são "boas"ou não.

Neste trabalho, objetivou-se primeiramente apresentar condições de otimalidade de segunda ordem para problemas de otimização multiobjetivo por meio da teoria dos cones tanto em casos gerais quanto no caso diferenciável e a partir de uma condição imposta à direção crítica no problema obteve-se condições de otimalidade de primeira ordem. Em seguida descreve-se versões de John e Kuhn-Tucker para problemas multiobjetivo com restrições. Essas condições coincidem (John e Kunh-Tucker) quando há hipótese de regularidade na solução já que com essas hipóteses obtém-se regularidade sobre as restrições ativas e consequentemente pelo menos um lambida envolvido no processo é positivo (veja Teorema 4.10).

Vários exemplos enriqueceram este trabalho de forma a elucidar a teoria dos cones para problemas de otimização multiobjetivo. Utilizou-se uma ordem parcial no $\mathbb{R}^n$ de forma a definir as soluções do problema multiobjetivo. Métodos clássicos de otimização multiobjetivo e métricas de desempenho e convergência foram colocados nessa dissertação a fim de enriquecer o texto. Também apresenta-se condições de otimalidade para problemas escalares com essa teoria. Por fim utilizou-se condições de regularidade e de qualificação para a função objetivo envolver-se no processo de determinação de condições necessárias, as condições de KKT.

Sugestões para trabalhos futuros

Com o intuito de progredir na pesquisa, abordando aspectos não estudados na presente dissertação, faz-se, em seguida, algumas sugestões e considerações para trabalhos futuros.

- Investigar condições suficientes de otimalidade para problemas de otimização multiobjetivo com restrições de igualdade e de desigualdade.

- Explorar condições de otimalidade para problemas de otimização multiobjetivo no contexto da teoria dos cones em problemas não-diferenciáveis.

Referências Bibliográficas

[1] Abadie, J. *On the KuhnTucker theorem*. In: Abadie, J. (ed.) Nonlinear Programming, pp. 2136. North-Holland, Amsterdam, 1967.

[2] Aghezzaf, B. *Second-order necessary conditions of the KuhnTucker type in multiobjective programming problems*. Control Cybern.28, p.213 à 224, 1999.

[3] Aghezzaf, B., Hachimi, M. *Second-order optimality conditions in multiobjective optimization problems*. J. Optim. Theory Appl.102, p.37 à 50, 1999.

[4] Aghezzaf, B., Hachimi, M. *New Results on Second-Order Optimality Conditions in Vector Optimization Problems*. J. Optim. Theory Appl.135, p.117 à 133, 2007.

[5] Arroyo, J.E.C., *Heurísticas e metaheurísticas para otimização combinatória multiobjetivo*. Tese (Doutorado em Engenharia Elétrica)- Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 2002.

[6] Azuma, Regina Mitsue, *Otimização multiobjetivo em problema de estoque e roteamento gerenciados pelo fornecedor*. Campinas, SP: [S.N] 2011.

[7] Bazaraa M.S., Sherali H.D., Shetty C.M., *Nonlinear Programing: Theory and Algorithms*, Second Edition, John Wiley & Sons, Inc., New York, 1993.

[8] Beyer, H.G., *The theory of evolution strategies*. Natural Computing Series. Springer-Verlag, 2001.

[9] Bigi, G., Castellani, M. *Second-order optimality conditions for differentiable multiobjective problems*. RAIRO Oper. Res.34, 411426, 2000.

[10] Cambini, A., Martein, L., Vlach, M. *Second-order tangent sets and optimality conditions*. Research Report, Japan Advanced Studies of Science and Technology, Hokuriku, Japan 1997.

[11] Coello Coello, C. A., *A comprehensive survey of evolutionary-based multiobjective optimization techniques*. Knowledge and information systems, 1, p.269 à 308, 1999.

[12] Coello Coello, C.A., *Evolutionary Multi-Objective Optimization: a historical view of the field*. Computacional Intelligence Magazine, IEE, p.28 à 36, 2006.

[13] Charnes A., Cooper W.W.. *Management Models and Industrial Applications os Linear Programming*, v.1, John Wiley & Sons, Inc., New York, 1961.

[14] Chen G.Y., *Conditions of Nondominated Solucions in Multicriteria Decision Making*. Journal of Mathematical Analysis and Applications 104, n.1, p.38-46, 1984.

[15] Deb, K., *Multi-Objective Optimization Using Evolucionary Algorithms*. Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons, Chicester, 2001.

[16] Deb, K.Pratap, A., Agarwal, S., & Meyarivan, T.. *A fast and elitist multiobjective genetic algorithm: NSGA-II*. Ieee transactions on evolutionary computation, 6(2), p.182 à 197, 2002.

[17] Edgeworth, F.Y., *Mathematical Physics*, P.Keagan, London, England, 1881.

[18] Ferreira, P.A.V., *Introdução à Programação Multiobjetivo*. [Campinas], 1999. Disponível em: <http://www.dt.fee.unicamp.br/ valente/ia527.html>. Acesso em: 18 nov. 2012.

[19] Fonseca C.M. & Fleming P.J., *Genetic Algorithms for Multiobjective Optimization: formulation, discussion e generalização*. Pages 416-423 of:Proceedings of the 5th international conference and genetic algorithms. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

[20] Guignard, M. *Generalized KuhnTucker conditions for mathematical programming problems in a Banach space*. SIAM J. Control7, 232 à 241, 1987.

[21] Haimes Y.Y., Lasdon L.S., Wismer D.A.. *On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization*, IEE Transactions on Systems, Man, and Cybernetics 1, p.297-297, 1971.

[22] Holland J.H., *Adaptation in Natural and Artificial Systems*. Ann Arbor, MIT Press, 1975.

[23] Horn J., Nafpliotis N., & Goldberg, D.E., *A Niched Pareto Genetic Algorithm for Multiobjective Optimization*. Evolutionary Computation 1, p.82-87, 1994.

[24] Hwang C.L., Masud A.S.M.. *Multiple Objective Decision Making - Methods and Aplications*, A State-of-the-Art Survey, Lecture Notes in Economics and Mathematical Systems 164, Springer-Verlag, Berlin, Heidelberg, 1979.

[25] Izmailov, A.; Solodov, I., *Otimizacão-volume 1: Condições de otimalidade, Elementos de Análise Conveza e de Dualidade*, 2.ed. Rio de Janeiro: IMPA, 2009.

[26] Izmailov, A.; Solodov, I., *Otimizacão-volume 2: Métodos Computacionais*. Rio de Janeiro: IMPA, 2007.

[27] Janhn J.. *Parametric Approximation Problem Arising in Vector Optimization*. Journal of Optimization Theory and Applications 54, n.3, p.503-516, 1987.

[28] Lin, J.G.. *Maximal vectors and multiobjective optimization*. J. Optim. Theory Appl.18, 4164, 1976.

[29] Marusiac, I. *On Fritz John Type Optimality Criterion in Multiobjective Optimization*, Analyse Numerique et Theorie del' Approximation, Vol.11, pp.109-114, 1982.

[30] Mavrotas G., *Generation os efficient solutions in Multiobjective Mathematical Programming problems using GAMS. Effective implemation of the $\epsilon -$ constraint method*. Lecturer, Laboratory of Industrial and Energy Economics, School of Chemical Engineering National Technical University of Athens, Zografou Campus, Athens 15780, Greece.

[31] Miettinen, K., *Nonlinear Multiobjective Optimization*. Boston, Massachusets: Kluwer Academic Publishers, 1999.

[32] Michalewicz Z., *Genetic Algorithms, Numerical Optimization and Constraints*. Proceedings of the 6th International Conference on Genetic Algorithms, Pittsburgh, 151-158, 1995.

[33] Pareto, V., *Manual de economia política*. Tradução de João Guilherme Vargas Netto. [S.I.]: Editora Nova Cultural, 1996.

[34] Pascoletti A., Serafini P.. *Scalarizing Vector Optimization Problems*. Journal of Optimization Theory and Applications 42. No. 4, p.499-524 1984.

[35] Penot, J.P. *Second-order conditions for minimization problems with constraints*. SIAM J. Control Optim.37, p.303318, 1998.

[36] Romero, C. *A General Structure of Achievement Function for a Goal Programing Model*. European Journal of Operational Research 153, v.3, p.675-686, 2004.

[37] Sawaragi Y., Nakayama H., Tanino T. *Theory of Multiobjective Optimization*, Academic Press, Inc., Orlando, Florida, 1985.

[38] Silva Neto, G.C., *Um Método para Solução de Problemas de Otimização Multiobjetivo em Ambiente Fuzzy*. Manaus, 2011. Dissertação de Mestrado. PPGM - Universidade Federal do Amazonas.

[39] Singh, C.. *Optimality conditions in multiobjective differentiable programming*. J. Optim. Theory Appl.53, 115123, 1987.

[40] Srinivas, N., & Deb, K., *Multiobjective Optimization using Nondominated Sorting in Genethic Algorithms*. Evolutionary Computation 2, p.221-242, 1994.

[41] Syswerda G., *Uniform Crossover in Genetic Algorithms.* Proceedings of the Third International Conference on Genetic Algorithms, Morgan Kaufmann Publishers, 2-9, 1989.

[42] Tapia C.G., Murtagh B.A.. *The Use of Preference Criteria in Interacive Multiobjective Mathematical Programing.* Asia-Pacific Journal of Operacional Research 6. No.2, p.131-147, 1989.

[43] Zitzler, E., & Thiele, L., *Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach.* Ieee trans. evolutionary computation, 3(4), p.257 à 271, 1999.