



COPPE/UFRJ

INFORMAÇÕES TRADE-OFF EM PROBLEMAS DE OTIMIZAÇÃO
MULTIOBJETIVO

Jesus Ossian da Cunha Silva

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

Orientador: Paulo Roberto Oliveira

Rio de Janeiro

Abril de 2009

INFORMAÇÕES TRADE-OFF EM PROBLEMAS DE OTIMIZAÇÃO
MULTIOBJETIVO

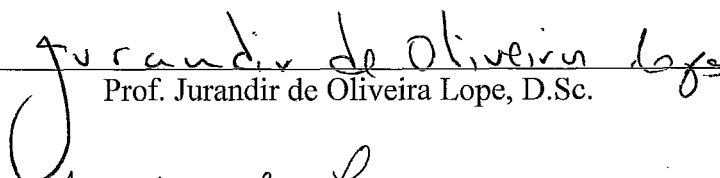

Jesus Ossian da Cunha Silva

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

Aprovada por:



Prof. Paulo Roberto Oliveira, Dr. Ing.

Prof. Jurandir de Oliveira Lope, D.Sc.

Prof. Márcia Helena Costa Fampa, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

ABRIL DE 2009

Silva, Jesus Ossian da Cunha

Informações Trade-off em Problemas de Otimização Multiobjetivo/ Jesus Ossian da Cunha Silva. – Rio de Janeiro: UFRJ/COPPE, 2009.

X, 68 p.: il.; 29,7 cm.

Orientador: Paulo Roberto Oliveira

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia de Sistemas e Computação, 2009.

Referencias Bibliográficas: p. 66-68.

1. otimização multiobjetivo. 2. trade-off. 3. tomada de decisões. 4. multiplicadores de Karush-Kuhn-Tucker. 5. cone de direções trade-off. I. Oliveira, Paulo Roberto. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*A todos que me ajudaram ao
longo desses anos.
A meu avô Alberto Cunha (in
memoriam).*

Agradecimentos

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro.

Agradeço ao Prof. Paulo Roberto Oliveira pela orientação durante o mestrado no PESC-COPPE-UFRJ.

Agradeço aos professores Márcia Fampa e Jurandir Oliveira por terem aceito participar da banca de defesa desta dissertação de mestrado.

Agradeço aos professores e amigos da UFPI, Barnabé Pessoa Lima, Gilvan Lima, João Xavier e Paulo Santos.

Agradeço aos amigos de turma durante o mestrado, Jurair, Michael, Paco e Thiago.

Agradeço aos amigos que conheci no Rio em especial Fabio Coutinho, Leizer Lima, Taisa e Vinicius Ramos.

Agradeço aos professores e funcionários do PESC-COPPE-UFRJ que de alguma forma contribuíram para a realização desse trabalho.

E sem dúvidas, agradeço a minha família, mãe, pai e manos.

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

INFORMAÇÕES TRADE-OFF EM PROBLEMAS DE OTIMIZAÇÃO MULTIOBJETIVO

Jesus Ossian da Cunha Silva

Abril/2009

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Problemas de otimização multiobjetivo caracterizam-se por terem como solução um conjunto de pontos não comparáveis entre si. Caso se deseje uma solução única é necessário alguma informação auxiliar que possa ajudar a decidir qual o melhor entre os pontos soluções. Essa informação pode ser obtida da própria estrutura do problema, seja com auxílio de programas de computadores ou de grupos de pessoas. Ela deve ser avaliada por uma pessoa ou grupo que tenham amplos conhecimentos do problema de modo que se possa determinar o melhor ponto solução.

O principal objetivo desse trabalho é o estudo de ferramentas matemáticas e metodologias que possam auxiliar no processo de tomada de decisões em problemas de otimização multiobjetivo de forma que possamos chegar a uma solução desejável.

Em otimização os multiplicadores de Karush-Kuhn-Tucker e cones são ferramentas matemáticas importantes na análise das condições de otimalidade.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

INFORMATION TRADE-OFF IN MULTIOBJECTIVE OPTIMIZATION
PROBLEMS

Jesus Ossian da Cunha Silva

April/2009

Advisor: Paulo Roberto Oliveira

Department: Systems Engineering and Computer Science

Multiobjective optimization problems are characterized by having as a solution set of points not comparable with each other. If a single solution is needed some information may help to decide on the best among points solutions. Such information can be obtained from the very structure of the problem with the aid of program computers or groups of people. This information should be evaluated by a person or group who have extensive knowledge of the problem in order to determine the best solution point.

The main objective of this work is the study of mathematical tools and methodologies that can help in decision-making in multiobjective optimization problems, so we can reach a desirable solution.

In optimization of the multiplier Karush-Kuhn-Tucker and cones are important mathematical tools in the analysis of the conditions of optimality.

Sumário

Lista de Figuras	x
1 Introdução	1
2 Otimização Multiobjetivo e Tomada de Decisões	3
2.1 Relação de ordem	4
2.2 Otimalidade de Pareto	5
2.3 Função Utilidade	7
2.4 Tomada de decisão	8
2.5 Um exemplo	9
2.6 Classificação dos métodos	10
3 Geração do conjunto solução	13
3.1 Método de ponderação	13
3.2 Método ϵ -restrição	17
3.2.1 O ϵ adequado	21
3.2.2 Condições de Otimalidade para o método ϵ -restrição	23
4 Trade-off e os multiplicadores de Karush-Kuhn-Tucker	25
4.1 Taxa <i>trade-off</i>	25
4.2 <i>Trade-off</i> como multiplicadores de Karush-Kuhn-Tucker	29
5 Direções trade-off em problemas de otimização multiojetivo convexos	40
5.1 Otimalidade de Pareto	40
5.2 Cone tangente e cone normal	41
5.3 Cone de direções <i>trade-off</i>	45

5.3.1	Direções trade-off em MOP convexos	47
5.3.2	Um exemplo de cálculo de direções <i>trade-off</i>	48
6	Conclusões e Trabalhos Futuros	55
7	Apêndice	57
7.1	Condições de otimalidade	57
7.2	Convexidade	59
7.3	Teorema da Função Implícita e Teorema da Sensibilidade	62
	Referências Bibliográficas	66

Lista de Figuras

2.1	Pareto ótimo próprio	6
2.2	Pareto ótimo em S	10
2.3	Pareto ótimo em Z	10
3.1	Geometria do método ponderação	16
3.2	Geometria do método ϵ -restrição	17
3.3	Gráfico do $P_1(\epsilon)$	20
3.4	Gráfico do exemplo 3.2	22
4.1	Relação entre taxa marginais de substituição e taxa trade-off	28
4.2	Solução gráfica para $P_1^0(2, 5; 3)$	35
4.3	Solução gráfica para $P_1^1(1, 5; 1, 5)$	36
4.4	Solução gráfica para $P_1^2(2, 51; 3)$	39
5.1	Cone tangente do conjunto Z	42
5.2	Cone normal do conjunto Z	43
5.3	Conjunto Pareto ótimo exemplo 3.1	49
5.4	Região definida pelas desigualdades 5.4 e 5.5	51
7.1	Gráfico do exemplo 7.1	59

Capítulo 1

Introdução

Em situações reais, é usual que os modelos matemáticos tenham a função objetivo definida não por uma única função, mas por um vetor constituído de várias funções conflitantes entre si. Estes modelos são definidos em espaços tipo \mathbb{R}^k , onde temos k funções componentes do vetor definido como objetivo. Eles são denominados problemas de otimização multiobjetivo(MOP).

Outra característica é o fato de não existir uma solução única para os mesmos. Além disto, as soluções são dadas por um conjunto onde seus elementos não possuem relação de ordem completa entre si. Diante dessa situação é preciso algum instrumento para estabelecer alguma relação de ordem, de forma que seja possível comparar os pontos em questão. Supondo que se queira um único ponto como solução, surge assim uma questão: qual é a melhor solução?

Esse novo problema é denominado problema de tomada de decisões multiobjetivo(MDMP) . Nele necessitamos de informações que possam auxiliar na escolha da melhor solução de acordo com algum conceito estabelecido. Nos problemas de tomada de decisões multiobjetivo precisamos de alguém para determinar essas informações, no caso o analista(pessoa ou programa de computador) que será o responsável por essa tarefa, em seguida decisões devem ser tomadas por uma pessoa(ou grupo) que será denominado decidor. Nesse trabalho denominaremos essas informações de *trade-off*.

Individuos tentam escolher a melhor decisão dentro de um conjunto de alternativas possíveis. Historicamente, a noção de "melhor" tem sido definida de modos diferentes em diferentes campos. Em economia onde a otimização multiobjetivo

provavelmente se originou, o "melhor" refere-se às decisões tomadas por compradores e vendedores (micro-economia) e governantes (macro-economia), o qual otimize simultaneamente vários critérios.

O crescimento dessa área manifestou-se principalmente nos EUA com contribuições de STADLER, STEUER e muitos outros. Os japoneses têm dado uma grande contribuição particularmente no que se refere aos aspectos teóricos (SAWARAGI, NAKAYAMA e TANINO). Outro grupo forte nessa área se encontra na Finlândia (KAISA e MÄKELÄ etc), com interesse em otimização multiobjetivo, métodos iterativos, tomada de decisões multicritério e metaheurística.

No capítulo 2 definimos o problema de otimização multiobjetivo e soluções Pareto. Comentamos ainda relações de ordem em espaços tipo \mathbb{R}^k , função utilidade e tomada de decisões multiobjetivo.

No capítulo 3 definimos dois métodos para resolução de problemas multiobjetivo e estabelecemos uma comparação entre o conjunto solução dos mesmos e o conjunto solução Pareto ótimo do problema de otimização multiobjetivo.

No capítulo 4 definimos taxa *trade-off* levando em consideração funções, pontos viáveis e direções viáveis e depois estabelecemos uma relação entre essas taxas e os multiplicadores de Karush-Kuhn-Tucker, o que é de grande utilidade no processo de tomada de decisões.

No capítulo 5 definiremos cones, direções viáveis, cones tangente, cone normal e cone de direções *trade-off*, e os caracterizamos. Relacionaremos cones tangentes e cones normais com conjunto Pareto ótimo. O conteúdo de estudo nesse capítulo é baseado na teoria de Análise Convexa de ROCKAFELLAR [1].

Capítulo 2

Otimização Multiobjetivo e Tomada de Decisões

Este capítulo se baseia em MIETTINEN [2], HAIMES AND CHANKONG [3] e FERREIRA [4]. Definimos o problema de otimização multiobjetivo e soluções Pareto, comentamos sobre relações de ordem em espaços tipo \mathbb{R}^k , função utilidade e tomada de decisões multiobjetivo

Problemas de otimização multiobjetivo caracterizam-se por possuírem a função objetivo definida de \mathbb{R}^n em \mathbb{R}^k , onde $n \geq 1$ e $k \geq 1$.

Denotamos um problema de otimização multiobjetivo (MOP) o seguinte

Problema 2.1 *[MOP]*

$$\begin{aligned} \min \quad & \{f_1(x), \dots, f_k(x)\} \\ \text{s. a } & x \in S, \end{aligned}$$

onde temos k funções objetivos $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto viável $S \neq \emptyset$. Denotamos $x \in S \subset \mathbb{R}^n$ como vetor de decisão, onde os x_i componentes são denominados variáveis de decisão.

A imagem do conjunto viável é denotada por $Z \subset \mathbb{R}^k$ e a definimos como conjunto objetivo viável. Os elementos de Z são denominados vetores objetivos e denotados por $z = f(x) = (f_1(x), f_2(x), \dots, f_k(x))^T$, onde cada $z_i = f_i(x)$ é denominada variável objetivo. Assim temos que $Z = f(S)$.

2.1 Relação de ordem

Definição 2.1 . Um conjunto Z é completamente ordenado de acordo com a relação de preferência " \leq " se dados dois elementos z^1 e $z^2 \in Z \subset \mathbb{R}^k$ é sempre verdade que $z^1 \leq z^2$ ou $z^2 \leq z^1$ com relação a " \leq " satisfazendo as seguintes propriedades

1. $z^1 \leq z^1$;
2. $z^1 \leq z^2$ e $z^2 \leq z^3$ então $z^1 \leq z^3$;
3. $z^1 \leq z^2$ e $z^2 \leq z^1$ então $z^1 = z^2$.

Em problemas de otimização multiobjetivo temos em geral a inexistência de soluções ótimas no sentido usual, isto é, não existe $x^* \in S$ tal que $f(x^*) \leq f(x)$ para todo $x \in S \cap B(x^*; r)$, onde $B(x^*; r) := \{x \in X : \|x^* - x\| < r\}$, qualquer que seja $r > 0$.

Abaixo caracterizamos o conjunto $S \subset \mathbb{R}^n$ de acordo com a relação de ordem entre seus elementos.

Definição 2.2 . Para um ponto qualquer $x^* \in \mathbb{R}^n$ temos

$$S_{<}(x^*) := \{x \mid f(x) \leq f(x^*) \text{ e } f(x) \neq f(x^*)\},$$

$$S_{\geq}(x^*) := \{x \mid f(x) \geq f(x^*)\},$$

$$S_{\sim}(x^*) := \{x \mid f(x) \not\leq f(x^*) \text{ e } f(x) \not\geq f(x^*)\}.$$

Observe que,

1. $\mathbb{R}^n = S_{<}(x^*) \cup S_{\geq}(x^*) \cup S_{\sim}(x^*)$, para todo $x^* \in \mathbb{R}^n$;
2. $S_{<}(x^*)$ e $S_{\geq}(x^*)$ contém pontos de S que dominam e são dominados por x^* de acordo com a relação " \leq ";
3. $S_{\sim}(x^*)$ contém todos os pontos não-comparáveis com relação a x^* ;
4. O ponto $x^* \in S$ candidato a solução desejável deverá ser tal que $S \cap S_{<}(x^*) = \emptyset$.

2.2 Otimalidade de Pareto

No processo de obtenção de solução de problemas de otimização escalar costumamos trabalhar no espaço das variáveis de decisão. No caso de problemas de otimização multiobjetivo é mais interessante trabalhar no espaço dos objetivos. Uma das razões é que o espaço objetivo costuma ter dimensão inferior ao espaço decisão, e além do mais, tem-se um maior controle no espaço objetivo ao longo do procedimento de obtenção de solução.

A seguir definiremos o conjunto solução dos problemas de otimização multiobjetivo.

Definição 2.3 . *Dado $x^* \in S$, temos que x^* é uma solução Pareto ótimo do Problema 2.1, se não existe outro $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ e no mínimo uma das desigualdades é estrita. Um vetor $f(x^*)$ é Pareto ótimo se o correspondente vetor decisão $x^* \in S$ é Pareto ótimo. Denote S^* o conjunto Pareto ótimo.*

Abaixo apresentamos algumas variantes da definição de Pareto ótimo.

Definição 2.4 . *Um vetor decisão $x^* \in S$ é Pareto ótimo fraco se não existe outro $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$. Um vetor objetivo $f(x^*)$ é Pareto ótimo fraco se o vetor correspondente $x^* \in S$ é Pareto ótimo fraco.*

Definição 2.5 . *Um vetor decisão $x^* \in S$ é Pareto ótimo próprio se ele é Pareto ótimo e se existe um $M > 0$ tal que para cada f_i e cada $x \in S$ satisfazendo $f_i(x) < f_i(x^*)$, existe no mínimo um f_j tal que $f_j(x) > f_j(x^*)$ e*

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

para todo $i, j = 1, \dots, k$ e $i \neq j$. Um vetor objetivo $f(x^)$ é chamado Pareto ótimo próprio se o vetor correspondente $x^* \in S$ é Pareto ótimo próprio. Uma solução Pareto ótimo que não é Pareto ótimo próprio é denominada solução Pareto ótimo impróprio.*

Exemplo 2.1 *Considere o seguinte problema(HAIMES AND CHANKONG [3])*

$$\begin{aligned} & \text{minimizar } ((x - 2)^2 + 1, (x - 4)^2 + 2) \\ & \text{sujeito a } -x \leq 0 \end{aligned}$$

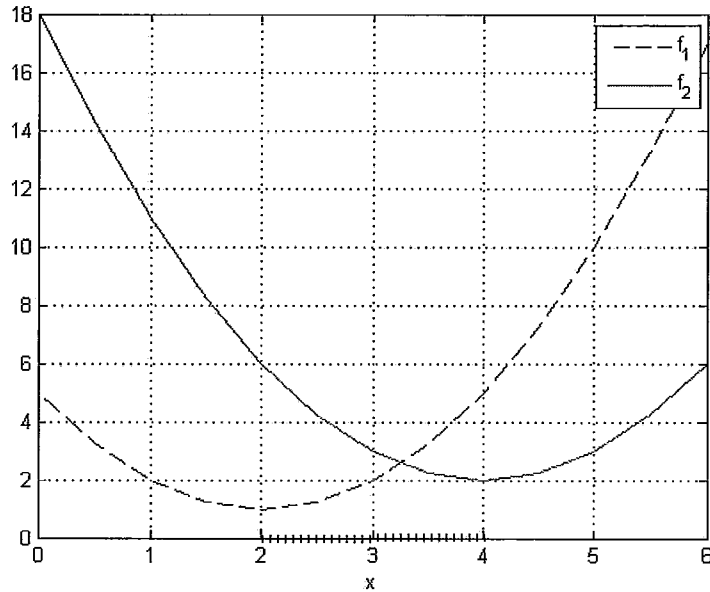


Figura 2.1: Pareto ótimo próprio

Denote $f_1 = (x - 2)^2 + 1$ e $f_2 = (x - 4)^2 + 2$. Temos que o conjunto Pareto ótimo está definido no intervalo $[2, 4]$ (Figura 2.1). O ponto $x = 2$ é um ponto Pareto ótimo impróprio já que para algum $x = 2 + h$, $h > 0$ temos que $f_2(x) < f_2(2)$, mas a taxa

$$\frac{f_2(x) - f_2(2)}{f_1(x) - f_1(2)} \rightarrow \infty \text{ com } h \searrow 0.$$

De modo análogo o ponto $x = 4$ é também Pareto ótimo impróprio enquanto os pontos no intervalo $(2, 4)$ são Pareto ótimo próprio.

Otimalidade de Pareto pode também ser definida em uma vizinhança $B(x^*; r)$ de x^* , onde $r > 0$.

Definição 2.6 . Um vetor decisão $x^* \in S$ é Pareto ótimo local se existe $r > 0$ tal que x^* é Pareto ótimo em $S \cap B(x^*; r)$.

Em problemas de otimização objetivo escalar, não há, em geral, soluções globais, e a condição usual para a existência de tais soluções é a convexidade. Também aqui necessitamos desta hipótese.

Dizemos que um problema de otimização multiobjetivo é convexo se S é um conjunto convexo e todas as f_j , $j = 1, \dots, k$ são funções convexas. O resultado a seguir foi retirado de MIETTINEN [2]

Teorema 2.1 . Se o problema de otimização multiobjetivo é convexo então toda solução de Pareto ótimo local é também Pareto ótimo global.

Prova 2.1 . Seja $x^* \in S$ Pareto ótimo local. Assim existe algum $r > 0$ e uma vizinhança $B(x^*; r)$ de x^* , tal que existe um $x \in S \cap B(x^*; r)$ para o qual não se tem $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$, e para no mínimo um índice j , $f_j(x) < f_j(x^*)$.

Assuma que x^* não é Pareto ótimo global. Neste caso existe algum outro vetor $x^0 \in S$ tal que

$$f_i(x^0) \leq f_i(x^*) \text{ para todo } i = 1, \dots, k, \text{ e} \quad (2.1)$$

$$f_j(x^0) < f_j(x^*) \text{ para algum } j. \quad (2.2)$$

Tome $\bar{x} = \beta x^0 + (1 - \beta)x^*$ onde $0 < \beta < 1$, tal que $\bar{x} \in B(x^*; r)$. Pela convexidade de S temos que $\bar{x} \in S$.

Pela convexidade das funções e equação 2.1 temos

$$\begin{aligned} f_i(\bar{x}) &\leq \beta f_i(x^0) + (1 - \beta)f_i(x^*) \\ &\leq \beta f_i(x^*) + (1 - \beta)f_i(x^*) \\ &= f_i(x^*) \end{aligned} \quad (2.3)$$

para todo $i = 1, \dots, k$. Devido a x^* ser Pareto ótimo local e $\bar{x} \in B(x^*; r)$, temos que $f_i(\bar{x}) = f_i(x^*)$ para todo i .

A equação 2.3 e $f_i(\bar{x}) = f_i(x^*)$ para todo i implica em $f_i(x^*) \leq \beta f_i(x^0) + (1 - \beta)f_i(x^*)$ para todo $i = 1, \dots, k$.

$$f_i(x^*) = f_i(\bar{x}) \leq \beta f_i(x^0) + (1 - \beta)f_i(x^*)$$

Usando o fato de $\beta > 0$ obtemos que $f_i(x^*) \leq f_i(x^0)$ para todo i . Na equação 2.2 temos $f_j(x^*) > f_j(x^0)$ para algum j , ou seja, temos uma contradição. Logo x^* é Pareto ótimo global.

c.q.d

2.3 Função Utilidade

Na seção 2.1 definimos relações de ordem em um subconjunto do \mathbb{R}^k . É possível estabelecer uma ordem de preferência entre os elementos de um subconjunto $Z \subset \mathbb{R}^k$ utilizando o conceito de função utilidade.

Definição 2.7 . *Uma função $U : \mathbb{R}^k \rightarrow \mathbb{R}$ que representa uma ordem de preferência entre os vetores objetivos é denominada função utilidade.*

Sejam dados z^1 e $z^2 \in Z$ dois diferentes vetores objetivos. Se ao aplicamos U em Z tal que $U(z^1) > U(z^2)$, então temos uma ordem de preferência onde z^1 é preferível a z^2 . Se $U(z^1) = U(z^2)$, então a escolha entre z^1 e z^2 é indiferente.

2.4 Tomada de decisão

Todo ponto Pareto ótimo é uma solução aceitável para o problema de otimização multiobjetivo. Sabemos também que cada uma das variáveis objetivos pode ser satisfeita por diferentes vetores decisão e sem uma relação de ordem bem definida.

O objetivo agora passa a ser a escolha de um ponto entre os vários pertencentes ao conjunto Pareto ótimo, o problema MDMP-Problema de Tomada de Decisões Multiobjetivo .

Resolver o MDMP é encontrar um ponto x^* Pareto ótimo, tal que x^* seja preferência do decisor.

Seguindo a abordagem utilizada em HAIMES AND CHANKONG [5] o processo solução do MDMP deve seguir três fases:

1. Gerar o conjunto solução(Pareto ótimo);
2. Obter uma informação significativa para tomada de decisão;
3. Usar a informação obtida na fase (2) para a seleção de uma solução final baseada em um critério de preferência.

No processo de solução dos problemas de tomada de decisões multiobjetivo necessitamos de um decisor e um analista. O decisor é a pessoa(ou grupo) que necessita ter conhecimento profundo do problema e que pode expressar uma preferência entre as soluções. Ele é o responsável pela solução final.

O analista pode ser uma pessoa(ou grupo) ou um programa de computador, responsável pela parte matemática do processo. O analista gera a informação a ser utilizada pelo decisor na seleção da melhor solução de acordo com sua preferência. Assim temos que o processo de tomada de decisões se dá através de uma cooperação entre decisor e analista.

No processo citado acima a fase (2) é realizada pelo analista ou programa de computador e a fase (3) é realizada pelo decididor.

Estaremos interessados apenas nas fases (1) e (2), sendo nosso foco principal as informações que podem ser geradas para ajudar ao decididor a escolher a melhor solução Pareto ótimo do problema de otimização multiobjetivo. Denominamos essas informações como "informações *trade-offs*".

2.5 Um exemplo

Analisemos o exemplo abaixo com duas funções definidas no vetor objetivo, ou seja, um vetor bi-objetivo. Trata-se de uma adaptação de exemplos de HAIMES AND CHANKONG [3].

$$\begin{aligned} & \text{minimizar } \{x, (x - 2)^2\} \\ & \text{sujeito a } x \geq 0, x \in \mathbb{R} \end{aligned}$$

Uma variável decisão $x \in [0, 2]$ é solução Pareto ótimo do problema, já que neste intervalo quando $f_1(x) = x$ cresce, $f_2(x) = (x - 2)^2$ decresce (Figura 2.2). Temos que o conjunto Pareto ótimo no espaço decisão é determinado pelo conjunto $\{x \mid x \in \mathbb{R}, 0 \leq x \leq 2\}$ e o conjunto Pareto ótimo no espaço objetivo é determinado pelo conjunto $\{f \mid f = (f_1, f_2), f_2 = (f_1 - 2)^2, 0 \leq f_1 \leq 2\}$. Temos que o conjunto Pareto ótimo no espaço objetivo traça a curva $f_2 = f_2(f_1)$, onde $f_2()$ é uma função estritamente decrescente de f_1 . Temos ainda que $\frac{df_2(f_1)}{df_1}$ deve ser negativa em todo ponto no conjunto Pareto ótimo no espaço objetivo, devido a um decréscimo em um objetivo poder ser alcançado somente com um acréscimo em outro objetivo. A quantidade $\frac{df_2(f_1)}{df_1}$ reflete a taxa *trade-off*. Para o problema com mais de dois objetivos (problema de minimização), o conjunto Pareto ótimo no espaço objetivo deve ser uma hipersuperfície na fronteira de Z com a propriedade $\frac{\partial f_k}{\partial f_j} < 0$ para no mínimo um par de $1 \leq j \neq k \leq m$.

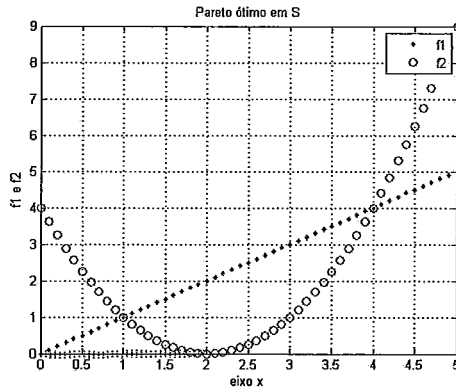


Figura 2.2: Pareto ótimo em S

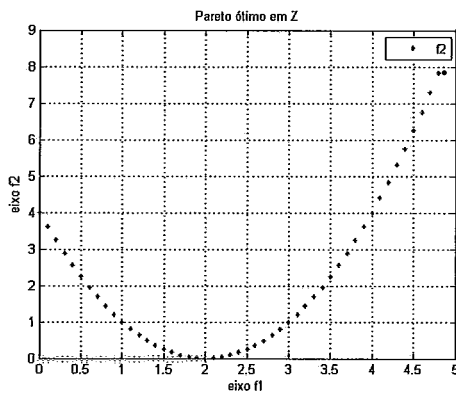


Figura 2.3: Pareto ótimo em Z

2.6 Classificação dos métodos

Métodos de otimização multiobjetivo podem ser classificados de vários modos, de acordo com diferentes critérios. HWANG e MASUD [6] classificam os métodos de acordo com a participação do decisor no processo solução.

Método não-preferência. A opinião do decisor não é levada em consideração no processo. O problema de otimização multiobjetivo é resolvido usando algum método relativamente simples e a solução obtida é apresentada ao decisor. O decisor pode aceitar ou rejeitar a solução. Essa classe de métodos é adequada para situações onde o decisor não tem expectativas sobre a solução e fica satisfeito com alguma solução ótima. A ordem de trabalho nessa classe de métodos se dá só pelo analista.

Método a posteriori. Após o conjunto Pareto ótimo ou parte dele ser gerado, ele

é apresentado ao decisor, que estabelecerá uma preferência entre as soluções. A desvantagem nessa classe de método é que o processo de geração de soluções é usualmente de alto custo computacional. Por outro lado é solicitado ao decisor seleção sobre um amplo conjunto de alternativas. Uma importante questão é como apresentar as alternativas ao decisor de um modo efetivo. A sequência do trabalho aqui se dá através de analista, decisor.

Método a priori. Nessa classe de método o decisor deve especificar sua preferência, expectativas e opinião antes do processo solução. A dificuldade é que o decisor não necessariamente conhece de antemão o que é possível atingir no problema e o quão realistas são suas expectativas. A sequência do processo segue a ordem: decisor, analista.

Método interativo. Nessa classe assume-se que o decisor possui capacidade para cooperação. Nela é possível se gerar e avaliar parte do conjunto Pareto ótimo, e o decisor pode especificar e corrigir suas preferências e seleção em um processo contínuo. Durante esse processo o decisor toma um melhor conhecimento do problema e o mesmo não precisa conhecer alguma estrutura de preferência global. Em adição, o decisor pode assumir que possui uma maior confiança na solução final já que está inteiramente envolvido em todo processo de solução.

O decisor deve trabalhar junto com o analista ou interagir com um software computacional. Podemos dizer que o analista procura determinar a estrutura de preferência ao decisor de um modo interativo. Após diversas interações, alguma informação é fornecida para o decisor e ele é convidado a responder algumas perguntas ou fornecer algum outro tipo de informação. O trabalho no método interativo segue a sequência analista, decisor, analista, decisor, ..., analista, decisor, analista, Após um número razoável e finito de iterações, cada iteração do método deverá produzir uma solução satisfatória ao decisor onde este deve estar convicto de que não existem soluções consideravelmente melhores. Os passos básicos no algoritmo interativo são:

1. encontrar a solução viável inicial,
2. interação com o decisor, e

3. obter uma nova solução ou um novo conjunto de novas soluções. Se a nova solução ou uma das soluções anteriores é aceitável ao decisor, o processo para. Caso contrário, volta-se para o passo 2).

Um tipo de problema a ser resolvido no momento de conceber um método interativo é qual tipo de dado deve ser usado na interação com o decisor. Ele deve ser significativo e fácil para o decisor compreender. Por outro lado os dados fornecidos para o decisor devem ser obtidos de forma fácil pelo analista e conter informações sobre o sistema. Excesso de informações não devem ser utilizadas e as obtidas a partir do decisor devem ser utilizadas de forma eficiente. Para garantir o maior benefício possível obtido a partir do método interativo, o decisor tem que encontrar o método válido e aceitável, e ele deve ser capaz de utilizá-lo de forma correta.

Capítulo 3

Geração do conjunto solução

Este capítulo baseia-se nos trabalhos HAIMES AND CHANKONG [5], MIETTINEN[2] e HAIMES AND CHANKONG [3]. Nele definimos dois métodos para resolução de problemas multiobjetivo e estabelecemos uma comparação entre o conjunto solução dos mesmos e o conjunto solução Pareto ótimo do problema de otimização multiobjetivo

Em otimização multiobjetivo utilizam-se métodos de escalarização ou de aproximação durante o processo solução. Estaremos interessados apenas em métodos de escalarização. Para detalhes sobre métodos de aproximações veja MIETTINEN[2]. Em métodos de escalarização a idéia é transformar o problema de otimização multiobjetivo em um problema de otimização escalar, o qual possui métodos de resolução bem desenvolvidos. Neste capítulo realizaremos um estudo sobre o método de ponderação e o método ϵ -restrição, assim como relacionamos o conjunto solução dos mesmos com o conjunto Pareto ótimo do MOP.

3.1 Método de ponderação

No método de ponderação associamos a cada componente da função objetivo um coeficiente peso e minimizamos a soma dos objetivos ponderados. Suponha os coeficientes de ponderação $w_i \in \mathbb{R}$ tal que $w_i \geq 0$ para todo $i = 1, \dots, k$ onde $\sum_{i=1}^k w_i = 1$. Assim temos que o problema de otimização multiobjetivo(Problema 2.1) transforma-se no seguinte

Problema 3.1 .[Problema de ponderação]

$$\begin{aligned} & \text{minimizar } \sum_{i=1}^k \omega_i f_i(x) \\ & \text{sujeito a } x \in S, \end{aligned}$$

onde $\omega_i \geq 0$ para todo $i = 1, \dots, k$ e $\sum_{i=1}^k \omega_i = 1$.

Em seguida caracterizamos as relações entre as soluções do problema de ponderação e as soluções Pareto do problema de otimização multiobjetivo. Serviram como fonte para os resultados abaixo, HAIMES AND CHANKONG [5] e MIETTINEN[2].

Teorema 3.1 . A solução do problema de ponderação é Pareto ótimo fraco.

Prova 3.1 . Seja $x^* \in S$ uma solução do problema de ponderação. Suponha que x^* não seja Pareto ótimo fraco. Nesse caso existe um $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$. De acordo com as hipóteses do conjunto dos coeficientes de ponderação, $\omega_j > 0$ para no mínimo um j . Temos que

$$\sum_{i=1}^k \omega_i f_i(x) < \sum_{i=1}^k \omega_i f_i(x^*),$$

um absurdo, já que por hipótese x^* é uma solução do problema de ponderação. Logo x^* é Pareto ótimo fraco.

c.q.d

Teorema 3.2 . A solução do problema de ponderação é Pareto ótimo se os coeficientes de ponderação são todos positivos.

Prova 3.2 . Seja $x^* \in S$ uma solução do problema de ponderação com todos coeficientes de ponderação positivos. Suponha que x^* não seja Pareto ótimo, assim como no Teorema 3.1 temos um absurdo. Logo x^* é Pareto ótimo.

c.q.d

Teorema 3.3 . Se a solução do problema de ponderação for única, então é Pareto ótimo.

Prova 3.3 . Seja $x^* \in S$ uma solução única do problema de ponderação. Suponha que x^* não é Pareto ótimo. Nesse caso, existe uma solução $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ e $f_j(x) < f_j(x^*)$ para no mínimo um índice j . Usando a hipótese de $\omega_i \geq 0$ para $i = 1, \dots, k$, temos

$$\sum_{i=1}^k \omega_i f_i(x) \leq \sum_{i=1}^k \omega_i f_i(x^*) \quad (3.1)$$

Por outro lado a unicidade de x^* significa,

$$\sum_{i=1}^k \omega_i f_i(x^*) < \sum_{i=1}^k \omega_i f_i(\bar{x}) \quad (3.2)$$

para todo $\bar{x} \in S$. Como as desigualdades (3.1) e (3.2) são contraditórias entre si, temos o absurdo e x^* é Pareto ótimo.

c.q.d

Teorema 3.4 . A solução do problema de ponderação é Pareto ótimo próprio se todos os coeficientes de ponderação são positivos.

Prova 3.4 . Seja $x^* \in S$ uma solução do problema de ponderação com todos os coeficientes de ponderação positivos. Suponha que x^* não é Pareto ótimo. Então temos que para algum $i \in \{1, \dots, k\}$ e $x \in S$ tal que $f_i(x^*) > f_i(x)$

$$f_i(x^*) - f_i(x) > M(f_j(x) - f_j(x^*))$$

para todo j tal que $f_j(x^*) < f_j(x)$, onde

$$M = (k - 1) \max_{i,j} \left(\frac{\omega_j}{\omega_i} \right) \quad (3.3)$$

. Podemos então escrever

$$f_i(x^*) - f_i(x) > (k - 1) \frac{\omega_j}{\omega_i} (f_j(x) - f_j(x^*)).$$

Multiplicando $\frac{\omega_i}{k - 1} > 0$ em ambos os lados obtemos,

$$\frac{\omega_i}{k - 1} (f_i(x^*) - f_i(x)) > \omega_j (f_j(x) - f_j(x^*)) (> 0 \geq \omega_l (f_l(x) - f_l(x^*))).$$

onde l difere de índice i fixo e dos índices j , os quais foram especificados pela equação

3.3. Realizando uma soma sobre todos os $j \neq i$ obtemos

$$\frac{(k-1)\omega_i}{k-1}(f_i(x^*) - f(x)) \geq \sum_{\substack{j=1 \\ j \neq i}} \omega_j(f_j(x) - f_j(x^*))$$

ou seja,

$$\begin{aligned} \omega_i f_i(x^*) - \omega_i f(x) &\geq \sum_{\substack{j=1 \\ j \neq i}} \omega_j f_j(x) - \sum_{\substack{j=1 \\ j \neq i}} \omega_j f_j(x^*) \\ \omega_i f_i(x^*) + \sum_{\substack{j=1 \\ j \neq i}} \omega_j f_j(x^*) &\geq \sum_{\substack{j=1 \\ j \neq i}} \omega_j f_j(x) + \omega_i f(x) \\ \sum_{j=1}^k \omega_j f_j(x^*) &> \sum_{j=1}^k \omega_j f_j(x) \end{aligned}$$

Trata-se de uma contradição o fato de que x^* é uma solução do problema de ponderação. Logo x^* é Pareto ótimo próprio.

c.q.d

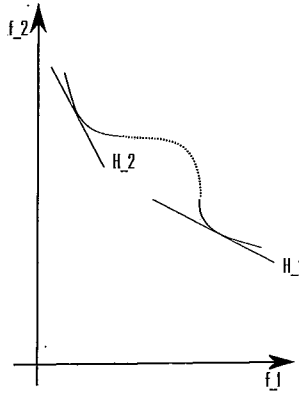


Figura 3.1: Geometria do método ponderação

O método ponderação funciona bem no caso em que todas as funções componentes do vetor objetivo são convexas, caso contrário somente parte do conjunto Pareto ótimo será gerado. Isto se deve ao fato do método ponderação gerar apenas pontos do conjunto Pareto ótimo que admitem hiperplanos-suporte, sendo que esses que estão na fronteira do conjunto imagem(Figura 3.1).

3.2 Método ϵ -restrição

No método ϵ -restrição um componente do vetor objetivo é selecionado para ser otimizado e todos os outros componentes objetivos são convertidos em restrições com a determinação de um limite superior para cada um deles. Temos que o problema de otimização multiobjetivo (Problema 2.1) transforma-se no seguinte,

Problema 3.2 .[Problema ϵ -restrição]

$$\text{minimizar } f_l(x)$$

$$\text{sujeito a } f_j(x) \leq \epsilon_j \text{ para todo } j = 1, \dots, k, j \neq l,$$

$$x \in S,$$

onde $l \in \{1, \dots, k\}$, $S_l(\epsilon) = \{x \mid f_j(x) \leq \epsilon_j, j \neq l\}$

e $\epsilon \in \epsilon = \{(\epsilon_1, \dots, \epsilon_{k-1}, \epsilon_{k+1}, \dots, \epsilon_m)^T \mid S_l(\epsilon) \neq \emptyset\}$. Para x ser viável ao problema ϵ -restrição é necessário que $x \in S \cap S_l(\epsilon)$.

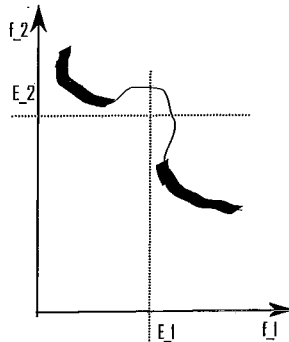


Figura 3.2: Geometria do método ϵ -restrição

A geração do ϵ_j não é tarefa fácil, pois o ponto obtido pode não ser Pareto ótimo e os valores do ϵ_j podem tornar o problema inviável. Na figura 3.2 por exemplo caso seja estabelecido simultaneamente as restrições $f_1 \leq \epsilon_1$ e $f_2 \leq \epsilon_2$ não existe solução viável, embora tomando cada restrição de forma isolada exista solução viável.

A seguir caracterizamos a relação entre as soluções do problema ϵ -restrição e as soluções Pareto do problema de otimização multiobjetivo. Os resultados abaixo

baseiam-se nos trabalhos HAIMES AND CHANKONG [5] e MIETTINEN [2]. Deixaremos de fazer algumas demonstrações devido a facilidade de construção das mesma, mas indicamos fontes para consultas caso seja necessário.

Teorema 3.5 . *A solução do problema ϵ -retrição é Pareto ótimo fraco.*

Prova 3.5 . *Seja x^* solução do problema ϵ -retrição. Suponha que x^* não é Pareto ótimo fraco. Segue que existe um $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$, ou seja, $f_j(x) < f_j(x^*) \leq \epsilon_j$ para todo $j = 1, \dots, k, j \neq l$. Assim x é viável com respeito ao problema ϵ -retrição. Temos ainda que $f_l(x) < f_l(x^*)$, o que contradiz a hipótese de x^* ser uma solução do problema ϵ -retrição. Logo x^* é Pareto ótimo fraco.*

c.q.d

Teorema 3.6 . *Um ponto $x^* \in S$ é Pareto ótimo do problema de otimização multi-objetivo se e somente se ele é solução do problema ϵ -retrição para todo $l = 1, \dots, k$, onde $\epsilon_j = f_j(x^*)$ para $j = 1, \dots, k, j \neq l$.*

Prova 3.6 . *A demonstração é de fácil construção, veja por exemplo Teorema 3.2.2 em MIETTINEN [2].*

c.q.d

Teorema 3.7 . *Um ponto $x^* \in S$ é Pareto ótimo se ele é solução única do problema ϵ -retrição para algum l com $\epsilon_j = f_j(x^*)$ para $j = 1, \dots, k, j \neq l$.*

Prova 3.7 . *Seja x^* uma solução única do problema ϵ -retrição para algum l , e x^* não seja Pareto ótimo. Logo existe um outro $\bar{x} \in S$ tal que $f_i(\bar{x}) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ e para no mínimo um índice j temos $f_j(\bar{x}) < f_j(x^*)$. A unicidade de x^* significa que para todo $x \in S$ tal que $f_i(x) \leq f_i(x^*)$, $i = 1, \dots, k$, $i \neq l$, vale $f_l(x^*) < f_l(x)$. Tem-se uma contradição com relação às desigualdades anteriores. Logo x^* é Pareto ótimo.*

c.q.d

Teorema 3.8 . *A solução única do problema ϵ -retrição é Pareto ótimo para algum dado vetor limite superior $\epsilon = (\epsilon_1, \dots, \epsilon_{l-1}, \epsilon_{l+1}, \dots, \epsilon_k)^T$.*

Prova 3.8 . A demonstração é de fácil construção, veja por exemplo Teorema 3.2.4 MIETTINEN [2].

c.q.d

Teorema 3.9 . O ponto $x^* \in S$ é Pareto ótimo local do problema de otimização multiobjetivo se e somente se x^* é uma solução local do problema ϵ -restrição para todo $l = 1, \dots, k$.

Prova 3.9 . Suponha que exista um l para o qual x^* não é solução local do problema ϵ -restrição, então para todo $r > 0$ existe $x \in S \cap B(x^*; r)$ tal que $f_l(x) < f_l(x^*)$ e $f_j(x) \leq f_j(x^*)$ para todo $j \neq l$. Portanto x^* não é Pareto ótimo local do problema ϵ -restrição.

Reciprocamente, se x^* é solução local do problema ϵ -restrição para todo $l = 1, \dots, k$, então para cada l existe $r_l > 0$ tal que x^* é o mínimo do problema ϵ -restrição em $S \cap B(x^*; r_l)$. Tome $r = \min \{r_1, \dots, r_k\} > 0$ fazendo x^* o mínimo do problema ϵ -restrição em $S \cap B(x^*; r)$ para todo $l = 1, \dots, k$. Portanto não existe $x \in S \cap B(x^*; r)$ tal que $f(x)_l \leq f_l(x^*)$ para todo $l = 1, \dots, k$. Logo x^* é Pareto ótimo local do problema de otimização multiobjetivo.

c.q.d

Definição 3.1 x^* é uma solução estrita do problema ϵ -restrição para algum l se não existe um $x \in S$ tal que $f_l(x) < f_l(x^*)$

Teorema 3.10 . Se x^* é uma solução estrita do problema ϵ -restrição para algum l então x^* é Pareto ótimo local do problema de otimização multiobjetivo.

Prova 3.10 . Veja Teorema 3 em HAIMES AND CHANKONG [5].

c.q.d

Exemplo 3.1 Considere o seguinte exemplo HAIMES AND CHANKONG [3],

$$\begin{aligned} & \text{minimizar } ((x_1 - 3)^2 + (x_2 - 2)^2, x_1 + x_2, x_1 + 2x_2) \\ & \text{sujeito a } -x_1 \leq 0 \\ & \quad \quad -x_2 \leq 0 \end{aligned}$$

Selecione $(x_1 - 3)^2 + (x_2 - 2)^2$ para ser minimizada e defina $x_1 + x_2$ e $x_1 + 2x_2$ como restrições limitadas por ϵ_2 e ϵ_3

$$\begin{aligned}
 P_1(\epsilon) \text{ minimizar } & (x_1 - 3)^2 + (x_2 - 2)^2 \\
 \text{sujeito a } & x_1 + x_2 \leq \epsilon_2 \\
 & x_1 + 2x_2 \leq \epsilon_3 \\
 & -x_1 \leq 0 \\
 & -x_2 \leq 0
 \end{aligned}$$

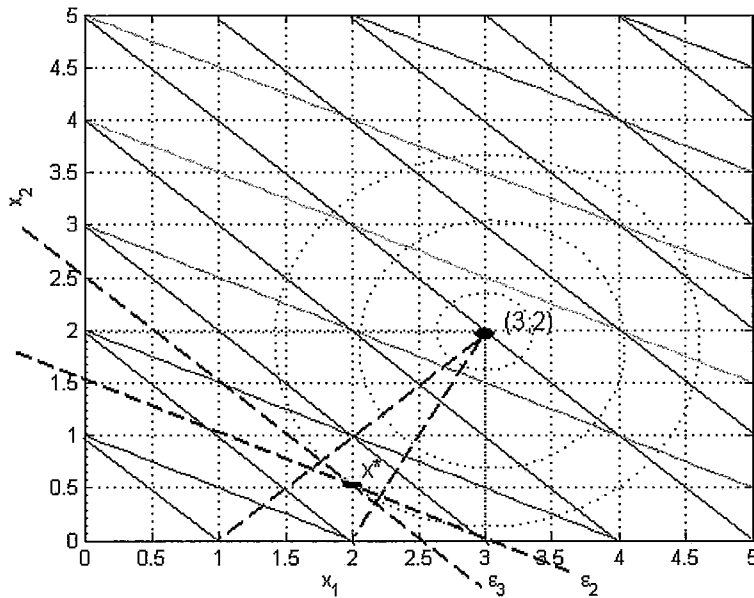


Figura 3.3: Gráfico do $P_1(\epsilon)$

Graficamente temos que a função objetivo do $P_1(\epsilon)$ define um círculo de centro no ponto $(3, 2)$. Resolver esse problema é equivalente a se encontram um ponto viável o mais próximo do ponto $(3, 2)$. De acordo com região viável determinada pelos limitantes ϵ_2 e ϵ_3 tem-se que x^* é uma solução do exemplo 3.3.

Observe que o conjunto Pareto ótimo é determinado por pontos no interior e na fronteira do triângulo definido pelos pontos extremos $(1, 0)$, $(2, 0)$ e $(3, 2)$ e o segmento entre $(0, 0)$ e $(1, 0)$.

3.2.1 O ϵ adequado

Na formulação do problema ϵ -restrição necessitamos de um vetor adequado $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_k)$ para limitar os componentes do vetor objetivo do problema de otimização multiobjetivo definidos como restrições no problema ϵ -restrição. Em FERREIRA [4] temos um novo problema originário do problema de otimização multiobjetivo e um método que fornecerá o vetor ϵ adequado para a formulação do problema ϵ -restrição.

Definição 3.2 . *Uma solução utópica $f^{**} = (f_1^{**}, f_2^{**}, \dots, f_k^{**})$ do problema de otimização multiobjetivo é definida como*

$$f_i^{**} = f_i(x^*), \quad i = 1, \dots, k$$

onde,

$$x^* = \arg \min_{x \in S} f_i(x)$$

É necessário encontrar valores adequados a inicialização do vetor $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_k)$. Resolvendo os k problemas mono-objetivos podemos obter valores ótimos individuais para cada objetivo f_i^* , $i = 1, \dots, k$, os quais compõem o objetivo f^{**} correspondente à solução utópica do problema.

Paralelamente ao determinamos o vetor f^{**} podemos selecionar os ótimos individuais de cada f_i e os aplicar em cada f_j , $j = 1, \dots, k$, $j \neq i$, determinando o vetor $f^0 = (f_1^0, f_2^0, \dots, f_k^0)$, onde o valor de cada f_i^0 , $i = 1, \dots, k$ correspondente a pior imagem de cada f_i , $i = 1, \dots, k$.

Em cada problema ϵ -restrição os limites ϵ_i das restrições referente a f_i , $i = 1, \dots, k$ são determinados através de um gerador de números aleatórios com distribuição de probabilidades uniforme atendendo à restrição $f^{**} \leq \epsilon \leq f^0$.

Exemplo 3.2 *Considere o exemplo 3.1. Encontremos a solução utópica,*

$$\begin{aligned} \text{minimizar } f_1 &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{sujeito a } &-x_1 \leq 0, \\ &-x_2 \leq 0. \end{aligned}$$

onde temos que $x^* = (3, 2)$ com $f_1^{**}(x^*) = 0$, $f_2^0(x^*) = 5$ e $f_3^0(x^*) = 7$.

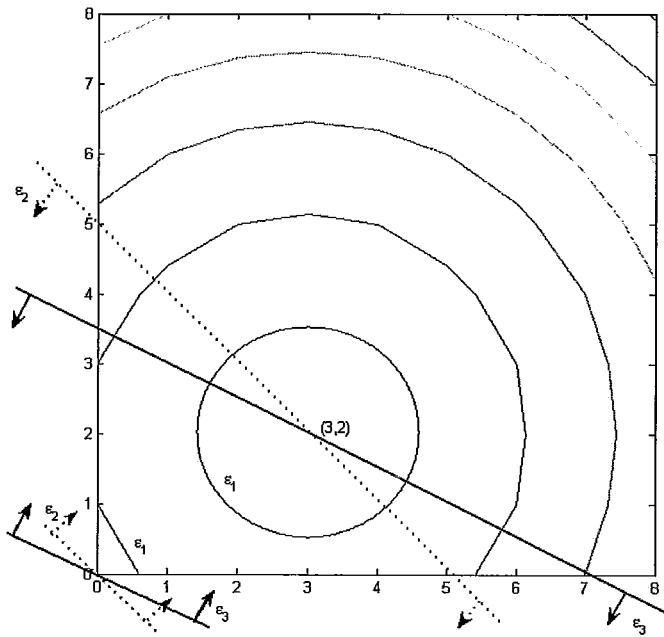


Figura 3.4: Gráfico do exemplo 3.2

$$\begin{aligned} & \text{minimizar } f_2 = x_1 + x_2 \\ & \text{sujeito a } -x_1 \leq 0, \\ & \quad \quad -x_2 \leq 0. \end{aligned}$$

onde temos que $x^* = (0, 0)$ com $f_1^0(x^*) = 13$, $f_2^{**}(x^*) = 0$ e $f_3^0(x^*) = 0$.

$$\begin{aligned} & \text{minimizar } f_2 = x_1 + 2x_2 \\ & \text{sujeito a } -x_1 \leq 0, \\ & \quad \quad -x_2 \leq 0. \end{aligned}$$

onde temos que $x^* = (0, 0)$ com $f_1^0(x^*) = 13$, $f_2^0(x^*) = 0$ e $f_3^{**}(x^*) = 0$.

Então temos $f^{**} = (0, 0, 0)$. Segue que $(0, 0, 0) \leq (\epsilon_1, \epsilon_2, \epsilon_3) \leq (13, 5, 7)$.

Computacionalmente o método ϵ -restrição não é simples devido ao grande número de operações, além disso temos que o mesmo possui restrições mais complexas que o problema original(MOP).

3.2.2 Condições de Otimalidade para o método ϵ -restrição

Suponha que a região viável S do problema de otimização multiobjetivo é da forma

$$S = \left\{ x \in \mathbb{R}^n \mid g(x) = \left(g_1(x), \dots, g_m(x) \right)^T \leq 0 \right\}$$

Associado ao problema ϵ -restrição temos a função Lagrangeana,

$$L(x, \lambda, \mu) = f_l(x) + \sum_{\substack{j=1 \\ j \neq l}}^k \lambda_j (f_j(x) - \epsilon_j) + \sum_{i=1}^m \mu_i g_i(x)$$

onde $\lambda \in \mathbb{R}^{k-1}$ e $\mu \in \mathbb{R}^m$ são os multiplicadores de Lagrange do problema ϵ -restrição.

Em seguida formulamos as condições necessárias de Karush-Kuhn-Tucker aplicadas ao método ϵ -restrição.

Observação 3.1 . *Suponha que a função objetivo e todas as restrições do problema ϵ -restrição são de classe C^1 em $x^* \in S$, um ponto regular das restrições do problema. Uma condição necessária para que x^* seja solução do problema ϵ -restrição é que existam vetores $0 \leq \lambda \in \mathbb{R}^{k-1}$ e $0 \leq \mu \in \mathbb{R}^m$ tais que*

1. $\nabla f_l(x^*) + \sum_{\substack{j=1 \\ j \neq l}}^k \lambda_j \nabla (f_j(x^*) - \epsilon_j) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) = 0$
2. $\lambda_j (f_j(x^*) - \epsilon_j) = 0$ para todo $j = 1, \dots, j \neq l$ e $\mu_i g_i(x^*) = 0$ para todo $i = 1, \dots, m$.

Os λ_j e μ_i são denominados multiplicadores de Karush-Kuhn-Tucker quando os mesmo estão associados às condições de otimalidade de Karush-Kuhn-Tucker.

Teorema 3.11 . *Suponha que a função objetivo e todas as restrições do problema ϵ -restrição são de classe C^1 em $x^* \in S$, o qual é um ponto regular das restrições do problema ϵ -restrição. Temos o seguinte*

1. *Se x^* é Pareto ótimo próprio, então x^* resolve o problema ϵ -restrição para algum f_l com $\epsilon_j = f_j(x^*)$, para $j = 1, \dots, k, j \neq l$ com todos os multiplicadores de Karush-Kuhn-Tucker associados às restrições $f_j(x) \leq \epsilon_j$ para $j = 1, \dots, k, j \neq l$, sendo positivos.*
2. *Se o MOP é convexo, então x^* é Pareto ótimo próprio se ele também é solução do problema ϵ -restrição com os multiplicadores de Karush-Kuhn-Tucker associado com as restrições $f_j(x) \leq \epsilon_j$ para $j = 1, \dots, k, j \neq l$, sendo todos positivos.*

Prova 3.11 Veja Teorema 4.13 em HAIMES AND CHANKONG [3].

c.q.d

Abaixo enunciamos as condições suficientes de segunda ordem aplicadas ao problema ϵ -restrição.

Observação 3.2 . Suponha que a função objetivo e as restrições do problema ϵ -restrição são de classe C^2 em $x^* \in S$, um ponto regular das restrições do problema. Uma condição suficiente para x^* ser uma solução do problema ϵ -restrição é que existam vetores $0 \leq \lambda \in \mathbb{R}^{k-1}$ e $0 \leq \mu \in \mathbb{R}^m$ tais que as condições de otimalidade da observação 3.1 sejam satisfeitas e a matriz Hessiana da correspondente função Lagrangeana

$$\nabla^2 f_l(x^*) + \sum_{\substack{j=1 \\ j \neq l}}^k \lambda_j \nabla^2 (f_j(x^*) - \epsilon_j) + \sum_{i=1}^m \nabla^2 \mu_i g_i(x^*)$$

seja definida positiva no conjunto $\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d = 0 \text{ para todo } i \text{ tal que } \mu_i > 0\}$

Suponha que x^* seja solução do problema ϵ -restrição para algum $\epsilon^* \in \epsilon$ e denotemos λ_{kj}^* , $j = 1, \dots, k$ e $j \neq k$, e μ_i^* , $i = 1, \dots, m$ os multiplicadores de Karush-Kuhn-Tucker associados às restrições f_j 's e g_i 's respectivamente. Isto é, dizemos que λ_{kj}^* e μ_i^* são os multiplicadores de Karush-Kuhn-Tucker associados ao problema ϵ -restrição em x^* se $(x^*, \lambda_{k1}^*, \dots, \lambda_{k(k-1)}^*, \lambda_{k(k+1)}^*, \dots, \lambda_{k(n)}^*, \mu_1^*, \dots, \mu_m^*)$ satisfaz as condições de Karush-Kuhn-Tucker para o problema ϵ -restrição.

Capítulo 4

Trade-off e os multiplicadores de Karush-Kuhn-Tucker

A expressão *trade-off* define uma situação onde ocorre um conflito de escolhas. Pode caracterizar uma ação econômica que visa a resolução de um problema acarretando outro. Taxa *trade-off* está relacionada a mudanças no valor da função objetivo quando movemos de uma solução para outra, através do aumento de uma variável objetivo e decréscimo de alguma outra variável objetivo.

Este capítulo baseia-se nos trabalhos HAIMES AND CHANKONG [5] e [3].

4.1 Taxa *trade-off*

O termo *trade-off* reflete uma situação de mudança onde sacrificamos uma variável objetivo para obter ganho em outra. No nosso caso utilizaremos taxa *trade-off* como informação que ajudará a responder à seguinte questão: "Qual é a mudança relativa no valor da variável objetivo f_i quando o valor da variável objetivo f_j é alterada por uma unidade?".

Consideramos dois tipos de *trade-off*: objetivo, que fornece informação sobre a real relação entre as funções objetivos, e subjetivo, que é uma informação determinada de forma implícita pelo decisor e reflete sua preferência.

A taxa *trade-off* pode ser obtida através de um quociente envolvendo dois componentes da função objetivo e dois pontos distintos do conjunto Pareto ótimo.

CHANKONG AND HAIMES [3] estabelecem as seguintes definições com relação

à taxa *trade-off*.

Definição 4.1 . Sejam $x^0, x^* \in S^*$ e as correspondentes imagens dos objetivos $f(x^0) = (f_1(x^0), \dots, f_k(x^0))$ e $f(x^*) = (f_1(x^*), \dots, f_k(x^*))$. Taxa *trade-off* entre x^0 e x^* envolvendo os objetivos f_i e f_j é definida como

$$T_{ij}(x^0, x^*) = \frac{f_i(x^0) - f_i(x^*)}{f_j(x^0) - f_j(x^*)}$$

onde $f_j(x^0) \neq f_j(x^*)$.

Podemos estender a definição 4.1 considerando o caso em que $f_j(x^0) = f_j(x^*)$,

$$\overline{T_{ij}(x^0, x^*)} := \begin{cases} -\infty & , \text{ se } f_i(x^0) - f_i(x^*) < 0 \text{ e } f_j(x^0) - f_j(x^*) = 0 \\ 0 & , \text{ se } f_i(x^0) - f_i(x^*) = 0 \text{ e } f_j(x^0) - f_j(x^*) = 0 \\ \infty & , \text{ se } f_i(x^0) - f_i(x^*) > 0 \text{ e } f_j(x^0) - f_j(x^*) = 0 \\ T_{ij}(x^0, x^*) & , \text{ caso contrário} \end{cases}$$

Taxas *trade-off* podem ser divididas em duas classes: *trade-off* parcial e *trade-off* total.

Definição 4.2 . $T_{ij}(x^0, x^*)$ é chamada *trade-off* parcial envolvendo f_i e f_j entre x^0 e x^* se $f_l(x^0) = f_l(x^*)$ para todo $l = 1, \dots, k$ e $l \neq j, i$. Quando $f_i(x^0) \neq f_i(x^*)$ para no mínimo um $l \in \{1, \dots, k\}$, $l \neq i, j$, $T_{ij}(x^0, x^*)$ é chamada *trade-off* total envolvendo f_i e f_j entre x^0 e x^* .

Em alguns casos pode ser interessante definir *trade-off* na vizinhança de alguma variável decisão $x^* \in S^*$. Abaixo definimos taxa *trade-off* utilizando conceitos como direções viáveis, limite e diferenciabilidade.

Definição 4.3 . Dado $x^* \in S^*$. Suponha que exista uma direção viável d (definição 7.5 do Apêndice). Se o limite

$$\begin{aligned} t_{ij}(x^*, d) &= \lim_{\alpha \searrow 0} T_{ij}(x^* + \alpha d, x^*) \\ &= \lim_{\alpha \searrow 0} \frac{f_i(x^* + \alpha d) - f_i(x^*)}{f_j(x^* + \alpha d) - f_j(x^*)} \end{aligned}$$

existe, denominamos $t_{ij}(x^*, d)$ taxa *trade-off* em x^* na direção d envolvendo os objetivos f_i e f_j .

Se f_i e f_j são ambas de classe C^1 em x^* temos,

$$t_{ij}(x^*, d) = \frac{\nabla f_i(x^*)^T d}{\nabla f_j(x^*)^T d}$$

Assim como na definição 4.1 é possível estabelecer uma classificação para taxa *trade-off* considerando a definição 4.3.

Definição 4.4 . Taxa *trade-off* $t_{ij}(x^*, d)$ é denominada taxa *trade-off* parcial se $f_l(x^*) = f_l(x^* + \alpha d)$ para todo $\alpha \in [0, \alpha^0]$, para algum $\alpha^0 > 0$ e para todo $l = 1, \dots, k$ onde $l \neq j, i$. Caso contrário $t_{ij}(x^*, d)$ é chamado taxa *trade-off* total.

Observe que $t_{ij}(x^*, d)$ não é definida quando $\nabla f_j(x^*)^T d = 0$. Podemos estender a definição 4.3 como foi feito na definição 4.1.

Para se definir *trade-off* não é necessário que os pontos considerados sejam Pareto ótimo, é necessário apenas que sejam viáveis. De qualquer forma estaremos interessados no conjunto Pareto ótimo, pois nele nenhum objetivo pode ser melhorado sem que ocorra uma perda no valor de no mínimo um outro objetivo. Além do mais desejamos que os pontos a serem explorados possuam informações relacionadas aos métodos solução para problemas de otimização multiobjetivo.

Observação 4.1 . Segundo HAIMES AND CHANKONG [5] em termos de *trade-off* temos que em uma solução Pareto ótimo, no mínimo uma das taxas *trade-off* envolvendo os objetivos f_i e f_j deve ser negativa, onde $i, j \in \{1, \dots, k\}$ e $i \neq j$. Em outras palavras para se melhorar o valor de um objetivo no mínimo um outro objetivo deve ser degradado (Exemplo 2.5).

Trade-off subjetivo está relacionado ao fato do decisor ter conhecimento de uma determinada função utilidade $U : \mathbb{R}^k \rightarrow \mathbb{R}$. Assume-se que essa função é monótona decrescente . Agora o decisor deve resolver o problema,

Problema 4.1

$$\begin{aligned} & \text{maximizar } U(f(x)) \\ & \text{sujeito a } x \in S \end{aligned}$$

A solução do Problema 4.1 irá expressar a preferência do decisor.

Definição 4.5 . Taxa marginal de substituição envolvendo f_i e f_j em x é denotada por m_{ij} e indica o quanto o valor de f_i deve ser melhorado para compensar a perda

de uma unidade no valor do objetivo f_j enquanto os outros objetivos não são modificados. Se as derivadas parciais de U com respeito a f_j e f_i existem em x , temos que a taxa marginal de substituição é denotada por

$$m_{ij} = \frac{\frac{\partial U(f(x))}{\partial f_j}}{\frac{\partial U(f(x))}{\partial f_i}}$$

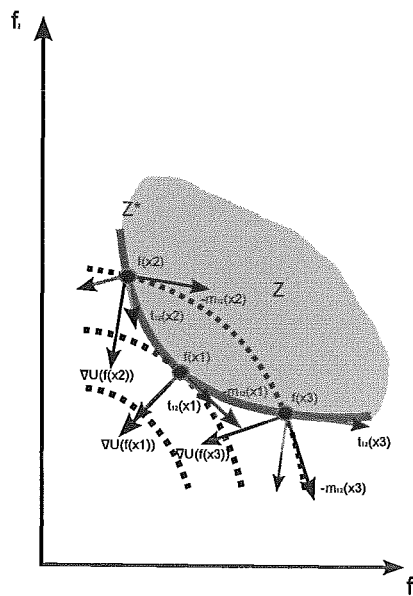


Figura 4.1: Relação entre taxa marginais de substituição e taxa trade-off

Considerando a taxa marginal de substituição entre cada par de objetivos somos capazes de determinar no espaço objetivo um hiperplano tangente ao conjunto de nível da função utilidade U em algum x . O hiperplano separa o espaço objetivo em dois subespaços onde um contém direções sobre x onde a função utilidade U é crescente e outro subespaço contém direções onde a função utilidade é decrescente. Isto ocorre devido ao vetor gradiente da função utilidade indicar a direção de maior crescimento da função.

Suponha que o problema possui uma superfície Pareto ótima convexa. Seja $x^* \in S$ um ponto Pareto ótimo e $t_{ij}(x^*)$ a correspondente taxa *trade-off* parcial envolvendo f_i e f_j . Se em algum ponto temos $-m_{ij}(x^*) = t_{ij}(x^*)$ para todo $i, j = 1, \dots, k, i \neq j$, então o ponto $x^* \in S$ pode ser considerado a melhor solução para o decisor (SAKAWA [7]). Isto se deve ao fato do conjunto de vetores objetivos Pareto ótimo

estarem sempre localizados na fronteira do conjunto dos vetores objetivos viáveis. Se a tangente à fronteira for perpendicular ao gradiente da curva de nível temos no caso da superfície Pareto convexo que não é possível se mudar para outra direção viável no espaço objetivo de tal modo que o valor de U cresça (Figura 4.1).

4.2 Trade-off como multiplicadores de Karush-Kuhn-Tucker

Consideremos o problema ϵ -restrição (Problema 3.2) e denotemos o conjunto das restrições por $S = \{x \mid x \in \mathbb{R}^n, g_i(x) \leq 0, i = 1, \dots, m\}$. O conjunto dos ϵ 's $\in \mathbb{R}^{k-1}$ os quais satisfazem o problema ϵ -restrição será denotado por ϵ .

Estudaremos as condições de otimalidade referentes ao problema ϵ -restrição de acordo com os multiplicadores de Lagrange.

O conceito chave entre os multiplicadores de Karush-Kuhn-Tucker λ_{kj}^* e taxa trade-off entre f_k e f_j em $x^* \in S^*$ é a interpretação da sensibilidade entre os multiplicadores.

Em seguida citamos um resultado de grande valia em otimização.

Teorema 4.1 (Teorema da Sensibilidade) . Dado $\epsilon^* \in \epsilon$. O ponto $x^* \in S$ resolve o problema ϵ -restrição com λ_{kj}^* sendo o multiplicador de Karush-Kuhn-Tucker associado à restrição $f_j(x) \leq \epsilon_j^*, j \neq k$, se

1. x^* é um ponto regular das restrições do problema ϵ -restrição;
2. as condições suficientes de segunda ordem são satisfeitas em x^* , e
3. não existem restrições degeneradas em x^* , com $\lambda_{kj}^* = -\frac{\partial f_k(x^*)}{\partial \epsilon_j}$ para todo $j \neq k$.

Prova 4.1 . Veja Teorema da Sensibilidade em LUENBERGER [8].

c.q.d

Abaixo dois corolários que serão de extrema valia na demonstração do Teorema 4.2, para maiores detalhes veja HAIMES AND CHANKONG [3].

Corolário 4.1 . Se x^* resolve o problema ϵ -restrição e satisfaz as hipóteses do Teorema 4.1, então existe uma vizinhança $B(\epsilon^*; R)$ de ϵ^* , $R > 0$ tal que para todo $\epsilon \in B(\epsilon^*; R)$, $x(\epsilon^*)$ resolve localmente e de forma única o problema ϵ -restrição e é uma função de classe C^1 com $x(\epsilon^*) = x^*$.

Corolário 4.2 . Suponha as hipóteses do teorema 4.1 satisfeitas, então existe uma vizinhança $B(\epsilon^*; R)$ de ϵ^* , $R > 0$, tal que para cada j com $\lambda_{kj}^* > 0$, $f_j(x(\epsilon)) = \epsilon_j$ para todo $\epsilon \in B(\epsilon^*; R)$.

O teorema abaixo é um dos principais resultados desse trabalho, o mesmo tem como fonte HAIMES AND CHANKONG [5], com algumas adaptações. Considere S^* o conjunto de pontos $x \in S$ e Z^* o conjunto dos $z \in Z$ os quais são Pareto ótimo com relação ao espaço decisão e espaço objetivo respectivamente.

Teorema 4.2 . Suponha que x^* resolve problema ϵ -restrição (Problema 3.2) para algum $\epsilon^* \in \epsilon$, com

1. x^* sendo um ponto regular com respeito às restrições ativas do problema ϵ -restrição,
2. as condições suficientes de segunda-ordem satisfeitas em x^* , e
3. todas as restrições ativas em x^* não-degeneradas.

Assuma sem perda de generalidade que $\lambda_{kj}^* > 0$ para todo $j = 1, \dots, p$ e $\lambda_{kj}^* = 0$ para todo $j = p + 1, \dots, k - 1$. Então,

a). Se $\epsilon_j^* = f_j(x^*)$ para todo $j = 1, \dots, k - 1$, então existe uma vizinhança $B(x^*; r)$ de x^* , $r > 0$ e uma função $\bar{x}(\epsilon) : \mathbb{R}^{k-1} \rightarrow \mathbb{R}^n$ de classe C^1 definida em alguma vizinhança $B(\epsilon^*; R) \subset \mathbb{R}^{k-1}$ de ϵ^* , $R > 0$ tal que

$$S^* \cap B(x^*; r) \subseteq \bar{x}(B(\epsilon^*; R)) \subseteq S^* \quad (4.1)$$

b). Suponha $p = k - 1$ e $B(x^*; r)$ são obtidos como em a). Sejam

$$Z_k^* = \{(f_1, \dots, f_k)^T \mid f_j = f_j(x), j = 1, \dots, k, x \in S^* \cap B(x^*; r)\}$$

e

$$Z_{k-1}^* = \{(f_1, \dots, f_{k-1})^T \mid f_j = f_j(x), j = 1, \dots, k - 1, x \in S^* \cap B(x^*; r)\}.$$

Então existe uma função \bar{f}_k de classe C^1 definida em Z_{k-1}^* tal que para cada

$$(f_1, \dots, f_k)^T \in Z^*, \quad f_k = \bar{f}_k(f_1, \dots, f_{k-1})^T.$$

Além disso

$$\frac{\partial \bar{f}_k}{\partial f_j}(f_1(x^*), \dots, f_{k-1}(x^*)) = -\lambda_{kj}^* \quad (4.2)$$

para cada $j = 1, \dots, k-1$.

c). Seja $1 \leq p < k-1$, onde $Z_\epsilon^* = \{(f_1, \dots, f_k)^T \mid f_j = f_j(x), j = 1, \dots, k, x \in \bar{x}(B(\epsilon^*; R))\}$. Então existem funções de classe C^1 , $\bar{f}_{p+1}, \dots, \bar{f}_{k-1}$ e \bar{f}_k definidas em $B(\epsilon^*; R)$ tais que para cada $(f_1, \dots, f_k)^T \in Z_\epsilon^*$ temos,

$$\bar{f}_j = \bar{f}_j(f_1, \dots, f_p, \epsilon_{p+1}^*, \dots, \epsilon_{k-1}^*), \text{ para todo } j = p+1, \dots, k.$$

Além disso para cada $l = 1, \dots, p$ temos

$$\left. \frac{\partial \bar{f}_k}{\partial f_l} \right|_{\epsilon=\epsilon^*} = -\lambda_{kl}^* = \frac{\nabla f_k(x^*)^T d_l^*}{\nabla f_l(x^*)^T d_l^*}$$

onde d_l^* é a direção $\frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}$. Também temos que para cada $j = p+1, \dots, k-1$,

$$\left. \frac{\partial \bar{f}_j}{\partial f_l} \right|_{\epsilon=\epsilon^*} = \nabla f_j(x^*)^T \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l} \quad (4.3)$$

Prova 4.2 .

a). Seja $B(\epsilon^*; R_\epsilon) \equiv \{\epsilon \mid \epsilon \in \mathbb{R}^{k-1}, \|\epsilon - \epsilon^*\| \leq R_\epsilon\}$, onde $R_\epsilon > 0$. Pela continuidade de f_j , para cada $1 \leq j \leq k-1$ em x , existe um $r_j > 0$ tal que

$$\|x - x^*\| < r_j \Rightarrow |f_j(x) - f_j(x^*)| < R_\epsilon$$

Seja $r \equiv \min_{1 \leq j \leq k-1} r_j$, assim temos o conjunto $B(x^*; r) \equiv \{x \mid x \in \mathbb{R}^n, \|x - x^*\| < r\}$.

Considere alguma $\varphi \in S^* \cap B(x^*; r)$. Defina $\bar{\epsilon}_j = f_j(x)$ para todo $j = 1, \dots, k-1$. Como $\varphi \in B(x^*; r)$ e $\epsilon_j^* = f_j(x^*)$ para todo $j = 1, \dots, k-1$, segue que $\bar{\epsilon} \in B(\epsilon^*; R)$. Portanto, pela definição de $B(\epsilon^*; R)$ (Corolário 4.1), o problema $\bar{\epsilon}$ -restrição tem uma solução local única \bar{x} em função de $\bar{\epsilon}$. Além disso, já que $\varphi \in S^*$ pelo Teorema (3.9) temos que φ é também solução local do problema de otimização multiobjetivo. Segue que $\varphi = \bar{x}$, o que prova a primeira parte de a).

Pelo Corolário(4.1) temos que existe uma vizinhança $B(\epsilon^*; R)$ de ϵ^* tal que para cada $\epsilon \in B(\epsilon^*; R)$, o problema ϵ -restrição possui uma solução local \bar{x} de classe C^1 definida em $B(\epsilon^*; R)$. Segue do Teorema(3.10) que $\bar{x}(B(\epsilon^*; R)) \subseteq S^*$.

b). De a) temos que para algum $\varphi \in S^* \cap B(x^*; r)$ existe $\bar{\epsilon} \in B(\epsilon^*; R)$ tal que $\varphi = \bar{x}(\bar{\epsilon})$. Portanto para todo $\varphi \in S^* \cap B(x^*; r)$, $f_k(\varphi) = f_k(\bar{x}(\bar{\epsilon})) \equiv \bar{f}_k(\bar{\epsilon})$, onde $\bar{f}_k()$ é uma função de ϵ definida em $B(\epsilon^*; R)$. Devido a $f_k()$ e $\bar{x}()$ serem funções de classe C^1 temos que $\bar{f}_k()$ também é de classe C^1 . Pelo Corolário(4.2) e pela definição de $\bar{x}()$, temos para cada $\varphi \in S^* \cap B(x^*; r)$, $f_j(x) = \bar{\epsilon}_j$ para todo $j = 1, \dots, k-1$. Portanto, $f_k(\varphi) = \bar{f}_k(f_1(\varphi), \dots, f_{k-1}(\varphi))$ para cada $\varphi \in S^* \cap B(x^*; r)$. Consequentemente para cada $(f_1, \dots, f_k) \in Z_k^*$, $f_k = \bar{f}_k(f_1, \dots, f_{k-1})$ como pedido.

Além disso temos que

$$\begin{aligned} \frac{\partial \bar{f}_k}{\partial f_j}(f_1^*, \dots, f_k^*) &\equiv \frac{\partial \bar{f}(\epsilon^*)}{\partial \epsilon_j} \\ &= -\lambda_{kj}^* \end{aligned}$$

onde a igualdade é dada pelo Teorema 4.1.

c) Pela definição de $\bar{x}()$ e $\bar{x}(B(\epsilon^*; R))$, para cada $\varphi \in \bar{x}(B(\epsilon^*; R))$ existe $\bar{\epsilon} \in B(\epsilon^*; R)$ tal que $\varphi = \bar{x}(\bar{\epsilon})$. Portanto $f_j(\varphi) = f_j(\bar{x}(\bar{\epsilon})) \equiv \bar{f}_j(\bar{\epsilon})$ para todo $j = p+1, \dots, k$. Além disso, pelo Corolário 4.2 $f_l(\varphi) = \bar{\epsilon}_l$ para todo $l = 1, \dots, p$. Portanto para cada $j = p+1, \dots, k$,

$$f_j(\varphi) = \bar{f}_j(f_1(\varphi), \dots, f_p(\varphi), \bar{\epsilon}_{p+1}, \dots, \bar{\epsilon}_{k-1})$$

para todo $\varphi \in \bar{x}(B(\bar{\epsilon}))$, o qual mostra que $\bar{f}_j()$, $j = p+1, \dots, k$, são as funções requeridas.

Observe que para algum $1 \leq l \leq p$

$$\begin{aligned} \frac{\partial \bar{f}_k}{\partial f_l} \Big|_{\epsilon=\epsilon^*} &\equiv \frac{\partial \bar{f}_k}{\partial \epsilon_l} \Big|_{\epsilon=\epsilon^*} \\ &= -\lambda_{kl}^* \end{aligned}$$

onde a igualdade é dada pelo Teorema 4.1.

Além do mais, já que para cada $1 \leq l \leq p$, $f_l(\bar{x}(\epsilon)) = \epsilon_l$ para todo $\epsilon \in B(\epsilon^*; R)$, pela regra da cadeia temos

$$1 = \frac{\partial \bar{f}_l}{\partial \epsilon_l} \Big|_{\epsilon=\epsilon^*} = \nabla f_l(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l} \quad (4.4)$$

Também

$$\left. \frac{\partial \bar{f}_k}{\partial \epsilon_l} \right|_{\epsilon=\epsilon^*} = \nabla f_k(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}$$

Portanto

$$-\lambda_{kl}^* = \frac{\nabla f_k(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}}{\nabla f_l(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}} = \frac{\nabla f_k(x^*) d_l^*}{\nabla f_l(x^*) d_l^*} \quad (4.5)$$

Finalmente, para cada $1 \leq l \leq p$ e $p+1 \leq j \leq k-1$, pela regra da cadeia,

$$\begin{aligned} \left. \frac{\partial \bar{f}_j}{\partial f_l} \right|_{\epsilon=\epsilon^*} &\equiv \left. \frac{\partial \bar{f}_j}{\partial \epsilon_l} \right|_{\epsilon=\epsilon^*} \\ &= \nabla f_j(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}. \end{aligned}$$

o que completa a prova.

c.q.d

O teorema acima faz um estudo das propriedades do conjunto Pareto ótimo local, ambos no espaço decisão e no espaço objetivo HAIMES AND CHANKONG [5].

Em a) temos que a superfície de Pareto ótimo na vizinhança de x^* pode ser completamente especificada parametricamente em termos de ϵ 's. Significa que existe uma correspondência entre cada solução Pareto ótimo local em uma vizinhança de x^* e uma vizinhança de ϵ^* . Embora tal descrição paramétrica não seja geralmente conhecida, o conhecimento de que existe pode ser muito útil no desenvolvimento de teorias ou algoritmos.

Em b) temos uma análise similar à feita em a) agora no espaço objetivo. Diante de algumas hipóteses, existem exatamente $k-1$ graus de liberdade para especificar um ponto pertencente à superfície Pareto ótimo na vizinhança de $f(x^*)$ no espaço objetivo. De fato podemos escolher f_1, \dots, f_{k-1} sobre $B(\epsilon^*; R)$ e calcular f_k através de $\bar{f}_k(f_1, \dots, f_{k-1})$, resultando nos pontos (f_1, \dots, f_k) pertencentes à superfície Pareto ótimo local no espaço objetivo.

Temos $-\lambda_{kj}^* = \frac{\partial \bar{f}_k}{\partial f_j}(f_1^*, \dots, f_{k-1}^*)$ (para algum $1 \leq j \leq k-1$), o qual representa a taxa de mudança de f_k por uma unidade de mudança em f_j enquanto todas os outros componentes objetivos não mudam. Assim $-\lambda_{kj}^*$ representa a taxa *trade-off* parcial envolvendo f_k e f_j em x^* .

O item c) é uma extensão da item b). Em b) é requerido que todas as restrições $f_j(x) \leq \epsilon_j^*, j = 1, \dots, k-1$ sejam ativas em x^* e os λ_{kj}^* sejam não-degenerados para

todo $j = 1, \dots, k - 1$. Suponha que estas hipóteses sejam violadas, ou seja, existem p restrições ativas não-degeneradas, onde $1 \leq p < k - 1$. Então o grau de liberdade para especificar um ponto Pareto ótimo local na vizinhança de $f(x^*)$ no espaço objetivo passa a ser exatamente o número de restrições ativas não-degeneradas. Em particular podemos escolher valores de f_1, \dots, f_p (em alguma vizinhança de $f_1(x^*), \dots, f_p(x^*)$), e então determinamos f_{p+1}, \dots, f_{k-1} através de alguma regra específica (no caso \bar{f}) definidas em termos de f_1, \dots, f_p e finalmente calcular f_k através de uma função específica sobre f_1, \dots, f_{k-1} . Os pontos resultantes (f_1, \dots, f_k) devem ser uma solução Pareto ótimo local no espaço objetivo.

As regras específicas $\bar{f}_{p+1}, \dots, \bar{f}_{k-1}$ são geralmente um mapeamento definida em $B(\epsilon^*; R)$ e cada $f_j, j = p + 1 \dots k - 1$, é escolhida sobre $f_j \in \bar{f}_j(f_1, \dots, f_p, \epsilon_{p+1}^*, \dots, \epsilon_{k-1}^*)$.

No item c) consideramos somente soluções Pareto ótimo locais. Para permanecer com este subconjunto, uma pequena mudança de δf_l no valor do objetivo ativo $f_l (l = 1, \dots, p)$ com todos os outro objetivos ativos permanecendo inalterados será feito, mudando (aproximadamente) f_k até $-\lambda_{kj}^* \delta f_l$ bem como uma mudança (aproximada) em cada $f_j, p + 1 \leq j \leq k - 1$, até $\nabla f_j(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l} \delta f_l$. Neste caso $-\lambda_{kj}^*$ representa a taxa *trade-off* total entre f_k e f_j em x^* na direção $\frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}$.

Sob o ponto de vista prático b) e c) provam um conveniente modo para estimar informações *trade-off* necessárias para continuação do processo iterativo de análise *trade-off* dos objetivos. Para ver como essa informação pode ser utilizada, suponha que tenhamos gerado um ponto Pareto ótimo $x^* \in S^*$ por meio da resolução de problema ϵ^* -restrição e com ele obtemos $-\lambda_{kj}^*, j = 1, \dots, k - 1$. Se o caso b) é aplicado, então para todo $j = 1, \dots, k - 1$, podemos perguntar ao decididor:

Dado $f_j = f_j(x^)$ para todo $j = 1, \dots, k$, como decrescer f_k até λ_{kj}^* unidades enquanto f_j crescer uma unidade e todos os outros objetivos permanecem inalterados?*

Se ocorre o caso c), isto é, $\lambda_{kj}^* > 0, j = 1, \dots, p$ e $\lambda_{kj}^* = 0, j = p + 1, \dots, k - 1$, então para todo $j = 1, \dots, p$ podemos perguntar ao decididor:

Dado $f_j = f_j(x^)$ para todo $j = 1, \dots, k$ como decrescer f_k até λ_{kj}^* unidades e mudar cada $f_j, j = p + 1, \dots, k - 1$ até $\nabla f_j(x^*) \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_j}$ unidades por um aumento unitário em f_i com todos os outros componentes objetivos $f_i, i = 1, \dots, p, i \neq l$ permanecendo inalterados?*

O decisor deve responder expressando sua preferência de acordo com a taxa *trade-off* estabelecida em uma escala ordinal. Embora existam p questões a serem respondidas ($p = k - 1$ no caso **b**)) o decisor somente necessita responder uma de cada vez.

A seguir aplicamos o Teorema 4.2 no exemplo 3.1. A mesma análise é feita em HAIMES AND CHANKONG [5]. Aqui tentaremos ser um pouco detalhista.

Exemplo 4.1 Considere o problema $P_1(\epsilon)$ do exemplo 3.1 onde $\epsilon = (\epsilon_2, \epsilon_3)$. Seja $\epsilon_2 = 2,5$ e $\epsilon_3 = 3$. O ponto $x^0 = (2; 0,5)$ é solução do problema, já que x^0 é um ponto regular, os gradientes das restrições ativas $\nabla f_2(x^0) = (1, 1)^T$ e $\nabla f_3(x^0) = (1, 2)^T$ são linearmente independentes e as condições de Karush-Kuhn-Tucker são satisfeitas (Figura 4.2).

Em particular,

$$\nabla f_1(x^0) + \lambda_{12}^0 \nabla f_2(x^0) + \lambda_{13}^0 \nabla f_3(x^0) = 0$$

o que implica que $\lambda_{12}^0 = 1$ e $\lambda_{13}^0 = 1$.

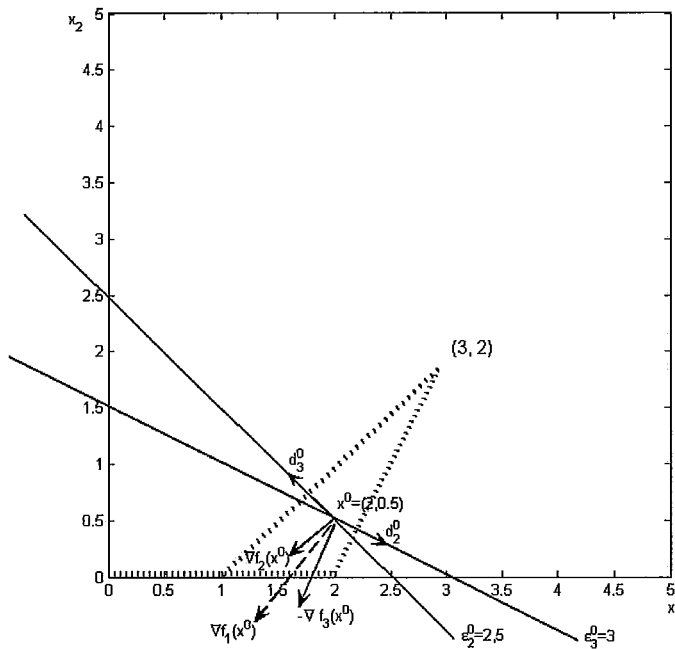


Figura 4.2: Solução gráfica para $P_1^0(2, 5; 3)$

Para este exemplo temos o seguinte função lagrangeana,

$$L(x, \lambda, \mu) = (x_1 - 3)^2 + (x_2 - 2)^2 + \lambda_{12}(x_1 + x_2 - \epsilon_2) + \lambda_{13}(x_1 + 2x_2 - \epsilon_3) - \mu_1 x_1 - \mu_2 x_2.$$

Calculando gradiente e hessiana,

$$\nabla L(x, \lambda, \mu) = (2(x_1 - 3) + \lambda_{12} + \lambda_{13} - \mu_1, 2(x_2 - 2)^2 + \lambda_{12} + 2\lambda_{13} - \mu_2),$$

e

$$H(x, \lambda, \mu) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

No ponto $x^0 = (2; 0, 5)$ temos que $\nabla L(x^0, \lambda, \mu) = (0, 0)$ e $H(x^0, \lambda, \mu)$ é definida positiva.

Sejam $\epsilon_2^1 = 1,5$ e $\epsilon_3^1 = 1,5$ onde denotamos o problema por $P_1^1(\epsilon)$. O ponto $x^1 = (1, 5, 0)$ é solução do $P_1^1(\epsilon)$, mas x^1 não é um ponto regular, já que as restrições $f_1(x^1), f_2(x^1)$ e $g_2 = -x_2$ são ativas e não é possível que 3 vetores sejam linearmente independentes no plano \mathbb{R}^2 (Figura 4.3).

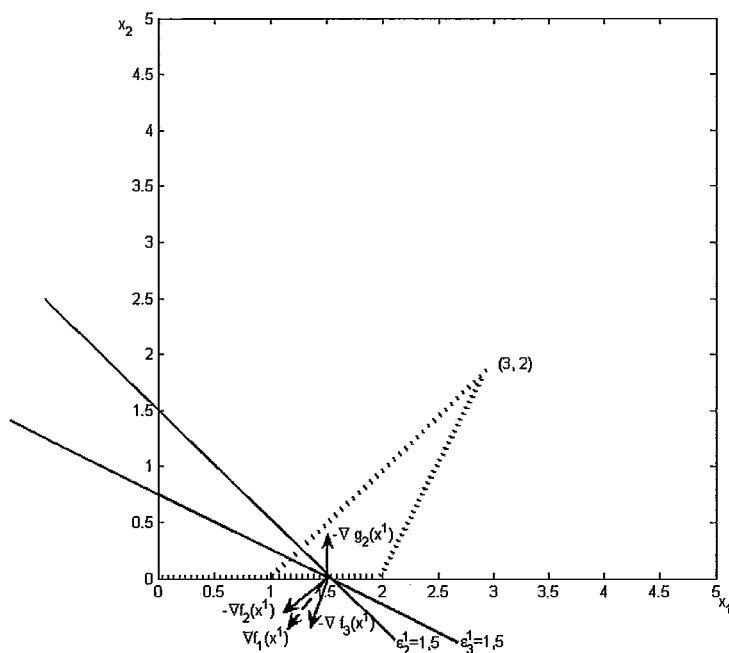


Figura 4.3: Solução gráfica para $P_1^1(1, 5; 1, 5)$

É possível obter vários valores para $\epsilon = (\epsilon_1, \epsilon_2)$ de forma que ainda assim conseguimos soluções Pareto ótimo para $P_1(\epsilon)$. Temos que para $B(\epsilon^0; R)$ onde $R > 0$, existe uma função $\bar{x} : B(\epsilon^0; R) \rightarrow \mathbb{R}^k$ definida pela regra: para cada $\epsilon \in B(\epsilon^0; R)$, ϵ resolve $P_1(\epsilon)$ (Teorema 4.2).

Após encontrar essa vizinhança calculamos $\bar{x}(\epsilon^0)$, $d_2^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_2}$ e $d_3^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_3}$.

Para algum $\epsilon \in B(\epsilon^0; R)$, a solução única de $P_1(\epsilon)$ é sempre um ponto na interseção das duas restrições, $f_2(x) \leq \epsilon_2$ e $f_3(x) \leq \epsilon_3$ (Figura 4.2). Assim para algum $\epsilon \in B(\epsilon^0; R)$, temos que $\bar{x}(\epsilon)$ é dado por,

$$f_2(x) = x_1 + x_2 = \epsilon_2,$$

$$f_3(x) = x_1 + 2x_2 = \epsilon_3.$$

Calculando $\bar{x}(\epsilon)$,

$$\bar{x}(\epsilon) = \begin{bmatrix} \bar{x}_1(\epsilon) \\ \bar{x}_2(\epsilon) \end{bmatrix} = \begin{bmatrix} 2\epsilon_2 - \epsilon_3 \\ \epsilon_3 - \epsilon_2 \end{bmatrix}.$$

Calculando d_2^0 e d_3^0 (Figura 4.2),

$$d_2^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

$$d_3^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_3} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Em seguida examinamos $f_1(f_2, f_3)$. Para todo $\epsilon \in B(\epsilon^0; R)$ as desigualdades $f_2 \leq \epsilon_2$ e $f_3 \leq \epsilon_3$ são sempre ativas na solução ótima $\bar{x}(\epsilon)$ de $P_1(\epsilon)$. Para cada $\epsilon \in B(\epsilon^0; R)$ no caso de $P_1(\epsilon_2^0, \epsilon_3^0)$,

$$f_1(\bar{x}(\epsilon)) = [(2\epsilon_2^0 - \epsilon_3^0) - 3]^2 + [(\epsilon_3^0 - \epsilon_2^0) - 3]^2,$$

$$f_1(f_2, f_3) = [(2f_2 - f_3) - 3]^2 + [(f_3 - f_2) - 3]^2.$$

Segue que,

$$\begin{aligned} \frac{\partial f_1(x^0)}{\partial f_2} &= 2(2f_2 - f_3 - 3)2 - 2(f_3 - f_2 - 2) \\ &= 10f_2 - 6f_3 - 8 \\ &= 10(x_1^0 + x_2^0) - 6(x_1^0 + 2x_2^0) - 8 \\ &= 10(2 + 0, 5) - 6(2 + 1) - 8 \\ &= -1 \\ &= -\lambda_{12}^0. \end{aligned}$$

e

$$\begin{aligned}
\frac{\partial f_1(x^0)}{\partial f_3} &= -2(2f_2 - f_3 - 3) - 2(f_3 - f_2 - 2) \\
&= -6f_2 + 4f_3 + 2 \\
&= -6(x_1^0 + x_2^0) + 4(x_1^0 + 2x_2^0) + 2 \\
&= -6(2, 5) + 4(3) + 2 \\
&= -1 \\
&= -\lambda_{13}^0.
\end{aligned}$$

O resultado acima possibilita ao decididor mudar de ponto ao invés de ficar com o ponto x^0 . Ele pode decidir, por exemplo, permanecer no nível de f_3 e alterar o nível de f_2 , por exemplo, aumentar f_2 por no máximo $\delta\epsilon_2$, onde $0 < \delta\epsilon_2 \leq 1$. De acordo com essa decisão a melhor alternativa seria $x^1 = \bar{x}(\epsilon_2 + \delta\epsilon_2, \epsilon_3)$, onde

$$\begin{aligned}
f_2(x^1) &= \epsilon_2^0 + \delta\epsilon_2^0, \\
f_3(x^1) &= \epsilon_3^0.
\end{aligned}$$

e

$$\begin{aligned}
f_1(x^1) &= [2f_2(x^0) - f_3(x^0) - 3]^2 + [f_3(x^0) - f_2(x^0) - 2]^2, \\
&\approx f_1(x^0) + \lambda_{12}\delta\epsilon_2^0
\end{aligned}$$

Movendo de x^0 para x^2 ocorre um movimento na direção d_2^0 no espaço de decisão (Figura 4.2). Por exemplo, se $\delta\epsilon_2^0 = 0,01$, a nova solução Pareto ótimo é (Figura 4.4),

$$\begin{aligned}
\begin{bmatrix} \bar{x}_1^2(\epsilon) \\ \bar{x}_2^2(\epsilon) \end{bmatrix} &= \begin{bmatrix} 2\epsilon_2^2 - \epsilon_3^2 \\ \epsilon_3^2 - \epsilon_2^2 \end{bmatrix} \\
&= \begin{bmatrix} 2(2,51) - 3 \\ 3 - 2,51 \end{bmatrix} \\
&= \begin{bmatrix} 2,02 \\ 0,49 \end{bmatrix}
\end{aligned}$$

com $f_2(x^2) = 2,51$, $f_3(x^2) = 3$ e

$$\begin{aligned} f_1(x^2) &= [2f_2(x^2) - f_3(x^2) - 3]^2 + [f_3(x^2) - f_2(x^2) - 2]^2 \\ &= [2(2,51) - 3 - 3]^2 + [3 - 2,51 - 2]^2 \\ &= [-0,98]^2 + [-1,51]^2 \\ &= 3,2405 \end{aligned}$$

Assim a mudança de $f_1(x^0) = 3,25$ para $f_1(x^2) = 3,2405$ a qual é igual a $f_1(x^2) - f_1(x^0) = -0,0095$ unidades é dada por $-\lambda_{12}^0 \delta \epsilon_2^0 \approx -0,01$ unidades, justificando o uso de $-\lambda_{12}^0$ como uma taxa trade-off aproximada em x^0 .

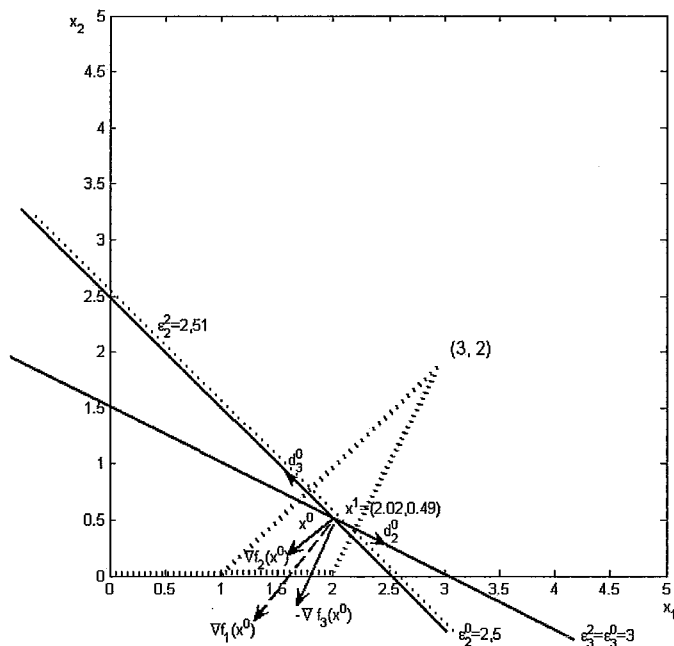


Figura 4.4: Solução gráfica para $P_1^2(2, 51; 3)$

Capítulo 5

Direções trade-off em problemas de otimização multiojetivo convexos

Direções *trade-off* exibem taxas relativas a mudanças de um ponto Pareto ótimo para outro ponto Pareto ótimo. Definiremos direções *trade-off* em cones, os quais determinam uma região de confiança onde podemos nos movimentar afim de obter uma melhor solução. Os cones em questão são definidos em pontos pertencentes ao conjunto Pareto ótimo.

Por notação temos que $z \prec y$ para $z, y \in \mathbb{R}^k$, significa que $z_i < y_i$ para todo $i = 1, \dots, k$, e correspondentemente, $z \preceq y$ significa que $z_i \leq y_i$ para todo i .

Este capítulo baseia-se em ROCKAFELLAR [1], HENING AND BUCHANAN [9], MÄKELÄ AND NEITTAANMÄKI [10] e MÄKELÄ AND MIETTINEN [11]. Nele utilizamos octante \mathbb{R}_-^k para estabelecer algumas propriedades e procuramos detalhar algumas provas e exemplos.

5.1 Otimalidade de Pareto

Nesta seção faremos um estudo do conjunto Pareto ótimo de acordo com o espaço dos objetivos HENING AND BUCHANAN[9]. Temos que Z denota o conjunto objetivo imagem.

Definição 5.1 . Um vetor objetivo $z \in Z$ é Pareto ótimo se não existe outro vetor objetivo $z' \in Z$ tal que $z'_i \leq z_i$ para todo $i = 1, \dots, k$ e $z'_j < z_j$ para no mínimo um índice j . O conjunto Pareto ótimo é denotado por

$$P(Z) = \{z \in Z \mid (z + \mathbb{R}_-^k) \cap Z = \emptyset\}.$$

Definição 5.2 . Um vetor objetivo $z \in Z$ é Pareto ótimo fraco se não existe outro vetor objetivo $z' \in Z$ tal que $z'_i < z_i$ para todo $i = 1, \dots, k$. O conjunto Pareto ótimo fraco de Z é denotado por

$$FP(Z) = \{z \in Z \mid (z + \text{int } \mathbb{R}_-^k) \cap Z = \emptyset\}.$$

Definição 5.3 . O conjunto Pareto ótimo próprio de Z é denotado como

$$PP(Z) = \{z \in Z \mid (z + C \setminus \{0\}) \cap Z = \emptyset\}$$

para algum cone convexo C tal que $\mathbb{R}_-^k \setminus \{0\} \subset \text{int } C$.

Relacionando as três definições acima temos a seguinte propriedade $PP(Z) \subset P(Z) \subset FP(Z) \subset Z$.

Podemos considerar as definições 5.1, 5.2 e 5.3 em uma vizinhança $B(z; R)$ de $z \in Z$, $R > 0$.

5.2 Cone tangente e cone normal

Abaixo estabelecemos algumas definições geométricas referentes a cones. Essas definições são tomadas em um ponto $z \in Z$.

Definição 5.4 . O cone de direções viáveis do conjunto $Z \subset \mathbb{R}^k$ em $z \in Z$ é denotado por

$$D(Z) = \{d \in \mathbb{R}^k \mid \text{existe } t > 0 \text{ tal que } z + td \in Z\}$$

Tomando o fecho do cone de direções viáveis $D(Z)$ nós obtemos o cone tangente ao conjunto Z .

Definição 5.5 . O cone tangente ao conjunto $Z \subset \mathbb{R}^k$ em $z \in Z$ é denotado por

$$T(Z) = \{d \in \mathbb{R}^k \mid \text{existe } t_j \searrow 0 \text{ e } d_j \rightarrow d \text{ tal que } z + t_j d_j \in Z\}.$$

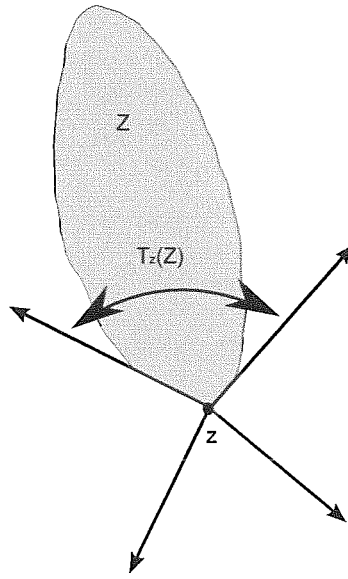


Figura 5.1: Cone tangente do conjunto Z

Definição 5.6 . *O cone normal do conjunto convexo Z em $z \in Z$ é o conjunto*

$$N(Z) = T(Z)^0 = \{y \in \mathbb{R}^k \mid y^T d \leq 0 \text{ para todo } d \in T(Z)\}$$

Os elementos pertencentes a $N(Z)$ e $T(Z)$ são chamados vetores normais e tangentes, respectivamente.

É possível se verificar que $T(Z)$ e $N(Z)$ são polares entre si.

$$\begin{aligned} T(Z)^0 &= \{y \mid y^T d \leq 0 \text{ para todo } d \in T(Z)\} \\ &= \{y \mid y^T d \leq 0 \text{ para todo } d \in D(Z)\} \\ &= \{y \mid y^T (z' - z) \leq 0 \text{ para todo } z' \in Z\} \\ &= N(Z) \end{aligned}$$

No capítulo 7 estabelecemos alguns resultados referentes a cone normal e cone tangente. Abaixo citamos um resultado que será utilizado posteriormente.

Lema 5.1 . $0 \in D(Z)$ e $Z \subset z + D(Z) \subset z + T(Z)$.

A seguir caracterizamos o conjunto Pareto ótimo de acordo com cone de direções viáveis $D(Z)$ e o octante não-positivo \mathbb{R}_-^k . Os mesmos se baseiam nos trabalhos de HENING AND BUCHANAN [9] e MÄKELÄ AND MIETTINEN [11] com algumas adaptações.

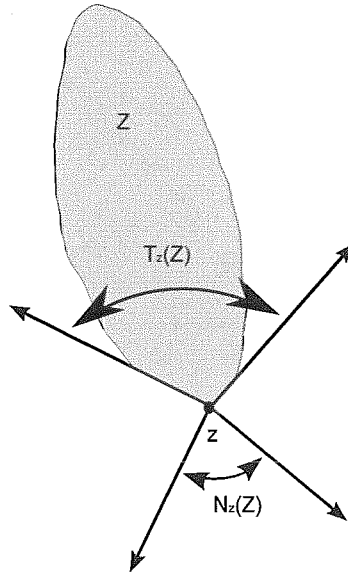


Figura 5.2: Cone normal do conjunto Z

Teorema 5.1 . O vetor objetivo $z \in P(Z)$ se, e somente se, $D(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$.

Prova 5.1 . Suponha que $z \in P(Z)$ e que exista um $d \in D(Z) \cap \mathbb{R}_-^k \setminus \{0\}$. Pela definição de $D(Z)$ e cone temos que existe um $t > 0$ tal que $z + td \in Z$ e $td \in \mathbb{R}_-^k \setminus \{0\}$. Assim temos que $z + td \in (z + \mathbb{R}_-^k \setminus \{0\}) \cap Z$, ou seja, $z \notin P(Z)$, o que é uma contradição. Logo temos que $D(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$.

Reciprocamente suponha que $D(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$ e $z \notin P(Z)$. Então existe $d \in \mathbb{R}_-^k \setminus \{0\}$ tal que $z + d \in Z$. Tomando $t = 1$ temos que $d \in D(Z)$, o que é uma contradição. Logo $z \in P(Z)$.

c.q.d

Teorema 5.2 . O vetor objetivo $z \in PP(Z)$ se, e somente se, $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$.

Prova 5.2 . Seja $z \in PP(Z)$, ou seja,

$$(z + C \setminus \{0\}) \cap Z = \emptyset \quad (5.1)$$

para algum cone C tal que $\mathbb{R}_-^k \setminus \{0\} \subset \text{int } C$. Suponha que exista um $d \in T(Z) \cap \mathbb{R}_-^k \setminus \{0\}$, então existe $t_j \searrow 0$ e $d_j \rightarrow d$ tal que $z + t_j d_j \in Z$. Já que $d \in \mathbb{R}_-^k \setminus \{0\} \subset \text{int } C$

e $d_j \rightarrow d$, existe um $n_1 > 0$ tal que $d_j \in \text{int } C \subset C \setminus \{0\}$ para todo $j \geq n_1$. Devido a C ser um cone e $t_j > 0$, temos que $t_j d_j \in C \setminus \{0\}$ para todo $j \geq n_1$. Então temos que $z + t_j d_j \in (z + C \setminus \{0\}) \cap Z$, uma contradição a equação (5.1). Portanto $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$.

Reciprocamente, assumamos que $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$. Pelo Teorema 7.7 existe um cone convexo C tal que $\mathbb{R}_-^k \setminus \{0\} \subset \text{int } C$ e $C \setminus \{0\} \cap T(Z) = \emptyset$. Pelo Lema 5.1 temos que $Z \subset z + T(Z)$, ou seja, $(z + C \setminus \{0\}) \cap Z = \emptyset$, logo $z \in PP(Z)$.

c.q.d

Teorema 5.3 . Se $D(Z)$ é fechado em algum ponto $z \in P(Z)$, então $z \in PP(Z)$.

Prova 5.3 . Seja $D(Z)$ um conjunto fechado em algum $z \in P(Z)$, usando o fato $D(Z) = \text{cl}D(Z) = T(Z)$ e o Teorema 5.1 temos que $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$. Pelo Teorema 5.2, $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$ se, e somente se, $z \in PP(Z)$.

c.q.d

Teorema 5.4 . As propriedades abaixo são equivalentes.

1. $z \in FP(Z)$.
2. $T(Z) \cap \text{int} \mathbb{R}_-^k = \emptyset$.
3. $D(Z) \cap \text{int} \mathbb{R}_-^k = \emptyset$.

Prova 5.4 . Suponha que $z \in FP(Z)$ e $T(Z) \cap \text{int } \mathbb{R}_-^k \neq \emptyset$. Por definição temos que existe $d \in \mathbb{R}_-^k$ e $d_j \rightarrow d$, $t_j \searrow 0$ tal que $z + t_j d_j \in Z$. Devido a $\text{int } \mathbb{R}_-^k$ ser aberto, temos que existe um $n > 0$ tal que $d_j \in \text{int } \mathbb{R}_-^k$ para todo $j \geq n$. Temos ainda que $\text{int } \mathbb{R}_-^k$ é um cone, onde para $t_j > 0$ segue que $t_j d_j \in \text{int } \mathbb{R}_-^k$ para todo $j \geq n$, ou seja $(z + \text{int } \mathbb{R}_-^k) \cap Z \neq \emptyset$, o que é uma contradição. Logo $z \in FP(Z)$.

Usando o fato de que $T(Z) = \text{cl}D(Z)$ temos que (2) implica em (3).

Suponha que $D(Z) \cap \text{int } \mathbb{R}_-^k = \emptyset$ e $z \notin FP(Z)$, então existe $d \in \text{int } \mathbb{R}_-^k$ tal que $z + d \in Z$. Tomando $t = 1$, temos que $d \in D(Z)$, uma contradição. Logo $z \in FP(Z)$.

c.q.d

Podemos ainda estabelecer propriedades que relacionem conjunto Pareto com cone normal assim como foi feito nos Teoremas 5.1 e 5.2 com relação a cone tangente.

Teorema 5.5 .

1. Se $z \in P(Z)$ então $N(Z) \cap \mathbb{R}_-^k \neq \emptyset$.
2. $z \in PP(Z)$ se, e somente se, $N(Z) \cap \mathbb{R}_-^k \setminus \{0\} \neq \emptyset$.

Prova 5.5 . Veja por exemplo MÄKELÄ AND MIETTINEN [11]

c.q.d

5.3 Cone de direções *trade-off*

No capítulo 2 definimos o conjunto imagem Z e o conjunto Pareto ótimo $P(Z)$ de um problema de otimização multiobjetivo. Tomando $T(Z)$, considere $PT(Z)$ o cone de direções de Pareto para Z . Pela definição de conjunto Pareto temos que $PT(Z) \cap \mathbb{R}_-^k = \emptyset$. Dado um ponto $z \in Z$ Pareto ótimo, alguma direção $d \in D(Z)$ pode levar em Z . Contudo $z + td$ pode ser dominado para todo $t > 0$. Se $d \in PT(Z)$ temos que $z + td$ não é dominado, e neste caso temos que $\{z + td \mid t \in \mathbb{R}\}$ é uma reta tangente a $P(Z)$ no ponto z . Se $T(Z) = D(Z)$ temos que $z + td$ poderá ser atualizado em Z para algum intervalo dado. Denominaremos $PT(Z)$ como cone de direções trade-off para o conjunto Z no ponto z .

Teorema 5.6 $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$ se, e somente se, $PT(Z) \neq \emptyset$.

Prova 5.6 Demonstração semelhante à do Teorema 5.2, para maiores detalhes veja por exemplo Corolário 5.1 em HENIG [12].

c.q.d

Teorema 5.7 . $PT(Z) \neq \emptyset$ se, e somente se, $z \in PP(Z)$.

Prova 5.7 . Pelo teorema 5.2 o vetor objetivo $z \in PP(Z)$ se, e somente se, $T(Z) \cap \mathbb{R}_-^k \setminus \{0\} = \emptyset$ o qual pelo Teorema 5.6 é valido se, e somente se, $PT(Z) \neq \emptyset$.

c.q.d

Quando um dos cones, normal ou tangente é dado de forma explicita como cone poliedral por exemplo, é possível encontrar o outro cone.

Teorema 5.8 . Seja $H \in \mathbb{R}^{k,s}$ e $S \in \mathbb{R}^{m,s}$. Os cones $\{y = Hw \mid Sw \geq 0, w \geq 0, w \in \mathbb{R}^s\}$ e $\{d \in \mathbb{R}^k \mid d^T H E \leq 0\}$ são polar um ao outro, onde E é a matriz cujas colunas são pontos extremos de $W = \{w \in \mathbb{R}^s \mid Sw \geq 0, w \geq 0, \sum w_i = 1\}$.

Prova 5.8 . Pelo Teorema 7.9, todo $w \in W$ pode ser escrito como $w = Eu$ onde $u \geq 0$ e $\sum u_i = 1$, tal que

$$\begin{aligned} \{y = Hw \mid Sw \geq 0, w \geq 0\} &= \{y = \alpha Hw \mid w \in W, \alpha \geq 0\} \\ &= \left\{ y = \alpha H E u \mid u \geq 0, \sum u_i = 1, \alpha \geq 0 \right\} \\ &= \{y = H E u \mid u \geq 0\}. \end{aligned}$$

Pelo Lema de Farka's (Teorema 7.10) temos que $\{y = H E u \mid u \geq 0\}$ e $\{d \in \mathbb{R}^k \mid d^T H E \leq 0\}$ são polar um ao outro.

c.q.d

O próximo teorema estabelece um método para encontrar uma direção *trade-off* sobre um dado cone tangente $T(Z)$ (HENING AND BUCHANAN)[9].

Teorema 5.9 . Se $d \in PT(Z)$ então $d^T y \leq 0$ para todo $y \in N(Z)$ e $d^T y = 0$ para algum $y \in N(Z) \cap \mathbb{R}_-^k$.

Prova 5.9 . Seja $d \in PT(Z)$. Pelo fato de $d \in T(Z) = N(Z)^0$ temos que $d^T y \leq 0$ para todo $y \in N(Z)$. Existe algum $y \geq 0$ com $y^T d \geq y^T d'$ para todo $d' \in T(Z)$. Temos $y^T d \leq 0$, caso contrário $\{y^T d \mid d' \in T(Z)\}$ é ilimitado. Usando o fato de $\{0\} \in T(Z)$ temos que $0 = y^T d \geq y^T d'$ para todo $d' \in T(Z)$.

c.q.d

Considerando $T(Z) = \{d \in \mathbb{R}^k \mid d^T H \leq 0\}$, então $N(Z) = \{y = Hw \mid w \geq 0\}$ e se $d \in PT(Z)$, então $d^T H w \leq 0$ para todo $w \geq 0$. Consequentemente $d^T h_q \leq 0$ para toda coluna h_q de H . Pelo Teorema 5.9, temos que $d^T h_q = 0$ para algum $Hw > 0$. Portanto $d^T h_q = 0$ para no mínimo uma coluna h_q de H .

Suponha que $N(Z) = \{\alpha y \mid \alpha \geq 0\}$ para algum $y \in \mathbb{R}^k$, então temos que $T(Z) = \{d \in \mathbb{R}^k \mid d^T y \leq 0\}$ e $PT(Z) = \{d \in \mathbb{R}^k \mid d^T y = 0\}$ de acordo com o teorema 5.9. Assim temos que o conjunto de direções *trade-off* é dado por $\{d \mid d^T y = 0\}$, ou seja, $\{z + d \mid d^T y = 0\}$ é tangente a Z em z .

5.3.1 Direções trade-off em MOP convexos

Considere o problema de otimização multiobjetivo (Problema 2.1). Assuma que cada f_i , $i = 1, \dots, k$ é convexa e o conjunto viável S referente ao conjunto restrição do problema é da seguinte forma

$$S = \{x \in \mathbb{R}^n \mid g_j(x) \leq 0, j = 1, \dots, m\}$$

onde cada g_j é convexa para $j = 1, \dots, m$.

Definição 5.7 . Um vetor $d \in \mathbb{R}^k$ é denominado subgradiente de f no ponto x se $f(x') \geq f(x) + d^T(x' - x)$ para todo $x' \in \mathbb{R}^k$. Denotamos $\partial f(x)$ conjunto dos subgradientes em x .

Devido a convexidade de todos g_j e f , temos que $\partial f(x) \neq \emptyset$ em todo $x \in S$ ROCKAFELLAR [1].

Teorema 5.10 . Suponha que $\text{int}(S) \neq \emptyset$ e seja $z \in Z$. Então $y \in N(Z)$ se, e somente se, para algum $x \in \mathbb{R}^n$ com $f(x) = z$, existe $\mu \in \mathbb{R}^m$ tal que

1. $0 \in (y_1 \partial f_1(x) + \dots + y_k \partial f_k(x) - \mu_1 \partial g_1(x) - \dots - \mu_m \partial g_m(x))$,
2. $\mu_j \geq 0, g_j(x) \leq 0, \mu_j g_j(x) = 0, j = 1, \dots, m$.

Prova 5.10 . Por definição temos que $y \in N(Z)$ se, e somente se, para algum $x \in S$ com $f(x) = z$ temos que $y^T f(x) \geq y^T f(x')$ para todo $x' \in S$. Assim $y \in N(Z)$ se, e somente se, para algum $x \in S$ com $f(x) = z$, x resolve

$$\min \{y^T f(x) \mid x \in S\} = \min \{y^T f(x) \mid g_1(x) \leq 0, \dots, g_m(x) \leq 0\}$$

O resultado segue do Teorema de Karush-Kuhn-Tucker ROCKAFELLAR [1].

c.q.d

Sejam f_i para $i = 1, \dots, k$ e g_j para $j = 1, \dots, m$, funções de classe C^1 . Pelo Teorema 5.10 temos que para um ponto $z \in P(Z)$ é possível encontrar o conjunto $N(Z)$ resolvendo o seguinte sistema de equações para algum $x \in S^* = \{x \in S \mid f(x) = z\}$,

$$\sum_i^k y_i \nabla f_i(x) - \sum_J \mu_j \nabla g_j(x) = 0, \quad y \geq 0, \quad \mu \geq 0, \quad \text{onde } J = \{j \mid g_j(x) = 0\}$$

Corolário 5.1 $z \in PP(Z)$ se, e somente se, existe $x \in S$ com $z = f(x)$, $y > 0$ e $\mu \in \mathbb{R}^m$ o qual satisfaz as condições Teorema 5.10.

Exemplo 5.1 Considere o seguinte exemplo (HENING AND BUCHANAN [9]) de problema de otimização linear multiobjetivo (MOLP).

$$\min Cx, \text{ sujeito a } Ax \leq b, \quad x \geq 0$$

onde $x \in \mathbb{R}^n$, $C \in \mathbb{R}^{k,n}$, $A \in \mathbb{R}^{m,n}$ e $b \in \mathbb{R}^m$. De acordo com o Teorema 5.10, $y \in N(Z)$ se, e somente se, existe $x \in \mathbb{R}^n$, $\mu \in \mathbb{R}^m$ e $\nu \in \mathbb{R}^n$ tal que $Cx = z$, $C^T y - A^T \mu + I\nu = 0$, $\mu \geq 0$, $\nu \geq 0$, $(Ax - b)^T \mu = 0$ e $x^T \nu = 0$.

5.3.2 Um exemplo de cálculo de direções trade-off

Utilizando-se o Teorema 5.8 é possível se encontrar $T(Z)$. Denotemos $F = \nabla f_i$ e $G = \nabla g_j$ e sem perda de generalidade assumamos que $\text{rank}(F) = \min(k, n)$ e $\text{rank}(F, G) = k$.

Consideremos o caso onde $\text{rank}(F) = n < k$. Neste caso existe uma partição $F = [F^1, F^2]$ tal que

$$\begin{aligned} N(Z) &= \{y \in \mathbb{R}^k \mid Fy - G\mu = 0, \mu \geq 0\} \\ &= \left\{ y \mid \begin{bmatrix} F^1 & F^2 \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} - G\mu = 0, \mu \geq 0 \right\} \\ &= \{y \mid F^1 y^1 + F^2 y^2 - G\mu = 0, \mu \geq 0\} \\ &= \{y \mid F^1 y_1 = -F^2 y_2 + G\mu = 0, \mu \geq 0\} \\ &= \{y \mid y^1 = -(F^1)^{-1} F^2 y^2 + (F^1)^{-1} G\mu = 0, \mu \geq 0\} \end{aligned} \quad (5.2)$$

onde F^1 possuir inversa.

Substituindo y^1 obtida em 5.2 temos,

$$\begin{aligned}
N(Z) &= \left\{ y \mid \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \begin{bmatrix} -(F^1)^{-1}F^2y^2 + (F^1)^{-1}G\mu \\ y^2 \end{bmatrix}, y^2 \geq 0, \mu \geq 0 \right\} \\
&= \left\{ y \mid y = \begin{bmatrix} -(F^1)^{-1}F^2 & (F^1)^{-1}G \\ I & 0 \end{bmatrix} \begin{bmatrix} y^2 \\ \mu \end{bmatrix}, y^2 \geq 0, \mu \geq 0 \right\}
\end{aligned}$$

Através do Teorema 5.8 temos,

$$T(Z) = \left\{ d \mid d^T \begin{bmatrix} -(F^1)^{-1}F^2 & (F^1)^{-1}G \\ I & 0 \end{bmatrix} \leq 0 \right\}$$

Analiseemos o caso citado acima através do exemplo 3.1 onde temos $n = 2$, $k = 3$ e $m = 2$. Em HENING AND BUCHANAN [9] existe uma análise do mesmo exemplo para o caso de maximização do MOP. Analizaremos o caso de minimização de modo detalhado. Segue que

$$F = \begin{bmatrix} 2(x_1 - 3) & 1 & 1 \\ 2(x_2 - 2) & 1 & 2 \end{bmatrix} \text{ e } G = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

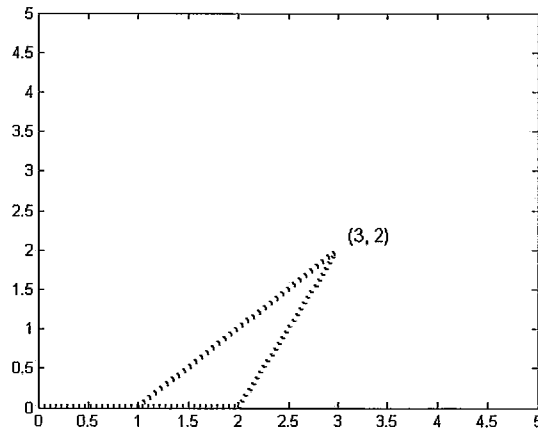


Figura 5.3: Conjunto Pareto ótimo exemplo 3.1

Como já comentado no exemplo 3.1 temos que o conjunto Pareto ótimo é determinado pelo triângulo definido pelos pontos extremos $(1,0)$, $(2,0)$ e $(3,2)$ e o

segmento que conecta os pontos $(0, 0)$ e $(1, 0)$. O conjunto Pareto ótimo impróprio é determinado pelo segmento fechado que conecta $(2, 0)$ e $(3, 2)$ e o segmento aberto que conecta $(1, 0)$ e $(3, 2)$ (veja Figura 5.3). Os cones normal e tangente são calculados para cada ponto Pareto ótimo e as direções *trade-off* são calculadas para cada ponto Pareto próprio, quando existem.

Observe que $n = 2 < k = 3$, assim podemos aplicar o caso em questão. Particionemos F em,

$$F^1 = \begin{bmatrix} 2(x_1 - 3) \\ 2(x_2 - 2) \end{bmatrix} \text{ e } F^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\text{onde } (F^2)^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Dividiremos nossa análise em 4 casos

Caso 1: Seja $x_1 > 0$ e $x_2 > 0$. De acordo com a determinação das partições de F temos

$$N(Z) = \left\{ y \mid y = \begin{bmatrix} I & 0 \\ -(F^2)^{-1}F^1 & (F^2)^{-1}G \end{bmatrix} \begin{bmatrix} y^1 \\ \mu \end{bmatrix}, y^1 \geq 0, \mu \geq 0 \right\}$$

Como $x_1 > 0$ e $x_2 > 0$ temos que o multiplicador $\mu = (\mu_1, \mu_2) = (0, 0)$. Segue

$$\begin{aligned} N(Z) &= \left\{ y \mid y = \begin{bmatrix} 1 & 0 & 0 \\ - \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2(x_1 - 3) \\ 2(x_2 - 2) \end{bmatrix} \\ \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} y_1 \\ 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ y \mid y = \begin{bmatrix} 1 & 0 & 0 \\ -4x_1 + 2x_2 + 8 & -2 & 1 \\ 2x_1 - 2x_2 - 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ 0 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

Pelo Teorema 5.8,

$$T(Z) = \left\{ d \in \mathbb{R}^3 \mid d^T \begin{bmatrix} 1 & 0 & 0 \\ -4x_1 + 2x_2 + 8 & -2 & 1 \\ 2x_1 - 2x_2 - 2 & 1 & -1 \end{bmatrix} \leq 0 \right\} \quad (5.3)$$

Pelo Corolário 5.1, $z \in PP(Z)$ com $x_1 > 0$ e $x_2 > 0$ se, e somente se, $y_1 > 0$, $(-4x_1 + 2x_2 + 8)y_1 > 0$ e $(2x_1 - 2x_2 - 2)y_1 > 0$, ou seja

$$-4x_1 + 2x_2 + 8 > 0 \quad (5.4)$$

$$2x_1 - 2x_2 - 2 > 0. \quad (5.5)$$

As desigualdades 5.4 e 5.5 são satisfeitas para (x_1, x_2) no interior do triângulo determinado pelos pontos extremos $(1, 0)$, $(2, 0)$ e $(3, 2)$ (Figura 5.4). Então

$$PT(Z) = \{d \in \mathbb{R}^3 \mid d_1 + (-4x_2 + 2x_2 + 8)d_2 + (2x_1 - 2x_2 - 2)d_3 = 0\}$$

No interior do triângulo qualquer $d \in PT(Z)$ é direção *trade-off*.

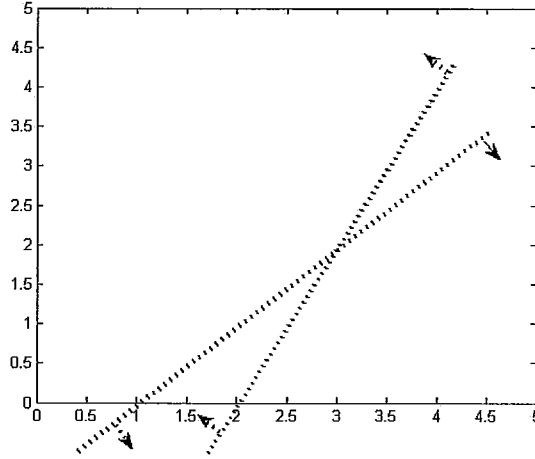


Figura 5.4: Região definida pelas desigualdades 5.4 e 5.5

Caso 2: Seja $x_1 > 0$ e $x_2 > 0$.

$$\begin{aligned} N(Z) &= \left\{ y \mid \begin{bmatrix} 2(x_1 - 3) & 1 & 1 \\ -4 & 1 & 2 \end{bmatrix} y - \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ y \mid y = \begin{bmatrix} 1 & 0 & 0 \\ -4x_1 + 8 & -2 & 1 \\ 2x_1 - 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ 0 \\ \mu_2 \end{bmatrix}, \mu_2 \geq 0 \right\} \end{aligned}$$

Pelo Teorema 5.8,

$$T(Z) = \left\{ d \in \mathbb{R}^3 \mid d^T \begin{bmatrix} 1 & 0 & 0 \\ -4x_1 + 8 & -2 & 1 \\ 2x_1 - 2 & 1 & -1 \end{bmatrix} \leq 0 \right\} \quad (5.6)$$

Pelo Corolário 5.1 $z \in PP(Z)$ com $x_1 > 0$ e $x_2 = 0$ se, e somente se, $y_1 > 0$, $(-4x_1 + 8)y_1 - \mu_2 > 0$ e $(2x_1 - 2)y_1 - \mu_2 > 0$, ou seja

$$-4x_1 + 8 - \frac{\mu_2}{y_1} > 0 \quad (5.7)$$

$$2x_1 - 2 - \frac{\mu_2}{y_1} > 0 \quad (5.8)$$

Quando $1 < x_1 < 2$ e $x_2 = 0$ temos que

$$\begin{aligned} PT(Z) &= \left\{ d \in \mathbb{R}^3 \mid \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -4x_1 + 8 & -2 & 1 \\ 2x_1 - 2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} \\ &= \{ d \in \mathbb{R}^3 \mid d_1 + (8 - 4x_1)d_2 + (2x_1 - 2)d_3 = 0 \}. \end{aligned}$$

Caso 3: Considere $x_1 = 0$ e $x_2 > 0$. Temos,

$$\begin{aligned} N(Z) &= \left\{ y \mid \begin{bmatrix} -6 & 1 & 1 \\ 2(x_2 - 2) & 1 & 2 \end{bmatrix} y - \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ 0 \end{bmatrix} = 0 \right\} \\ &= \left\{ y \mid y = \begin{bmatrix} 1 & 0 & 0 \\ 2x_2 + 8 & -2 & 1 \\ -2x_2 - 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ \mu_1 \\ 0 \end{bmatrix}, \mu_1 \geq 0 \right\} \end{aligned}$$

Pelo Teorema 5.8,

$$T(Z) = \left\{ d \mid d^T \begin{bmatrix} 1 & 0 & 0 \\ 2x_2 + 8 & -2 & 1 \\ -2x_2 - 2 & 1 & -1 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Pelo Corolário 5.1 $z \in PP(Z)$ com $x_1 = 0$ e $x_2 > 0$ se, e somente se, $y_1 > 0$, $(2x_2 + 8)y_1 - 2\mu_1 > 0$ e $(-2x_2 - 2)y_1 + \mu_1 > 0$, com $\mu_1 \geq 0$, ou seja

$$y_1 > 0, \quad (5.9)$$

$$2x_2 + 8 - 2\frac{\mu_1}{y_1} > 0, \quad (5.10)$$

$$-2x_2 + 2 + \frac{\mu_1}{y_1} > 0 \text{ e} \quad (5.11)$$

$$\mu_1 \geq 0. \quad (5.12)$$

As desigualdades 5.9-5.12 não tem solução para $y_1 > 0$, ou seja, não existe direções *trade-off*(Figura 5.3).

Caso 4: Considere $x_1 = 0$ e $x_2 = 0$. Então

$$\begin{aligned} N(Z) &= \left\{ y \mid \begin{bmatrix} -6 & 1 & 1 \\ -4 & 1 & 2 \end{bmatrix} y - \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \mu \geq 0 \right\} \\ &= \left\{ y \mid y = \begin{bmatrix} 1 & 0 & 0 \\ -8 & -2 & 1 \\ -2 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ \mu_1 \\ \mu_2 \end{bmatrix}, \mu_1 \geq 0, \mu_2 \geq 0 \right\} \end{aligned}$$

Pelo Teorema 5.8,

$$T(Z) = \left\{ d \in \mathbb{R}^3 \mid d^T \begin{bmatrix} 1 & 0 & 0 \\ 8 & -2 & 1 \\ -2 & 1 & -1 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Pelo Corolário 5.1, $z \in PP(Z)$ com $x_1 = 0$ e $x_2 = 0$ se, e somente se, $y_1 > 0$, $8y_1 - 2\mu_1 + \mu_2 > 0$, $-2y_1 + \mu_1 - \mu_2 > 0$, $\mu_1 \geq 0$ e $\mu_2 \geq 0$.

Observe que $(0, 0)$ é Pareto ótimo próprio (Figura 5.3) e a única direção direção *trade-off* é

$$\begin{aligned} PT(Z) &= \{d_1 + 8d_2 - 2d_3 = 0, d_3 = d_2 \leq 0, d_1 \geq 0\} \\ &= \left(1, \frac{-1}{6}, \frac{-1}{6}\right). \end{aligned}$$

Capítulo 6

Conclusões e Trabalhos Futuros

Em virtude do que foi mencionado é possível se resolver problemas de otimização multiobjetivo através de métodos de otimização escalar, por exemplo o método ϵ -restrição que apesar de não poder ser capaz de gerar todo conjunto Pareto ótimo determina soluções em função do ϵ dado. Após resolver o problema utiliza-se informações *trade-off* para decidir qual a melhor solução Pareto ótimo.

A determinação da melhor solução é feita de forma interativa (processo de tomada de decisões) com a participação do analista e decisor. O analista fornece informações *trade-off* e o decisor de posse dessas informações toma decisões com o objetivo de encontrar uma melhor solução.

Percebe-se que os multiplicadores de Karush-Kuhn-Tucker, ferramentas matemáticas de grande utilidade em otimização são utilizados nesse processo. Através da análise desses multiplicadores o decisor chega a uma solução desejável. Os multiplicadores podem ser interpretados como taxas, as quais determinam o ganho em uma determinada variável objetivo com uma perda em outra, enquanto as outras permanecem inalteradas. Os multiplicadores de Karush-Kuhn-Tucker são interpretados nesse caso como informações *trade-off*.

Uma outra ferramenta matemática importante em tomada de decisões para problemas de otimização multiobjetivo convexos é o cone tangente, o qual pode ser utilizado para se determinar pontos Pareto ótimo. Assim como o cone tangente o seu polar, o cone normal, também pode ser utilizado para determinar pontos Pareto ótimo. Um tipo especial de cone é o de direções *trade-off*, que é determinado em pontos Pareto ótimo. Através dele se determina se um ponto Pareto ótimo é ou não

uma solução desejável em um conjunto Pareto ótimo.

Pretendo continuar a pesquisar em problemas de otimização multiobjetivo com ênfase em tomada de decisões, cone de direções *trade-off* para problemas não-convexos e aplicações relacionadas a problemas de otimização multiobjetivo. Tenho interesse em problemas de otimização de portfólio e controle ótimo.

Capítulo 7

Apêndice

Neste capítulo apresentamos, sem demonstração, alguns resultados que utilizamos em nossa dissertação. O mesmo se baseia nos seguintes trabalhos ROCKAFELLAR [1], LUENBERGER [8], MÄKELÄ AND NEITTAANMÄKI [10], LIMA [13] e HENIG [14].

7.1 Condições de otimalidade

Considere o problema da seguinte forma

Problema 7.1

$$\begin{aligned} & \text{minimizar } f(x) \\ & \text{sujeito a } g_j(x) \leq 0, \text{ para todo } j = 1, \dots, m \end{aligned}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$, são funções objetivos e restrições respectivamente.

Denotamos $J(x^*) \equiv \{j \mid g_j(x^*) = 0\}$ onde $j = 1, \dots, m$, o conjunto de índices para os quais as restrições g_j do Problema 7.1 são ativas em $x^* \in \mathbb{R}^n$

Definição 7.1 . O ponto $x^* \in S$ é denominado ponto regular se o gradiente das restrições ativas em x^* são lineamente independentes.

Definição 7.2 . Uma restrição ativa é denominada não-degenerada se $\lambda_i > 0$, para cada $j \in J(x^*)$.

Em seguida apresentamos as condições de otimalidade referentes ao Problema 7.1.

Teorema 7.1 .[Condições Necessárias de KKT] *Seja x^* um ponto regular das restrições do Problema (7.1). Para x^* ser um mínimo local do Problema 7.1 é necessário existir um conjunto de multiplicadores de Karush-Kuhn-Tucker $\mu \in \mathbb{R}^m$ tal que,*

$$\begin{aligned} \nabla f(x^*) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x^*) &= 0 \\ \mu_j g_j(x^*) &= 0, \text{ para todo } j = 1, \dots, m \\ \mu_j &\geq 0, \text{ para todo } j = 1, \dots, m \end{aligned}$$

Prova 7.1 . *Veja LUENBERGER [8].*

Exemplo 7.1 *Considere o seguinte problema*

$$\begin{aligned} \text{minimizar } f &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{sujeito a } g_1 &= x_1 + x_2 \leq 2, \\ g_2 &= x_1^2 + (x_2 - 1)^2 \leq 1, \\ -x_1 &\leq 0, \\ -x_2 &\leq 0 \end{aligned}$$

Geometricamente o Teorema 7.1 afirma que para x^ ser o mínimo local do exemplo 7.1 se tem que $-\nabla f(x^*)$ pertence ao cone positivo formado pelo gradiente das restrições ativas.*

Se tem que $(1,1)$ é solução ótima do exemplo 7.1, é um ponto regular das restrições ativas já que $\nabla g_1(1,1) = (1,1)$ e $\nabla g_2(1,1) = (2,0)$ são linearmente independentes. Além disso $-\nabla f(1,1) = (4,1)$ pertence ao cone de positivo formado por $\nabla g_1(1,1)$ e $\nabla g_2(1,1)$. De fato $-\nabla f(1,1) = \nabla g_1(1,1) + \frac{3}{2}\nabla g_2(1,1)$.

Devemos dividir o conjunto de índices do conjunto $J(x^*)$ em dois subconjuntos separáveis $J_+(x^*) = \{j \mid \mu_j > 0\}$ e $J_0(x^*) = \{j \mid \mu_j = 0\}$ de acordo com o multiplicador ótimo de Karush-Kuhn-Tucker μ_j relacionado as restrições g_j , para cada $j \in J(x^*)$.

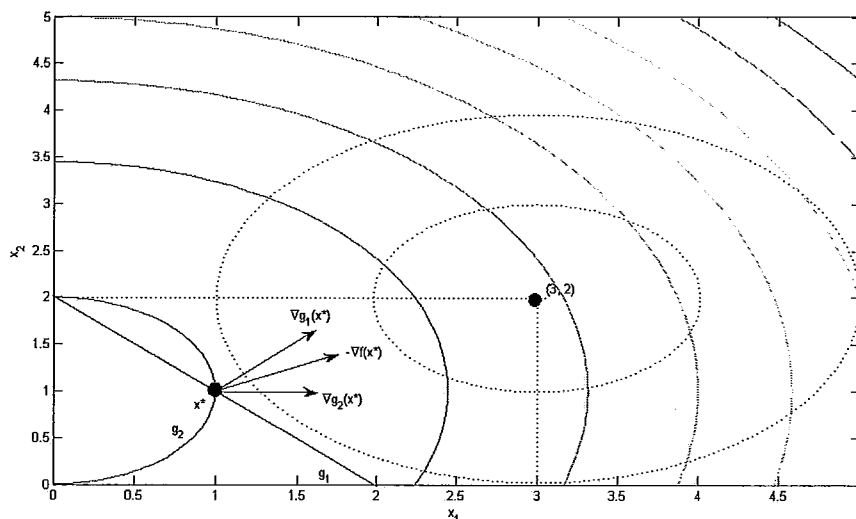


Figura 7.1: Gráfico do exemplo 7.1

Teorema 7.2 .[Condições Suficiente de KKT de Segunda Ordem] Assuma que f e g_j , $j = 1, \dots, m$ são de classe C^2 e ∇^2 é a matrix hessiana. Se as condições do Teorema 7.1 ocorrem em x^* e a matrix

$$\nabla^2 f(x^*) + \sum_{j \in J(x^*)} \mu_j \nabla^2 g_j(x^*)$$

é definida positiva para todos vetores $d \neq 0$ os quais satisfazem as condições

$$\nabla g_j(x^*)^T d = 0, \text{ para todo } j \in J_+(x^*)$$

$$\nabla g_j(x^*)^T d \leq 0, \text{ para todo } j \in J_0(x^*)$$

então x^* é mínimo local do Problema 7.1.

Prova 7.2 . Veja LUENBERGER [8].

7.2 Convexidade

Definição 7.3 . Um conjunto $D \in \mathbb{R}^n$ é convexo se para quaisquer $x, y \in D$ e $0 \leq \alpha \leq 1$, tem-se que $\alpha x + (1 - \alpha)y \in D$.

O conjunto vazio, o conjunto unitário e o próprio \mathbb{R}^n são exemplos triviais de conjuntos convexo. O ponto $\alpha x + (1 - \alpha)y$, onde $0 \leq \alpha \leq 1$ é denominado combinação convexa de x e y .

Definição 7.4 . Um conjunto $C \in \mathbb{R}^n$ é denominado cone quando $d \in C$ implica que $td \in C$ para todo $t \geq 0$.

Se C é um cone não vazio, necessariamente $0 \in C$. Como exemplo de cones temos o \mathbb{R}^n , qualquer subespaço do \mathbb{R}^n como por exemplo o octante não positivo \mathbb{R}_-^n .

Abaixo alguns resultados referente a cone normal e tangente. Estabelecemos algumas características referentes ao cone tangente de Z .

Teorema 7.3 . O cone tangente $T(Z)$ também pode ser escrito como, $T(Z) = cl \{d \in \mathbb{R}^k \mid \text{existe } t > 0 \text{ tal que } z + td \in Z\}$.

Prova 7.3 . Veja por exemplo Teorema 2.2.7 em MÄKELÄ AND NEITTAANMÄKI [10].

c.q.d

Teorema 7.4 . O cone tangente $T(Z)$ do conjunto convexo Z é um cone convexo fechado contendo o zero.

Prova 7.4 . Veja por exemplo Teorema 2.2.4 em MÄKELÄ AND NEITTAANMÄKI [10].

c.q.d

Teorema 7.5 O cone normal pode ser escrito como

$$N(Z) = \{y \in \mathbb{R}^k \mid (z' - z)^T y \leq 0 \text{ para todo } z' \in Z\}.$$

Prova 7.5 . Veja Teorema 2.2.8 por exemplo em MÄKELÄ AND NEITTAANMÄKI [10].

c.q.d

Um resultado natural que segue da definição de dualidade entre cone tangente e cone normal é que o cone normal possui propriedades similares às de cone tagente. Deste modo temos o seguinte teorema.

Teorema 7.6 . *O cone normal $N(Z)$ do conjunto Z é um cone convexo fechado contendo o zero.*

Prova 7.6 . *Segue da definição de cone normal.*

c.q.d

Teorema 7.7 . *Seja D um cone convexo fechado agudo e A um cone fechado tal que $A \cap D = \{0\}$. Então existe uma sequência $\{D_j\}$ de cones convexos agudo fechados separando estritamente A e D .*

Prova 7.7 . *Veja Teorema da Separação por exemplo em HENIG [14].*

Definição 7.5 . *O vetor $d \in \mathbb{R}^n$ é uma direção viável em relação ao conjunto D no ponto x^* , quando existe um $\gamma > 0$ tal que $x^* + td \in D$ para todo $0 \leq t \leq \gamma$.*

Definição 7.6 . *Se $D \in \mathbb{R}^n$ é um conjunto convexo, diz-se que uma função $f : D \rightarrow \mathbb{R}$ é convexa em D quando para quaisquer $x, y \in D$ e $0 \leq \alpha \leq 1$, tem-se*

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

f é estritamente convexa quando a desigualdade acima é estrita para todo $x \neq y$ e $0 < \alpha < 1$.

Definição 7.7 . *O epígrafo da função $f : D \rightarrow \mathbb{R}$ é o conjunto*

$$E_f = \{(x, c) \in D \times \mathbb{R} \mid f(x) \leq c\}.$$

Teorema 7.8 . *Seja $D \in \mathbb{R}^n$ um conjunto convexo. Uma função $f : D \rightarrow \mathbb{R}$ é convexa em D se e somente, se o epígrafo de f é um conjunto convexo em $D \times \mathbb{R}$.*

Prova 7.8 . *Veja ROCKAFELLAR [1].*

Teorema 7.9 . *Um conjunto convexo limitado e fechado é o envoltório convexo de seus pontos extremos.*

Prova 7.9 . *Veja ROCKAFELLAR [1].*

Teorema 7.10 [Lema de Farkas']. Uma desigualdade $\langle a_0, x \rangle \leq 0$ é uma consequência do sistema

$$\langle a_i, x \rangle \leq 0, \quad i = 1, \dots, m$$

se e somente, se existe números reais não-negativos $\lambda_1, \dots, \lambda_m$ tal que

$$\sum_{i=1}^m \lambda_i a_i = a_0.$$

Prova 7.10 . Veja Lema de Farkas' ROCKAFELLAR [1].

7.3 Teorema da Função Implícita e Teorema da Sensibilidade

Teorema 7.11 .[Teorema da Função Implícita] Dada a função $f : U \rightarrow \mathbb{R}$, de classe C^k ($k \geq 1$) no aberto $U \subset \mathbb{R}^{n+1}$, ou seja $(x_0, y_0) \in U$ tal que $f(x_0, y_0) = c$ e $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Existe uma bola $B(x_0; r)$ e um intervalo $I = (y_0 - R; y_0 + R)$ com as seguintes propriedades:

1. $B \times \bar{I} \subset U$ e $\frac{\partial f}{\partial y}(x, y) \neq 0$ para todo $(x, y) \in B \times \bar{I}$;
2. Para todo $x \in B$ existe um único $y = \xi(x) \in I$ tal que $f(x, y) = f(x, \xi(x)) = c$.

A função $\xi : B \rightarrow I$, assim definida é de classe C^k e suas derivadas parciais em cada ponto $x \in B$ são dadas por

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}.$$

Prova 7.11 . Veja Teorema da Função Implícita por exemplo em LIMA [13].

Teorema 7.12 .[Teorema da Sensibilidade] Seja f e g_i funções de classe C^2 e considere o problema,

Problema 7.2

minimizar $f(x)$

sujeito a $g_j(x) \leq c_j$, para todo $j = 1, \dots, m$

Suponha que x^* é uma solução local para do problema quando $c_j = 0$, para todo $j = 1, \dots, m$. Suponha ainda que x^* satisfaz as seguintes condições,

1. O ponto x^* é um ponto regular;
2. Condições suficiente de segunda-ordem são satisfeita em x^* ;
3. As restrições ativas são não-degeneradas em x^* .

então existe uma função de classe C^1 , $\bar{x} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ definida em uma vizinhança de $0 \in \mathbb{R}^m$ tal que $\bar{x}(0) = x^*$ e tal que para todo c na vizinhança de 0 , um vetor $\bar{x}(c)$ é uma solução local para o Problema 7.2. Além disso os multiplicadores de Karush-Kuhn-Tucker ótimos relacionados as restrições na solução x^* tem a seguinte interpretação

$$\frac{\partial f(\bar{x}(0))}{\partial c_j} = -\lambda_j, \text{ para todo } j = 1, \dots, m$$

Prova 7.12 . Veja Teorema da Sensibilidade em LUENBERGER[8].

Notações

\mathbb{N}	: o conjunto dos números naturais;
\mathbb{R}	: o conjunto dos números reais;
\mathbb{R}^n	: o espaço euclidiano de dimensão n ;
\mathbb{R}_-	: o ortante não positivo do \mathbb{R}^n ;
$\langle x, y \rangle$: produto interno euclidiano em $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$;
$\ x\ $: norma euclidiana;
$\ x\ _p$: norma-p de $x \in \mathbb{R}^n$ ($p \geq 1$);
$\ x\ _\infty$: norma-infinito de $x \in \mathbb{R}^n$;
$B(x; \delta)$: bola aberta de centro x e raio δ ;
$B[x; \delta]$: bola fechada de centro x e raio δ ;
$\text{int}X$: interior do conjunto X ;
$\text{cl}X$: o fecho do conjunto X ;
MOP	: problemas de otimização multiobjetivo;
MDMP	: problemas de tomada de decisões multiobjetivo;
S	: subconjunto viável pertencente ao \mathbb{R}^n ;
Z	: subconjunto imagem pertencente ao \mathbb{R}^k ;
x	: vetor decisão ;
z	: vetor objetivo ;
U	: função utilidade;
$\nabla f(x)$: gradiente de f no ponto x ;
$\nabla^2 f(x)$: hessiana de f no ponto x ;
$T_{ij}(x^0, x^*)$: taxa trade-off entre x^0 e x^* envolvendo os objetivos f_i e f_j ;
$t_{ij}(x^*, d)$: taxa trade-off em x^* na direção d envolvendo os objetivos f_i e f_j ;
$m_{ij}(x^*, d)$: taxa marginal de substituição envolvendo f_i e f_j ;

- $L(x, \lambda, \mu)$: função lagrangeana aplicada no ponto x e os multiplicadores λ e μ ;
- $P(Z)$: conjunto Pareto ótimo de Z ;
- $FP(Z)$: conjunto Pareto ótimo fraco de Z ;
- $PP(Z)$: conjunto Pareto ótimo próprio de Z ;
- $D(Z)$: cone de direções viáveis de Z definido em z ;
- $T(Z)$: cone tangente de Z definido em z ;
- $N(Z)$: cone normal do conjunto Z definido em z ;
- C : cone convexo;
- C^k : conjunto das funções que possuem derivadas de ordem k e são contínuas.

Referências Bibliográficas

- [1] ROCKAFELLAR, R. T., *Convex Analysis*. Princeton Univ. Press: Princeton, New Jersey, EUA, 1970.
- [2] MIETTINEN, K., *Nonlinear Multiobjective Optimization*. v. 12. *International Series in Operations Research and Management Science*, Kluwer Academic Publishers, Dordrecht, 1999.
- [3] HAIMES, Y. H., CHANKONG, V., *Multiobjective decision making: theory and methodology*. Dover: Mineola, USA, 1983.
- [4] FERREIRA, P. A. V., *Otimização Multiobjetivo: Teoria e Aplicações*, Ph.D. Thesis, FEEC-UNICAMP, 1998.
- [5] HAIMES, Y. H., CHANKONG, V., “Kuhn-Tucker Multipliers as Trade-Offs in Multiobjective Decision-Making Analysis”, *Automatica*, v. 15, pp. 59–72, 1979.
- [6] HWANG, C., MASUD, A., *Multiple Objective Decision Making - Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag: Berlin, Heidelberg, 1979, Lecture Notes in Economics and Mathematical Systems 164.
- [7] SAKAWA, M., “Interactive multiobjective decision making by the sequential proxy optimization technique:SPOT”, *European Journal of Operational Research*, v. 1, n. 9, pp. 386–396, 1982.
- [8] LUENBERGER, D. E., *Linear and Nonlinear Programming*. Second edition ed. Addison-Wesley: Reading, MA, USA, 1984.

- [9] HENIG, M. I., BUCHANAN, J., “Tradeoff Directions in Multiobjetivos Optimization Problems”, *Mathematical Programming*, v. 78, n. 3, pp. 357–374, 1997.
- [10] MÄKELÄ, M. M., NEITTAANMÄKI, P., *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co.: Sigapore, 1992.
- [11] MIETTIENTE, K., MÄKELÄ, M. M., “On cone characterizations of weak, proper and Pareto optimality in multiobjective optimization”, *Mathematical Methods of Operations Research*, v. 1, n. 53, pp. 233–245, 2001.
- [12] HENIG, M. I., “Proper efficiency with respect to cones”, *Journal of Optimization Theory and Applications*, v. 36, pp. 387–407, 1982.
- [13] LIMA, E. L., *Análise Real*. v. 2. *Coleção Matemática Universitária*, IMPA, 2004.
- [14] HENIG, M. I., “A Cone Separation Theorem”, *Journal of Optimization Theory and Applications*, v. 36, pp. 451–455, 1982.
- [15] BOULOS, P., *Introdução ao Cálculo*. v. 3. Edgard Blücher: São Paulo, 1983.
- [16] CLARKE, F. H., *Optimization and Nonsmooth Analysis*. John Wiley Sons: New York, EUA, 1983.
- [17] DA CUNHA, N., POLAK, E., “Constrained minimization under vector-valued criteria in finite dimensional spaces”, *Journal of Mathematical Analysis and Applications*, v. 19, n. 1, pp. 103–124, 1967.
- [18] GEOFFRION, A., “Proper efficiency and the theory of vector maximization”, *Journal of Mathematical Analysis and Applications*, v. 22, pp. 613–630, 1972.
- [19] IZMAILOV, A., SOLODOV, M., *Otimização - volume 1: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. IMPA: Rio de Janeiro, RJ, 2005.

- [20] LIMA, E. L., *Análise Real*. v. 1. *Coleção Matemática Universitária*, IMPA, 2004.
- [21] LUC, D. T., *Theory of Vector Optimization*. Springer-Verlag: Berlim, Heidelberg, 1989, Lecture Notes in Economics and Mathematical Systems.
- [22] MACULAN, N., FAMPA, M. C., *Otimização linear*. UNB: Brasília, 2006.
- [23] MANGASARIAN, O. L., *Nonlinear Programming*. McGraw-Hill, 1969.
- [24] MARTINEZ, J. M., SANTOS, S. A., *Métodos Computacionais de Otimização*. IMECC-UNICAMP: São Paulo, 1995.
- [25] ROCKAFELLAR, R. T., *La théorie des sub-gradients et ses applications à l'optimisation*. Les Presses de l'Université de Montréal: Montréal, 1979.