Teoria, métodos e aplicações de otimização multiobjetivo

Phillipe Rodrigues Sampaio

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Mestrado em Ciência da Computação
Orientador: Prof. Dr. Ernesto G. Birgin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, março de 2011

Teoria, métodos e aplicações de otimização multiobjetivo

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Phillipe Rodrigues Sampaio
e aprovada pela Comissão Julgadora.

Comissão Julgadora:

- Prof. Dr. Ernesto Julián Goldberg Birgin (orientador) - IME-USP
- Prof. Dr. Walter Figueiredo Mascarenhas - IME-USP
- Prof. Dr. Luis Mauricio Graña Drummond - UFRJ

Agradecimentos

Primeiramente, a Deus, por ter sempre guiado minha vida diligentemente, ajudando-me em todos os momentos difíceis dessa jornada e fazendo-se sempre presente.

À minha família, pelo esforço que fizeram para que eu pudesse chegar até aqui, e pelo apoio e amor incomensuráveis, expressados, constantemente, durante todo esse tempo.

Ao meu orientador, professor Ernesto G. Birgin, pela excelente orientação ao longo do trabalho e pelo conhecimento transmitido nesses anos de pesquisa.

Ao professor Paulo Silva, pelas excelentes aulas ministradas e sugestões a respeito deste trabalho.

Aos meus amigos da Comunidade Evangélica Bethfagé, pelo caloroso suporte, em suas diversas formas.

Aos meus novos e grandes amigos da Assembléia de Deus Catumbi, que tão bem me receberam, de forma que lhes sou eternamente grato.

Aos meus colegas de laboratório, pelos ótimos momentos que tivemos nessa caminhada.

Resumo

Problemas com múltiplos objetivos são muito frequentes nas áreas de Otimização, Economia, Finanças, Transportes, Engenharia e várias outras. Como os objetivos são, geralmente, conflitantes, faz-se necessário o uso de técnicas apropriadas para obter boas soluções. A área que trata de problemas deste tipo é chamada de Otimização Multiobjetivo.

Neste trabalho, estudamos os problemas dessa área e alguns dos métodos existentes para resolvê-los. Primeiramente, alguns conceitos relacionados ao conjunto de soluções são definidos, como o de eficiência, no intuito de entender o que seria a melhor solução para este tipo de problema. Em seguida, apresentamos algumas condições de otimalidade de primeira ordem, incluindo as do tipo Fritz John para problemas de Otimização Multiobjetivo. Discutimos ainda sobre algumas condições de regularidade e total regularidade, as quais desempenham o mesmo papel das condições de qualificação em Programação Não-Linear, propiciando a estrita positividade dos multiplicadores de Lagrange associados às funções objetivo.

Posteriormente, alguns dos métodos existentes para resolver problemas de Otimização Multiobjetivo são descritos e comparados entre si. Ao final, aplicamos a teoria e métodos de Otimização Multiobjetivo nas áreas de *Compressed Sensing* e Otimização de Portfolio. Exibimos então testes computacionais realizados com alguns dos métodos discutidos envolvendo problemas de Otimização de Portfolio e fazemos uma análise dos resultados.

Palavras-chave: Otimização Multiobjetivo, Programação Não-Linear, Compressed Sensing, Otimização de Portfolio.

Abstract

Problems with multiple objectives are very frequent in areas such as Optimization, Economy, Finance, Transportation, Engineering and many others. Since the objectives are usually conflicting, there is a need for appropriate techniques to obtain good solutions. The area that deals with problems of this type is called Multiobjective Optimization.

The aim of this work is to study the problems of such area and some of the methods available to solve them. Firstly, some basic concepts related to the feasible set are defined, for instance, efficiency, in order to comprehend which solution could be the best for this kind of problem. Secondly, we present some first-order optimality conditions, including the Fritz John ones for Multiobjective Optimization. We also discuss about regularity and total regularity conditions, which play the same role in Nonlinear Multiobjective Optimization as the constraint qualifications in Nonlinear Programming, providing the strict positivity of the Lagrange multipliers associated to the objective functions.

Afterwards, some of the existing methods to solve Multiobjective Optimization problems are described and compared with each other. At last, the theory and methods of Multiobjective Optimization are applied into the fields of Compressed Sensing and Portfolio Optimization. We, then, show computational tests performed with some of the methods discussed involving Portfolio Optimization problems and we present an analysis of the results.

Keywords: Multiobjective Optimization, Nonlinear Programming, Compressed Sensing, Portfolio Optimization.

Sumário

Lista de Abreviaturas **ix**

Lista de Figuras **xi**

Lista de Tabelas **xiii**

1 Introdução **1**
1.1 Notação 2

2 O problema **3**
2.1 Formulação matemática 3
2.2 Ordens parciais e cones convexos 3
2.3 Conceitos básicos de Otimização Multiobjetivo 5

3 Condições de otimalidade **11**
3.1 Condições necessárias de primeira ordem 11
3.2 Condições de regularidade 16

4 Métodos para Otimização Multiobjetivo **23**
4.1 Classificação dos métodos 23
4.2 Método dos pesos 24
4.2.1 Descrição do método 25
4.2.2 Relações com o problema multiobjetivo 26
4.2.3 Vantagens e desvantagens 29
4.3 Método NISE 30
4.3.1 Descrição do método 30
4.3.2 Vantagens e desvantagens 32
4.4 Método da ϵ-Restrição 32
4.4.1 Descrição do método 33
4.4.2 Relações com o problema multiobjetivo 33
4.4.3 Vantagens e desvantagens 37
4.5 Método do gradiente 38
4.5.1 Descrição do método 39
4.5.2 Vantagens e desvantagens 40
4.6 Método das direções viáveis 40
4.6.1 Descrição do método 41
4.6.2 Vantagens e desvantagens 42

5 Aplicações e experimentos computacionais **43**
5.1 Compressed Sensing 43
5.1.1 Compressão de sinais 44
5.1.2 Basis pursuit 45
5.2 Otimização de Portfolio 46
5.2.1 Modelo de Markowitz e Otimização Multiobjetivo 48
5.2.2 Fronteira eficiente Risco-Retorno 50
5.2.3 Testes computacionais utilizando o modelo de Markowitz 51
5.2.4 Modelo com 3 funções objetivo 66

6 Conclusões **69**
6.1 Trabalhos futuros 69

A Dados dos experimentos computacionais **71**
A.1 Problema 1 71
A.2 Problema 2 72
A.3 Problema 3 72
A.4 Problema 4 73
A.5 Problema 5 74

Referências Bibliográficas **75**

Índice Remissivo **78**

Lista de Abreviaturas

POM	Problema de Otimização Multiobjetivo.
TDMC	Tomada de Decisão Multicritério.
FJPOM	Condições de Fritz John para Problemas de Otimização Multiobjetivo.
KKT	Condições de otimalidade de Karush-Kuhn-Tucker.
CQA	Condição de Qualificação de Abadie.
CRA	Condição de Regularidade de Abadie.
CRAG	Condição de Regularidade de Abadie Generalizada.
CQMF	Condição de Qualificação de Mangasarian-Fromovitz.
CRMF	Condição de Regularidade de Mangasarian-Fromovitz.
CQG	Condição de Qualificação de Guignard.
CRG	Condição de Regularidade de Guignard.
NISE	*Noninferior Set Estimation.*
CS	*Compressed Sensing.*
MV	Modelo da Média-Variância.
MVA	Modelo da Média-Variância-Assimetria.

Lista de Figuras

1.1 Relações $\leq$ e $<$ no $\Re^2$. 2

2.1 Tipos de eficiência. 6
2.2 Eficiência própria. 8
2.3 Exemplo de solução propriamente eficiente local. 9
2.4 Espaço objetivo. 9

3.1 Exemplo 3.1.1. 13
3.2 Exemplo 3.2.1. 20

4.1 Classificação dos métodos. 24
4.2 Método dos pesos. 26
4.3 Soluções ótimas e inferiores: A e B. 26
4.4 Desvantagens do método dos pesos. 30
4.5 Método NISE. 31
4.6 Erro máximo possível. 31
4.7 O método NISE não está bem definido quando o conjunto viável no espaço objetivo não é convexo. 32
4.8 Método da ϵ-restrição. 34
4.9 Exemplo em que o método dos pesos e o NISE não conseguem aproximar a fronteira eficiente (linha mais grossa), enquanto que o método da ϵ-restrição consegue variando o parâmetro ϵ. Os únicos pontos que os dois primeiros métodos conseguem obter são os minimizadores A e B das funções objetivo. 38

5.1 Soluções A e B obtidas usando as normas L_1 e L_2, respectivamente. 46
5.2 Fronteira eficiente para problemas de Otimização de Portfolio. 51
5.3 Fronteira eficiente do Problema 1 obtida com o método dos pesos e variando δ em $[0, \infty)$. 54
5.4 Fronteira eficiente do Problema 1 aproximada usando o método dos pesos com número de problemas resolvidos igual a 10 e a 50. 55
5.5 Fronteira eficiente do Problema 1 aproximada com o método da ϵ-Restrição e acréscimos de 0.01 e 0.001 unidades ao retorno esperado inicial. 56
5.6 Fronteira eficiente do Problema 1 aproximada com o método NISE e valores da constante multiplicadora iguais a 0.1 e 0.01. 57
5.7 Aproximação da fronteira eficiente do Problema 1 obtida pelo método das direções viáveis com multistart. 59

5.8 Exemplo em que o método das direções viáveis modificado não consegue encontrar uma solução crítica. 60
5.9 Localização na fronteira eficiente das soluções obtidas pelo método das direções viáveis. 60
5.10 Erro do método Z = $\int_a^b (f(x) - g(x))dx$. 61
5.11 Fronteira eficiente do Problema 1. 62
5.12 Fronteira eficiente do Problema 2. 63
5.13 Fronteira eficiente do Problema 3. 63
5.14 Desempenho do método dos pesos no Problema 2 com 50 subproblemas resolvidos. Número de soluções diferentes encontradas × valores do peso w. 64
5.15 Fronteira eficiente do Problema 4. 65
5.16 Fronteira eficiente do Problema 5. 65
5.17 Assimetria presente na função densidade de probabilidade de dois portfolios com médias e variâncias iguais. 66

Lista de Tabelas

5.1 Histórico dos retornos anuais 47
5.2 Investimentos de cada problema. 53
5.3 Desempenho do método dos pesos no Problema 1. 55
5.4 Desempenho do método da ϵ-restrição no Problema 1. 56
5.5 Desempenho do método NISE no Problema 1. 57
5.6 Desempenho dos métodos para o Problema 1. 61
5.7 Desempenho dos métodos para o Problema 2. 62
5.8 Desempenho dos métodos para o Problema 3. 62
5.9 Desempenho dos métodos para o Problema 4. 64
5.10 Desempenho dos métodos para o Problema 5. 64

Capítulo 1

Introdução

Dentro da área de otimização, existem problemas nos quais há vários objetivos a serem alcançados ao invés de apenas um. Geralmente, esses objetivos são conflitantes entre si, e, raramente, existe uma solução que seja ótima para todos os objetivos simultaneamente. Essas são as características de um problema de *Otimização Multiobjetivo* (POM), também conhecido como problema de *Otimização Multicritério* ou *Multiatributo*. O processo de otimização, entretanto, pertence a outro mais geral, chamado de *Tomada de Decisão Multicritério* (TDMC). Essa distinção entre os dois processos se deve, principalmente, às diferentes áreas que trabalham com problemas com múltiplos objetivos usando suas ferramentas específicas, como, por exemplo, Pesquisa Operacional e Economia.

Como o processo de TDMC é muito abrangente, é possível distinguir os problemas existentes e as respectivas áreas que tratam deles em dois tipos. O primeiro, chamado de *Análise de Decisão Multicritério*, é uma área que lida com problemas em que o conjunto de soluções viáveis é discreto, pré-determinado e finito. No segundo tipo, chamado de Otimização Multiobjetivo, as soluções viáveis não são explicitamente conhecidas, mas, geralmente, representadas por funções de restrição [Miettinen, 1999].

Nosso objetivo é estudar os problemas da área de Otimização Multiobjetivo e os métodos existentes para resolvê-los. No Capítulo 2, apresentamos o problema generalizado com respeito a cones convexos e ordens parciais induzidas por eles. Focamos, porém, no caso em que o cone convexo é dado pelo ortante não-negativo do $\Re^p$, onde p é o número de funções objetivo a serem otimizadas. Definimos ainda alguns conceitos relacionados ao conjunto de soluções, como o de *eficiência*, no intuito de entender o que seria a melhor solução para este tipo de problema.

No Capítulo 3, apresentamos algumas condições necessárias de otimalidade de primeira ordem, incluindo as do tipo Fritz John para problemas de Otimização Multiobjetivo, encontradas em [Bigi, 1999] e [Luc, 1989]. Dissertamos ainda sobre condições de regularidade e total regularidade encontradas em [Pappalardo e Bigi, 1999], [Bigi, 1999] e [Maeda, 1994], as quais desempenham o mesmo papel das condições de qualificação em Programação Não-Linear, propiciando a estrita positividade dos multiplicadores de Lagrange associados às funções objetivo. Alguns teoremas encontrados em [Chankong e Haimes, 1982] relacionando condições do tipo Fritz John para um problema de Otimização Multiobjetivo e as condições de Karush-Kuhn-Tucker (KKT) para o problema escalarizado são discutidos no Capítulo 4, bem como as relações entre as soluções dos dois problemas.

Alguns métodos encontrados em [Cohon, 1978] e [Fliege e Svaiter, 2000] para resolução de problemas de Otimização Multiobjetivo são descritos no Capítulo 4. Discutimos sobre suas vantagens e desvantagens, sempre fornecendo exemplos de casos onde o uso de cada um pode não ser adequado.

No Capítulo 5, aplicamos a teoria e métodos de Otimização Multiobjetivo nas áreas de *Compressed Sensing* e Otimização de Portfolio. Exibimos então testes computacionais realizados com alguns dos métodos discutidos envolvendo problemas de Otimização de Portfolio e fazemos uma análise desses resultados. Ao final, apresentamos nossas conclusões sobre o tema estudado com base no desenvolvimento do trabalho e fornecemos sugestões para trabalhos futuros.

1.1 Notação

Denotamos por $\Re$ o conjunto dos números reais e por $\Re^n$, o conjunto de vetores reais n-dimensionais. Para todo vetor $v \in \Re^n$, usamos v_i para indicar sua i-ésima componente.

Vetores no $\Re^n$ serão vistos como vetores coluna. Para todo vetor $v \in \Re^n$, v' representa a transposta de v, sendo um vetor linha n-dimensional.

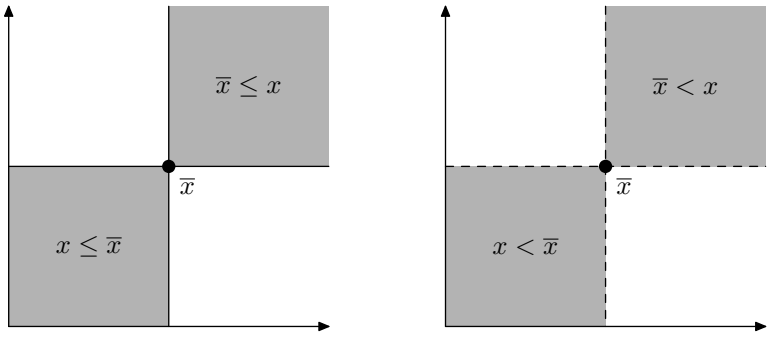
O conjunto formado por todos os números reais não-negativos será denotado por $\Re_+$, e o conjunto formado por todos os números reais estritamente positivos, por $\Re_{++}$. Também faremos uso do negativo de $\Re_+$, i.e., $-\Re_+ = \{-a \mid a \in \Re_+\}$, e do negativo de $\Re_{++}$, $-\Re_{++} = \{-a \mid a \in \Re_{++}\}$.

Sejam $x, y \in \Re^n$. A relações $\leq$ e $<$ em $\Re^n$ serão definidas da seguinte forma:

$$x \leq y \Longleftrightarrow x_i \leq y_i, \ \forall\, i = 1, \ldots, n. \tag{1.1}$$

$$x < y \Longleftrightarrow x_i < y_i, \ \forall\, i = 1, \ldots, n. \tag{1.2}$$

As Figuras 1.1(a) e 1.1(b) ilustram os casos (1.1) e (1.2), respectivamente.



(a) $\{x \mid x \leq \overline{x}\} = \overline{x} - \Re^2_+$ (b) $\{x \mid x < \overline{x}\} = \overline{x} - \Re^2_{++}$

Figura 1.1: *Relações $\leq$ e $<$ no $\Re^2$.*

Sejam $f : \Re^n \to \Re^p$ e $f_i : \Re^n \to \Re$, $i = 1, \ldots, p$, funções diferenciáveis tais que $f(x) = (f_1(x), \ldots, f_p(x))'$. O *Jacobiano* de f em x será denotado por $J_f(x)$, sendo uma matriz $p \times n$, com entradas

$$(J_f(x))_{ij} = \frac{\partial f_i}{\partial x_j}(x).$$

O vetor gradiente de cada f_i será denotado por

$$\nabla f_i = \left(\frac{\partial f_i}{\partial x_1}, \ldots, \frac{\partial f_i}{\partial x_n}\right)'.$$

Seja o conjunto $C \subseteq \Re^n$. Denotaremos o fecho, a envoltória convexa e o interior de C, por $\mathrm{cl}(C)$, $\mathrm{conv}(C)$ e $\mathrm{int}(C)$, respectivamente.

Capítulo 2

O problema

Neste capítulo, apresentaremos o problema generalizado de Otimização Multiobjetivo. Discutiremos ainda a relação entre cones convexos pontudos e ordens parciais induzidas por eles, definindo o problema de Otimização Multiobjetivo generalizado sob a ótica da Análise Convexa, estudado com mais profundidade em [Luc, 1989] e [Sawaragi *et al.*, 1985]. Além disso, apresentaremos os conceitos básicos de Otimização Multiobjetivo encontrados em [Cohon, 1978], [Kuhn e Tucker, 1951], [Geoffrion, 1968], [Luc, 1989], [Sawaragi *et al.*, 1985] e [Zadeh, 1963]. Para facilitar a compreensão de algumas definições, ilustramo-las com exemplos gráficos.

2.1 Formulação matemática

A formulação de um POM generalizado, a ser adotada de agora em diante, é descrita da seguinte forma:

$$\begin{array}{ll} \min & f(x) = (f_1(x), f_2(x), \ldots, f_p(x))' \\ \text{s.a:} & g_1(x) \leq 0, \ldots, g_m(x) \leq 0, \end{array} \tag{2.1}$$

onde p é o número de funções objetivo, cada $g_i : \Re^n \mapsto \Re$ é uma função de restrição e $f : \Re^n \mapsto \Re^p$ é uma função conhecida como *função multiobjetivo*. Cada $f_k : \Re^n \mapsto \Re$ define uma função objetivo. O conjunto viável do problema será definido, portanto, como $X = \{x \mid g_i(x) \leq 0,\ i = 1, \ldots, m\}$. Note que, se $p = 1$, (2.1) torna-se um problema comum de um único objetivo, podendo ser resolvido com os métodos apropriados para este tipo de problema. Logo, estamos interessados em $p > 1$.

2.2 Ordens parciais e cones convexos

Nesta seção, dissertaremos um pouco sobre cones convexos pontudos e as ordens parciais induzidas por eles. A partir desta relação, definiremos o que vem a ser um problema generalizado de Otimização Multiobjetivo com respeito a cones. Estudos mais aprofundados de Otimização Multiobjetivo sob o aspecto teórico da Análise Convexa podem ser encontrados em [Sawaragi *et al.*, 1985] e [Luc, 1989].

Uma relação binária A em $\Re^p$ é um subconjunto do produto $\Re^p \times \Re^p$. Assim, diz-se que um elemento $x \in \Re^p$ está relacionado com $y \in \Re^p$ por A se $(x, y) \in A$.

Definição 2.2.1: Seja A uma relação binária em $\Re^p$. Dizemos que ela é

(1) *reflexiva*, se $(x, x) \in A$ para todo $x \in \Re^p$;

(2) *anti-simétrica*, se $(x, y) \in A$ e $(y, x) \in A$ implica que $x = y$ para $x, y \in \Re^p$;

(3) *transitiva*, se $(x, y) \in A$ e $(y, z) \in A$ implica que $(x, z) \in A$ para $x, y, z \in \Re^p$;

(4) *conectada*, se $(x, y) \in A$ ou $(y, x) \in A$ para todo $x, y \in \Re^p$ com $x \neq y$;
(5) *linear*, se $(x, y) \in A$ implica que $(tx + z, ty + z) \in A$ para $x, y, z \in \Re^p$ e $t > 0$.

Uma *ordem parcial* é uma relação binária que possui as propriedades (1), (2) e (3). Se, além disso, a propriedade (5) é satisfeita, dizemos que a ordem parcial é linear. Quando uma ordem parcial satisfaz a propriedade (4), ela é chamada de *ordem total*. Para exemplificar, consideremos a relação "menor ou igual" em $\Re$. Tal relação é uma ordem total. Quando lidamos, porém, com vetores no $\Re^p$, uma ordem total não é obtida tão naturalmente. Entretanto, $\Re^p$ é dotado de várias ordens parciais, como, por exemplo, aquela definida em (1.1).

Definição 2.2.2: $C \subseteq \Re^p$ é um cone se $tx \in C$ sempre que $x \in C$, para todo $t > 0$. Dizemos que um cone C é *pontudo* se $C \cap -C = \{0\}$.

A definição acima de cone, usada em [Rockafellar, 1997] e em [Bertsekas *et al.*, 2003], implica que um cone pode não conter o vetor nulo. Em outras definições, tem-se $t \geq 0$ em vez de $t > 0$.

Temos a seguinte propriedade para cones quando há convexidade envolvida.

Proposição 2.2.1: Seja C um cone. Então, C é convexo se, e somente se, $C + C \subseteq C$.

Existe ainda uma relação conhecida entre cones convexos pontudos e ordens parciais lineares, como mostra a proposição abaixo.

Proposição 2.2.2: Seja A uma relação binária em $\Re^p$. A é uma ordem parcial linear se, e somente se, existe um cone convexo pontudo $C \subseteq \Re^p$ tal que

$$(x, y) \in A \Longleftrightarrow y - x \in C. \tag{2.2}$$

Prova: ($\Longrightarrow$) Sejam A uma ordem parcial linear e $C = \{x \in \Re^p \mid (0, x) \in A\}$. A propriedade (5) garante que C é um cone. Dados quaisquer $x, y \in C$, por definição de C, temos que $(0, x) \in A$ e $(0, y) \in A$. Pela propriedade (5) de linearidade, temos $(-y, 0) \in A$ e, pela transitividade, obtemos $(-y, x) \in A$. Pela linearidade de A, temos $(0, x + y) \in A$; logo, $x + y \in C$. Portanto, pela Proposição 2.2.1, C é convexo. Se $x \in C \cap -C$, então $(0, x) \in A$ e $(0, -x) \in A$. Como A é linear, temos $(x, 0) \in A$. Como A é anti-simétrica, segue que $x = 0$. Portanto, C é pontudo.
($\Longleftarrow$) Dado um cone convexo pontudo $C \in \Re^p$, mostraremos que a relação binária A estabelecida em (2.2) é uma ordem parcial linear. A propriedade (5) segue de forma imediata da definição de C. Como $0 \in C$, A é reflexiva. Dados $x, y, z \in C$, suponha que $(x, y) \in A$ e $(y, z) \in A$. Por (2.2), temos $y - x \in C$ e $z - y \in C$. Pela convexidade de C e pela Proposição 2.2.1, $z - x \in C$ e, portanto, $(x, z) \in A$. Logo, A é transitiva. Como $(x, y) \in A$ e $(y, x) \in A$ implica que $y - x \in C \cap -C$, podemos concluir que A é anti-simétrica pelo fato de que C é pontudo. $\square$

Denotando $(x, y) \in A$ por $x \leq_C y$, podemos reescrever (2.2) por

$$x \leq_C y \Longleftrightarrow y - x \in C. \tag{2.3}$$

Para exemplificar o resultado da proposição acima, considere o cone convexo pontudo C como sendo

o ortante não-negativo $\Re^p_+$. A ordem parcial linear definida em (1.1) é uma ordem parcial induzida por $\Re^p_+$.

As próximas duas definições são fundamentais para entender a natureza de um problema generalizado de Otimização Multiobjetivo.

Definição 2.2.3: Sejam $Y \subseteq \Re^p$, $C \subseteq \Re^p$ um cone convexo pontudo e $\leq_C$ a ordem parcial linear induzida por C. Um ponto $x \in Y$ é o ponto *minimal ideal* de Y se $x \leq_C y$ para todo $y \in Y$.

Definição 2.2.4: Sejam $Y \subseteq \Re^p$, $C \subseteq \Re^p$ um cone convexo pontudo e $\leq_C$ a ordem parcial linear induzida por C. Um ponto $x \in Y$ é um ponto *minimal* de Y se não existe $y \in Y \backslash \{x\}$ tal que $y \leq_C x$.

Quando a ordem em questão é total, os conceitos de ponto minimal ideal e ponto minimal coincidem. Isso ocorre, por exemplo, quando o cone convexo pontudo é o conjunto dos números reais não-negativos, $\Re_+$, e a ordem parcial linear induzida é a de $\leq$. Entretanto, ao considerar ordens parciais, um ponto minimal ideal pode não existir. Essa é a razão do conceito de ponto minimal ser importante ao lidar com ordens deste tipo.

Denote por $\min_C Y$ o conjunto dos pontos minimais de um conjunto $Y \subseteq \Re^p$ e por $f(X)$, o conjunto formado pelos valores obtidos pela avaliação da função multiobjetivo f nos pontos $x \in X$. Um problema generalizado de Otimização Multiobjetivo com respeito a cones consiste, portanto, em encontrar os pontos $x \in X$ tal que $f(x) \in \min_C f(X)$. Na literatura, ele é conhecido como problema de Otimização Vetorial. Em particular, o presente trabalho dedica-se ao caso em que $C = \Re^p_+$, e, portanto, a ordem parcial definida em (1.1) é considerada. Assim, na próxima seção, introduziremos conceitos básicos de Otimização Multiobjetivo considerando que $C = \Re^p_+$.

2.3 Conceitos básicos de Otimização Multiobjetivo

A partir daqui, particularizamos o cone convexo pontudo em estudo como sendo o ortante não-negativo $C = \Re^p_+$. A ordem parcial linear induzida é, portanto, a definida em (1.1). Como foi dito, queremos encontrar os pontos $x \in X$ tal que $f(x) \in \min_C f(X)$. Os pontos x que satisfazem essa condição possuem um nome específico, como mostra a definição seguinte.

Definição 2.3.1: Uma solução $x^* \in X$ é *eficiente* se não existe outro ponto $x \in X$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$.

Eficiência é um conceito equivalente ao de *Pareto-otimalidade* e *não-inferioridade*. Também está estreitamente ligado ao de *não-dominância*, sendo que este lida com a avaliação da função multiobjetivo numa solução viável em vez da própria solução em sua definição. Dizer que uma solução x^* é eficiente equivale a afirmar que não existe outro ponto viável x tal que os valores das funções objetivo avaliadas nele sejam menores ou iguais, sendo estritamente menor em, pelo menos, uma das funções. São soluções deste tipo em que estamos interessados ao tentar resolver um POM.

Usando a definição de ponto minimal da seção anterior, temos que uma solução x^* é eficiente se $f(x^*)$ é minimal de $f(X)$. Quando um ponto viável não satisfaz a Definição 2.3.1, ele é chamado de *ineficiente*.

Há ainda variações do conceito de eficiência, as quais estão definidas a seguir.

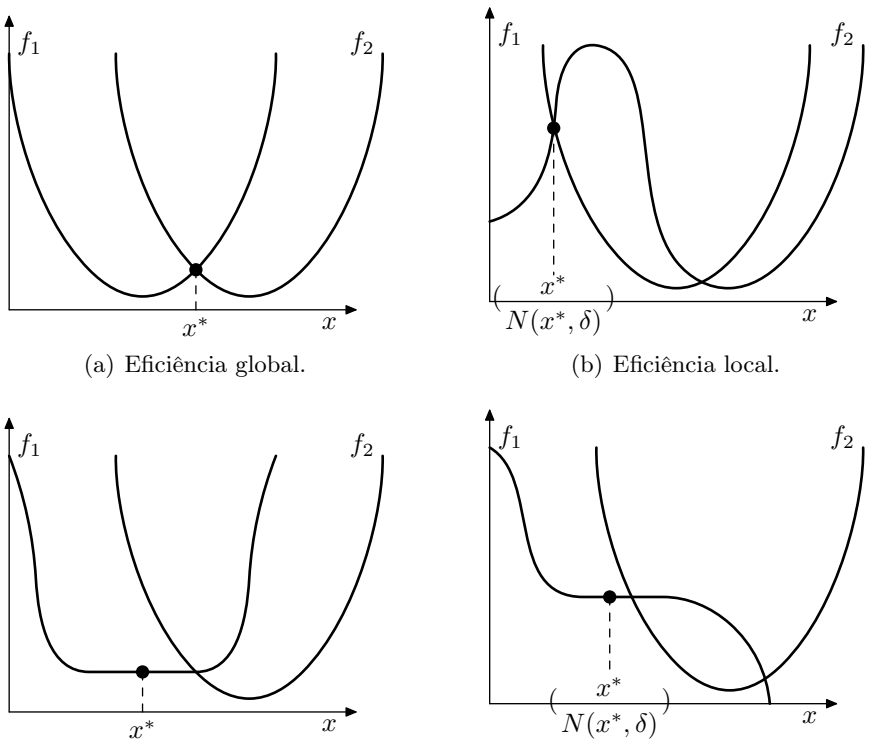
Definição 2.3.2: Uma solução $x^* \in X$ é *eficiente local* se existe $\delta > 0$ tal que x^* é eficiente em $X \cap N(x^*, \delta)$, onde $N(x, \delta) = \{y \mid y \in \Re^n, \|x - y\| < \delta\}$.

Quando estivermos nos referindo tanto a soluções eficientes quanto a soluções eficientes locais em algum momento, denotaremos as primeiras por soluções eficientes globais. Claramente, vemos que, se todas as funções objetivo são convexas, qualquer solução eficiente local é também uma solução eficiente global.

Definição 2.3.3: Uma solução $x^* \in X$ é *fracamente eficiente* se não existe um outro ponto $x \in X$ tal que $f(x) < f(x^*)$.

Definição 2.3.4: Uma solução $x^* \in X$ é *fracamente eficiente local* se existe $\delta > 0$ tal que x^* é eficiente fraca em $X \cap N(x^*, \delta)$.

É fácil ver que, se uma solução minimiza alguma função objetivo que, por sua vez, é convexa, então essa solução é fracamente eficiente. A Figura 2.1 mostra exemplos dos quatro tipos de eficiência definidos anteriormente. Observe que, nos exemplos, as soluções eficientes locais, fracamente eficientes e fracamente eficientes locais não são soluções eficientes globais. Além disso, a solução fracamente eficiente não é uma solução eficiente local e a solução fracamente eficiente local não é uma solução fracamente eficiente.



(a) Eficiência global.

(b) Eficiência local.

(c) Eficiência fraca.

(d) Eficiência fraca local.

Figura 2.1: *Tipos de eficiência.*

Há ainda os conceitos de eficiência *própria* e *imprópria*, introduzidos em [Kuhn e Tucker, 1951],

e, depois, modificados em [Geoffrion, 1968] e discutidos por vários outros autores. Em [Kuhn e Tucker, 1951], os autores mostraram um problema de maximização com dois objetivos no qual uma solução eficiente admite um ganho de primeira ordem em uma função objetivo ao custo da perda de segunda ordem em outra função objetivo. Tal problema, considerando que estamos minimizando funções objetivo, é dado por minimizar $f(x) = (f_1(x), f_2(x))'$, onde $f_1(x) = -x$, $f_2(x) = x^2 - 2x$, sujeito a $x \geq 0$. Qualquer ponto x tal que $x \geq 1$ é uma solução eficiente para o problema. Porém, para qualquer $d > 0$, considerando o ponto $x^* = 1$, temos que $\nabla f_1(x^*)d < 0$ e $\nabla f_2(x^*)d = 0$. Em outras palavras, temos $J_f(x^*)d \leq 0$ e $J_f(x^*)d \neq 0$. Entretanto, temos que $\nabla^2 f_1(x^*)d = 0$ e $\nabla^2 f_2(x^*)d > 0$ para $d > 0$. Assim, vemos que x^* admite um decréscimo de primeira ordem em f_1 ao custo do aumento de segunda ordem em f_2. Para evitar tais tipos de soluções eficientes, Kuhn e Tucker definiram o conceito de eficiência própria. Para apresentá-lo, precisamos definir o seguinte conjunto.

Definição 2.3.5: O *cone linearizado* de X em $x \in X$ é o conjunto dado por

$$L_{\leq}(X, x) = \{y \mid \nabla g_i(x)'y \leq 0,\ i \in A(x)\},$$

onde $A(x) = \{i \mid g_i(x) = 0\}$ é conjunto dos índices das restrições ativas em x.

O cone linearizado de X em $x \in X$ é, portanto, uma aproximação do conjunto X em torno de x, obtida pela linearização das funções de restrição ativas em x. Tal conjunto desempenhará um papel importante nas condições de otimalidade a serem discutidas no Capítulo 3.

Lembramos agora a condição de qualificação de Kuhn e Tucker [Kuhn e Tucker, 1951], a qual é usada na definição de eficiência própria dos mesmos autores.

Definição 2.3.6: Diz-se que o conjunto de restrições X satisfaz a *condição de qualificação de Kuhn e Tucker* (CQKT) se, para cada $x \in X$ tal que $A(x) \neq \emptyset$, toda direção d satisfazendo $\nabla g_j(x)'d \leq 0$, $j \in A(x)$, é tangente a um arco diferenciável contido em X.

Definição 2.3.7: [Kuhn-Tucker] Suponha que as funções f_k, $k = 1, \ldots, p$, e g_i, $i = 1, \ldots, m$, são diferenciáveis e que CQKT é satisfeita. Uma solução eficiente $x^* \in X$ é *propriamente eficiente* se um dos dois casos ocorre:

(1) Se $A(x^*) = \emptyset$, $J_f(x^*)d \leq 0$ e $J_f(x^*)d \neq 0$ para nenhum $d \in \Re^n$;

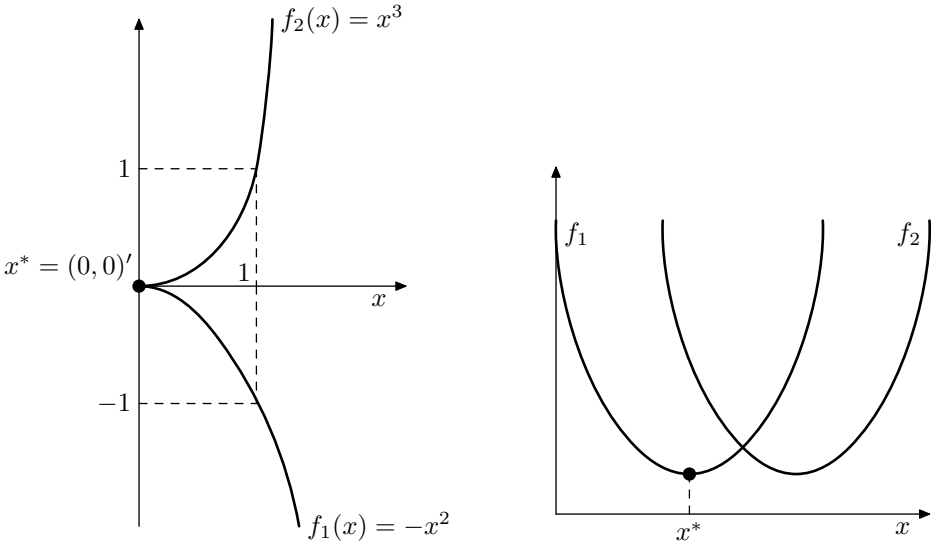
(2) Se $A(x^*) \neq \emptyset$, $J_f(x^*)d \leq 0$ e $J_f(x^*)d \neq 0$ para nenhum $d \in L_{\leq}(X, x^*)$.

Posteriormente, em [Geoffrion, 1968], o autor apresentou um problema com dois objetivos, onde uma solução eficiente é própria, conforme a definição de Kuhn e Tucker, mas apresenta uma anomalia. Tal problema é minimizar $f(x) = (f_1(x), f_2(x))'$, onde $f_1(x) = -x^2$, $f_2(x) = x^3$, sujeito a $x \geq 0$. Na Figura 2.2(a), as funções objetivo estão representadas considerando apenas as partes que satisfazem a restrição $x \geq 0$. Claramente, podemos ver que $x^* = 0$ é uma solução eficiente para o problema. Além disso, como todas as funções são diferenciáveis e vale a condição de qualificação de Kuhn e Tucker, ela também é propriamente eficiente segundo a definição feita anteriormente. Entretanto, movendo-se para um outro ponto x positivo e suficientemente pequeno, o decréscimo em f_1 pode ser arbitrariamente grande em relação ao aumento em f_2. Para excluir soluções eficientes

de tal tipo, Geoffrion modificou a definição de eficiência própria, restringindo ainda mais o número de soluções eficientes consideradas próprias. Esta nova definição de eficiência própria, que segue logo abaixo, será aquela que usaremos neste trabalho.

Definição 2.3.8: [Geoffrion] Uma solução eficiente $x^* \in X$ é *propriamente eficiente* se existe $M > 0$ tal que, para cada $i = 1, \ldots, p$ e cada $x \in X$ satisfazendo $f_i(x) < f_i(x^*)$, existe pelo menos um $j \neq i$ tal que $f_j(x) > f_j(x^*)$ e $f_i(x^*) - f_i(x) \leq M(f_j(x) - f_j(x^*))$. Uma solução eficiente que não é propriamente eficiente é chamada de *impropriamente eficiente*.

Observe que a definição de Geoffrion não supõe que vale qualquer condição de qualificação. Geoffrion também prova em seu artigo que, supondo que as funções f_k e g_i são diferenciáveis e vale a condição de qualificação de Kuhn e Tucker, toda solução eficiente que é propriamente eficiente segundo a Definição 2.3.8 também o é segundo a Definição 2.3.7. O inverso, porém, nem sempre é verdade, como mostra o exemplo que vimos anteriormente. A Figura 2.2(b) mostra um exemplo de uma solução propriamente eficiente segundo a Definição 2.3.8.



(a) Kuhn e Tucker (b) Geoffrion

Figura 2.2: *Eficiência própria.*

Definição 2.3.9: Uma solução eficiente $x^* \in X$ é *propriamente eficiente local* se existe $\delta > 0$ tal que x^* é propriamente eficiente em $X \cap N(x^*, \delta)$.

A Figura 2.3 mostra um exemplo de uma solução x^* impropriamente eficiente, mas que é uma solução propriamente eficiente local para uma determinada vizinhança de x^*. Ela é impropriamente eficiente, pois a função objetivo f_2 decresce ilimitadamente enquanto que a função objetivo f_1 varia apenas dentro de um intervalo para $x \geq x_0$.

Além disso, em [Isermann, 1974], o autor mostrou que, quando o problema multiobjetivo possui funções objetivo e restrições lineares, todas as soluções eficientes são propriamente eficientes. Além dos conceitos já apresentados, existem outros como, por exemplo, *super eficiência*, introduzido em [Borwein e Zhuang, 1993], e *eficiência estrita*, apresentado em [Jiménez, 2002], mas que não farão parte do nosso escopo.

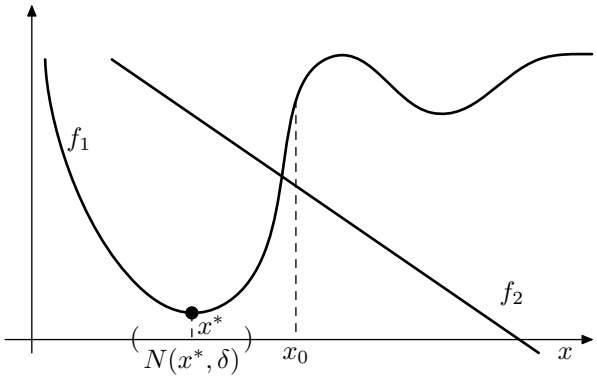


Figura 2.3: *Exemplo de solução propriamente eficiente local.*

O conjunto de soluções eficientes de um POM é chamado de *fronteira eficiente*. Na literatura, a fronteira eficiente é muitas vezes chamada de *fronteira de Pareto*. Dependendo do problema, podemos enxergar tal conjunto graficamente, no *espaço objetivo*. O espaço objetivo difere do *espaço de decisão*, que é definido pelas variáveis de decisão do problema, da seguinte forma: os eixos são representados pelas funções objetivo do problema, logo, cada ponto nesse espaço possui suas coordenadas definidas pelos valores das funções objetivo avaliadas em um ponto no espaço de decisão.

Seja o conjunto $\overline{X} = \{f(x) \mid x \in X\}$. Tal conjunto corresponde ao conjunto viável no espaço objetivo. Na Figura 2.4, podemos ver o conjunto viável de um problema com dois objetivos no espaço objetivo, assim como a fronteira eficiente do problema representada nesse espaço. Os pontos A e B são exemplos de soluções eficientes, e C, de uma solução inferior. Em rigor, A e B são pontos no espaço de decisão, logo, eles deveriam estar representados como $f(A)$ e $f(B)$ no espaço objetivo. Porém, para facilitar a notação, representaremos os pontos do espaço de decisão no espaço objetivo sem ter que avaliar a função multiobjetivo neles. A parte do gráfico onde a linha é mais grossa constitui a fronteira eficiente. Fixando um ponto nessa curva, não é possível se mover para outro ponto viável decrescendo simultaneamente as duas funções objetivo, ou pelos menos decrescer o valor de uma função e deixar a outra de igual valor. É fácil ver também que C é uma solução ineficiente, pois podemos melhorar os valores da funções objetivo individuais movendo-nos para a esquerda e/ou para baixo.

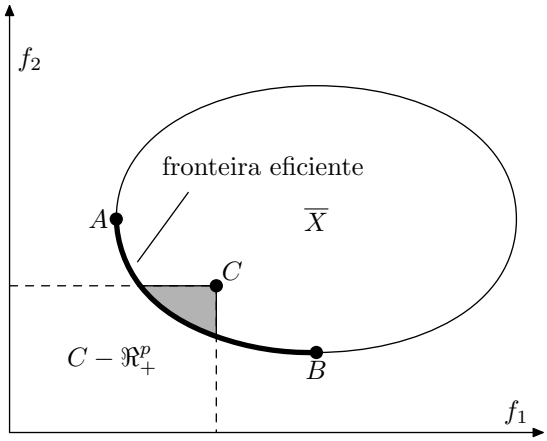


Figura 2.4: *Espaço objetivo.*

Pela Figura 2.4, podemos facilmente concluir que um ponto viável x é eficiente se, e somente

se, $(f(x) - \Re^p_+) \cap \overline{X} = \{f(x)\}$; ou seja, não existe nenhuma direção partindo de $f(x)$ no sentido do ortante não-positivo do $\Re^p$ que conduza a um ponto viável no espaço objetivo diferente de $f(x)$. Examinando o ponto C, vemos que existem várias direções partindo de C no sentido do ortante não-positivo que conduzem a outros pontos viáveis; movendo-se para qualquer um deles, poder-se-ia melhorar o valor de alguma função objetivo, mantendo as outras com valor, no máximo, o mesmo.

Muitas vezes, desejamos obter não apenas uma solução eficiente, mas toda a fronteira eficiente, ou, pelo menos, uma aproximação dela. Ainda que haja várias soluções eficientes encontradas no problema, sabemos que apenas uma será escolhida no final. A essa solução escolhida se dá o nome de *solução de melhor compromisso*. A razão entre a quantidade que deve ser aumentada para um objetivo a fim de que haja diminuição de um outro objetivo é denominada de *tradeoff*. Os tradeoffs e as soluções eficientes são dois tipos de informações importantes para que o tomador de decisão possa escolher a solução de melhor compromisso.

Utilizando o raciocínio de Zadeh [Zadeh, 1963], dado um ponto viável x, podemos definir três conjuntos disjuntos no espaço de decisão $\Re^n$:

$$\begin{aligned} Q^{<}(x) &= \{y \mid f(y) \leq f(x), f(y) \neq f(x)\}, \\ Q^{\geq}(x) &= \{y \mid f(y) \geq f(x)\}, \\ Q^{\sim}(x) &= \{y \mid f(y) \not\leq f(x) \text{ e } f(y) \not\geq f(x)\}. \end{aligned}$$

O primeiro conjunto, $Q^{<}(x)$, contém todos os pontos que são melhores que x, ou seja, todos os vetores que possuem valores das funções objetivo avaliadas neles menores ou iguais aos de x, sendo estritamente menor em, pelo menos, uma das funções. Dizemos que um ponto y deste tipo é *superior* em relação a x. O segundo conjunto, $Q^{\geq}(x)$, contém todos os pontos que são *inferiores* a x. O terceiro conjunto contém todos os pontos não comparáveis a x.

Zadeh salientou que qualquer vetor do $\Re^n$ pertence a um desses três conjuntos. Logo, temos que $\Re^n = Q^{<}(x) \cup Q^{\geq}(x) \cup Q^{\sim}(x)$. Observe, então, que para uma solução viável x ser eficiente é necessário e suficiente que se tenha $X \cap Q^{<}(x) = \emptyset$.

Capítulo 3

Condições de otimalidade

Na área de Otimização Multiobjetivo, muitos artigos têm mostrado condições necessárias de primeira ordem ao longo dos anos. Entre estas, condições do tipo Fritz John têm sido analisadas para estes problemas. Diferentemente do caso de problemas com um único objetivo, há várias possibilidades para os valores dos multiplicadores de Lagrange associados às funções objetivo. O vetor de multiplicadores pode ser inteiramente igual ao vetor zero ou não, mas isso não garante que todos os multiplicadores serão diferentes de zero. Quando um multiplicador associado a alguma função objetivo é igual ao zero, esta função não desempenha papel algum na condição de otimalidade, o que não é desejado. As condições de qualificação em Programação Não-Linear, envolvendo apenas as funções de restrição, permitem que o vetor de multiplicadores associados às funções objetivo seja diferente de zero. Entretanto, tais condições não propiciam a estrita positividade de cada multiplicador. Para tal feito, é necessário usar condições que envolvam as funções objetivo. Estas condições são chamadas de condições de regularidade e são um tipo de extensão das condições de qualificação usuais. Em [Bigi, 1999], [Pappalardo e Bigi, 1999] e [Maeda, 1994], várias condições de regularidade são discutidas. Particularmente, em [Bigi, 1999] e [Pappalardo e Bigi, 1999], os autores apresentam uma classificação concernente à regularidade do problema multiobjetivo.

Apresentaremos, neste capítulo, condições necessárias de primeira ordem encontradas em [Bigi, 1999], [Luc, 1989] e [Maeda, 1994], incluindo as condições do tipo Fritz John para problemas de Otimização Multiobjetivo. Apresentaremos também algumas condições de regularidade e total regularidade encontradas em [Bigi, 1999], [Pappalardo e Bigi, 1999] e [Maeda, 1994].

Para os resultados deste capítulo, supomos todas as funções objetivo e de restrição continuamente diferenciáveis.

3.1 Condições necessárias de primeira ordem

Consideremos primeiro o caso onde o problema multiobjetivo é irrestrito, ou seja, $X = \Re^n$. Temos o seguinte resultado.

Proposição 3.1.1: Se $x^* \in \Re^n$ é uma solução fracamente eficiente local, então não existe $d \in \Re^n$ tal que

$$\nabla f_i(x^*)'d < 0, \quad \forall\, i = 1, \ldots, p. \tag{3.1}$$

Prova: Suponha que x^* é uma solução fracamente eficiente local e que (3.1) ocorre para alguma direção $d \in \Re^n$. Então, temos que

$$J_f(x^*)d < 0.$$

Pela diferenciabilidade das funções objetivo e pelo Teorema do Valor Médio, temos que, para todo $\epsilon > 0$, existe um $s \in (0,1)$ tal que

$$f(x^* + \epsilon d) = f(x^*) + \epsilon J_f(x^* + s\epsilon d)d.$$

Como todas as funções objetivo são continuamente diferenciáveis, temos que, para todo $\epsilon > 0$ suficientemente pequeno, $J_f(x^* + s\epsilon d)d < 0$, e, portanto, $f(x^* + \epsilon d) < f(x)$. Isso contradiz o fato de x^* ser uma solução fracamente eficiente local. $\square$

Consideremos agora o problema multiobjetivo restrito. Definiremos a seguir um conjunto importante para o desenvolvimento das condições de otimalidade adiante.

Definição 3.1.1: Seja $Y \subseteq \Re^q$. O *cone tangente* a Y em $y \in \mathrm{cl}(Y)$ é dado por

$$T(Y, y) = \{w \in \Re^q \mid \exists \{w_n\} \to w,\ \exists \{t_n\} \downarrow 0 \text{ t.q. } y + t_n w_n \in Y\}.$$

O cone tangente é uma aproximação do conjunto Y ao redor do ponto y por retas tangentes a tal conjunto. Ele sempre é fechado, sendo convexo se assim for o conjunto Y. O cone tangente, porém, pode não ser convexo. As condições necessárias que apresentaremos adiante estão direta ou indiretamente relacionadas a tal cone.

Proposição 3.1.2: Se $x^* \in X$ é uma solução fracamente eficiente local, então

$$T(f(X \cap N), f(x^*)) \cap -\Re^p_{++} = \emptyset, \tag{3.2}$$

para alguma vizinhança N de x^*.

Prova: Suponha que $x^* \in X$ é uma solução fracamente eficiente local. Por contradição, dado $n \in \mathbb{N}$, suponha que existe $z_n \in T(f(X \cap N(x^*, 1/n)), f(x^*)) \cap -\Re_{++}$. Logo, existem sequências $\{z_k\} \to z_n$, $\{t_k\} \downarrow 0$ e $x_k \in X \cap N(x^*, 1/n)$ tais que $f(x^*) + t_k z_k = f(x_k)$ para qualquer $k \in \mathbb{N}$. Seja $\bar{k}$ grande o suficiente para se ter $z_{\bar{k}} \in -\Re^p_{++}$ e defina $x_n := x_{\bar{k}}$. Portanto, tem-se $\{x_n\} \to x^*$ e $f(x^*) - f(x_n) \in \Re^p_{++}$, contradizendo a suposição inicial. $\square$

O resultado acima pode não ser muito útil na prática devido à dificuldade de se verificar a condição (3.2). Para $p > 1$, a condição acima não é suficiente, como é mostrado no exemplo seguinte.

Exemplo 3.1.1: Considere $n = p = 2$, f como sendo a função identidade e $X = \{x \in \Re^2 \mid -x_1^2 \leq x_2 \leq -x_1^3\ ;\ x_1 \leq 0\}$. Seja $x^* = (0,0)'$. Como podemos ver na Figura 3.1, tem-se $T(f(X \cap N), f(x^*)) = -\Re_+ \times \{0\}$ para qualquer vizinhança N de x^*, e, portanto, a condição (3.2) é satisfeita. Entretanto, como f é a função identidade, x^* não é uma solução fracamente eficiente local.

Proposição 3.1.3: Se (3.2) é satisfeita para alguma vizinhança N de $x^* \in X$, então a condição

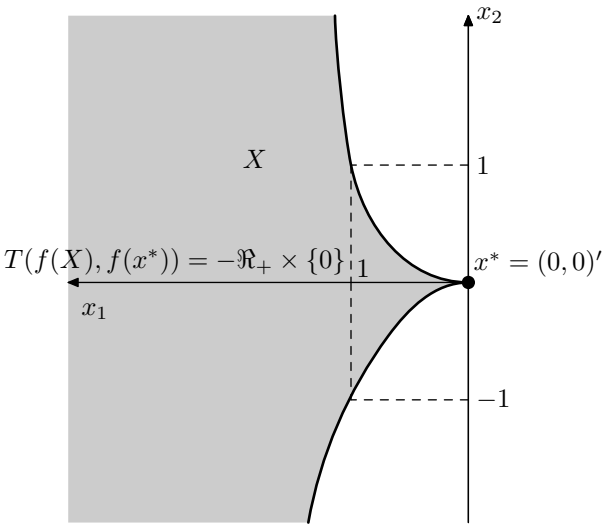


Figura 3.1: *Exemplo 3.1.1.*

abaixo

$$\max_k [\nabla f_k(x^*)' w] \geq 0 \tag{3.3}$$

é satisfeita para todo $w \in T(X, x^*)$.

Prova: Seja $w \in T(X, x^*)$. Logo, existem sequências $\{w_n\} \to w$ e $\{t_n\} \downarrow 0$ tais que $x_n := x^* + t_n w_n \in X$. Além disso, $x_n \in N$ para n suficientemente grande. Seja $d_n := t_n^{-1}(f(x_n) - f(x^*))$. Temos, então, $d_n = J_f(x^*)w_n + \epsilon_n$, com $\{\epsilon_n\} \to 0$. Assim, temos $d_n \to J_f(x^*)w$ e $f(x_n) = f(x^*) + t_n d_n$, o que garante que $J_f(x^*)w \in T(f(X \cap N), f(x^*))$. Assim, (3.3) segue de (3.2). $\square$

Combinando as duas últimas proposições, obtemos o seguinte resultado, que será usado por um dos métodos de resolução de problemas de Otimização Multiobjetivo a serem discutidos no próximo capítulo.

Proposição 3.1.4: Se $x^* \in X$ é uma solução fracamente eficiente local, então (3.3) é satisfeita para todo $w \in T(X, x^*)$.

As próximas condições usam conceitos de cone tangente e cone linearizado. Lembramos que o cone linearizado de X em um ponto $x^* \in X$ é dado por

$$L_{\leq}(X, x^*) = \{y \mid \nabla g_i(x^*)' y \leq 0,\ i \in A(x^*)\},$$

onde $A(x^*) = \{i \mid g_i(x^*) = 0\}$ é conjunto dos índices das restrições ativas em x^*.

Definição 3.1.2: O *cone linearizado fraco* de X em $x^* \in X$ é o conjunto dado por

$$L_{<}(X, x^*) = \{y \mid \nabla g_i(x^*)' y < 0,\ i \in A(x^*)\}.$$

O cone linearizado fraco de X em $x^* \in X$ é um conjunto aberto. Além disso, temos que $L_{<}(X, x^*) \subseteq \operatorname{int}(L_{\leq}(X, x^*))$.

Proposição 3.1.5: Seja $x^* \in X$. Então, temos

$$L_<(X, x^*) \subseteq T(X, x^*) \subseteq L_\leq(X, x^*). \tag{3.4}$$

Prova: Primeiramente, provaremos que $T(X, x^*) \subseteq L_\leq(X, x^*)$. Seja $w \in T(X, x^*)$. Então, existem sequências $\{w_n\} \to w$ e $\{t_n\} \downarrow 0$ tais que $x_n := x^* + t_n w_n \in X$. Portanto, para todo $i \in A(x^*)$, temos

$$0 \geq t_n^{-1}(g_i(x_n) - g_i(x^*)) = \nabla g_i(x^*)' w_n + \epsilon_n,$$

com $\{\epsilon_n\} \to 0$. Tomando o limite, obtemos $\nabla g_i(x^*)' w \leq 0$.

Provaremos, agora, que $L_<(X, x^*) \subseteq T(X, x^*)$. Seja $w \in L_<(X, x^*)$, $\{t_n\} \downarrow 0$ e $x_n := x^* + t_n w_n$. Considere $i \in A(x^*)$. Temos, então, para $w_n \to w$,

$$t_n^{-1} g_i(x_n) = t_n^{-1} g_i(x^*) + \nabla g_i(x^*)' w_n + \epsilon_n \to \nabla g_i(x^*)' w < 0.$$

Para $i \notin A(x^*)$, pela continuidade de g_i, temos $g_i(x_n) \to g_i(x^*) < 0$. Portanto, existe n_0 suficientemente grande tal que, para todo $n \geq n_0$, $x_n \in X$. Considerando as sequências com índices iniciais iguais a n_0, podemos concluir que $w \in T(X, x^*)$. $\square$

As proposições 3.1.5 e 3.1.4 combinadas fornecem o resultado subseqüente relacionado com a impossibilidade de resolução de um dado sistema homogêneo linear.

Proposição 3.1.6: Se $x^* \in X$ é uma solução fracamente eficiente local, então o sistema

$$\begin{array}{rcll} \nabla f_k(x^*)' w & < & 0, & k = 1, \ldots, p, \\ \nabla g_i(x^*)' w & < & 0, & \forall\, i \in A(x^*), \end{array} \tag{3.5}$$

não possui solução $w \in \Re^n$.

Como é bem delineado em [Bigi, 1999], a proposição acima em conjunto com o Teorema da Alternativa de Motzkin [Mangasarian, 1969] propiciam o desenvolvimento de condições do tipo Fritz John para Otimização Multiobjetivo. O Teorema da Alternativa de Motzkin está enunciado logo abaixo. É fácil perceber a ligação entre ele e as condições necessárias de Fritz John.

Proposição 3.1.7: [Teorema da Alternativa de Motzkin] Sejam a_k, b_i, $c_j \in \Re^n$, com $k \in I^< \neq \emptyset$, $i \in I^\leq$ e $j \in I^=$, onde $I^<$, $I^\leq$ e $I^=$ são conjuntos de índices finitos. Então, o sistema

$$\begin{array}{rcll} a_k' w & < & 0, & k \in I^<, \\ b_i' w & \leq & 0, & i \in I^\leq, \\ c_j' w & = & 0, & j \in I^=, \end{array} \tag{3.6}$$

não possui solução se, e somente se, existem números λ_k, $k \in I^<$, não todos nulos, σ_i, $i \in I^\leq$, e θ_j, $j \in I^=$, tais que

$$\begin{array}{c} \displaystyle\sum_{k \in I^<} \lambda_k a_k + \sum_{i \in I^\leq} \sigma_i b_i + \sum_{j \in I^=} \theta_j c_j = 0, \\ \lambda_k \geq 0,\ k \in I^<,\ \sigma_i \geq 0,\ i \in I^\leq. \end{array} \tag{3.7}$$

Proposição 3.1.8: [Condições de Fritz John para POM (FJPOM)] Se x^* é uma solução fracamente eficiente local, então existem $\lambda \in \Re^p$ e $\mu \in \Re^m$ não todos nulos tais que

$$\sum_{k=1}^{p} \lambda_k \nabla f_k(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0, \tag{3.8}$$

$$\mu_i g_i(x^*) = 0, \; i = 1, \ldots, m, \tag{3.9}$$

$$\lambda_k \geq 0, \; k = 1, \ldots, p, \; \mu_i \geq 0, \; i = 1, \ldots, m. \tag{3.10}$$

Prova: A prova segue da Proposição 3.1.6, aplicando o Teorema da Alternativa de Motzkin (Proposição 3.1.7) e fazendo $\mu_i = 0$ para todo $i \notin A(x^*)$. □

Observe que o resultado acima não implica que o vetor λ é diferente de zero. As condições de qualificação, quando satisfeitas no problema, levam a tal conclusão. Tomaremos, como exemplo, a condição de qualificação de Abadie [Abadie, 1967], umas das condições mais fracas em Programação Não-Linear.

Definição 3.1.3: O conjunto de restrições X satisfaz a *condição de qualificação de Abadie* (CQA) em $x^* \in X$ se

$$L_{\leq}(X, x^*) \subseteq T(X, x^*).$$

Satisfeita a condição de qualificação de Abadie, pela Proposição 3.1.5, temos que $L_{\leq}(X, x^*) = T(X, x^*)$. Pode-se obter um resultado ainda mais forte que o da Proposição 3.1.6 com esta condição.

Proposição 3.1.9: Suponha que CQA é satisfeita. Se $x^* \in X$ é uma solução fracamente eficiente local, então o sistema

$$\begin{array}{rcll} \nabla f_k(x^*)'w & < & 0, & k = 1, \ldots, p, \\ \nabla g_i(x^*)'w & \leq & 0, & \forall\, i \in A(x^*), \end{array} \tag{3.11}$$

não possui solução $w \in \Re^n$.

Prova: Por contradição, suponha que existe $w \in \Re^n$ tal que w resolve (3.11). A condição de qualificação de Abadie implica que $w \in T(X, x^*)$. Pela Proposição 3.1.3, $\nabla f_k(x^*)'w \geq 0$ para, pelo menos, um índice $k \in \{1, \ldots, p\}$, contradizendo (3.11). □

Combinada com o Teorema da Alternativa de Motzkin (Proposição 3.1.7), a proposição acima garante que o vetor de multiplicadores associados às funções objetivo é diferente do vetor zero, como está enunciado logo abaixo.

Proposição 3.1.10: Suponha que CQA é satisfeita. Se x^* é uma solução fracamente eficiente, então existem vetores $\lambda \geq 0$, não nulo, e $\mu \geq 0$ satisfazendo as condições de Fritz John para POM.

3.2 Condições de regularidade

Como vimos na seção anterior, as condições de qualificação garantem a existência de um vetor não nulo de multiplicadores de Lagrange associados às funções objetivo. Entretanto, tais condições não garantem a estrita positividade de cada um dos multiplicadores. É necessário, portanto, usar as funções objetivo na construção de condições que produzam esse resultado. Em [Maeda, 1994], o autor propôs condições envolvendo as funções objetivo que garantem a existência de multiplicadores diferentes de zero associados a cada função objetivo. Como o termo original *"constraint qualification"* diz respeito apenas às funções de restrição, as condições que também envolvem as funções objetivo são chamadas de *condições de regularidade*. Como há várias possibilidades de valores e existência para o vetor de multiplicadores de Lagrange associados às funções objetivo, Pappalardo e Bigi propuseram uma classificação concernente à regularidade do problema, introduzindo o conceito de *total regularidade* [Pappalardo e Bigi, 1999].

Denote por $M(x)$ o conjunto dos vetores dos multiplicadores de Fritz John associados a um ponto $x \in X$, i.e., $M(x) = \{(\lambda, \mu) \mid \text{vale } (3.8)\text{-}(3.10)\}$. Esta é a definição da classificação propriamente dita.

Definição 3.2.1: Dado $x \in X$ tal que $M(x) \neq \emptyset$, diz-se que x é

(1) *regular fraco*, se $\exists\ (\lambda, \mu) \in M(x)$ tal que $\lambda \neq 0$;

(2) *regular*, se $\exists\ (\lambda, \mu) \in M(x)$ tal que $\lambda_k > 0$, $k = 1, \ldots, p$;

(3) *totalmente regular fraco*, se $\forall\ (\lambda, \mu) \in M(x)$, tem-se $\lambda \neq 0$;

(4) *totalmente regular*, se $\forall\ (\lambda, \mu) \in M(x)$, tem-se $\lambda_k > 0$, $k = 1, \ldots, p$.

Claramente, vemos que (4) implica (3) e (2), os quais implicam (1). Em [Maeda, 1994], foi provado que, se alguma das condições de regularidade apresentadas no artigo for satisfeita num ponto eficiente, então este é regular. Para apresentar as condições de regularidade encontradas em tal artigo, faremos antes algumas definições.

Definição 3.2.2: O conjunto das direções para $x^* \in X$ que não promovem aumento de f até primeira ordem é dado por

$$D(x^*) = \{d \in \Re^n \mid J_f(x^*)d \leq 0\}.$$

Definição 3.2.3: O conjunto dos pontos viáveis melhores ou de mesma qualidade de $x^* \in X$ é dado por

$$X^{\leq} = \{x \in X \mid f(x) \leq f(x^*)\}.$$

Definição 3.2.4: [Maeda] O problema satisfaz a *condição de regularidade de Abadie* (CRA) em $x^* \in X$ se

$$[L_{\leq}(X, x^*) \cap D(x^*)] \subseteq T(X^{\leq}, x^*))).$$

Como foi dito anteriormente, as condições de regularidade envolvem as funções objetivo para

garantir a estrita positividade dos multiplicadores de Lagrange associados às funções objetivo. Assim, a condição de regularidade de Abadie envolve as funções objetivo no lado direito através da própria definição do cone tangente de $X^{\leq}$ em x^*. No lado esquerdo, adicionamos o conjunto $D(x^*)$ devido ao cone linearizado de X em x^*, por definição, não trazer nenhuma informação do conjunto $X^{\leq}$, o qual é o ponto-chave para tal condição por carregar informação sobre as funções objetivo. Em outras palavras, temos, no lado esquerdo, o cone linearizado de $X^{\leq}$ em x^*.

Definição 3.2.5: O conjunto dos pontos viáveis superiores ou de mesma qualidade de $x^* \in X$ quando a função f_i é removida de f é dado por

$$X_i^{\leq} = \{x \in X \mid f_k(x) \leq f_k(x^*),\ k = 1, \ldots, p,\ k \neq i\}.$$

Definição 3.2.6: [Maeda] O problema satisfaz a *condição de regularidade de Abadie generalizada* (CRAG) em $x^* \in X$ se

$$[L_{\leq}(X, x^*) \cap D(x^*)] \subseteq \bigcap_{i=1}^{p} T(X_i^{\leq}, x^*))).$$

Pelo fato do cone tangente ser isótono com respeito à inclusão ($Y' \subseteq Y$ implica que $T(Y', y) \subseteq T(Y, y)$), pode-se concluir que, se a CRA é satisfeita em $x^* \in X$, também o é a CRAG.

Antes de apresentar a condição de regularidade de Guignard, lembraremos a condição de qualificação de Guignard para Programação Não-Linear.

Definição 3.2.7: O conjunto de restrições X satisfaz a *condição de qualificação de Guignard* (CQG) em $x^* \in X$ se

$$L_{\leq}(X, x^*) \subseteq \mathrm{cl}(\mathrm{conv}(T(X, x^*))). \tag{3.12}$$

Como o cone linearizado de X em x^* é sempre convexo, ele não pode ser igual a um conjunto que não é convexo. Assim, a condição de qualificação de Guignard estende a condição de qualificação de Abadie para cones tangentes $T(X, x^*)$ não convexos, sendo, portanto, mais fraca. O conjunto do lado direito de (3.12) é, por sua vez, chamado de *cone pseudotangente* de X em x^*. Definiremos agora a condição de regularidade de Guignard encontrada em [Maeda, 1994].

Definição 3.2.8: [Maeda] O problema satisfaz a *condição de regularidade de Guignard* (CRG) em $x^* \in X$ se

$$[L_{\leq}(X, x^*) \cap D(x^*)] \subseteq \bigcap_{i=1}^{p} \mathrm{cl}(\mathrm{conv}(T(X_i^{\leq}, x^*))).$$

A condição da Definição 3.2.8 é mais fraca que a da Definição 3.2.6. Facilmente, concluímos que, se a CRAG é satisfeita em $x^* \in X$, também o é a CRG.

O inverso da CRG sempre ocorre, como é provado em [Maeda, 1994]; logo, tem-se a igualdade

quando tal condição de regularidade é satisfeita. Tal condição de regularidade pode ainda ser melhor entendida através do seguinte resultado envolvendo ela e a resolução de um sistema. A resolução desse sistema, por sua vez, está relacionada à existência de multiplicadores de Lagrange estritamente positivos associados às funções objetivo através de um teorema da alternativa, como será mostrado em sequência. Analogamente, podemos seguir este raciocínio com outras condições de regularidade apresentadas adiante.

A prova da proposição abaixo possui uma pequena modificação em relação à original encontrada em [Maeda, 1994] devido às diferentes definições de cone tangente. Porém, as duas definições são equivalentes. Para um estudo sobre outras definições de cone tangente, veja [Giorgi e Guerraggio, 1992].

A CRG permite provar um resultado mais forte do que o da Proposição 3.1.9, como pode ser visto abaixo.

Proposição 3.2.1: Se $x^* \in X$ é uma solução eficiente e a CRG é satisfeita em x^*, então o sistema

$$\begin{aligned} \nabla f_k(x^*)'d &\leq 0, \quad k = 1, \ldots, p, \\ \nabla f_k(x^*)'d &< 0, \quad \text{para pelo menos um } k, \\ \nabla g_i(x^*)'d &\leq 0, \quad \forall\, i \in A(x^*), \end{aligned}$$

não possui solução $d \in \Re^n$.

Prova: Por contradição, suponha que exista $d \in \Re^n$ que seja solução do sistema acima. Logo, temos que $d \in L_{\leq}(X, x^*) \cap D(x^*)$. Sem perda de generalidade, podemos supor que

$$\begin{aligned} &\nabla f_1(x^*)'d < 0, \\ &\nabla f_k(x^*)'d \leq 0, \quad k = 2, \ldots, p. \end{aligned}$$

Pela Definição 3.2.8, temos que $d \in \text{cl}(\text{conv}(T(X_1^{\leq}, x^*)))$. Portanto, existe uma sequência $\{d_m\} \subseteq \text{conv}(T(X_1^{\leq}, x^*))$ tal que $\lim_{m\to\infty} d_m = d$. Para cada d_m, $m = 1, 2, \ldots$, existem números K_m, $\lambda_{mk} \geq 0$ e vetores $d_{mk} \in T(X_1^{\leq}, x^*)$, $k = 1, 2, \ldots, K_m$, tais que

$$\sum_{k=1}^{K_m} \lambda_{mk} = 1, \quad \sum_{k=1}^{K_m} \lambda_{mk} d_{mk} = d_m.$$

Para cada $m = 1, 2, \ldots,$ e $k = 1, 2, \ldots, K_m$, como $d_{mk} \in T(X_1^{\leq}, x^*)$, pela definição de cone tangente, existem sequências $\{d_{mk}^n\} \to d_{mk}$ e $\{t_{mk}^n\} \downarrow 0$ tais que

$$x^* + t_{mk}^n d_{mk}^n \in X_1^{\leq}.$$

Portanto, para todo n, temos que

$$\begin{aligned} f_i(x^* + t_{mk}^n d_{mk}^n) &\leq f_i(x^*), \quad i = 2, 3, \ldots, p, \\ g_j(x^* + t_{mk}^n d_{mk}^n) &\leq 0, \quad j \in A(x^*). \end{aligned}$$

Por outro lado, como $x^* \in X$ é uma solução eficiente, para todo n, temos que

$$f_1(x^* + t_{mk}^n d_{mk}^n) \geq f_1(x^*).$$

Pelas três desigualdades acima, segue que

$$\begin{aligned}
&\nabla f_1(x^*)'d_{mk} \geq 0,\\
&\nabla f_k(x^*)'d_{mk} \leq 0, \quad k = 2, \ldots, p,\\
&\nabla g_j(x^*)'d_{mk} \leq 0, \quad j \in A(x^*).
\end{aligned}$$

Pela linearidade e continuidade do produto interno, obtemos

$$\begin{aligned}
&\nabla f_1(x^*)'d \geq 0,\\
&\nabla f_k(x^*)'d \leq 0, \quad k = 2, \ldots, p,\\
&\nabla g_j(x^*)'d \leq 0, \quad j \in A(x^*).
\end{aligned}$$

Contradição. □

Proposição 3.2.2: Se $x^* \in X$ é uma solução eficiente e a CRG é satisfeita em x^*, então x é regular.
Prova: Daremos apenas a idéia da prova. Se $x^* \in X$ é uma solução eficiente e a CRG é satisfeita em x^*, então o sistema da Proposição 3.2.1 não possui solução. Podemos, então, usar o Teorema da Alternativa de Tucker [Mangasarian, 1969] (p. 29-30) e obter o resultado desejado. □

Lembraremos agora a condição de qualificação de Mangasarian-Fromovitz, para apresentar a condição de regularidade de Mangasarian-Fromovitz. Como nosso problema (2.1) envolve apenas restrições de desigualdade, não há suposição de independência linear.

Definição 3.2.9: O problema com restrições de desigualdade satisfaz a *condição de qualificação de Mangasarian-Fromovitz* (CQMF) em $x^* \in X$ se o sistema

$$\nabla g_i(x^*)'d < 0, \quad \forall\, i \in A(x^*), \tag{3.13}$$

possui solução $d \in \Re^n$.

Definição 3.2.10: [Maeda] O problema satisfaz a *condição de regularidade de Mangasarian-Fromovitz* (CRMF) em $x^* \in X$ se $\nabla f_k(x^*)$, $k = 1, \ldots, p$, são linearmente independentes e o sistema

$$\begin{aligned}
&\nabla f_k(x^*)'d = 0, \quad k = 1, \ldots, p,\\
&\nabla g_i(x^*)'d < 0, \quad \forall\, i \in A(x^*),
\end{aligned}$$

possui solução $d \in \Re^n$.

Da mesma forma com as outras condições de regularidade, Maeda prova que, se CRMF é satisfeita em $x^* \in X$, também o é a CRG. O Exemplo 3.2.1 mostra, porém, que a condição de regularidade de Guignard, que é suficiente para garantir a regularidade, não é necessária para garantir a regularidade fraca.

Exemplo 3.2.1: Considere o seguinte problema irrestrito:

$$\begin{aligned}
\min \quad & f(x) = (-x^3, x^3)'\\
\text{s.a:} \quad & x \in \Re.
\end{aligned}$$

Como pode ser observado na Figura 3.2, todos os pontos em $\Re$ são eficientes. Considere $x = 0$. Temos, então,

$$\begin{aligned} [L_{\leq}(X, x) \cap D(x)] &= \Re, \\ T(X_1^{\leq}, x) &= \{x \in \Re \mid x \leq 0\}, \\ T(X_2^{\leq}, x) &= \{x \in \Re \mid x \geq 0\}. \end{aligned}$$

Portanto, CRG não é satisfeita em x. Entretanto, para qualquer $\lambda_1 > 0$ e $\lambda_2 > 0$, temos que

$$\lambda_1 \nabla f_1(x) + \lambda_2 \nabla f_2(x) = \lambda_1(-3(0)^2) + \lambda_2(3(0)^2) = 0,$$

onde $f_1(x) = -x^3$ e $f_2(x) = x^3$.

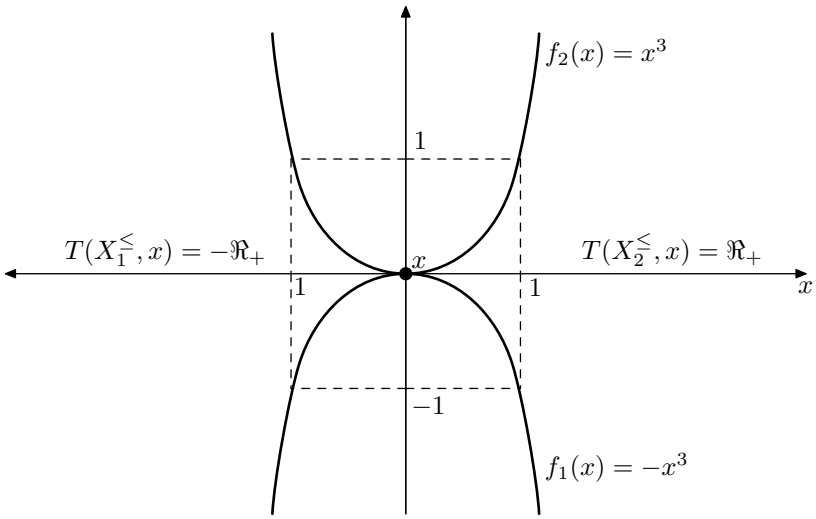


Figura 3.2: *Exemplo 3.2.1.*

Apesar do bom resultado que a Proposição 3.2.2 fornece, esta envolve tão somente soluções eficientes globais, regularidade e regularidade fraca. Posteriormente, Pappalardo e Bigi obtiveram resultados mais fortes envolvendo eficiência fraca local, total regularidade e total regularidade fraca [Pappalardo e Bigi, 1999]. Para a apresentar esses resultados, usamos as definições a seguir.

Definição 3.2.11: O conjunto dos pontos viáveis que não permitem que $x^* \in X$ seja eficiente fraco quando a função f_i é removida de f unido ao conjunto formado pelo ponto x^* é dado por

$$X_i^{<} = \{x \in X \mid f_k(x) < f_k(x^*),\ k = 1, \ldots, p,\ k \neq i\} \cup \{x^*\}.$$

Para efeito de seus resultados, Pappalardo e Bigi modificaram a condição de regularidade de Guignard em [Maeda, 1994], obtendo a seguinte condição de regularidade.

Definição 3.2.12: [Pappalardo e Bigi] O problema satisfaz a *condição de regularidade de Guignard* (CRG) em $x^* \in X$ se

$$[L_{\leq}(X, x^*) \cap D(x^*)] \subseteq \bigcap_{i=1}^{p} \mathrm{cl}(\mathrm{conv}(T(X_i^{<}, x^*))).$$

Proposição 3.2.3: Se $x^* \in X$ é uma solução fracamente eficiente local e a CRG é satisfeita em x^*, então, para cada $i \in \{1, \ldots, p\}$, o sistema

$$\begin{aligned} &\nabla f_i(x^*)'w < 0, \\ &\nabla f_k(x^*)'w \leq 0, \quad k = 1, \ldots, p, \ k \neq i, \\ &\nabla g_i(x^*)'w \leq 0, \quad \forall\, i \in A(x^*), \end{aligned} \tag{3.14}$$

não possui solução em $w \in \Re^n$.
Prova: Pode ser encontrada em [Bigi, 1999].

Usando a proposição acima, que trata da impossibilidade de resolução de p sistemas lineares, um para cada função objetivo, consegue-se chegar ao seguinte resultado envolvendo a regularidade e a eficiência fraca local.

Proposição 3.2.3: Se $x^* \in X$ é uma solução fracamente eficiente local e a CRG é satisfeita em x^*, então, x^* é regular.
Prova: Aplica-se o Teorema da Alternativa de Motzkin (Proposição 3.1.7) ao sistema (3.14) para cada $p \in \{1, \ldots, p\}$ e somam-se os multiplicadores de Lagrange resultantes. □

Para garantir a total regularidade, Pappalardo e Bigi modificaram a condição de regularidade de Mangasarian-Fromovitz em [Maeda, 1994], obtendo a seguinte condição de regularidade.

Definição 3.2.13: [Pappalardo e Bigi] O problema satisfaz a *condição de regularidade de Mangasarian-Fromovitz* (CRMF) em $x^* \in X$ se, para todo $s \in \{1, \ldots, p\}$, existe $d_s \in \Re^n$ tal que

$$\begin{aligned} &\nabla f_k(x^*)'d_s < 0, \quad k = 1, \ldots, p, \ k \neq s, \\ &\nabla g_i(x^*)'d_s < 0, \quad \forall\, i \in A(x^*). \end{aligned} \tag{3.15}$$

A condição de qualificação de Mangasarian-Fromovitz, assim como a condição de regularidade de Mangasarian-Fromovitz modificada por Pappalardo e Bigi, são necessárias e suficientes para garantir total regularidade fraca e total regularidade, respectivamente, como é demonstrado logo abaixo.

Proposição 3.2.4: Seja $x^* \in X$ tal que $M(x^*) \neq \emptyset$. Então,

(1) x^* é totalmente regular fraco se, e somente se, CQMF é satisfeita em x^*.

(2) x^* é totalmente regular se, e somente se, CRMF é satisfeita em x^*.

Prova: (1) ($\Longleftarrow$) Seja $(\lambda, \mu) \in M(x^*)$, e suponha, por contradição, que $\lambda = 0$. Logo, pelas condições FJPOM, devemos ter $\mu \neq 0$. Aplicando o Teorema da Alternativa de Motzkin (Proposição 3.1.7), como (3.13) possui solução $d \in \Re^n$, temos que ter $\mu = 0$, chegando numa contradição. Portanto, temos $\lambda \neq 0$.
($\Longrightarrow$) Seja $(\lambda, \mu) \in M(x^*)$. Por contradição, suponha que CQMF não é satisfeita, ou seja, (3.13) não possui solução $d \in \Re^n$. Pelo Teorema da Alternativa de Motzkin, existe $\mu \geq 0$, $\mu \neq 0$, tal que $\sum_{i=1}^m \mu_i g_i(x^*) = 0$. Logo, $(0, \mu) \in M(x^*)$, contradizendo a total regularidade fraca de x^*.
(2) A prova é análoga à do caso (1). □

Em [Maciel *et al.*, 2009], os autores relacionaram as diferentes condições de regularidade com a noção de independência linear positiva de conjuntos de gradientes, recentemente introduzida em Programação Não-Linear para uma aplicação de Programação Quadrática Sequencial [Qi e Wei, 2000] e provada ser uma condição de qualificação [Andreani *et al.*, 2005]. Entre outros resultados, foi provado em [Maciel *et al.*, 2009] que a condição de regularidade de independência linear positiva é necessária e suficiente para a total regularidade, sendo equivalente à condição de regularidade de Mangasarian-Fromovitz por Pappalardo e Bigi.

Capítulo 4

Métodos para Otimização Multiobjetivo

Neste capítulo, apresentamos alguns métodos para resolver problemas de Otimização Multiobjetivo encontrados em [Cohon, 1978], [Cohon *et al.*, 1979], [Zadeh, 1963] e [Haimes *et al.*, 1971]. Tais métodos consistem na escalarização do problema multiobjetivo. Chamamos de *escalarização* a transformação de um problema com vários objetivos em um outro com apenas um objetivo. Métodos que não envolvem escalarização também serão discutidos, a saber, o método de descida do gradiente para problemas de Otimização Multiobjetivo irrestritos [Fliege e Svaiter, 2000] e o método de direções viáveis [Fliege e Svaiter, 2000] para o caso restrito. Há outros métodos que dispensam escalarização, mas que não serão abordados neste trabalho, como, por exemplo, o método de Newton [Fliege *et al.*, 2009] e o método Simplex Multiobjetivo [Zeleny, 1974]. A menos que se mencione explicitamente, não se supõe aqui que as funções f_k e g_i são diferenciáveis.

4.1 Classificação dos métodos

Num processo de tomada de decisão, existem dois papéis distintos e importantes: o do analista, que gera as soluções eficientes do problema com base em informações fornecidas a ele; e do tomador de decisão, que receberá do analista as soluções geradas e escolherá aquela que mais lhe agrade - a solução de melhor compromisso. Quando lidamos com vários objetivos, pode haver ainda vários tomadores de decisão, cada um com seus objetivos próprios conflitantes. Tal tipo de problema é comumente encontrado na esfera política do governo, por exemplo.

Há diversos métodos para resolver problemas envolvendo mais de um objetivo. Cada um deles possui características e aplicações diferentes. A escolha de um método para resolver um POM deve ser realizada com cuidado, pois não existe um método melhor que todos os outros. Há de se analisar a intenção inicial do analista e do tomador de decisão no processo. Algumas vezes, precisamos encontrar a fronteira eficiente com exatidão, em outros casos, uma aproximação já basta. Cada tipo de problema possui suas características próprias; por isso, um método pode ser muito bom para um tipo de problema e ineficiente para outros. O tipo de análise que contempla a verificação da quantidade de tomadores de decisão no processo e as características peculiares do problema a ser tratado pode ser visto como uma *análise de contexto* do processo de tomada de decisão. Essa análise de contexto é uma das características que servem para classificar os tipos de métodos existentes para resolução de um POM.

Outra característica importante é a direção do fluxo de informação no processo. Há dois tipos comuns de fluxo de informação: do analista para o tomador de decisão (bottom-up) e do tomador de decisão para o analista (top-down). No primeiro tipo, o analista apenas gera as soluções eficientes e entrega-as ao tomador de decisão para que ele escolha a que melhor lhe agrada. Não

há preferência por algum objetivo ou solução *a priori* do tomador de decisão que é fornecida ao analista e incorporada ao método de resolução do problema. No segundo tipo, o tomador de decisão informa suas preferências ao analista, que as incorporará ao método para resolver o problema.

Os métodos do tipo bottom-up são popularmente conhecidos como *métodos de geração*. Este trabalho enfatiza métodos desse tipo.

Na Figura 4.1, podemos ver os tipos de métodos existentes distinguidos pelo fluxo de informação e análise de contexto do problema. Essa classificação dos métodos segue a de Cohon, em seu livro "*Multiobjective Programming and Planning*" [Cohon, 1978].

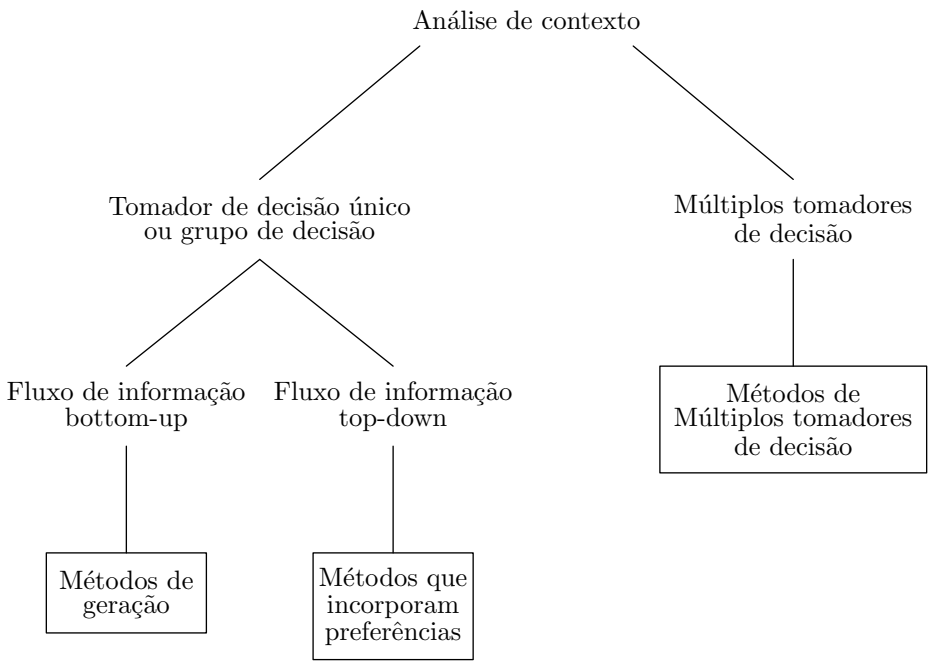


Figura 4.1: *Classificação dos métodos.*

4.2 Método dos pesos

Um dos métodos mais comuns para resolver um POM é o método dos pesos (em inglês, *the weighting method*), onde todas as funções objetivo são combinadas em uma única função objetivo usando um vetor de pesos $w \geq 0$, com $\|w\|_1 = 1$. Dessa forma, o problema original transforma-se num problema de um único objetivo com as restrições originais, ao qual nos referiremos adiante como *problema dos pesos* ou como $\mathrm{P}(w)$, quando quisermos nos referir a um vetor de pesos w. O método dos pesos serve para obter uma aproximação da fronteira eficiente e foi experimentado, primeiramente, em Zadeh [Zadeh, 1963]. A vantagem deste método é sua simplicidade.

Apesar do método dos pesos ser definido aqui como um método de geração, ele pode incorporar preferências de tomadores de decisão também. A principal diferença na abordagem deste método que o classifica em uma das categorias descritas na seção anterior está na escolha do vetor w. Se o vetor de pesos w for fornecido pelo tomador de decisão ao analista, ele deixa de ser um método de geração. Em outras palavras, o fluxo de informação é o principal fator classificatório do método e o ponto de partida para o uso da técnica. Neste trabalho, supomos que o vetor de pesos w é escolhido pelo analista.

4.2.1 Descrição do método

O problema (2.1) escalarizado, usando este método com um dado vetor de pesos $w \geq 0$ tal que $\|w\|_1 = 1$, transforma-se no seguinte problema:

$$\begin{aligned} \mathrm{P}(w): \quad \min \quad & \sum_{k=1}^{p} w_k f_k(x) \\ \text{s.a:} \quad & g_i(x) \leq 0, \quad i = 1, \ldots, m. \end{aligned} \tag{4.1}$$

Satisfazendo algumas condições discutidas na próxima subseção, cada solução ótima do problema escalarizado (4.1) é uma solução eficiente para o problema (2.1). A solução ótima também seria a solução de melhor compromisso se o vetor w fosse escolhido *a priori* pelo tomador de decisão.

Para enxergar melhor o fato de que a solução ótima de (4.1) é uma solução eficiente, podemos tomar um exemplo simples de um problema com apenas duas funções objetivo. Neste caso, o problema (4.1) torna-se, para um dado $w \geq 0$, com $\|w\|_1 = 1$,

$$\begin{aligned} \mathrm{P}(w): \quad \min \quad & w_1 f_1(x) + w_2 f_2(x) \\ \text{s.a:} \quad & g_i(x) \leq 0, \qquad i = 1, \ldots, m. \end{aligned} \tag{4.2}$$

Sabemos que as curvas de nível da função objetivo de (4.2) são dadas por

$$w_1 f_1 + w_2 f_2 = c, \tag{4.3}$$

onde c é uma constante qualquer. Se $w_2 \neq 0$, podemos ainda reescrever (4.3) da seguinte forma:

$$f_2 = \frac{1}{w_2}(c - w_1 f_1). \tag{4.4}$$

Como $w \geq 0$, podemos ver em (4.4) que a função objetivo de (4.2) terá sempre um coeficiente angular não-positivo. Combinando esse fato com o de que o problema é de minimização, concluímos que a solução ótima cairá sempre na parte mais à esquerda do conjunto viável no espaço objetivo, onde encontra-se a fronteira eficiente, como vimos na Seção 2.3. Para diferentes valores de w, poderemos encontrar diferentes soluções eficientes. Entretanto, pode acontecer de uma única solução eficiente ser a solução ótima de problemas distintos para escolhas diferentes do vetor w.

Suponha que o conjunto das soluções viáveis de (4.2) no espaço objetivo é representado na Figura **??**. Claramente, podemos ver na figura que diferentes escolhas de w podem acarretar na descoberta de soluções eficientes distintas.

O método dos pesos consiste, portanto, em resolver (4.1) para vetores w distintos afim de obter uma aproximação da fronteira eficiente a partir das soluções ótimas encontradas até que haja uma razoável distribuição delas, ou seja, não exista um espaço muito grande entre as soluções descobertas. Geralmente, os valores iniciais para w são dados pelos vetores canônicos $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$, que possuem 1 na i-ésima entrada e 0 nas outras.

Na Figura 4.3, podemos ver um caso onde uma solução ótima de um problema dos pesos com dois objetivos não é uma solução eficiente. Neste caso, a escolha de $w = (w_1, w_2)'$ foi a de $w_1 = 0$ e $w_2 > 0$; ou seja, estamos tão somente minimizando a função f_2. Para este exemplo, as soluções A, B e C são ótimas, mas apenas C é uma solução eficiente. Veremos na próxima subseção que,

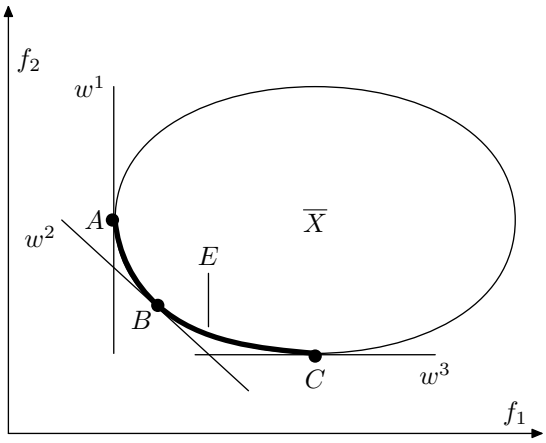


Figura 4.2: *Método dos pesos.*

para uma solução ótima x de um $\mathrm{P}(w)$ ser uma solução eficiente, deve-se ter um dos dois casos: $w > 0$ ou x é solução ótima única de $\mathrm{P}(w)$. Na subseção subsequente, também serão dados maiores detalhes sobre as relações entre as soluções de um problema dos pesos e um POM, bem como as relações entre as condições de Fritz John para os dois tipos de problemas.

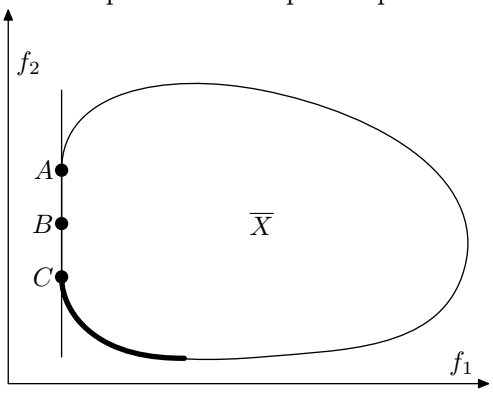


Figura 4.3: *Soluções ótimas e inferiores: A e B.*

4.2.2 Relações com o problema multiobjetivo

Apresentaremos agora algumas relações envolvendo soluções eficientes para o problema multiobjetivo e soluções ótimas para o problema dos pesos, bem como as relações entre as condições de Fritz John para os dois tipos de problemas.

Proposição 4.2.1: Se x^* é solução ótima única de $\mathrm{P}(w)$ para um dado vetor $w \geq 0$, com $\|w\|_1 = 1$, então x^* é uma solução eficiente para o problema multiobjetivo.
Prova: Seja x^* a solução ótima única de $\mathrm{P}(w)$ para um dado vetor $w \geq 0$, com $\|w\|_1 = 1$. Temos, então,

$$\sum_{k=1}^{p} w_k f_k(x^*) < \sum_{k=1}^{p} w_k f_k(x), \tag{4.5}$$

para todo $x \in X$. Por contradição, suponha que x^* é inferior. Logo, deve existir $x \in X$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Como $w \geq 0$ e $f(x) \leq f(x^*)$, tem-se $\sum_{k=1}^{p} w_k f_k(x) \leq \sum_{k=1}^{p} w_k f_k(x^*)$. Porém, isso contradiz (4.5). $\square$

Proposição 4.2.2: Se x^* é solução ótima de $\mathrm{P}(w)$ para um dado vetor $w > 0$, com $\|w\|_1 = 1$,

então x^* é uma solução eficiente para o problema multiobjetivo.
Prova: Seja x^* uma solução ótima de $\mathrm{P}(w)$ para um dado vetor $w > 0$, com $\|w\|_1 = 1$. Por contradição, suponha que x^* é inferior. Logo, deve existir $x \in X$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Como $w > 0$, se avaliarmos a função objetivo de $\mathrm{P}(w)$ em x, obteremos um valor menor do que o avaliado no ponto x^*. Porém, isso contradiz a otimalidade de x^*. □

Geoffrion [Geoffrion, 1968] mostrou ainda que as soluções ótimas de um problema dos pesos estão fortemente relacionadas às soluções propriamente eficientes. Dois dos resultados apresentados em seu artigo seguem abaixo.

Proposição 4.2.3: Se x^* é solução ótima de $\mathrm{P}(w)$ para um dado vetor $w > 0$, então x^* é uma solução propriamente eficiente para o problema multiobjetivo.
Prova: Seja x^* uma solução ótima de $\mathrm{P}(w)$ para um dado vetor $w > 0$. Mostraremos que x^* é uma solução propriamente eficiente com

$$M = (p-1) \max_{1 \leq i \leq p} \max_{j \neq i} \frac{w_j}{w_i},$$

supondo, naturalmente, que $p \geq 2$. Por contradição, suponha que para algum i e $x \in X$ tem-se

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

para todo j tal que $f_j(x^*) < f_j(x)$. Temos, então, que

$$f_i(x^*) - f_i(x) > \frac{(p-1)}{w_i} w_j(f_j(x) - f_j(x^*)) \quad \text{para todo } j \neq i.$$

Multiplicando a expressão acima por $w_i/(p-1)$ e somando para todo $j \neq i$, obtemos

$$w_i(f_i(x^*) - f_i(x)) > \sum_{j \neq i} w_j(f_j(x) - f_j(x^*)),$$

ou, equivalentemente,

$$\sum_{k=1}^{p} w_k f_k(x) < \sum_{k=1}^{p} w_k f_k(x^*),$$

o que contradiz a otimalidade de x^* em $\mathrm{P}(w)$. □

Proposição 4.2.4: Suponha que as funções f_i são convexas no conjunto viável convexo X. Se x^* é uma solução propriamente eficiente para o problema multiobjetivo, então ele também é solução ótima de $\mathrm{P}(w)$ para algum $w > 0$, com $\|w\|_1 = 1$.
Prova: Seja x^* uma solução propriamente eficiente e suponha que as funções f_i são convexas no conjunto viável convexo X. Temos que provar que x^* é solução ótima de $\mathrm{P}(w)$ para algum $w > 0$. Por definição de eficiência própria, existe um escalar $M > 0$ tal que, para cada i ($i = 1, \ldots, p$), o sistema

$$f_i(x) < f_i(x^*) \tag{4.6}$$

$$f_i(x) + Mf_j(x) < f_i(x^*) + Mf_j(x^*), \quad \text{para todo } j \neq i, \tag{4.7}$$

não admite solução $x \in X$. Pela convexidade das funções f_i, através do resultado descrito em [Berge e Gouila-Ho 1965] (p. 62), temos que para o i-ésimo sistema existem $w_j^i \geq 0$ $(j = 1, \ldots, p)$, com $\sum_{j=1}^p w_j^i = 1$, tal que

$$w_i^i f_i(x) + \sum_{\substack{j=1 \\ j\neq i}}^{p} w_j^i (f_i(x) + Mf_j(x)) \geq w_i^i f_i(x^*) + \sum_{\substack{j=1 \\ j\neq i}}^{p} w_j^i (f_i(x^*) + Mf_j(x^*))$$

ou, equivalentemente,

$$f_i(x) + M\sum_{\substack{j=1 \\ j\neq i}}^{p} w_j^i f_j(x) \geq f_i(x^*) + M\sum_{\substack{j=1 \\ j\neq i}}^{p} w_j^i f_j(x^*),$$

para todo $x \in X$. Somando a expressão acima para cada i, após rearranjar os termos, obtemos

$$\sum_{j=1}^{p} \left(1 + M\sum_{\substack{i=1 \\ i\neq j}}^{p} w_j^i\right) f_j(x) \geq \sum_{j=1}^{p} \left(1 + M\sum_{\substack{i=1 \\ i\neq j}}^{p} w_j^i\right) f_j(x^*), \tag{4.8}$$

para todo $x \in X$. Escolhendo o vetor de pesos $\bar{w}$ tal que $\bar{w}_j = (1 + M\sum_{i\neq j} w_j^i)$, obtemos $\bar{w} > 0$. Normalizando o vetor $\bar{w}$ em relação à norma $\|\cdot\|_1$, $\bar{w}_j = \bar{w}_j/\|\bar{w}\|_1$, obtemos $\|\bar{w}\|_1 = 1$. Como (4.8) vale para todo $x \in X$, temos que x^* é solução ótima de $\mathrm{P}(\bar{w})$. $\square$

Desta forma, o problema de encontrar soluções propriamente eficientes acaba se reduzindo ao problema de resolver $\mathrm{P}(w)$ para diferentes vetores $w > 0$, considerando que o problema possui todas as funções objetivo convexas num conjunto viável convexo. Há ainda outras condições necessárias e suficientes para uma solução eficiente ser propriamente eficiente, por exemplo, utilizando o conceito de estabilidade para problemas de Otimização Multiobjetivo escalarizados. Para maiores detalhes, veja [Benson e Morin, 2000].

As relações entre as condições de Fritz John (3.8)-(3.10) para o problema multiobjetivo e as condições KKT para o problema dos pesos são, de certa forma, uma justificativa teórica para o método da somas ponderada. Isso pode ser visto, claramente, através dos resultados que seguem. Para maior clareza, apresentamos as condições KKT para o problema dos pesos (KKT_w) mostrando que elas são necessárias, sob certas condições, para garantir a otimalidade no problema dos pesos.

Proposição 4.2.5: [Condições KKT para $\mathrm{P}(w)$ (KKT_w)] Suponha que $f_k(x)$, $k = 1, \ldots, p$, e $g_i(x)$, $i = 1, \ldots, m$, são funções continuamente diferenciáveis. Se $x \in X$ é um mínimo local de $\mathrm{P}(w)$ para algum $w \geq 0$, com $\|w\|_1 = 1$, e vale alguma condição de qualificação em x, então existem multiplicadores $\mu_i \geq 0$, $i = 1, \ldots, m$ tais que

$$\sum_{k=1}^{p} w_k \nabla f_k(x) + \sum_{i=1}^{m} \mu_i \nabla g_i(x) = 0, \tag{4.9}$$

$$\mu_i g_i(x) = 0, \quad i = 1, \ldots, m. \tag{4.10}$$

Proposição 4.2.6: Se um ponto $x \in X$ é regular fraco, então existe um vetor $w \geq 0$, com $\|w\|_1 = 1$, tal que x satisfaz as condições KKT_w (4.9)-(4.10) para o $\mathrm{P}(w)$ correspondente.
Prova: Seja $x \in X$ um ponto regular fraco. Seja $\lambda \geq 0$, o vetor não nulo de multiplicadores para as funções objetivo da condição FJPOM (3.8) normalizado em relação à norma $\|\cdot\|_1$. Tomando λ como o vetor de pesos para o problema dos pesos, temos que a condição KKT_w (4.9), onde $w = \lambda$, é idêntica à condição FJPOM (3.8). Como as outras condições FJPOM e KKT_w são iguais, a prova está concluída. □

Proposição 4.2.7: Se um ponto $x \in X$ satisfaz as condições KKT_w (4.9)-(4.10), com $\mu_i \geq 0$, $i = 1, \ldots, m$, para $\mathrm{P}(w)$, onde $w \geq 0$ e $\|w\|_1 = 1$, então x é regular fraco.
Prova: Seja $x \in X$ um ponto que satisfaz as condições KKT_w (4.9)-(4.10), com $w \geq 0$ e $\|w\|_1 = 1$. Por (4.9), (4.10) e $w \geq 0$, x satisfaz (3.8)-(3.10) com $\lambda = w$. □

Utilizando os resultados acima e outros resultados clássicos de Programação Não-Linear, podemos chegar às seguintes conclusões:

Colorário 4.2.8. Suponha que as funções objetivo e de restrição são continuamente diferenciáveis. Seja x uma solução ótima de $\mathrm{P}(w)$, para algum $w \geq 0$, com $\|w\|_1 = 1$, e suponha que alguma condição de qualificação é satisfeita em x. Então, x é regular fraco.

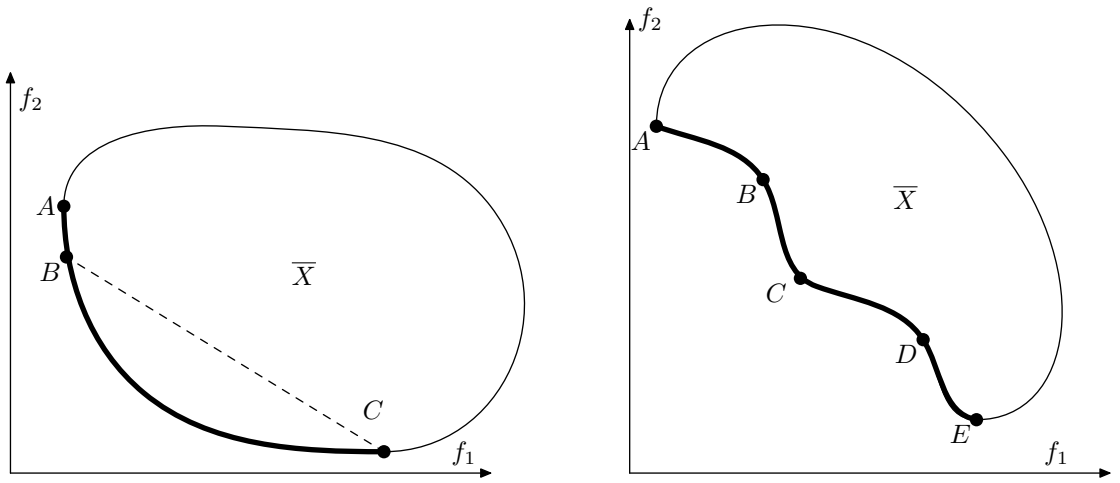
Colorário 4.2.9: Suponha que $x \in X$ é regular fraco e que todas as funções objetivo e de restrição são convexas. Então, x é solução ótima de $\mathrm{P}(w)$, para algum vetor $w \geq 0$, com $\|w\|_1 = 1$.

4.2.3 Vantagens e desvantagens

Como mencionado anteriormente, a vantagem do método dos pesos é a da simplicidade. Entretanto, existem muitas desvantagens concernentes a este método. A primeira é que não há garantias de uma boa área de cobertura da fronteira eficiente, pois não se sabe se as soluções eficientes encontradas estão bem distribuídas entre todas as outras soluções eficientes do problema. A Figura 4.4(a) mostra um exemplo de tal caso. A segunda desvantagem é ter que resolver um problema de Programação Não-Linear sempre que quisermos encontrar uma nova solução eficiente. Se tal problema já for difícil de se resolver por si só, o custo computacional será um fator preocupante para o analista.

Para problemas onde o conjunto viável não é convexo, em [Das e Dennis, 1997] e [Messac *et al.*, 2000a,b], os autores mostram ainda que este método é ineficaz para encontrar pontos em partes não-convexas da fronteira eficiente no espaço objetivo. A Figura 4.4(b) ilustra isso. Nesse exemplo, temos um problema com dois objetivos representado no espaço objetivo. Todos os pontos na parte em negrito do conjunto viável são soluções eficientes. Facilmente, podemos concluir que as soluções eficientes B e D não serão soluções ótimas de $\mathrm{P}(w)$, para qualquer $w \geq 0$. As únicas soluções eficientes que poderemos obter como soluções ótimas são aquelas situadas na envoltória convexa do conjunto viável no espaço objetivo, como, por exemplo, os pontos A, C e E. Em outras palavras, para qualquer escolha do vetor de pesos w, a função objetivo de $\mathrm{P}(w)$, linear no espaço objetivo, sempre atingirá o mínimo em um ponto pertencente à envoltória convexa de $\overline{X}$.

Outra dificuldade é achar w tal que $\mathrm{P}(w)$ tenha soluções ótimas, pois o problema pode ser ilimitado.



(a) Sem garantia de boa aproximação da fronteira eficiente.

(b) Ineficácia quando a fronteira eficiente não é convexa no espaço objetivo.

Figura 4.4: *Desvantagens do método dos pesos.*

4.3 Método NISE

O método NISE (*Noninferior Set Estimation*) é um método proposto em [Cohon *et al.*, 1979] para obter uma boa aproximação da fronteira eficiente para problemas onde o conjunto viável no espaço objetivo é convexo. Ele utiliza o método dos pesos; porém, há um controle da aproximação da fronteira eficiente a ser obtida, sendo mais vantajoso nesse sentido. Em [Cohon, 1978], um algoritmo que implementa este método é detalhadamente apresentado para a resolução de problemas com dois objetivos, e, com algumas modificações, para problemas com mais de dois objetivos. Por simplicidade, na apresentação do método neste trabalho, consideraremos problemas com apenas dois objetivos.

4.3.1 Descrição do método

Para explicar melhor o método, considere um problema com dois objetivos, f_1 e f_2, onde o conjunto viável $\overline{X}$ é convexo no espaço objetivo. Suponha que encontramos duas soluções eficientes, A e B. Então, o segmento de reta que conecta A e B no espaço objetivo é viável, podendo ser eficiente ou ineficiente. Se for eficiente, movendo-nos na direção do ortante não-positivo, estaremos fora do conjunto viável do problema. Caso seja ineficiente, movendo-se nessa mesma direção, haverá pontos correspondentes a soluções eficientes. A Figura 4.5 mostra tal exemplo. O segmento de reta que conecta A e B é ineficiente; portanto, movendo-nos na direção do ortante não-positivo, encontraremos a solução eficiente C.

Como em vários outros métodos que se utilizam de problemas escalarizados, o NISE inicia otimizando cada objetivo individualmente. Para o exemplo da Figura 4.5, as funções f_1 e f_2 otimizadas, individualmente, originam os pontos A e B. O próximo passo é encontrar uma nova solução eficiente. Para isso, o NISE utiliza o problema dos pesos com pesos $w_1 > 0$ e $w_2 > 0$ para f_1 e f_2 tal que $-w_1/w_2$ seja igual à inclinação do segmento de reta que conecta A e B - nesse caso, os pesos são estritamente positivos para que nenhuma função seja re-otimizada individualmente. A solução ótima do novo problema dos pesos corresponderá ao ponto C; tal solução, além de ser ótima, também é solução eficiente para o problema multiobjetivo.

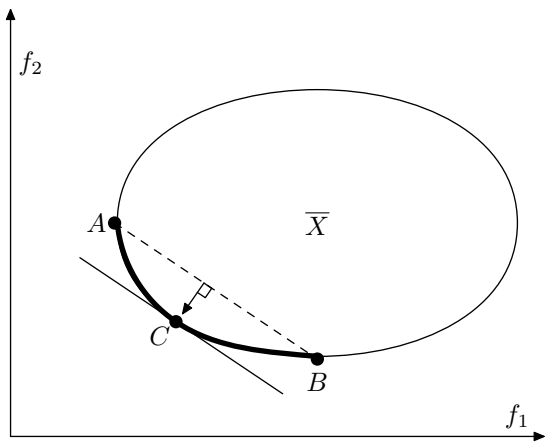


Figura 4.5: *Método NISE.*

A precisão da aproximação da fronteira eficiente é controlada através do uso de um valor pré-definido pelo analista, chamado de *erro máximo permitido*. A cada iteração do método, calcula-se o que é chamado de *erro máximo possível* para cada par de soluções eficientes adjacentes encontradas até então. O algoritmo pára quando tal erro para cada par é menor ou igual ao erro máximo permitido. Para maior compreensão do conceito de erro máximo possível, observemos a Figura 4.6. As inclinações das retas D e E foram obtidas através dos pesos usados nos problemas dos pesos para achar os pontos A e B, respectivamente. Sabemos que nenhum ponto viável está à esquerda da reta D nem abaixo da reta E; caso contrario, A e B não seriam pontos correspondentes às soluções ótimas dos problemas dos pesos considerados em iterações anteriores. Assim, o ponto C, situado na interseção das retas D e E, é um limitante inferior para a parte da fronteira eficiente localizada entre os pontos A e B. Para achar achar um limitante superior, basta pegar o segmento de reta que conecta A e B. Como $\overline{X}$ é convexo, os pontos nesse segmento correspondem a soluções viáveis do problema multiobjetivo; além disso, todos os pontos acima dessa reta correspondem a soluções ineficientes. O erro máximo possível será, portanto, a altura do triângulo formado pelos pontos A, B e C. Quando esta altura é igual a zero, isso significa que o segmento de reta que conecta A e B faz parte da fronteira eficiente.

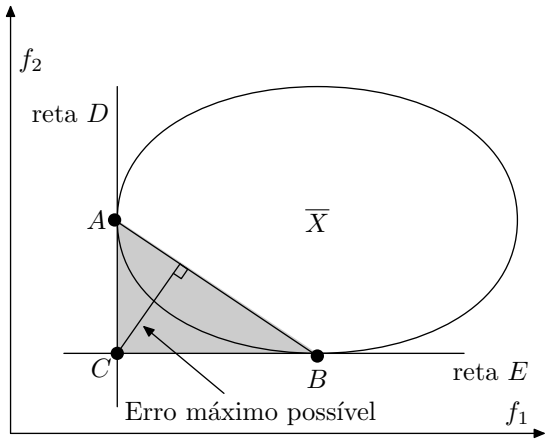


Figura 4.6: *Erro máximo possível.*

Basicamente, a cada iteração do método, os seguintes passos são realizados:

Passo 1. Se todos os erros máximos possíveis (alturas de todos os triângulos) obtidos pelos pares

de soluções eficientes adjacentes encontradas são menores que o erro máximo permitido, PARE. Caso contrário, vá para o Passo 2.

Passo 2. Dentre os pares de soluções eficientes com erro máximo possível maior que o erro máximo permitido, escolha um com o maior erro máximo possível e resolva um novo problema dos pesos onde os pesos $w_1 \geq 0$ e $w_2 \geq 0$ para f_1 e f_2 são tais que $-w_1/w_2$ seja igual à inclinação do segmento de reta que conecta as duas soluções no espaço objetivo. (A solução ótima encontrada é uma nova solução eficiente.) Retorne ao Passo 1.

4.3.2 Vantagens e desvantagens

A principal vantagem do método NISE reside no fato de que ele pode fornecer uma boa aproximação da fronteira eficiente, o que não é garantido pelo método dos pesos. Essa vantagem se deve à existência do controle da precisão da aproximação da fronteira eficiente. Contudo, por utilizar o método dos pesos para encontrar as soluções eficientes, o método NISE apresenta o mesmo problema de ter que resolver um problema de Programação Não-Linear sempre que se quiser encontrar uma nova solução eficiente. Além disso, apresenta a mesma desvantagem do método dos pesos em relação a problemas com conjunto viável não-convexo. Em caso de não-convexidade do conjunto viável no espaço objetivo, o método não está bem definido, pois o segmento de reta que conecta um par de soluções pode não ser um limitante inferior para a fronteira eficiente, como mostra a Figura 4.7.

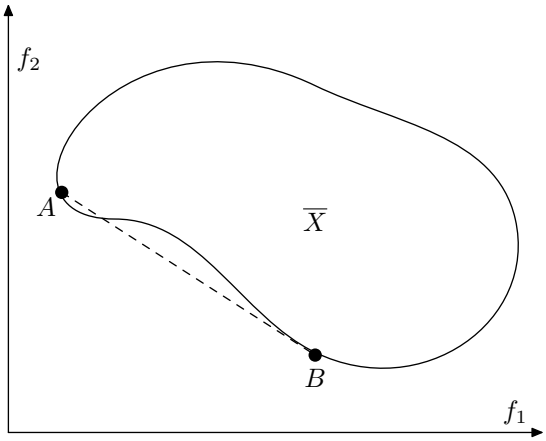


Figura 4.7: *O método NISE não está bem definido quando o conjunto viável no espaço objetivo não é convexo.*

4.4 Método da ϵ-Restrição

O método da ϵ-Restrição é bastante conhecido. Ele consiste em minimizar um único objetivo enquanto os outros são incorporados ao conjunto de restrições do problema, sendo restringidos pelos valores das componentes de um vetor $\epsilon \in \Re^{p-1}$, fornecido *a priori*. Esta abordagem de otimizar uma função objetivo enquanto outras são incorporadas às restrições parece ter sido, inicialmente, sugerida em [Marglin, 1967]. Posteriormente, Haimes *et al.* apresentaram uma nova formulação para um problema envolvendo duas funções objetivo [Haimes *et al.*, 1971]. Esta formulação, intitulada por eles de formulação da ϵ-restrição, é a origem do método que detém o mesmo nome. Assim como os outros métodos apresentados anteriormente, o método da ϵ-restrição é um método de geração

que visa obter uma aproximação da fronteira eficiente do problema.

4.4.1 Descrição do método

O problema (2.1) escalarizado com relação ao objetivo k, usando este método com um dado vetor $\epsilon \in \Re^{p-1}$, transforma-se no seguinte problema:

$$\begin{array}{lll} \mathrm{P}_k(\epsilon): & \min & f_k(x) \\ & \text{s.a:} & f_j(x) \leq \epsilon_j, \quad j = 1, \ldots, p, \quad j \neq k, \\ & & g_i(x) \leq 0, \quad i = 1, \ldots, m. \end{array} \tag{4.11}$$

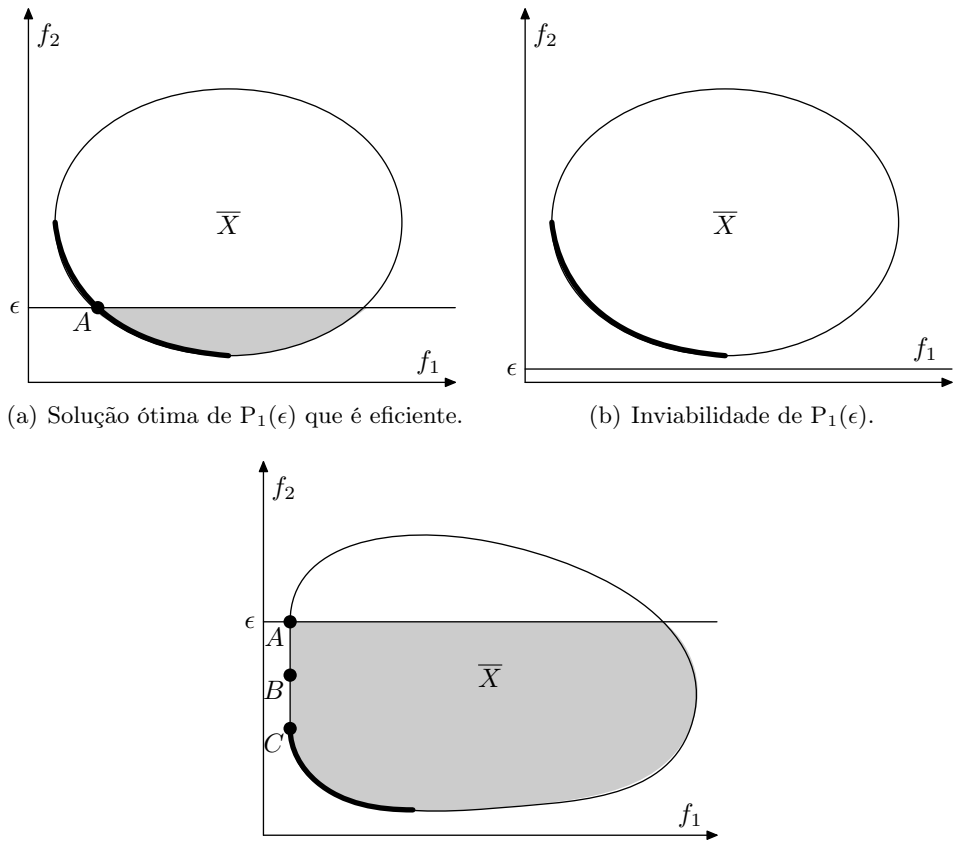
A função objetivo que será minimizada é escolhida *a priori*, assim como o vetor $\epsilon \in \Re^{p-1}$, o qual é considerado como parâmetro do método. Satisfazendo algumas condições que serão discutidas na próxima subseção, uma solução ótima do problema escalarizado (4.11) pode ser uma solução eficiente para o problema (2.1). Umas das condições necessárias para que isso ocorra está relacionada, antes de tudo, à viabilidade do problema. Para que soluções viáveis existam, é fundamental que o vetor $\epsilon \in \Re^{p-1}$ seja escolhido adequadamente, de forma que exista pelo menos um ponto x satisfazendo as restrições do problema. A Figura 4.8(a) mostra um exemplo de um problema com dois objetivos onde o método é aplicado. Neste caso, o vetor ϵ é unidimensional e possibilita a viabilidade. A parte sombreada do conjunto $\overline{X}$ corresponde ao conjunto viável do problema (4.11) no espaço objetivo. O ponto A na figura corresponde a solução ótima de (4.11). Além disso, o ponto A pertence à fronteira eficiente do problema multiobjetivo; portanto, a solução ótima de (4.11), neste caso, é uma solução eficiente. A Figura 4.8(b), por sua vez, mostra um exemplo do mesmo problema, sendo que, neste caso, a escolha de ϵ gera um problema escalarizado inviável. Outra condição necessária para que uma solução ótima de (4.11) seja eficiente é que todas as restrições relacionadas às funções objetivo sejam satisfeitas com igualdade. Se isso não acontecer e houver mais de uma solução ótima, algumas dessas soluções podem não ser eficientes. A Figura 4.8(c) mostra um exemplo de tal caso. Nessa figura, os pontos A, B e C são soluções ótimas de (4.11), mas apenas o ponto C é eficiente.

Neste método, o que fazemos é escolher alguma função objetivo f_k para ser minimizada e tentar obter uma aproximação da fronteira eficiente a partir das soluções ótimas de (4.11) para diferentes parâmetro de entrada ϵ.

4.4.2 Relações com o problema multiobjetivo

Apresentaremos agora algumas relações envolvendo soluções eficientes para o problema multiobjetivo, soluções ótimas para o problema dos pesos e soluções ótimas para o problema da ϵ-restrição, bem como as relações entre as condições de Fritz John para os três problemas. Relações envolvendo soluções propriamente eficientes e soluções ótimas para o problema da ϵ-restrição também serão discutidas.

Proposição 4.4.1: x^* é uma solução eficiente para o problema multiobjetivo se, e somente se, x^* resolve $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para todo $k = 1, \ldots, p$.
Prova: ($\Longrightarrow$) Seja x^* uma solução eficiente. Por definição, não existe $x \in X$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Logo, x^* resolve $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para todo $k = 1, \ldots, p$.
($\Longleftarrow$) Suponha que x^* resolve $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para todo $k = 1, \ldots, p$. Por contradição, suponha que x^* é inferior. Logo, deve existir algum ponto $x \in X$ tal que $f(x) \leq f(x^*)$



(a) Solução ótima de $\mathrm{P}_1(\epsilon)$ que é eficiente.

(b) Inviabilidade de $\mathrm{P}_1(\epsilon)$.

(c) Soluções ótimas de $\mathrm{P}_1(\epsilon)$ que são inferiores: A e B.

Figura 4.8: *Método da ϵ-restrição.*

e, para algum $\bar{k}$, $f_{\bar{k}}(x) < f_{\bar{k}}(x^*)$. Pelo fato de $f(x) \leq f(x^*)$, concluímos que x é solução viável para $\mathrm{P}_{\bar{k}}(\epsilon^*)$. Como $f_{\bar{k}}(x) < f_{\bar{k}}(x^*)$, x^* não resolve resolve $\mathrm{P}_{\bar{k}}(\epsilon^*)$. Absurdo. $\square$

Pelo resultado acima, temos, portanto, uma condição necessária e suficiente para uma solução ser eficiente. Todavia, na prática, esse resultado pode não ser de muita valia, dado que é necessário verificar a otimalidade de x^* para todas as p funções objetivo.

Proposição 4.4.2: Seja x^* uma solução ótima única de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para um determinado k. Então, x^* é uma solução eficiente para o problema multiobjetivo.
Prova: Suponha que x^* é solução única de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para um determinado k. Por contradição, suponha que x^* é inferior. Logo, deve existir algum ponto $x \in X$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. Pelo fato de $f(x) \leq f(x^*)$, concluimos que x é solução viável para $\mathrm{P}_k(\epsilon^*)$. Como $f(x) \leq f(x^*)$, tem-se, particularmente, que $f_k(x) \leq f_k(x^*)$. Logo, x é solução ótima de $\mathrm{P}_k(\epsilon^*)$. Porém, isso contradiz a unicidade de x^*. $\square$

Proposição 4.4.3: Seja x^* uma solução ótima de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para um determinado k. Suponha que as funções objetivo e de restrição são todas convexas. Então, existe um vetor $w \geq 0$ tal que x^* é uma solução ótima de $\mathrm{P}(w)$.
Prova: Suponha que x^* é solução ótima de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para um determinado k, e que o problema é convexo. Pelo Teorema de Gordan em [Mangasarian, 1969], considerando a

matriz A do enunciado do teorema como sendo $A = (\nabla f_1(x) \dots \nabla f_p(x))$, e pela otimalidade de x^*, existe $w \geq 0$ tal que o gradiente da função objetivo de $\mathrm{P}(w)$ avaliado em x^* é igual a 0. Logo, x^* é um ponto estacionário para $\mathrm{P}(w)$. A convexidade do problema garante que x^* é solução ótima para $\mathrm{P}(w)$. $\square$

Proposição 4.4.4: Seja x^* uma solução ótima de $\mathrm{P}(w)$ para um dado vetor $w \geq 0$, onde $w_k > 0$ para um determinado k. Então, x^* é solução ótima de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$.
Prova: Suponha que x^* é uma solução ótima de $\mathrm{P}(w)$ para um dado vetor $w \geq 0$, onde $w_k > 0$ para um determinado k. Por contradição, suponha que x^* não resolve $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$. Logo, existe $x \in X$ tal que $f_j(x) \leq f_j(x^*)$, $j \neq k$, e $f_k(x) < f_k(x^*)$. Como $w \geq 0$ e $w_k > 0$, temos que $\sum_{i=1}^p w_i f_i(x) < \sum_{i=1}^p w_i f_i(x^*)$. $\square$

Para maior clareza dos resultados subseqüentes, apresentamos as condições KKT para o problema da ϵ-restrição (KKT_ϵ), mostrando que elas são necessárias, sob certas condições, para garantir a otimalidade de $\mathrm{P}_k(\epsilon)$.

Proposição 4.4.5: [Condições KKT para $\mathrm{P}_k(\epsilon)$ (KKT_ϵ)] Suponha que $f_j(x)$, $j = 1, \dots, p$, e $g_i(x)$, $i = 1, \dots, m$, são funções continuamente diferenciáveis. Se $x \in X \cap \{x \mid f_j(x) \leq \epsilon_j,\ j = 1, \dots, p,\ j \neq k\}$ é um mínimo local de $\mathrm{P}_k(\epsilon)$ para um determinado k e algum $\epsilon \in \Re^{p-1}$ e vale alguma condição de qualificação em x, então existem multiplicadores $\lambda_j \geq 0$, $j = 1, \dots, p$, $j \neq k$, e $\mu_i \geq 0$, $i = 1, \dots, m$ tais que

$$f_k(x) + \sum_{j \neq k} \lambda_j \nabla f_j(x) + \sum_{i=1}^m \mu_i \nabla g_i(x) = 0, \tag{4.12}$$

$$\lambda_j (f_j(x) - \epsilon_j) = 0,\ j = 1, \dots, p,\ j \neq k, \tag{4.13}$$

$$\mu_i g_i(x) = 0, \quad i = 1, \dots, m. \tag{4.14}$$

Proposição 4.4.6: Suponha que x^* é uma solução propriamente eficiente do POM (2.1). Seja $k \in \{1, \dots, p\}$. Suponha que os gradientes das restrições de $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, ativas em x^* são linearmente independentes e que todas as funções f_i e g_i são continuamente diferenciáveis. Então, x^* resolve $\mathrm{P}_k(\epsilon^*)$ e satisfaz as condições KKT_ϵ (4.12)-(4.14) com $(\lambda, \mu) \geq 0$ para $\mathrm{P}_k(\epsilon^*)$. Além disso, todos os multiplicadores de Lagrange associados às restrições $f_j(x) \leq \epsilon_j$, $j \neq k$, de $\mathrm{P}_k(\epsilon^*)$ são estritamente positivos.
Prova: Suponha que x^* é uma solução propriamente eficiente. Como x^* é eficiente, pela Proposição 4.4.1, x^* resolve $\mathrm{P}_k(\epsilon^*)$, com $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para todo $k = 1, \dots, p$. Seja $k \in \{1, \dots, p\}$. Supondo que os gradientes das restrições de $\mathrm{P}_k(\epsilon^*)$ ativas em x^* são linearmente independentes e que todas as funções f_j e g_i são continuamente diferenciáveis, garante-se a existência de um conjunto de multiplicadores de Lagrange $\lambda_{kj}^* \geq 0$, $j \neq k$, e $\mu_i^* \geq 0$, para todo $i = 1, \dots, m$, tal que as condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$ sejam satisfeitas. Devemos mostrar que λ_{kj}^* deve ser estritamente positivo para todo $j \neq k$. Utilizando o resultado de um teorema em [Geoffrion, 1968] (Comprehensive Theorem), temos que, pela eficiência própria e propriedades de regularidade de x^*, existem $\lambda_j^0 > 0$, $j = 1, \dots, p$,

e $\mu_i^0 \geq 0$, $i = 1, \ldots, m$, tal que

$$\sum_{j=1}^{p} \lambda_j^0 \nabla f_j(x^*) + \sum_{i=1}^{m} \mu_i^0 \nabla g_i(x^*) = 0,$$

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

ou, equivalentemente,

$$\begin{aligned} \nabla f_k(x^*) + \sum_{j \neq k} (\lambda_j^0/\lambda_k^0) \nabla f_j(x^*) + \sum_{i=1}^{m} (\mu_i^0/\lambda_k^0) \nabla g_i(x^*) = 0, \\ (\mu_i^0/\lambda_k^0) g_i(x^*) = 0, \quad i = 1, \ldots, m. \end{aligned} \tag{4.15}$$

Como $\nabla f_j(x^*)$, $j \neq k$, e $\nabla g_i(x^*)$, tal que $g_i(x^*) = 0$, são linearmente independentes, temos que (4.15) possuirá solução única em termos de λ_j^0/λ_k^0, $j \neq k$, e μ_i^0/λ_k^0, $i = 1, \ldots, m$. Combinando isso com o fato de que (4.15) é nada mais que uma das condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$, a qual é satisfeita por λ_{kj}^*, $j \neq k$, e μ_i^*, $i = 1, \ldots, m$, temos que $\lambda_{kj}^* = \lambda_j^0/\lambda_k^0 > 0$ para todo $j \neq k$, como requerido. □

Apresentamos agora as relações entre as condições de Fritz John (3.8)-(3.10) para o problema multiobjetivo e as condições KKT_ϵ (4.12)-(4.14) para o problema da ϵ-restrição.

Proposição 4.4.7: Se $x^* \in X$ é regular fraco, então x^* satisfaz as condições KKT_ϵ (4.12)-(4.14) para $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para todo k tal que $\lambda_k > 0$, sendo este o multiplicador de Lagrange das condições FJPOM (3.8)-(3.10) associado à k-ésima função objetivo.
Prova: Suponha que $x^* \in X$ é regular fraco. Logo, por definição, existem multiplicadores $\lambda_j \geq 0$, $j = 1, \ldots, p$, não todos nulos e $\mu_i \geq 0$, $i = 1, \ldots, m$ tais que

$$\sum_{j=1}^{p} \lambda_j \nabla f_j(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0. \tag{4.16}$$

$$\mu_i g_i(x^*) = 0, \quad i = 1, \ldots, m. \tag{4.17}$$

Seja k um índice tal que $\lambda_k > 0$. Dividindo (4.16) por λ_k, obtemos

$$\nabla f_k(x^*) + \sum_{j \neq k} \bar{\lambda_j} \nabla f_j(x^*) + \sum_{i=1}^{m} \bar{\mu_i} \nabla g_i(x^*) = 0. \tag{4.18}$$

onde $\bar{\lambda_i} = \lambda_i/\lambda_k \geq 0$ e $\bar{\mu_j} = \mu_j/\lambda_k \geq 0$. Considerando que $f_i(x^*) = \epsilon^*$, $i \neq k$, por (4.18), concluímos que x^* satisfaz as condições KKT_ϵ (4.12)-(4.14) para $\mathrm{P}_k(\epsilon^*)$. □

Proposição 4.4.8: Se $x^* \in X$ satisfaz as condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$, então x^* é regular fraco.
Prova: Suponha que $x^* \in X$ satisfaz as condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$. Logo, existem multiplicadores $\lambda_j \geq 0$, $j = 1, \ldots, p$, $j \neq k$, e $\mu_i \geq 0$, $i = 1, \ldots, m$ tais que

$$f_k(x^*) + \sum_{j \neq k} \lambda_j \nabla f_j(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0, \tag{4.19}$$

$$\lambda_j (f_j(x^*) - \epsilon_j^*) = 0, \; j = 1, \ldots, p, \; j \neq k, \tag{4.20}$$

$$\mu_i g_i(x^*) = 0, \;\; i = 1, \ldots, m, \tag{4.21}$$

Fazendo $\lambda_k = 1$ e tomando o vetor de multiplicadores $\lambda = (\lambda_1, \ldots, \lambda_k, \ldots, \lambda_p)'$, podemos concluir que x^* é regular fraco. $\square$

Colorário 4.4.9: $x^* \in X$ é regular fraco se, e somente, se x^* satisfaz as condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$.

Valendo-se dos resultados acima concernentes às relações entre as condições FJPOM e as condições KKT_ϵ, apresentamos o que seria o inverso da Proposição 4.4.6.

Proposição 4.4.9: Suponha que x^* é uma solução ótima de $\mathrm{P}_k(\epsilon^*)$, onde $\epsilon_j^* = f_j(x^*)$, $j \neq k$, para um determinado k. Suponha ainda que todas as funções objetivo e de restrição são convexas e que os gradientes das funções de restrição de $\mathrm{P}_k(\epsilon^*)$ ativas em x^* são linearmente independentes (logo, x^* satisfaz as condições KKT_ϵ para $\mathrm{P}_k(\epsilon^*)$). Suponha, então, que os multiplicadores de Lagrange das condições KKT_ϵ associados às restrições $f_j(x) \leq \epsilon_j$, $j \neq k$, são estritamente positivos. Então, x^* é uma solução propriamente eficiente para o problema multiobjetivo.
Prova: O resultado segue através da Proposição 4.4.8, da Proposição 4.2.6 e da Proposição 4.2.3. A condição de que os pesos sejam estritamente positivos para que o resultado da Proposição 4.2.3 seja obtido é garantida pela suposição inicial de estrita positividade dos multiplicadores de Lagrange das condições KKT_ϵ associados às restrições $f_j(x) \leq \epsilon_j$, $j \neq k$. A convexidade é usada para passar da conclusão da Proposição 4.2.6 às hipóteses da Proposição 4.2.3. $\square$

4.4.3 Vantagens e desvantagens

A vantagem do método da ϵ-restrição reside no controle de cobertura da fronteira eficiente. Nesse ponto, este método é melhor que o método dos pesos, que não possui nenhuma garantia de uma boa cobertura. Entretanto, o método da ϵ-restrição apresenta a mesma desvantagem do método dos pesos: ter que resolver um problema de Programação Não-Linear sempre que quiser encontrar uma nova solução eficiente. A segunda desvantagem é inerente à escolha do vetor ϵ. Se tal vetor for escolhido de maneira inadequada, os problemas escalarizados a serem resolvidos podem ser inviáveis. A escolha do vetor ϵ, por sua vez, não é uma questão trivial e sua dificuldade aumenta à medida que cresce o número de funções objetivo que se deseja otimizar.

Quando o conjunto viável no espaço objetivo não é convexo, o método da ϵ-restrição também leva vantagem em relação aos dois anteriores. Um exemplo em que o método dos pesos e o NISE não conseguem aproximar a fronteira eficiente, enquanto que o método da ϵ-restrição o faz, é dado pelo seguinte problema:

$$\begin{array}{ll} \min & (x_1, x_2)' \\ \text{s.a:} & x_2 \geq -x_1^4 + 1, \\ & x_1 \geq 0,\ x_2 \geq 0. \end{array}$$

A Figura 4.9 mostra o conjunto viável e a fronteira eficiente do problema. Os únicos pontos que o método dos pesos e o NISE conseguem obter são aqueles que minimizam as funções objetivo individualmente, A e B, devido a não-convexidade do conjunto viável. Por outro lado, o método da ϵ-restrição consegue obter pontos em qualquer lugar da curva ao variar o parâmetro ϵ.

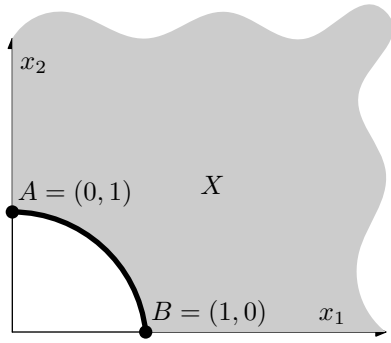


Figura 4.9: *Exemplo em que o método dos pesos e o NISE não conseguem aproximar a fronteira eficiente (linha mais grossa), enquanto que o método da ϵ-restrição consegue variando o parâmetro ϵ. Os únicos pontos que os dois primeiros métodos conseguem obter são os minimizadores A e B das funções objetivo.*

4.5 Método do gradiente

Em [Fliege e Svaiter, 2000], os autores propuseram um método para Otimização Multiobjetivo sem a necessidade de parâmetros e escalarização do problema multiobjetivo. Para tal feito, consideraram, primeiramente, o caso em que o problema multiobjetivo é irrestrito. O método proposto é uma extensão do método do gradiente para problemas de Programação Não-Linear irrestritos.

Como foi dito no Capítulo 3, uma condição necessária para um determinando ponto x ser fracamente eficiente local é a de que não existe $d \in \Re^n$ tal que

$$J_f(x)d < 0. \tag{4.22}$$

Chamaremos de soluções críticas aquelas soluções que satisfazem a condição (4.22). Se uma solução não é crítica, então existe uma direção $d \in \Re^n$ tal que

$$J_f(x))d < 0,$$

donde concluímos que d é uma direção de descida para a função multiobjetivo f. Assim, temos que, se uma solução é fracamente eficiente, então ela é crítica. Porém, nem toda solução crítica é fracamente eficiente.

A condição (4.22) estende a noção de criticalidade do caso de otimização escalar irrestrita (gradiente igual a zero).

4.5.1 Descrição do método

A idéia do método é escolher um x e verificar se a condição (4.22) é satisfeita. Se não, tenta-se encontrar uma direção de descida d para x e escolher o tamanho do passo para andar na direção de d, partindo de x.

Dado um ponto x, defina a matriz $A = J_f(x)$ e a função $f_x : \Re^n \longrightarrow \Re$ por

$$f_x(d) := \max\{(Ad)_i \mid i = 1, \ldots, p\}.$$

A direção de descida para um ponto x, se houver alguma, será dada como o resultado do seguinte problema irrestrito:

$$\begin{array}{ll} \min & f_x(d) + g(d) \\ \text{s.a:} & d \in \Re^n, \end{array} \tag{4.23}$$

onde $g : \Re^n \longrightarrow \Re$ é qualquer função própria, fechada e fortemente convexa de tal forma que $g(0) = 0$, $g(d) > 0$ $(d \neq 0)$ e $g(d) = o(\|d\|)$ $(d \longrightarrow 0)$. Em particular, vale para $g(d) = \frac{1}{2}\|d\|^2$.

Como f_x e g são funções próprias e fechadas, f_x é uma função convexa e g é fortemente convexa, a função objetivo de (4.23) é própria, fechada e fortemente convexa. Logo, o problema acima possui solução única. Para evitar a não-diferenciabilidade originada no primeiro termo da função objetivo, podemos reescrever (4.23) como

$$\begin{array}{ll} \min & \tau + \frac{1}{2}\|d\|^2 \\ \text{s.a:} & (Ad)_i \leq \tau, \quad i = 1, \ldots, p. \end{array} \tag{4.24}$$

A vantagem de lidar com (4.24) é que esse problema é convexo quadrático com restrições de desigualdades lineares.

Proposição 4.5.1: Sejam $d(x)$ e $\alpha(x)$ a solução ótima e o valor ótimo do problema (4.23), respectivamente. Se x é crítico, então $d(x) = 0$ e $\alpha(x) = 0$. Caso contrário, $\alpha(x) < 0$.
Prova: A prova segue da definição do problema.

Pode ser interessante, na prática, resolver o problema (4.23) de maneira inexata. Assim, considerando $g(d) = \frac{1}{2}\|d\|^2$, se um ponto x não é crítico, dizemos que ele é uma solução aproximada de (4.23) com uma tolerância $\sigma \in (0, 1]$ se

$$f_x(d) + \frac{1}{2}\|d\|^2 \leq \sigma\alpha(x). \tag{4.25}$$

Dessa forma, tem-se uma direção de descida assim que a condição (4.25) for satisfeita. A escolha do tamanho do passo é feita utilizando a condição suficiente de descida de Armijo. O método consiste, portanto, no seguinte algoritmo:

Algoritmo 1
Passo 1. Escolha $\beta \in (0, 1)$, $\sigma \in (0, 1]$ e $x^0 \in \Re^n$. Faça $k := 0$.
Passo 2. Se x^k é crítico, pare.
Passo 3. Encontre uma solução aproximada d de (4.23), com $x = x^k$ e tolerância σ.

Passo 4. Escolha o tamanho do passo $t_k \in (0, 1]$ como o máximo de

$$T_k = \{t = 1/2^j \mid j \in \mathbb{N},\ f(x^k\ +\ td^k)\ \leq\ f(x^k)\ +\ \beta t J_f(x^k)d^k\}.$$

Passo 5. Faça $x^{k+1} := x^k + t_k d^k$, $k := k + 1$ e vá para o Passo 2.

Obviamente, se o Algoritmo 1 termina num número finito de passos, o último ponto x^k é crítico. Para casos em que uma sequência infinita de pontos $\{x^k\}$ é gerada, Fliege e Svaiter provam que todo ponto de acumulação de tal seqüencia é crítico, supondo que f possui conjuntos de nível limitados, no sentido de que o conjunto $\{x \in \Re^n \mid f(x) \leq f(x^0)\}$ é limitado.

4.5.2 Vantagens e desvantagens

A vantagem deste método é que ele não demanda parâmetros de entrada ou ordem de prioridade atribuída às funções objetivo. Tampouco é necessário escalarizar o problema multiobjetivo. Além disso, utilizando a formulação do problema (4.24), tem-se um problema onde tudo é diferenciável, sendo quadrático convexo e com desigualdades lineares. Há ainda o fato de que soluções exatas para o problema (4.23) também não são necessárias. Uma desvantagem dele é que não sabemos *a priori* onde a solução crítica encontrada se situará na fronteira eficiente. Diferentemente dos métodos anteriores, este propõe obter apenas uma solução eficiente em vez de uma aproximação de toda a fronteira eficiente. Em caso de querer cobrir toda a Fronteira, seria necessário executar o método com *multistart*. Porém, mesmo fazendo isso, nada garante que os pontos obtidos pelo método serão diferentes entre si.

4.6 Método das direções viáveis

O método das direções viáveis foi proposto em [Fliege e Svaiter, 2000] para o caso restrito. Tal método é uma extensão do método de direções viáveis de Zoutendijk para Programação Não-Linear [Zoutendijk, 1960], sendo necessário, portanto, que o conjunto viável do problema possua interior não-vazio para garantir a convergência.

No Capítulo 3, vimos que uma condição necessária para uma solução $x \in X$ ser fracamente eficiente local para problemas de Otimização Multiobjetivo restritos é que o sistema

$$\nabla f_i(x)'d < 0,\ \forall\, i = 1, \ldots, p,\ d \in T(X, x), \tag{4.26}$$

não possua solução. Como definimos anteriormente, uma solução que satisfaz essa condição é chamada de crítica. Considere o cone $L_{\leq}(X, x)$. Suponha que a condição de qualificação de Abadie vale para todo $x \in X$, ou seja:

$$L_{\leq}(X, x) = T(X, x).$$

Temos, então, que a condição (4.26) torna-se equivalente à de que o sistema

$$\begin{array}{ll} \nabla f_i(x)'d < 0, & \forall\, i = 1, \ldots, p, \\ \nabla g_j(x)'d \leq 0, & \forall\, j \in A(x), \end{array} \tag{4.27}$$

não possua solução $d \in \Re^n$. Fazendo a matriz $\hat{A}$ ser matriz obtida ao adicionar as linhas $\nabla g_j(x)'$,

$j \in A(x)$, ao Jacobiano da f, obtemos a condição necessária equivalente a (4.27) de que não existe $d \in \Re^n$ tal que $\hat{A}d < 0$.

Dessa forma, podemos usar o método do gradiente da seção anterior considerando a matriz $\hat{A}$ no lugar de $A = J_f(x)$, tendo apenas o cuidado de controlar o tamanho do passo escolhido, de forma que a viabilidade sempre seja satisfeita.

4.6.1 Descrição do método

Seja $I = \{1, \ldots, p\}$, $\epsilon \in \Re_{++}$ e $x \in \Re^n$. Defina o conjunto $I_\epsilon := \{j \in I \mid g_j(x) \geq -\epsilon\}$. Para encontrar uma direção de descida viável, resolve-se o seguinte problema:

$$\begin{array}{lll} \min & \alpha & \\ \text{s.a:} & \nabla f_i(x)'d \leq \alpha, & i = 1, \ldots, p, \\ & \nabla g_j(x)'d \leq \alpha, & j \in I_\epsilon(x), \\ & \|d\|_\infty \leq 1. & \end{array} \tag{4.28}$$

Denote por $\alpha(x, \epsilon)$ o valor ótimo de (4.28) para um dado ϵ de entrada para o problema. Utilizando o lema de Farkas, temos que $\alpha(x, 0) = 0$ para $x \in X$ se, e somente se, existem vetores $\lambda \in \Re^p_+ \backslash \{0\}$ e $\mu \in \Re^m_+$ tal que

$$\sum_{k=1}^{p} \lambda_k \nabla f_k(x) + \sum_{i=1}^{m} \mu_i \nabla g_i(x) = 0,$$

$$\sum_{i=1}^{m} \mu_i g_i(x) = 0.$$

Assim, $\alpha(x, 0) = 0$ se, e somente se, x é um ponto regular fraco. O algoritmo para o caso restrito segue logo abaixo.

Algoritmo 2

Passo 1. Escolha $\beta \in (0, 1)$, $\epsilon_0 \in \Re_{++}$ e $x^0 \in \Re^n$. Faça $k := 0$.

Passo 2. Resolva (4.28), obtendo uma direção d^k e um valor ótimo $\alpha(x^k, \epsilon_k)$.

Passo 3. Se $\alpha(x^k, \epsilon_k) \geq -\epsilon_k$, faça $\epsilon_{k+1} := \epsilon_k/2$, $x^{k+1} := x^k$, $k := k + 1$ e vá para o passo 2.

Passo 4. Escolha um tamanho de passo $t_k \in (0, 1]$ tal que

$$f(x^k \ + \ t_k d^k) \ \leq \ f(x^k) \ + \ \beta t J_f(x^k) d^k$$

e

$$x^k \ + \ t d^k \in X.$$

Passo 5. Faça $x^{k+1} := x^k + t_k d^k$, $k := k + 1$ e vá para o Passo 2.

Proposição 4.6.1: Suponha que J_f e ∇g_i $(i \in I)$ são Lipschitz-contínuas e que a sequência de pontos $\{x^k\}$ gerada pelo algoritmo é limitada, então existe um ponto de acumulação y de $\{x^k\}$ tal que $\alpha(y, 0) = 0$.

Prova: A prova é idêntica à do teorema de convergência do método P1 de Zoutendijk.

4.6.2 Vantagens e desvantagens

As mesmas vantagens e desvantagens relacionadas ao método do gradiente são inerentes ao método das direções viáveis.

Capítulo 5

Aplicações e experimentos computacionais

Apresentaremos a seguir duas aplicações da teoria e métodos de Otimização Multiobjetivo. A primeira aplicação está associada à área de reconstrução de sinais esparsos, a qual obteve importantes resultados recentes. A segunda trata de problemas de Otimização de Portfolio e de como eles podem ser resolvidos utilizando métodos de Otimização Multiobjetivo. Nesta última aplicação, implementamos alguns métodos discutidos no Capítulo 4 e apresentamos resultados de testes computacionais realizados com eles.

5.1 Compressed Sensing

Recentemente, muitos trabalhos têm sido desenvolvidos na área de Processamento de Sinais envolvendo a aquisição e a reconstrução de sinais esparsos. O Teorema de Amostragem de Shannon-Nyquist afirma que para evitar a perda de informação ao representar digitalmente um sinal analógico, deve-se ter uma taxa de amostragem de pelo menos duas vezes a freqüência presente no sinal original. Geralmente, os sinais são compressíveis, ou seja, apresentam redundância ao se fazer uma amostragem pontual em algum domínio. Assim, após o processo de *amostragem*, onde há a discretização do sinal original num certo domínio, tenta-se reduzir o número de bits necessários para representá-lo digitalmente, sendo este processo chamado de *compressão*. As técnicas de amostragem seguida de compressão, por sua vez, levam em conta o teorema de Shannon-Nyquist, tendo-se mostrado eficiente em muitas pesquisas. Entretanto, tal teorema não considera a possível esparsidade do sinal original, sendo pessimista nos casos em que isto ocorre. A teoria de *Compressed Sensing* (CS) surgiu, pois, como uma nova abordagem na área de Processamento de Sinais ao propor a reconstrução de sinais esparsos utilizando uma taxa de amostragem menor do que a requerida no teorema de Shannon-Nyquist; ou seja, o objetivo é tentar obter um sinal já comprimido no ato do sensoriamento, daí a origem do nome. Resultados mostram que a reconstrução dos sinais esparsos utilizando a abordagem de CS é realizada perfeitamente com alta probabilidade quando certas condições são satisfeitas [Candès *et al.*, 2006].

A teoria de CS é também aplicada à área de Processamento de Imagens com grande similitude. Câmeras digitais comuns tendem a capturar um grande número de amostras de uma imagem, mas armazenam apenas uma versão comprimida dela, havendo, portanto, pouca eficiência. A idéia, então, é reconstruir imagens esparsas a partir de um número pequeno de amostras da imagem original. Apresentaremos adiante algumas abordagens encontradas em [Chen *et al.*, 2001] e [van den Berg e Friedlander, 2008] para lidar com o problema de otimização envolvido, estabelecendo ainda uma relação com a teoria e métodos no âmbito de Otimização Multiobjetivo. As informações abaixo sobre CS foram extraídas de [Baraniuk, 2007] e [Candès, 2006].

5.1.1 Compressão de sinais

A tarefa de transformação de um sinal analógico em um sinal digital compreende três processos principais: *amostragem*, onde o sinal analógico é discretizado em algum domínio (no tempo, por exemplo); *quantização*, onde a amplitude do sinal amostrado é discretizada no domínio em questão; *codificação*, onde o sinal é transformado em códigos utilizando, geralmente, a codificação binária.

Considere um sinal x discretizado no domínio do tempo e de comprimento finito. Podemos representar tal sinal como um vetor coluna $N \times 1$ com entradas reais x_i, $i = 1, \ldots, N$ - uma imagem, por exemplo, também pode ser representada por um longo vetor, com cada entrada correspondendo a um determinado pixel. Todo sinal no $\Re^N$ pode ser representado como uma combinação linear dos N vetores de uma base $\{\Psi_i\}_{i=1}^N$. Para simplificar, supomos que a base é ortonormal. Temos, então,

$$x = \sum_{i=1}^{N} s_i \Psi_i \quad \text{ou} \quad x = \Psi s, \tag{5.1}$$

onde Ψ é uma matriz $N \times N$ com os vetores Ψ_i como colunas e s é um vetor coluna com entradas $s_i = \Psi_i' x$. Assim, temos duas representações diferentes de um mesmo sinal: x, no domínio do tempo, e s, através de Ψ.

Definição 5.1.1: Diz-se que um sinal x é K-esparso se ele é uma combinação linear de apenas K vetores de uma base.

A definição acima equivale a dizer que K coeficientes s_i em (5.1) são não nulos e $(N - K)$ são iguais a zero. O caso de interesse ocorre quando se tem $K \ll N$. Um sinal é *compressível* quando a representação em (5.1) possui poucos coeficientes s_i grandes e muitos coeficientes pequenos.

Como foi dito no início do capítulo, as técnicas de amostragem seguida de compressão requerem uma amostragem completa do sinal, comprimindo-o posteriormente. Assim, para obter uma representação K-esparsa de um sinal compressível, são obtidas todas as N amostras do sinal no processo de aquisição dos dados. Os coeficientes s_i são, por sua vez, obtidos pelo sistema linear $s = \Psi' x$. Os K coeficientes maiores são localizados e codificados, enquanto que os $(N - K)$ coeficientes menores são descartados. A problemática envolvida reside, principalmente, no fato de que é realizada uma amostragem completa do sinal para a codificação de apenas poucos coeficientes. Mesmo K sendo um número pequeno, N pode ser um número grande, havendo, portanto, pouca eficiência para casos onde há esparsidade. Além disso, há ainda o custo de ter que localizar os K maiores coeficientes.

A abordagem de CS propõe a obtenção de $M < N$ medições no processo de amostragem. Sejam as M medições denotadas por $y_j = 1, \ldots, M$. A escolha dessas medições é feita utilizando alguma matriz $M \times N$ fixa Φ, de forma que se tem

$$y = \Phi x = \Phi \Psi s = \Theta s, \tag{5.2}$$

onde $\Theta = \Phi\Psi$ é uma matriz $M \times N$. Devido ao fato da matriz Φ ser fixa, dizemos que o método empregado é *não-adaptativo*. Os objetivos são, portanto, construir (a) uma *matriz de medições estável* Φ tal que a informação em qualquer sinal K-esparso ou compressível não seja prejudicada pela redução de dimensionalidade de $x \in \Re^N$ para $y \in \Re^M$ e (b) um algoritmo de reconstrução para recuperar x ou s a partir de apenas $M \approx K$ medições y.

Como $M \ll N$, o problema (5.2) é mal-posto. Entretanto, uma condição necessária e suficiente

para o problema ser bem-posto é que, para qualquer vetor v compartilhando com s as mesmas K entradas não nulas e para algum $\delta > 0$, tenha-se

$$1 - \delta \leq \frac{\|\Theta v\|_2}{\|v\|_2} \leq 1 + \delta. \tag{5.3}$$

Ou seja, a matriz Θ preserva o comprimento desses vetores K-esparsos particulares. Porém, os K coeficientes não nulos de s não são conhecidos *a priori*. Uma condição suficiente para a estabilidade da solução é que Θ satisfaça (5.3) para um vetor v $3K$-esparso arbitrário. Tal condição é denominada *Restricted Isometry Property* (RIP) [Candès *et al.*, 2006]. Um outro resultado, encontrado em [Candès *et al.*, 2006], afirma que, quando $M \geq cK \log(N/K)$, onde c é uma constante, Θ satisfaz a condição RIP com alta probabilidade, ou seja, sinais K-esparsos e compressíveis podem ser recuperados perfeitamente com alta probabilidade com apenas $M \geq cK \log(N/K) \ll N$ medições.

5.1.2 Basis pursuit

Dada a matriz Φ, desejamos resolver o seguinte problema:

$$\begin{aligned} \min \quad & \|s\|_0 \\ \text{s.a:} \quad & \Theta s = y, \end{aligned} \tag{5.4}$$

onde $\Theta = \Phi\Psi$ e $\|\cdot\|_0$ é definido por $\|s\|_0 = \sharp\{i \mid s_i \neq 0\}$. Todavia, o problema acima é combinatório e NP-completo, despertando o interesse pela busca de outras formas de obter uma representação esparsa do sinal x. Uma possibilidade seria utilizar a norma Euclidiana na função objetivo, resultando no seguinte problema de quadrados mínimos:

$$\begin{aligned} \min \quad & \|s\|_2^2 \\ \text{s.a:} \quad & \Theta s = y. \end{aligned} \tag{5.5}$$

Entretanto, a norma euclidiana não prima pela esparsidade da solução como mostra a Figura 5.1, sendo ineficaz para nosso propósito. Assim, surge uma outra possibilidade que é a de usar a norma L_1. O problema em que a norma L_1 é considerada é chamado de *basis pursuit* e é dado por:

$$\begin{aligned} \min \quad & \|s\|_1 \\ \text{s.a:} \quad & \Theta s = y. \end{aligned} \tag{5.6}$$

A Figura 5.1 também mostra que a norma L_1 é de melhor proveito que a norma L_2. Claramente, podemos observar na figura que a solução ótima B, encontrada no problema de quadrados mínimos, não é esparsa, ao contrário da solução A, encontrada utilizando a norma L_1.

Em aplicações reais, o sinal amostrado pode conter erros nas medições, ou seja, tem-se $\Theta s = y + \sigma$, onde σ é o erro derivado das medições. Assim, a restrição de (5.6) pode não ser satisfeita para nenhum vetor s, tornando o problema inviável. Nosso objetivo agora se divide em dois: minimizar $\|s\|_1$ e minimizar $\|\Theta s - y\|$. Temos, portanto, um problema de Otimização Multiobjetivo com duas funções objetivo. Quando se tem uma estimativa $\tau \approx \|\sigma\|$ do erro, uma abordagem razoável seria utilizar o método da ϵ-restrição para resolver o problema multiobjetivo. Considerando $f_1 = \|s\|_1$ e $f_2 = \|\Theta s - y\|_2$, o problema da ϵ-restrição com relação à função objetivo f_1 torna-se

$$\begin{aligned} \mathrm{P}_1(\tau): \quad \min \quad & \|s\|_1 \\ \text{s.a:} \quad & \|\Theta s - y\|_2 \leq \tau, \end{aligned} \tag{5.7}$$

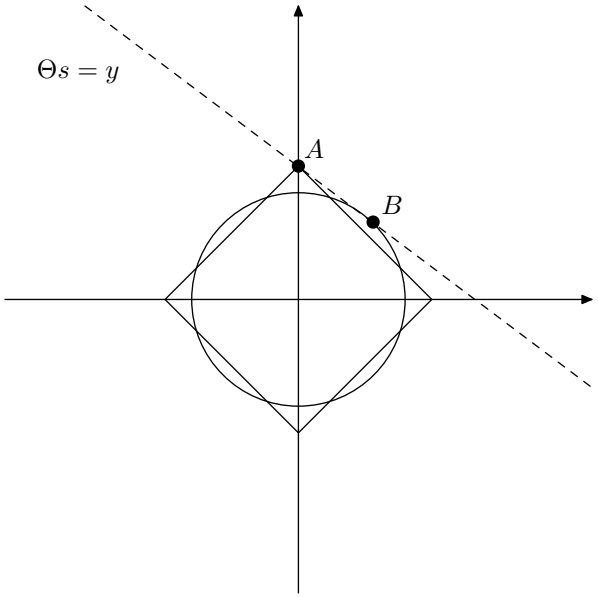


Figura 5.1: *Soluções A e B obtidas usando as normas L_1 e L_2, respectivamente.*

onde $\tau \geq 0$ é uma estimativa do erro presente nas medições.

Em [van den Berg e Friedlander, 2008], é apresentado um método que considera o problema da ϵ-restrição com relação à função objetivo f_2:

$$\begin{array}{rll} \mathrm{P}_2(\lambda): & \min & \|\Theta s - y\|_2 \\ & \text{s.a:} & \|s\|_1 \leq \lambda, \end{array} \tag{5.8}$$

onde $\lambda \geq 0$.

Outra possibilidade, proposta em [Chen *et al.*, 2001], é a utilização do método dos pesos, lidando com o seguinte problema ponderado:

$$\min \quad \tfrac{1}{2}\|\Theta s - y\|_2^2 + \gamma\|s\|_1. \tag{5.9}$$

Para mais detalhes sobre a área de *Compressed Sensing*, veja [Baraniuk, 2007] e [Candès, 2006].

5.2 Otimização de Portfolio

Considere um conjunto de n investimentos contendo ações de empresas, fundos mútuos, opções e outros derivativos. Esses investimentos possuem retornos futuros e riscos associados a eles. Desejamos saber o quanto devemos investir em cada um afim de que o retorno total futuro seja o máximo possível, e o risco, o mínimo possível. Em outras palavras, queremos otimizar nosso portfolio. Geralmente, os investimentos que propiciam o maior retorno esperado são aqueles que possuem maiores riscos. Na Teoria Moderna do Portfolio, introduzida em [Markowitz, 1952] e que será considerada neste trabalho, o investidor pode escolher uma troca entre o risco e retorno esperado do portfolio através da curva da fronteira eficiente.

As taxas de retorno futuro r_i dos investimentos são variáveis randômicas. Porém, não temos como prever o futuro; tudo o que temos são dados históricos dos investimentos de anos anteriores. Nesse modelo, consideramos os retornos do histórico como bons indicadores dos retornos futuros. A Tabela 5.1 apresenta os retornos entre os anos 1991 e 2002 de um fundo mútuo do índice S&P500,

Tabela 5.1: *Histórico dos retornos anuais*

Ano	$r_{SP\&500}$	r_{IBM}	r_{Sony}
1991	+0.2631	-0.2124	-0.1027
1992	+0.0446	-0.4336	-0.0037
1993	+0.0706	+0.1208	+0.4785
1994	-0.0154	+0.3012	+0.1348
1995	+0.3411	+0.2430	+0.1046
1996	+0.2026	+0.6584	+0.0772
1997	+0.3101	+0.3811	+0.3905
1998	+0.2700	+0.7624	-0.2028
1999	+0.1953	+0.1701	+2.9681
2000	-0.1014	-0.2120	-0.5109
2001	-0.1304	+0.4231	-0.3484
2002	-0.2337	-0.3570	-0.0808

das ações da IBM e das ações (ADR) da Sony. O retorno futuro esperado do investimento i, $\mathrm{E}(r_i)$, é, portanto, a média dos retornos desse investimento nos anos anteriores. Para os três investimentos citados anteriormente, temos $\mathrm{E}(r_{S\&P500}) = 0.101 = 10.1\%$, $\mathrm{E}(r_{IBM}) = 0.154 = 15.4\%$ e $\mathrm{E}(r_{Sony}) = 0.242 = 24.2\%$.

Por sua vez, o risco de cada investimento também precisa ser mensurado. Para calculá-lo, calculamos antes a variância de cada variável r_i:

$$\mathrm{Var}(r_i) = \frac{(r_{i1} - \mathrm{E}(r_i))^2 + \ldots + (r_{iT} - \mathrm{E}(r_i))^2}{T-1},$$

onde T é o número de anos considerados e r_{ij} é o retorno do investimento i no ano j. Para os três investimentos considerados na tabela, temos

$$\mathrm{Var}(r_{S\&P500}) = \frac{(+0.1620)^2 + \ldots + (-0.3348)^2}{11} = 0.0362;$$

$$\mathrm{Var}(r_{IBM}) = \frac{(-0.3361)^2 + \ldots + (-0.5105)^2}{11} = 0.1503;$$

$$\mathrm{Var}(r_{Sony}) = \frac{(+0.3348)^2 + \ldots + (-0.3228)^2}{11} = 0.8149.$$

Os desvios-padrão de cada investimento são, portanto: $\mathrm{desvio}(r_{S\&P500}) = \sqrt{\mathrm{Var}(r_{S\&P500})} = 19.0\%$, $\mathrm{desvio}(r_{IBM}) = \sqrt{\mathrm{Var}(r_{IBM})} = 38.8\%$ e $\mathrm{desvio}(r_{Sony}) = \sqrt{\mathrm{Var}(r_{Sony})} = 90.3\%$. O risco de cada investimento, finalmente, é considerado como sendo o desvio padrão do seu retorno.

Suponha que nosso portfolio seja composto por 66.6% do dinheiro investido na IBM e 33.3% investido na Sony. Para calcular o retorno futuro esperado dele, fazemos

$$\begin{aligned} \mathrm{E}(r_p) &= \mathrm{E}(66.6\% \cdot r_{IBM} + 33.3\% \cdot r_{Sony}) \\ &= 66.6\% \cdot \mathrm{E}(r_{IBM}) + 33.3\% \cdot \mathrm{E}(r_{Sony}) \\ &= 66.6\% \cdot 15.4\% + 33.3\% \cdot 24.2\% = 18.3\%. \end{aligned}$$

Para calcular o risco do nosso portfolio, calculamos antes a variância. Usamos o fato de que $\mathrm{Var}(r_p) = \mathrm{Cov}(r_p, r_p)$, onde $\mathrm{Cov}(r_i, r_j)$ é a covariância entre duas variáveis calculada como

$$\mathrm{Cov}(r_i, r_j) = \frac{(r_{i1} - \mathrm{E}(r_i)) \cdot (r_{j1} - \mathrm{E}(r_j)) + \ldots + (r_{iT} - \mathrm{E}(r_i)) \cdot (r_{jT} - \mathrm{E}(r_j))}{T-1}.$$

Após algumas contas, obtemos a seguinte fórmula para a variância do retorno de um portfolio qualquer $r_p = \sum_{i=1}^{n} x_i r_i$, onde x_i é a proporção investida no investimento i:

$$\mathrm{Var}(r_p) = \sum_{i=1}^{n} x_i^2 \cdot \mathrm{Var}(r_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_i x_j \cdot \mathrm{Cov}(r_i, r_j), \tag{5.10}$$

ou, equivalentemente,

$$\mathrm{Var}(r_p) = x' Q x, \tag{5.11}$$

onde Q é a matriz de covariância.

5.2.1 Modelo de Markowitz e Otimização Multiobjetivo

Seja $\mu = (\mu_1, \ldots, \mu_n)'$ o vetor composto pelos retornos esperados dos investimentos. Temos, então, dois objetivos: minimizar $f_1(x) = x'Qx$ e maximizar $f_2(x) = \mu' x$. Podemos, desta forma, utilizar a teoria e os métodos de Otimização Multiobjetivo para resolver problemas de Otimização de Portfolio.

O modelo da Média-Variância (MV) de Markowitz que usaremos pode ser descrito de várias formas, sendo uma delas a seguinte:

$$\begin{aligned} \max \quad & \mu' x - \delta x' Q x \\ \text{s.a:} \quad & \sum_{i=1}^{n} x_i = 1, \end{aligned} \tag{5.12}$$

onde $\delta \geq 0$ é uma constante de aversão ao risco, a qual pode ser variada para obter a fronteira eficiente. Claramente, podemos ver que tal modelo é exatamente uma aplicação do método dos pesos que vimos anteriormente. Uma outra formulação conhecida é dada por restringir o retorno esperado a um valor mínimo:

$$\begin{aligned} \min \quad & x' Q x \\ \text{s.a:} \quad & \mu' x \geq R, \\ & \sum_{i=1}^{n} x_i = 1, \end{aligned} \tag{5.13}$$

onde R é um valor mínimo que se deseja para o retorno esperado. Quando se conhece um valor para R dentro do intervalo onde está a fronteira eficiente, podemos trocar o menor ou igual por igual apenas, obtendo $\mu' x = R$. Tal problema é exatamente uma aplicação do método da ϵ-restrição. Como a variância de uma variável é sempre não-negativa, Q é uma matriz semi-definida positiva. Assim, este é um problema de Programação Quadrática Convexa, e, portanto, as condições KKT são condições de otimalidade necessárias e suficientes para este caso. Uma forma de resolver tal problema de Programação Quadrática seria, por exemplo, através do método Simplex de Wolf para Programação Quadrática [Wolf, 1959].

Uma outra possibilidade é restringir a outra função objetivo, obtendo o problema de Programação Não-Linear abaixo:

$$\begin{array}{ll} \max & \mu' x \\ \text{s.a:} & x'Qx \leq \bar{\sigma}^2, \\ & \displaystyle\sum_{i=1}^{n} x_i = 1, \end{array} \tag{5.14}$$

onde $\bar{\sigma}^2$ é um valor máximo desejado para a variância da solução. Restringir a variância do portfolio a algum valor pré-definido é um critério usado quando desejamos conhecer o *tracking error* de um portfolio em relação a algum índice escolhido previamente. O *Tracking error*, por sua vez, é uma medida de quão perto o portfolio está em relação a algum portfolio de *benchmark* x_B. O *tracking error ex-ante* do portfolio x em relação a x_B pode ser definido como:

$$\text{TE}(x) \ = \sqrt{(x - x_B)'Q(x - x_B)}.$$

Por sua vez, o retorno em excesso do portfolio x em relação ao x_B é dado por $\mu'(x - xB)$. Assim, o problema (5.14) pode ser substituído por outro envolvendo o retorno em excesso e o *tracking error*:

$$\begin{array}{ll} \max & \mu'(x - x_B) \\ \text{s.a:} & (x - x_B)'Q(x - x_B) \leq \overline{TE}^2, \\ & \displaystyle\sum_{i=1}^{n} x_i = 1, \end{array} \tag{5.15}$$

onde $\overline{TE}$ é um valor pré-definido. Em [Jorion, 2003], o autor observa que retornos históricos são medidas enormemente ruidosas para o retorno esperado e que os investidores podem não ser capazes ou mesmo desejarem determinar um retorno esperado mínimo aceitável e constantes de aversão ao risco. Segundo Jorion, é muito mais fácil restringir o risco do portfolio. Como a restrição do *tracking error* é quadrática e convexa, ela pode ser reescrita de tal forma que o problema (5.15) se transforme num problema de Otimização Cônica e técnicas para este tipo de problema possam ser usadas adequadamente.

Restrições do tipo $x \geq 0$ podem ser adicionadas ao problema se vendas a descoberto forem proibidas na solução. Como pode ser visto na fórmula da variância do portfolio (5.10), à medida que o número de investimentos cresce, o número de termos envolvendo a covariância entre variáveis é muito maior que o número de termos envolvendo a variância de uma variável. Logo, quem tende a dominar o somatório são as covariâncias. Quanto menor estas forem, menor será o risco do nosso portfolio. Por isso, é importante a diversificação do portfolio. Uma forma de tentar diversificar é adicionar restrições de limite a investimentos de um mesmo setor, por exemplo:

$$\sum_{i \text{ pertence ao setor } k} x_i \leq m_k.$$

Podemos ainda adicionar restrições envolvendo custos de transação. Tais restrições podem ser do tipo

$$x_1 + \ldots + x_n + t(x_1) + \ldots + t(x_n) \leq b,$$

onde t é uma função de x_i definida pelo corretor e b é o orçamento atual disponível. Observe que a função t pode ser convexa ou não, o que pode tornar o problema de Otimização de Portfolio

não-convexo. Custos de transação também podem ser considerados na função objetivo da seguinte forma:

$$\max \sum_{i=1}^{n} \mu_i x_i - \sum_{i=1}^{n} k_i |x_i - x_i^0|,$$

onde x^0 é o portfolio já existente.

5.2.2 Fronteira eficiente Risco-Retorno

Podemos reescrever (5.13) como

$$\begin{array}{ll} \min & x'Qx \\ \text{s.a:} & A'x = B, \end{array} \tag{5.16}$$

onde $A = (\mu \quad \bar{1})$, $\bar{1} = (1 \ldots 1)'$ e $B = \begin{pmatrix} R \\ 1 \end{pmatrix}$.

O sistema KKT para o problema acima é dado por:

$$\left\{ \begin{array}{rcl} 2Qx + A\lambda_0 & = & 0 \\ A'x & = & B, \end{array} \right. \tag{5.17}$$

onde $\lambda_0 = (\lambda_1 \quad \lambda_2)'$. Suponha que a matriz Q é definida positiva, ou seja, a variância do portfolio é sempre estritamente positiva. Definindo $\lambda = -1/2\lambda_0$ e resolvendo a primeira equação de (5.17) para x, obtemos

$$x = Q^{-1}A\lambda.$$

A segunda equação, por sua vez, torna-se

$$A'Q^{-1}A\lambda = B \implies \lambda = (A'Q^{-1}A)^{-1}B \equiv H^{-1}B,$$

onde $H = A'Q^{-1}A$ e $H' = (A'Q^{-1}A)' = A'(Q^{-1})'A = A'Q^{-1}A = H$; logo, H é uma matriz 2×2 simétrica. Usando as expressões acima para reescrever a variância, obtemos

$$\text{Var}(r_p) = x'Qx = x'QQ^{-1}A\lambda = x'A\lambda = (A'x)'H^{-1}B = B'H^{-1}B.$$

Sabemos ainda que

$$H \equiv \begin{pmatrix} a & b \\ b & c \end{pmatrix} \implies H^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}.$$

Defina $d \equiv det(H) = ac - b^2$. Como $H = A'Q^{-1}A$, é fácil ver que:

$$\begin{array}{rcl} a & = & \mu'Q^{-1}\mu, \\ b & = & \mu'Q^{-1}\bar{1} = \bar{1}'Q^{-1}\mu, \\ c & = & \bar{1}'Q^{-1}\bar{1}. \end{array}$$

Como Q é definida positiva, Q^{-1} também é; logo, $a > 0$ e $c > 0$. Como Q^{-1} é definidia positiva, temos que

$$(b\mu - a\bar{1})'Q^{-1}(b\mu - a\bar{1}) = bba - abb - abb + aac = a(ac - b^2) = ad > 0,$$

Como $a > 0$, temos $d > 0$. Usando H para reescrever a variância, obtemos

$$\begin{aligned} \mathrm{Var}(r_p) &= \frac{1}{d}\begin{pmatrix} R & 1 \end{pmatrix}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}\begin{pmatrix} R \\ 1 \end{pmatrix} \\ &= \frac{1}{d}(cR^2 - 2bR + a). \end{aligned} \tag{5.18}$$

Usualmente, a variância de um portfolio x_p é denotada por σ_p^2, e seu desvio padrão, por σ_p. Pela equação (5.18), concluímos que o gráfico da variância em função do retorno no espaço (σ_p^2, R) é uma parábola. Porém, no espaço (σ_p, R), o gráfico é o lado direito de uma hipérbole, como veremos a seguir. A fórmula para o gráfico no espaço (σ_p, R) é dada por:

$$\sigma_p^2 = \frac{cR^2 - 2bR + a}{d} = \frac{cR^2 - 2bR + d/c + b^2/c}{d}.$$

Dividindo o lado esquerdo por $1/c$ e o lado direito por c/c^2, obtemos

$$\frac{\sigma_p^2}{c} = \frac{cR^2 - 2bR + d/c + b^2/c}{d/c^2} = (R - b/c)^2 d/c^2 + 1,$$

ou, equivalentemente,

$$\frac{\sigma_p^2}{1/c} - \frac{(R - b/c)^2}{d/c^2} = 1.$$

Como podemos ver, a fórmula acima é de uma hipérbole cujo centro é $(0, b/c)$. A Figura 5.2 mostra a fronteira eficiente no espaço (σ_p, R), sendo sempre a parte superior do gráfico.

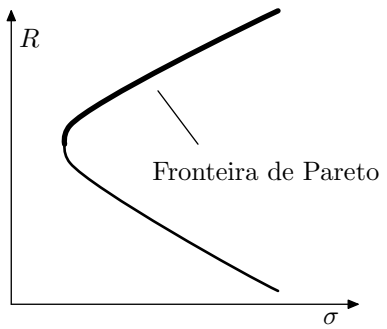


Figura 5.2: *Fronteira eficiente para problemas de Otimização de Portfolio.*

Uma possibilidade para escolha dentre os portfolios eficientes é através da *razão de Sharpe*, definida como a razão entre o retorno esperado pelo desvio padrão do portfolio. O portfolio com a maior razão, chamado de *portfolio tangencial*, fornece o maior retorno esperado por unidade de risco, sendo uma escolha razoável dentre os portfolios eficientes existentes.

5.2.3 Testes computacionais utilizando o modelo de Markowitz

No mundo real do mercado de ações, existem milhares de investimentos possíveis para o investidor escolher. Obviamente, você não considerará todos eles como entrada de qualquer método para obter um portfolio eficiente, pois isso acarretaria em calcular milhares de covariâncias entre os

investimentos, fora o importante fato de que o erro entre o modelo e a realidade iria ser enormemente amplificado. Sendo assim, é recomendado que se escolha sempre um número bem reduzido de investimentos possíveis como entrada de qualquer método. Seguindo esse pensamento, implementamos os 4 métodos de Otimização Multiobjetivo para problemas restritos e fizemos uma análise comparativa de seus resultados, a saber: método dos pesos, método da ϵ-restrição, método NISE e o método das direções viáveis. O método de direções viáveis foi ligeiramente modificado para que pudesse ser aplicado em problemas com restrições de desigualdade e igualdade.

Os métodos dos pesos e o da ϵ-restrição já são bastante conhecidos e usados para a resolução de problemas de Otimização de Portfolio. Nosso objetivo é, portanto, apresentar métodos de Otimização Multiobjetivo que não foram aplicados no contexto de Otimização de Portfolio e comparar os seus resultados. Nesta subseção, utilizamos apenas o modelo MV de Markowtiz para os experimentos.

O problema abordardo é:

$$\begin{aligned} \min \quad & (x'Qx, -\mu' x) \\ \text{s.a:} \quad & \sum_{i=1}^{n} x_i = 1, \\ & x \geq 0. \end{aligned} \tag{5.19}$$

Como temos apenas restrições lineares, isso implica que a condição de qualificação de Abadie é satisfeita para todo ponto viável. Logo, pela Proposição 3.1.10 do Capítulo 3, toda solução eficiente é regular fraca.

Esta subseção se divide em duas partes: a primeira trata cada método individualmente, comparando seus resultados para um determinado problema e valores diferentes dos parâmetros de entrada; a segunda trata de comparar os métodos entre si para 5 problemas, sob uma métrica a ser explicada mais na frente.

Todos os métodos foram implementados na linguagem C/C++ utilizando a biblioteca IlOG Concert Technology e o resolvedor CPLEX, sendo executados num processador Intel Core 2 Quad 2.40GHz com o sistema operacional Linux. Cada método recebe um arquivo de entrada com o número de investimentos, o número de períodos considerados no histórico de preços dos investimentos e os preços dos investimentos em cada um dos períodos. Devido ao tamanho dos problemas considerados, aspectos relacionados ao tempo de processamento dos métodos não são apresentados por terem se mostrado irrelevantes durante os experimentos. O Método da Barreira disponível no software CPLEX foi utilizado para resolver os subproblemas dentro de cada método de Otimização Multiobjetivo implementado.

Uma solução eficiente para o problema (5.19) também pode ser obtida utilizando o método Simplex para Programação Quadrática [Wolf, 1959]. Nesse artigo, o autor também mostra que toda a fronteira eficiente pode ser obtida facilmente utilizando a versão paramétrica desse método. Sendo assim, a implementação dos 4 métodos propriamente ditos tem como propósito o estudo comparativo do desempenho de cada um numa aplicação prática. Por outro lado, na Subseção 5.2.4, veremos problemas em que o modelo matemático escolhido possui três objetivos ao invés de apenas dois. Nesse caso, não há métodos conhecidos que encontrem facilmente toda a fronteira eficiente, já que não há garantias de que qualquer solução eficiente encontrada seja global devido à perda de convexidade no problema.

Tabela 5.2: *Investimentos de cada problema.*

Problema 1	Problema 2	Problema 3	Problema 4	Problema 5
GFSA3	ITSA4	LREN3	GOLL4	RDCD3
BBAS3	USIM5	GOLL4	VIVO4	BRTO4
VIVO4	VALE5	TAMM4	TCSL3	VALE5
TCSL3	TAMM4	GGBR4	BTOW3	ALLL3
TLPP4	AMBV4	EMBR3	GGBR4	TLPP4
ELPL5	ELET3	BBDC4	EMBR3	ITSA4
AMBV4	ITUB4	ELPL5	LAME4	BVMF3
PETR4	LAME4	ITUB4	ELPL5	CCRO3
EMBR3	PETR4	GFSA3	LREN3	USIM5
BBDC4	TNLP4	RDCD3	ALLL3	BTOW3

Foram testados 5 problemas, contendo cada um preços históricos mensais de 10 ações pertencentes ao índice IBOVESPA de outubro de 2007 a outubro de 2010, totalizando 27 investimentos possíveis. A Tabela 5.2 apresenta os investimentos de cada um dos problemas. Os preços das ações foram todos retirados do site Yahoo! Finanças e estão disponíveis no Apêndice A.

Como as taxas de retorno são multiplicativas sobre o tempo, usamos a média geométrica em vez da média aritmética para calculá-las, ou seja,

$$\mu_i = \left(\prod_{t=1}^{p} (1 + r_{it}) \right)^{1/p} - 1.$$

Em todos os problemas, usamos a formulação (5.12) para o método dos pesos e a formulação (5.13) para o método da ϵ-Restrição. Se um determinado método não otimiza as funções individualmente em seu processo, nós o fazemos antes de iniciar o método e incluímos suas soluções no conjunto de soluções obtidas.

5.2.3.1 Testes individuais

Método dos pesos

Como foi apresentado no Capítulo 4, o método dos pesos se utiliza de um vetor de pesos $w \geq 0$ tal que $\|w\| = 1$. Para problemas de Otimização de Portfolio, temos apenas dois pesos w_1 e w_2, de forma que temos a seguinte função objetivo ponderada:

$$w_1 \mu' x - w_2 x' Q x, \tag{5.20}$$

ou, equivalentemente,

$$w \mu' x - (1 - w) x' Q x, \tag{5.21}$$

onde w agora é um escalar. Dividindo (5.21) por w e fazendo $\delta = (1 - w)/w$, obtemos

$$\mu' x - \delta x' Q x, \tag{5.22}$$

que é exatamente a função objetivo do problema (5.12). Como foi discutido anteriormente, nenhuma distribuição do vetor de pesos garante uma boa distribuição de pontos na fronteira eficiente. Porém, usar uma distribuição uniforme de valores para w no intervalo $[0, 1]$ é mais natural que variar δ

indefinidamente em $[0, \infty)$.

A Figura 5.3 mostra uma aproximação da fronteira eficiente para o Problema 1 obtida variando δ em $[0, \infty)$. Os pontos na curva foram obtidos em ordem decrescente do valor do retorno esperado. Nesse exemplo, supomos que estamos dispostos a resolver 10 subproblemas ponderados para aproximar a fronteira eficiente. O valor inicial de δ é zero, sendo acrescido em cinco unidades a cada novo subproblema. Podemos ver na Figura 5.3 que, a partir de um certo valor de δ, a distância entre os novos pontos obtidos no espaço objetivo diminui bastante. A aproximação da parte superior da curva depende muito da variação de δ. Se a variação não for muito pequena e a curva for muito acentuada, a parte de cima não será bem aproximada. No caso da Figura 5.3, a aproximação da Fronteira não foi tão ruim; entretanto, para isso acontecer, tivemos que realizar alguns testes para descobrir qual o acréscimo necessário para a constante δ entre os subproblemas. Como, mais adiante, iremos comparar os métodos entre si com base num número fixo de subproblemas, testes adicionais para calcular parâmetros que obtêm uma aproximação melhor não serão permitidos. Assim, sem nenhuma informação sobre a fronteira eficiente, a forma mais natural de tentar aproximá-la é variando w em $[0, 1]$, que é o que faremos daqui em diante.

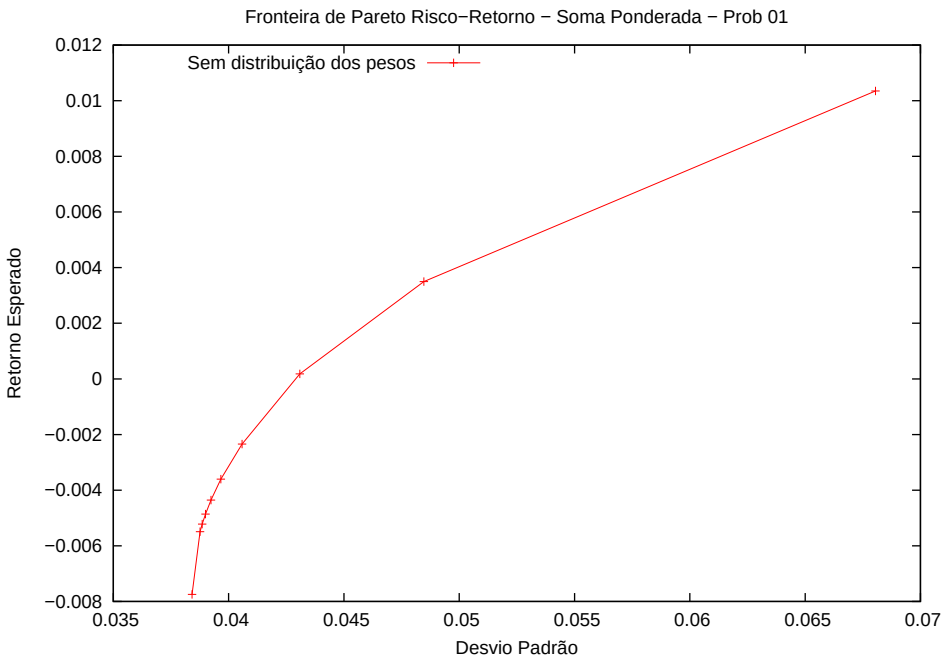


Figura 5.3: *Fronteira eficiente do Problema 1 obtida com o método dos pesos e variando δ em $[0, \infty)$.*

A Tabela 5.3 mostra o número de soluções eficientes diferentes encontradas para o Problema 1 quando estamos dispostos a resolver um número limitado de subproblemas. Consideramos uma distribuição uniforme de w no intervalo $[0, 1]$ para os experimentos.

Podemos ver que o número de soluções diferentes encontradas também não é igual ao número de subproblemas resolvidos neste caso. Ou seja, mesmo havendo uma distribuição uniforme para os valores de w, nada garante que os pontos gerados na curva serão diferentes entre si. Na verdade, não sabemos nada sobre esses pontos *a priori*, nem mesmo suas localizações aproximadas.

A Figura 5.4 mostra a aproximação da fronteira eficiente para o Problema 1 usando o método com número de subproblemas igual a 10 e a 50. Obviamente, quanto maior for o número de subproblemas resolvidos, melhor será a cobertura da fronteira eficiente. Podemos ver que a parte

Tabela 5.3: *Desempenho do método dos pesos no Problema 1.*

Subproblemas resolvidos	Soluções diferentes encontradas
10	06
20	10
30	15
40	20
50	24

inferior da curva é mal aproximada quando resolvemos apenas 10 subproblemas. Isso pode ser causado pelo número pequeno de soluções diferentes encontradas. Além disso, existe o fato da má distribuição desses pontos na curva. Para tentar obter mais pontos na parte inferior, é preciso usar uma malha mais fina para o peso w.

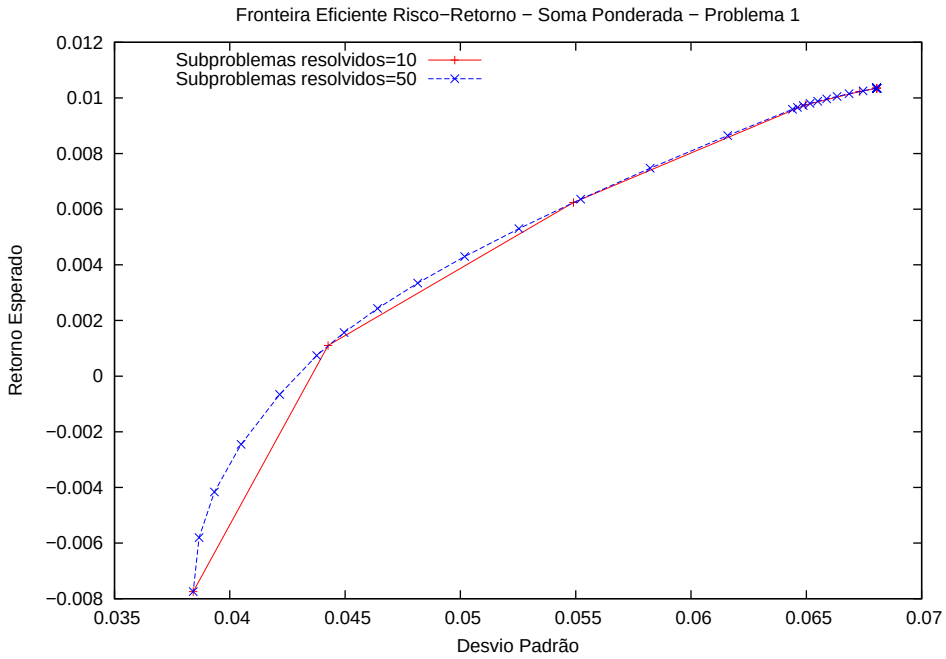


Figura 5.4: *Fronteira eficiente do Problema 1 aproximada usando o método dos pesos com número de problemas resolvidos igual a 10 e a 50.*

Método da ϵ-Restrição

Neste método, devemos restringir o retorno esperado do nosso portfolio a algum valor mínimo R. Entretanto, se escolhermos mal este valor, o problema pode se tornar inviável. Uma escolha possível para R pode ser o valor mínimo entre os retornos esperados dos investimentos, pois sabemos que aplicar todo o dinheiro no investimento com retorno mínimo é uma solução viável. Entretanto, essa idéia é ruim, pois este valor pode ser muito pessimista, e isso acarretaria na obtenção de muitas soluções iguais ao tentar minimizar a variância. Uma alternativa seria fazer R ser igual à média dos retornos dos investimentos. Porém, isso pode fazer com que parte da fronteira eficiente não seja obtida se R for maior que o valor mínimo de retorno da Fronteira. A melhor alternativa, a qual usamos para os experimentos, é minimizar a variância individualmente - sem considerar o retorno no problema - e avaliar o retorno da solução obtida. Esta se encontra na extremidade inferior da fronteira eficiente e o valor do seu retorno é um limite inferior para o retorno de qualquer

Tabela 5.4: *Desempenho do método da ϵ-restrição no Problema 1.*

Acréscimo	Subproblemas Resolvidos	Soluções diferentes encontradas
0.001	20	20
0.002	11	11
0.003	8	8
0.005	5	5
0.01	3	3

portfolio na curva. Assim, fazemos R ser igual ao retorno desta solução. Para obter novas soluções, acrescemos a R novos valores e resolvemos os novos subproblemas correspondentes, ou seja, a cada novo subproblema, fazemos $R = R$+acréscimo. O acréscimo é realizado até que o novo subproblema se torne inviável para o retorno esperado requerido. A escolha do valor de acréscimo é importante para o controle de cobertura da fronteira eficiente. É importante observar que, para otimizar portfolios, a escolha do vetor ϵ inicial não é difícil, mas isso nem sempre é verdade para outros tipos de problemas.

A Tabela 5.4 mostra o número de soluções eficientes diferentes encontradas para o Problema 1 e o número de subproblemas resolvidos para diferentes valores de acréscimo para R. Como podemos ver, o número de subproblemas resolvidos é sempre igual ao número de soluções diferentes obtidas. Isso se deve à escolha inicial de R que fizemos. Se o valor inicial de R fosse igual ao valor mínimo dos retornos dos investimentos, a diferença entre os números seria grande.

A Figura 5.5 mostra os pontos dispostos na fronteira eficiente para o Problema 1 obtidos usando o método com dois valores diferentes de acréscimo para o retorno esperado inicial: 0.01 e 0.001. Como podemos ver, obtemos pontos razoavelmente bem distribuídos nos dois casos. Naturalmente, quanto menor for o valor do acréscimo ao retorno esperado inicial, melhor será a aproximação da fronteira eficiente.

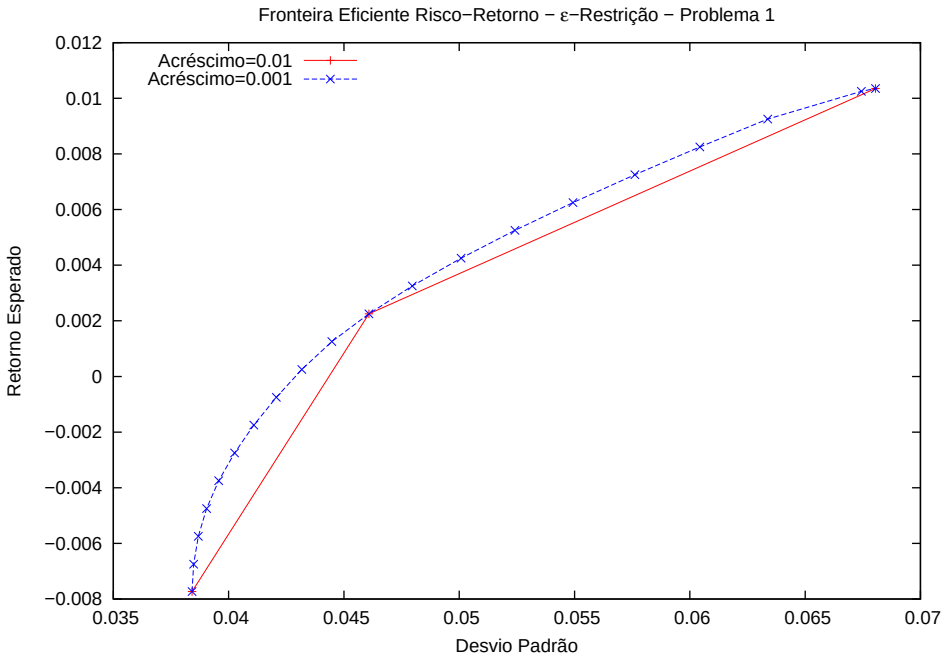


Figura 5.5: *Fronteira eficiente do Problema 1 aproximada com o método da ϵ-Restrição e acréscimos de 0.01 e 0.001 unidades ao retorno esperado inicial.*

Tabela 5.5: *Desempenho do método NISE no Problema 1.*

Constante	Subproblemas resolvidos	Soluções diferentes encontradas
0.5	3	3
0.1	5	5
0.05	7	7
0.03	10	10
0.02	10	10
0.01	15	15

Método NISE

No método NISE, o valor do erro máximo permitido é, geralmente, definido como sendo o valor do primeiro erro máximo possível calculado vezes alguma constante menor que um. O valor desta constante determinará quantos pontos na curva serão obtidos. Outra possibilidade é fazer o erro máximo permitido igual a zero e parar o método após um determinado número de iterações.

A Tabela 5.5 mostra o número de soluções eficientes diferentes encontradas para o Problema 3 e o número de subproblemas resolvidos para diferentes valores da constante multiplicadora.

A Figura 5.6 mostra os pontos dispostos na fronteira eficiente para o Problema 1 obtidos usando o método com dois valores diferentes da constante multiplicadora: 0.01 e 0.1. Os pontos na curva não foram obtidos em nenhuma ordem específica, visto que o método procura sempre encontrar pontos onde o erro máximo possível entre dois pontos já encontrados seja máximo. Como podemos observar, obtemos soluções bem distribuídas nos dois casos. Naturalmente, quanto menor for o valor da constante multiplicadora, menor será o erro máximo permitido e, portanto, melhor será a aproximação da fronteira eficiente.

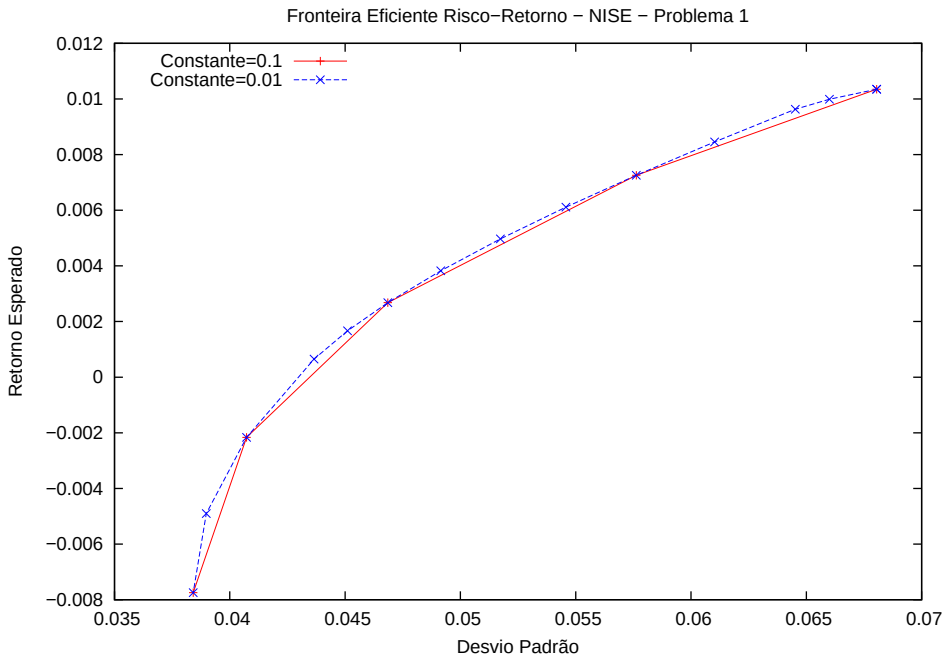


Figura 5.6: *Fronteira eficiente do Problema 1 aproximada com o método NISE e valores da constante multiplicadora iguais a 0.1 e 0.01.*

Método das direções viáveis

O método das direções viáveis, como foi descrito no Capítulo 4, é aplicado em problemas onde o conjunto viável é dado por restrições de desigualdade apenas. Para usá-lo nesta aplicação, nós o adaptamos para que ele pudesse ser usado em problemas onde o conjunto viável é representado por restrições de desigualdade e igualdade. Supomos aqui que as restrições de igualdade são lineares, de forma que a convergência do método continua sendo garantida. Mostraremos, mais na frente, um exemplo em que o método falha em encontrar uma solução crítica para um problema com restrições de igualdade não-lineares.

Considere o seguinte problema:

$$\begin{array}{ll} \min & f(x) = (f_1(x), f_2(x), \ldots, f_p(x))' \\ \text{s.a:} & g_1(x) \leq 0, \ldots, g_m(x) \leq 0, \\ & h_1(x) = 0, \ldots, h_l(x) = 0. \end{array} \tag{5.23}$$

Seja $X = \{x \mid g_j(x) \leq 0,\ j = 1, \ldots, m;\ h_i(x) = 0,\ i = 1, \ldots, l\}$. Como vimos no Capítulo 3, uma condição necessária para uma solução $x \in X$ ser fracamente eficiente local é a de que o sistema

$$\nabla f_k(x)'d < 0,\ \forall\, k = 1, \ldots, p,\ d \in T(X, x), \tag{5.24}$$

não possua solução. Supondo que a condição de qualificação de Abadie vale para todo $x \in X$, ou seja, $L_{\leq}(X, x) = T(X, x)$, temos que a condição (5.24) torna-se equivalente à de que o sistema

$$\begin{array}{ll} \nabla f_k(x)'d < 0, & \forall\, k = 1, \ldots, p, \\ \nabla g_j(x)'d \leq 0, & \forall\, j \in A(x), \\ \nabla h_i(x)'d = 0, & \forall\, i = 1, \ldots, l, \end{array} \tag{5.25}$$

não possua solução $d \in \Re^n$. Se este for o caso, dizemos que x é uma solução crítica.

Seja $I = \{1, \ldots, p\}$, $\epsilon \in \Re_+$ e $x \in \Re^n$. Defina o conjunto $I_\epsilon := \{j \in I \mid g_j(x) \geq -\epsilon\}$. Usando os resultados acima, resolvemos agora o seguinte subproblema para encontrar uma direção de descida viável para o atual ponto x de uma iteração:

$$\begin{array}{lll} \min & \alpha & \\ \text{s.a:} & \nabla f_k(x)'d \leq \alpha, & k = 1, \ldots, p, \\ & \nabla g_j(x)'d \leq \alpha, & j \in I_\epsilon(x), \\ & \nabla h_i(x)'d = 0, & i = 1, \ldots, l, \\ & \|d\|_\infty \leq 1. & \end{array} \tag{5.26}$$

Se $\alpha \geq 0$, x é uma solução crítica. Caso contrário, uma direção de descida viável é obtida.

A única modificação no subproblema que encontra uma direção de descida está no fato de que o novo ponto deve satisfazer as restrições de igualdade. No problema de Otimização de Portfolio que estamos lidando, temos apenas restrições lineares, o que implica que a condição de qualificação de Abadie é satisfeita para todo $x \in X$ e que a sequência de pontos geradas na execução do método satisfaz a viabilidade. Como as funções objetivo são convexas, se tivermos $\alpha \geq 0$, x será não apenas uma solução crítica, mas uma solução eficiente global.

O tamanho do passo t dado na direção d obtida no subproblema acima é dado pela seguinte fórmula:

$$t = \min\{t^*,\ t_{MAX}\},$$

onde t^* é o minimizador unidimensional da variância, uma função quadrática convexa, e t_{MAX} é o passo máximo que se pode dar para que se tenha $x + td \geq 0$. O passo máximo é calculado da seguinte maneira:

Como t é um número não-negativo, se $d_i \geq 0$, então $x_i + td_i \geq 0$. Caso contrário, devemos ter $t \leq -x_i/d_i$. Logo, temos

$$t = \min_i \left\{ \frac{-x_i}{d_i} \mid d_i < 0 \right\}.$$

Para os problemas utilizados em nossos experimentos, a modificação no método das direções viáveis para lidar com restrições de igualdade serviu para encontrar soluções eficientes globais. Na Figura 5.7, é mostrada a aproximação da fronteira eficiente para o Problema 1 obtida pelo método utilizando 10 pontos iniciais escolhidos aleatoriamente. Podemos ver que todas as 10 soluções diferentes encontradas são eficientes.

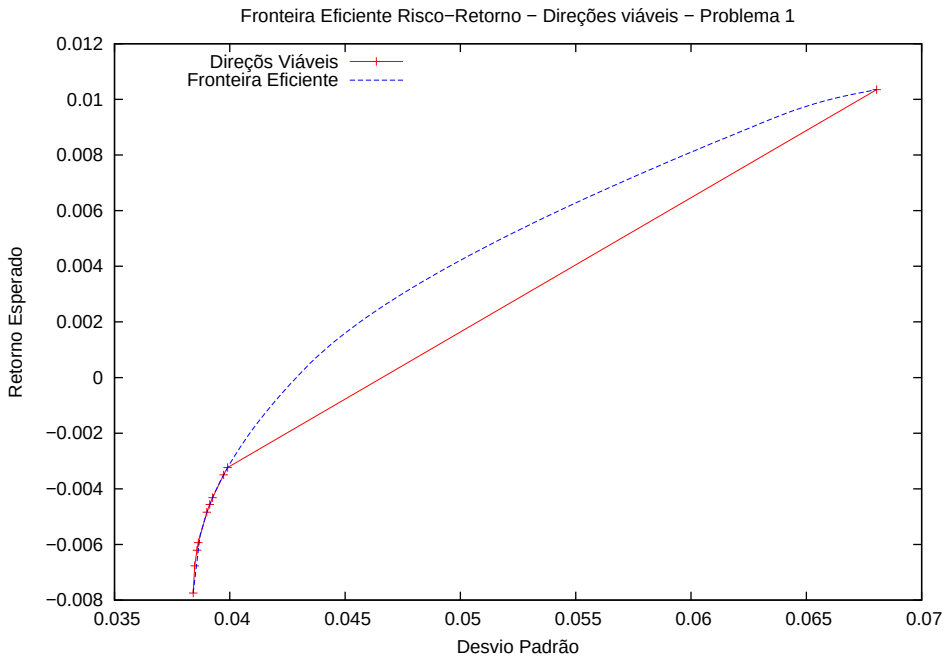


Figura 5.7: *Aproximação da fronteira eficiente do Problema 1 obtida pelo método das direções viáveis com multistart.*

A adaptação do método das direções viáveis para o caso em que há restrições de igualdade não-lineares não é imediata. O exemplo a seguir ilustra esse caso.

$$\begin{aligned} \min \quad & (x_1 + x_2, x_1 - x_2)' \\ \text{s.a:} \quad & x_1^2 + \frac{x_2^2}{4} = 1, \\ & x_1^2 + x_2^2 = 1, \\ & x \geq -2. \end{aligned} \tag{5.27}$$

As únicas soluções viáveis deste problema são os pontos $(-1, 0)$ e $(1, 0)$. O primeiro ponto, $(1, 0)$, é uma solução eficiente, enquanto que o ponto $(-1, 0)$ não é nem mesmo uma solução crítica. O cone linearizado do conjunto viável em qualquer dessas soluções é dado por $\{(0, y) \mid y \in \Re\}$. Os cones tangentes do conjunto viável nessas duas soluções são dados pelo mesmo conjunto, implicando que a

condição de qualificação de Abadie é satisfeita para todo ponto viável. Assim, temos um problema no qual, supostamente, o método das direções viáveis deveria encontrar uma solução crítica. Porém, escolhendo o ponto viável inicial como $x^0 = (-1, 0)$, o método nunca conseguirá sair desse ponto, apesar do outro ponto $(1, 0)$ ser eficiente. Este problema pode ser visto melhor na Figura 5.8.

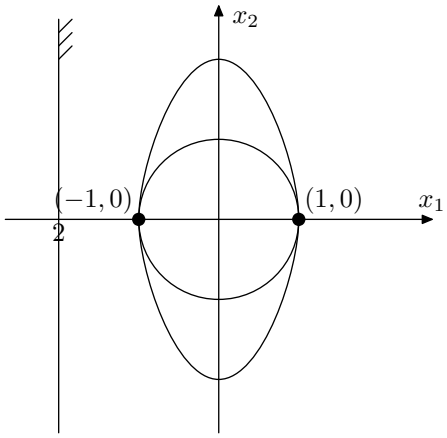


Figura 5.8: *Exemplo em que o método das direções viáveis modificado não consegue encontrar uma solução crítica.*

A falha em encontrar a solução crítica no problema anterior se deve à não-convexidade do conjunto viável. O fato dos problemas considerados em nossos experimentos serem todos convexos pode ter contribuído para a convergência do método.

Apesar de não sabermos *a priori* a localização exata na fronteira eficiente de uma solução obtida pelo método das direções viáveis, podemos ter uma idéia aproximada dela. Após escolher o ponto viável inicial x^0 para executar o método, sabemos que o percurso que ele fará no espaço objetivo será sempre de forma com que os valores das funções objetivo sejam todos decrescidos. Isso se deve ao fato de que as direções escolhidas são de descida para todas as funções objetivo. A Figura 5.9 mostra o percurso que um ponto inicial realiza até o ponto final obtido pelo método. A escolha do ponto inicial, portanto, influencia a localização do ponto final obtido pelo método, podendo servir para propósitos específicos do tomador de decisões ou mesmo para aproximar melhor a fronteira eficiente.

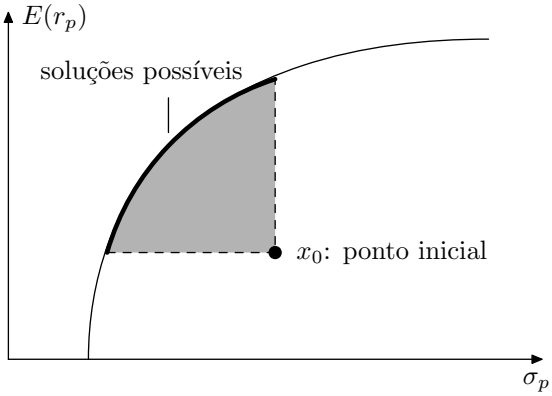


Figura 5.9: *Localização na fronteira eficiente das soluções obtidas pelo método das direções viáveis.*

Tabela 5.6: *Desempenho dos métodos para o Problema 1.*

Problema 1		
Método	Erro	Soluções diferentes encontradas
Pesos	0.238423	06
ϵ-Restrição	0.0393884	10
Nise	0.0289485	10
Direções Viáveis	1	10

5.2.3.2 Resultados comparativos entre os métodos

Para comparar os métodos, calculamos a área da figura situada entre a fronteira eficiente do problema e sua aproximação obtida pela interpolação linear das soluções no espaço objetivo fornecidas por cada método. O erro de um método é definido como o valor desta área, a qual é calculada através da regra trapezoidal de integração numérica. Se essa área for grande, significa que a aproximação não é boa; caso contrário, a curva está próxima da verdadeira. O cálculo do erro é mostrado melhor na Figura 5.10. Como estamos calculando as áreas das curvas usando integração numérica, resolvemos utilizar uma grande quantidade de pontos (em média, 200) obtidos pelo método da ϵ-restrição para estimar a fronteira eficiente verdadeira de cada problema. Se precisássemos da fronteira eficiente verdadeira, ela poderia ser obtida utilizando Programação Quadrática Paramétrica, por exemplo.

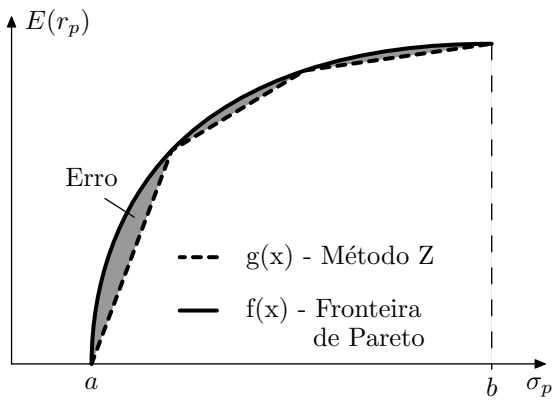


Figura 5.10: *Erro do método* $Z = \int_a^b (f(x) - g(x))dx$.

Supondo que estamos dispostos a resolver 10 subproblemas para obter uma aproximação da curva, comparamos os métodos entre si com base nos seus erros e no número de soluções diferentes encontradas. No caso do método das direções viáveis, estamos dispostos a executar o método com 10 pontos iniciais diferentes, escolhidos aleatoriamente. Em cada problema, dividimos o valor do erro de cada método pelo maior dos quatro para uma melhor comparação entre eles.

Como já foi dito sobre o método da ϵ-restrição, para cada subproblema, precisamos escolher o valor inicial R. O valor escolhido para cada problema será sempre igual ao retorno do portfolio que minimiza a variância, o qual é o valor mínimo de retorno que uma solução eficiente pode possuir.

Como mostram a Tabela 5.6 e a Figura 5.11, os métodos NISE e ϵ-Restrição possuem um desempenho muito superior ao dos pesos no Problema 1, o que já era esperado. Isso se deve principalmente à falta de controle de distribuição dos pontos na fronteira eficiente pelo do método dos pesos, o que não ocorre nos outros dois. Tal desempenho, como veremos adiante, repete-se em

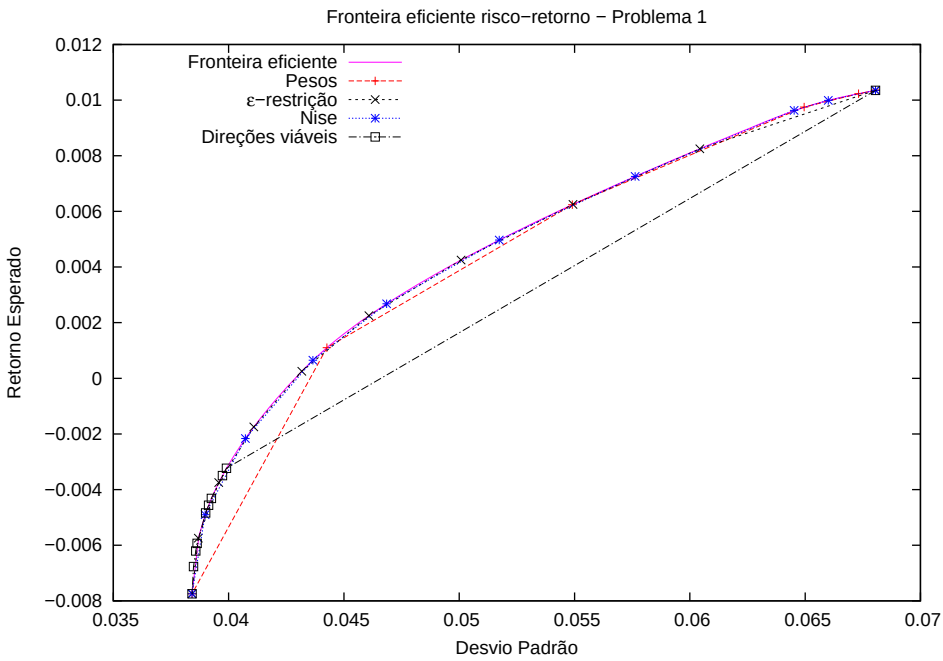


Figura 5.11: *Fronteira eficiente do Problema 1.*

Tabela 5.7: *Desempenho dos métodos para o Problema 2.*

Problema 2		
Método	Erro	Soluções diferentes encontradas
Pesos	0.136713	05
ϵ-Restrição	0.0259655	10
Nise	0.01638	09
Direções Viáveis	1	2

todos os outros problemas. Por sua vez, o método das direções viáveis apresenta o pior desempenho, o que acontecerá também em todos os outros problemas. A causa disso é a ausência de conhecimento *a priori* sobre a localização das soluções obtidas pelo método. A única informação que possuímos é a do percurso que o ponto inicial escolhido faz para chegar ao ponto final, como foi discutido anteriormente.

O valor inicial de R no método da ϵ-Restrição para o Problema 1 é -0.00774781, que é o valor mínimo de retorno que uma solução eficiente pode possuir neste problema. Ainda neste método, o valor de acréscimo escolhido para o retorno mínimo esperado é igual a 0.002, ou seja, para cada novo problema, fazemos $R = R + 0.002$. Tal valor foi definido de forma que o número de problemas resolvidos igualasse a 10, como combinamos inicialmente. Para o restantes dos problemas, seguimos o mesmo raciocínio para a escolha inicial de R e do valor de acréscimo.

Tabela 5.8: *Desempenho dos métodos para o Problema 3.*

Problema 3		
Método	Erro	Soluções diferentes encontradas
Pesos	0.110475	08
ϵ-Restrição	0.0556822	10
Nise	0.0756729	10
Direções Viáveis	1	10

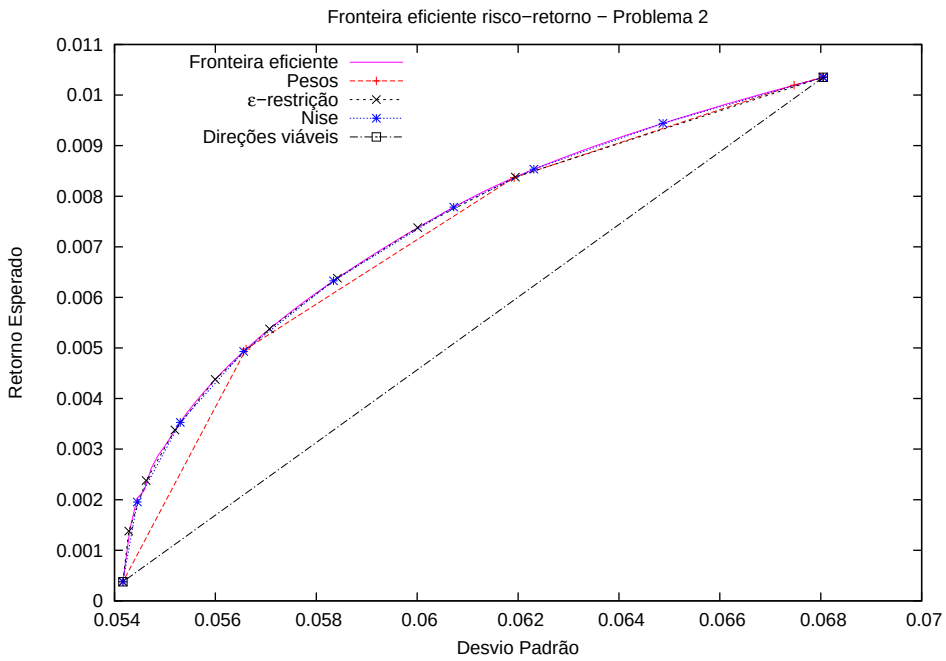


Figura 5.12: *Fronteira eficiente do Problema 2.*

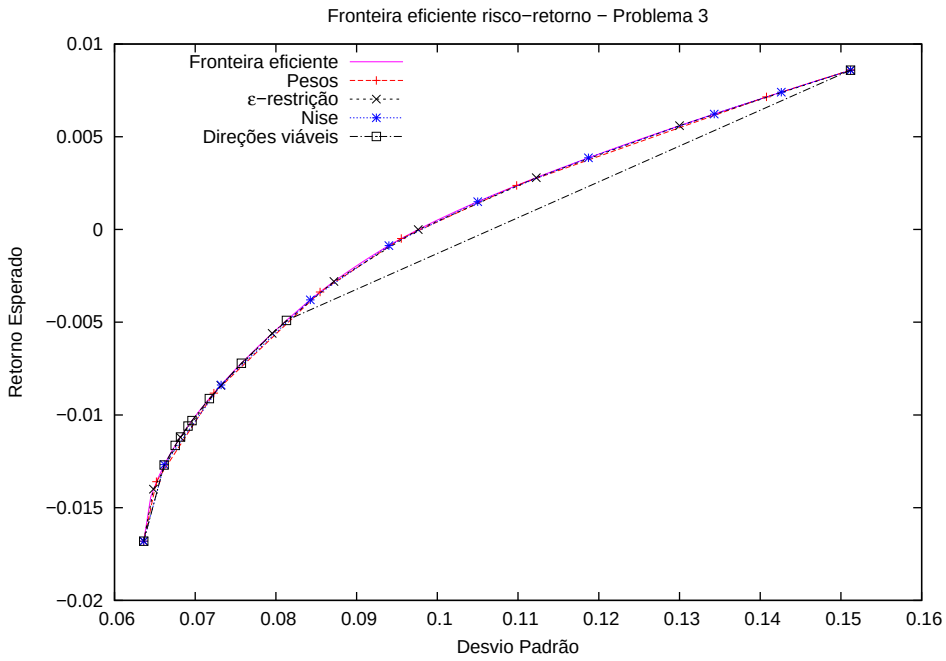


Figura 5.13: *Fronteira eficiente do Problema 3.*

Nos Problemas 1 e 2, o método NISE obteve erro menor que os outros três métodos. A sua superioridade em relação ao método dos pesos é natural, visto que o NISE aperfeiçoa o método dos pesos ao usá-lo dentro dele próprio. Entretanto, não há garantias de que ele também seja sempre melhor que o método da ϵ-Restrição. Como podemos ver na Figura 5.13, o erro do método da ϵ-Restrição no Problema 3 é menor que o do método NISE, apesar da diferença ser bem pequena.

No Problema 2, em particular, o método das direções viáveis obteve uma única solução, não importando qual era o ponto inicial aleatório. A solução obtida é a que minimiza a variância do

Tabela 5.9: *Desempenho dos métodos para o Problema 4.*

Problema 4		
Método	Erro	Soluções diferentes encontradas
Pesos	1	10
ϵ-Restrição	0.151038	10
Nise	0.1345	10
Direções Viáveis	0.816824	10

Tabela 5.10: *Desempenho dos métodos para o Problema 5.*

Problema 5		
Método	Erro	Soluções diferentes encontradas
Pesos	0.604715	10
ϵ-Restrição	0.193207	10
Nise	0.0809767	10
Direções Viáveis	1	10

portfolio. Como sempre otimizamos as funções objetivo individualmente antes de executar qualquer método, o número de soluções diferentes para o método das direções viáveis é igual a 2, sendo a segunda solução aquela que minimiza o retorno do portfolio.

Podemos ainda observar que o método dos pesos e o das direções viáveis conseguem aproximar melhor a fronteira eficiente no Problema 3 do que nos problemas anteriores. Para o método dos pesos, tal fato se deve à obtenção de um número maior de soluções diferentes no Problema 3, fazendo com que a distribuição dos pontos na curva seja melhor.

Para verificar a razão do método dos pesos não conseguir aproximar bem a parte inferior da fronteira eficiente no Problema 2, resolvemos este com o número de subproblemas igual a 50. A Figura 5.14 mostra o número de soluções diferentes encontradas ao variar o peso w no intervalo $[0, 1]$. Podemos ver que, para $w \in [0, 0.66]$, a solução obtida é sempre a mesma, apesar de 33 subproblemas diferentes terem sido resolvidos. Para obter pontos na parte da curva onde o método dos pesos obteve uma aproximação ruim, ou seja, quando $0.054 \leq \sigma_p < 0.057$, onde σ_p é o retorno esperado do portfolio, w deve estar dentro do intervalo $[0.88, 1]$. Temos, portanto, 66% do intervalo para w gerando uma única solução, enquanto que apenas 12% leva a pontos na parte inferior da curva. Esse exemplo mostra a dificuldade que o método dos pesos pode ter para gerar uma boa cobertura da fronteira eficiente.

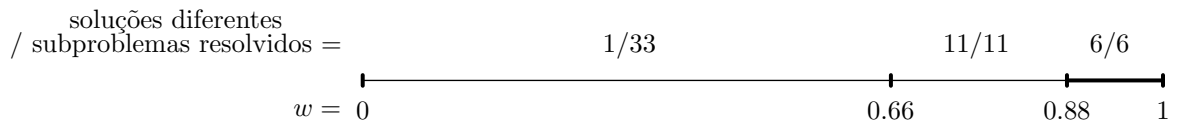


Figura 5.14: *Desempenho do método dos pesos no Problema 2 com 50 subproblemas resolvidos. Número de soluções diferentes encontradas × valores do peso w.*

Nos Problemas 4 e 5, apesar dos erros do método dos pesos e do método das direções viáveis serem maiores que os dos NISE e da ϵ-Restrição, podemos ver nas Figuras 5.15 e 5.16 que a aproximação da fronteira eficiente pelos dois primeiros métodos não é tão ruim nos dois problemas. Isso se deve ao formato da curva, a qual se aproxima de duas semi-retas conectadas. Casos como esses, em que a curva se assemelha a uma função linear por partes, favorecem o desempenho do método dos pesos e o das direções viáveis.

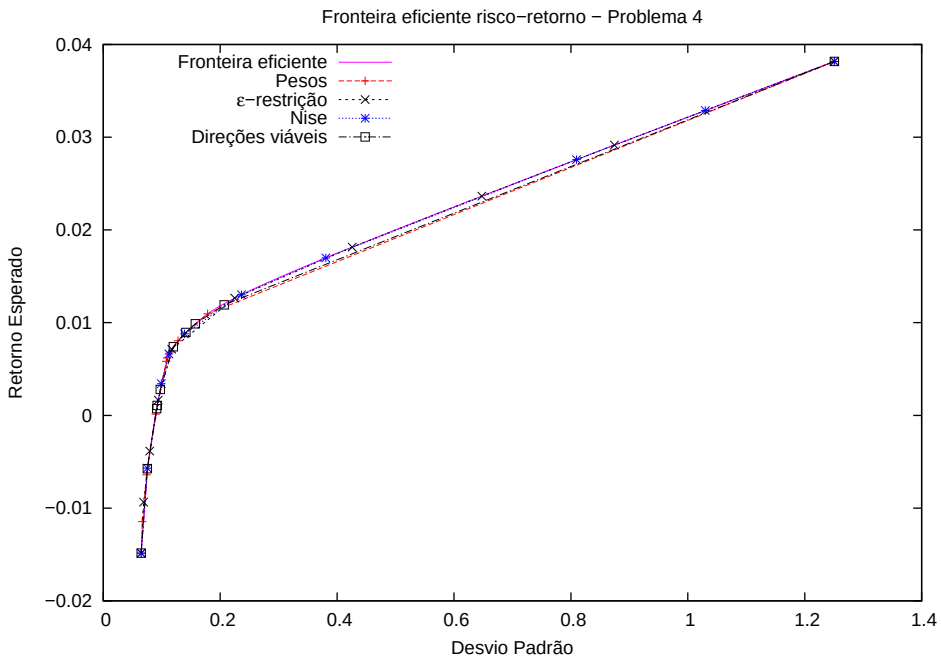


Figura 5.15: *Fronteira eficiente do Problema 4.*

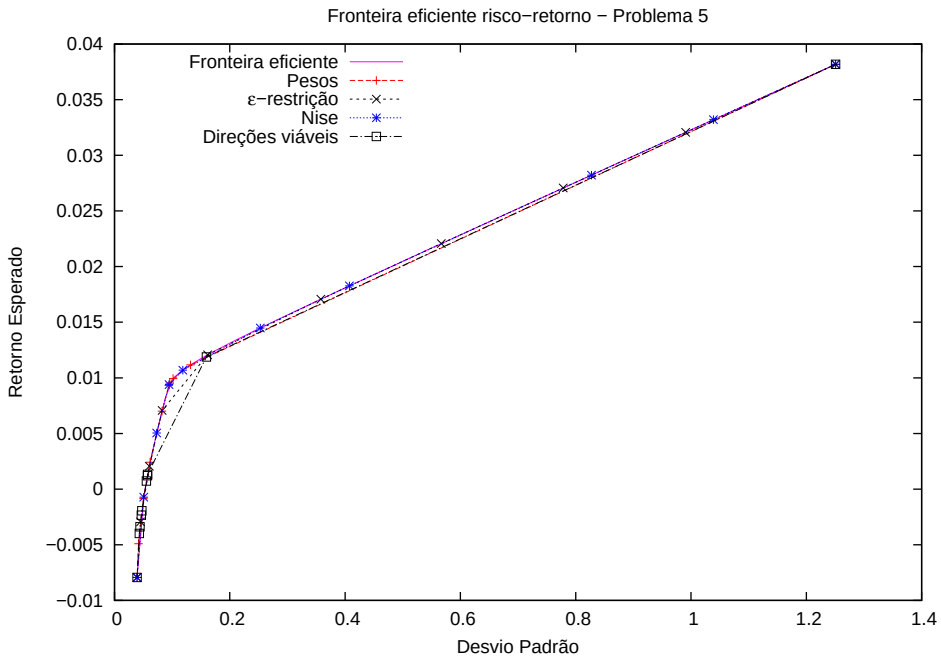


Figura 5.16: *Fronteira eficiente do Problema 5.*

Pelos resultados acima, podemos concluir que os métodos NISE e da ϵ-Restrição apresentam desempenho superior ao método dos pesos e ao método das direções viáveis quando estamos dispostos a resolver um número reduzido de subproblemas. Em particular, o método da ϵ-Restrição parecer ser um pouco mais complexo que os outros três devido à escolha dos seus parâmetros. Se não escolhermos um valor inicial para R razoável e um valor de acréscimo adequado para o mesmo, o problema pode tornar-se inviável ou podemos obter uma fraca aproximação da fronteira eficiente.

5.2.4 Modelo com 3 funções objetivo

É importante ressaltar que o modelo MV de Markowitz gerou muitas críticas por ser um modelo simplista. Outros modelos oriundos dele, como o modelo de Black-Litterman, tentam ser mais fiéis incorporando mais informações sobre o mercado. Detalhes sobre outros modelos de otimização em finanças podem ser encontrados em [Cornuéjols e Tütüncu, 2007]. Uma das críticas que é interessante ao nosso contexto diz respeito ao modelo de Markowitz tentar modelar o retorno de um investimento como uma função normalmente distribuída. Entretanto, estudos revelam que uma distribuição normal não é frequente nesta aplicação. Por isso, outros modelos tentam incorporar não apenas o primeiro e segundo momentos da distribuição de probabilidade (respectivamente, média e variância), mas também o terceiro momento: *assimetria*. Esta serve como uma medida da assimetria da distribuição de probabilidade do retorno do portfolio, sendo definida por

$$\mathrm{E}\left((r_p - \mu_p)^3\right).$$

Se seu valor for negativo, significa que a "cauda" do lado esquerdo da distribuição é mais longa que a do lado direito. Se for positivo, o inverso ocorre. Se for igual a zero, significa que os valores são relativamente distribuídos de forma uniforme, o que não implica numa distribuição simétrica. A Figura 5.17 mostra duas formas diferentes da distribuição para valores distintos do sinal. Nesse exemplo, temos dois portfolios com médias e variâncias iguais (a distribuição de probabilidade de um é a imagem espelhada do outro), mas com diferentes valores da assimetria.

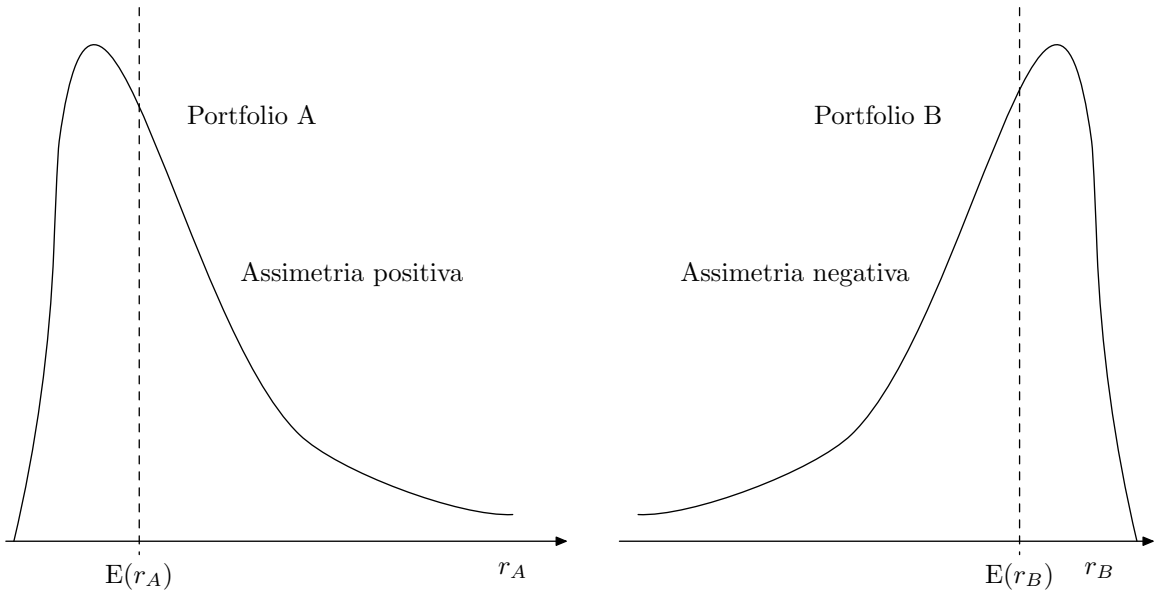


Figura 5.17: *Assimetria presente na função densidade de probabilidade de dois portfolios com médias e variâncias iguais.*

Os investidores preferem uma assimetria positiva, pois isso implica numa baixa probabilidade de obter um retorno negativo grande. Assim, em vez de otimizar apenas duas funções objetivo como é feito no modelo de Markowitz, temos agora três: maximizar o retorno, maximizar a assimetria e minimizar a variância; tal modelo é chamado de modelo da Média-Variância-Assimetria (MVA). Desta forma, os métodos de Otimização Multiobjetivo são bastante úteis e, de fato, têm sido usados para resolver estes problemas, como pode ser visto em [Lai, 1991] e [Wang e Xia, 2002].

O problema utilizando o modelo MVA é dado por:

$$
\begin{array}{rll}
\text{(Variância)} & \min & \displaystyle\sum_{i,j=1}^{n} x_i x_j \operatorname{Cov}(r_i, r_j) \\
\text{(Retorno)} & \max & \displaystyle\sum_{i=1}^{n} \mu_i x_i \\
\text{(Assimetria)} & \max & \displaystyle\sum_{i,j,k=1}^{n} x_i x_j x_k \operatorname{Coass}(r_i, r_j, r_k) \\
& \text{s.a:} & \displaystyle\sum_{i=1}^{n} x_i = 1, \\
& & x \geq 0,
\end{array}
\tag{5.28}
$$

onde

$$
\operatorname{Coass}(r_i, r_j, r_k) = \frac{1}{T} \sum_{t=1}^{T} (r_{it} - \mu_i)(r_{jt} - \mu_j)(r_{kt} - \mu_k)
$$

é a co-assimetria entre os retornos r_i, r_j e r_k.

Neste modelo, perdemos a convexidade devido à nova função objetivo referente à assimetria do portfolio. Além disso, o espaço objetivo se torna tridimensional, tornando-se ainda mais difícil obter uma visualização geométrica da fronteira eficiente. Uma forma de abordar o problema (5.28) é através de um método de Otimização Multiobjetivo conhecido como *Programação de Alvo* (originalmente, *Goal Programming*). Em tal método, o tomador de decisão especifica níveis de aspiração para as funções objetivo e quaisquer desvios desses níveis são minimizados. Uma função objetivo juntamente com um nível de aspiração forma um *alvo*. Na literatura, podemos ver essa técnica largamente aplicada no modelo MVA.

Existe ainda um outro modelo que adiciona o quarto momento da distribuição de probabilidade no problema, chamada de *curtose*. Esta é uma medida do achatamento da curva da função de distribuição. Para mais informações sobre os modelos de três e quatro objetivos, ver, por exemplo, [Kraus e Litzenberger, 1976] e [Lai *et al.*, 2006].

Capítulo 6

Conclusões

Neste trabalho, estudamos o problema de Otimização Multiobjetivo e os conceitos básicos dessa área no Capítulo 2. Algumas condições de otimalidade e regularidade encontradas na literatura foram discutidas no Capítulo 3, onde foi mostrado o progresso em relação à obtenção de multiplicadores de Lagrange estritamente positivos associados às funções objetivo. No Capítulo 4, apresentamos métodos existentes para resolver problemas de Otimização Multiobjetivo. Discutimos suas vantagens e desvantagens, procurando sempre exemplificar os casos em que cada um pode falhar. Vimos, através de um exemplo, que o método dos pesos e o NISE não servem para resolver problemas em que o conjunto viável no espaço objetivo não é convexo, enquanto que o método da ϵ-restrição consegue fazê-lo ao variar seu parâmetro. Por outro lado, esses três métodos propõem aproximar toda a fronteira eficiente, enquanto que o método do gradiente e o das direções viáveis visam obter apenas uma solução eficiente, sendo necessário o uso de *multistart* para tentar aproximar toda a Fronteira. Em particular, o método do gradiente e o das direções viáveis dão poucas informações sobre a localização da solução eficiente obtida na fronteira eficiente.

Também apresentamos aplicações da teoria e métodos de Otimização Multiobjetivo nas áreas de *Compressed Sensing* e Otimização de Portfolio. Nesta última, implementamos quatro métodos para resolver 5 problemas e analisamos seus desempenhos. Pudemos concluir que o método NISE obteve desempenho superior aos outros quatro no propósito de aproximar a fronteira eficiente, sendo seguido pelo método da ϵ-restrição e depois pelo método dos pesos. Tais resultados servem como estímulo ao uso de métodos de Otimização Multiobjetivo na área de Otimização de Portfolio.

6.1 Trabalhos futuros

Na parte prática, pretendemos estudar com mais profundidade a área de *Compressed Sensing*, analisando todos os resultados obtidos até o presente nessa área, para então aplicar a teoria e os métodos de Otimização Multiobjetivo discutidos neste trabalho. Uma possibilidade seria implementar o método do gradiente apresentado no Capítulo 4 e verificar a qualidade da solução, a qual poderia ser uma representação esparsa de uma imagem. Os problemas a serem abordados se dividem em duas classes: aqueles que podem ser tratados como problemas de Programação Linear e os que podem ser tratados como problemas de cone de segunda ordem. Em particular, o problema (5.7), onde é considerado haver erro nas medições, pode ser tratado como o seguinte problema:

$$
\begin{aligned}
\min \quad & \sum_i u_i \\
\text{s.a:} \quad & s - u \leq 0, \\
& -s - u \leq 0, \\
& \frac{1}{2}\left(\|\Theta s - y\|_2^2 - \tau^2\right) \leq 0.
\end{aligned}
\tag{6.1}
$$

Este problema pode então ser resolvido, por exemplo, usando o Método da Barreira disponível no resolvedor CPLEX. Igualmente, podemos usar o CPLEX para resolver a "versão linear" dos problemas da outra classe.

Na parte teórica, também estamos interessados no fortalecimento das condições de Fritz John para Otimização Multiobjetivo. Até um certo tempo, a linha de desenvolvimento clássica da teoria de multiplicadores de Lagrange se baseava no estabelecimento de condições de fácil verificação que implicavam na condição de qualificação de Abadie, que, por sua vez, garante a existência de multiplicadores de Lagrange através do Lema de Farkas. Apesar de ser intuitiva, tal linha não se estende ao caso em que um conjunto abstrato qualquer S é adicionado ao conjunto viável X do problema, que é composto por restrições de igualdade e desigualdade apenas. Em [Bertsekas *et al.*, 2003], as condições de Fritz John são generalizadas para o caso em que um conjunto abstrato adicional é considerado e é apresentada uma nova condição de qualificação, chamada de *pseudonormalidade*, que se aplica a tal caso. Nossa intenção é, portanto, seguir a mesma linha de desenvolvimento dentro da área de Otimização Multiobjetivo, dado que não encontramos na literatura nenhuma abordagem deste tipo.

Apêndice A

Dados dos experimentos computacionais

Os dados dos problemas da Tabela 5.2 usados em nossos experimentos estão dispostos neste apêndice da seguinte forma:

- preços do investimento 1 em ordem crescente dos meses considerados
- preços do investimento 2 em ordem crescente dos meses considerados
- ...
- preços do investimento 10 em ordem crescente dos meses considerados

A.1 Problema 1

```
30.92 33.45 33.19 29.70 33.05 29.01 36.49 36.35 27.66 27.00 23.34 24.10
15.26 8.76 10.49 11.70 9.91 11.65 18.98 17.80 16.39 23.80 27.31 26.68
26.14 29.04 28.24 24.37 13.70 12.29 11.84 11.00 10.80 13.20 12.10 13.05

31.30 31.00 30.40 29.25 28.27 23.11 28.90 32.54 26.15 25.10 23.79 22.75
14.77 14.30 14.68 14.20 13.85 16.87 18.46 21.16 21.18 23.64 26.16 31.23
28.21 30.90 29.70 28.10 29.77 29.85 29.99 25.98 24.65 30.54 28.48 32.13

40.84 42.56 37.64 40.60 40.64 40.92 48.40 45.04 41.16 36.08 33.68 31.16
23.56 29.00 28.24 33.00 38.70 30.55 35.27 39.90 37.32 42.50 43.25 44.70
43.20 52.53 54.48 53.49 49.10 48.30 46.20 49.97 46.40 47.05 42.20 46.20

11.99 9.55 8.40 8.59 9.35 7.26 7.65 6.74 5.98 5.44 6.05 7.00
5.50 6.90 4.91 6.75 6.54 5.90 6.56 7.66 6.70 5.53 5.45 5.87
5.50 6.36 7.15 7.00 7.35 7.00 6.10 7.01 7.30 6.95 6.99 7.26

59.45 47.50 45.20 48.88 48.32 44.50 43.81 45.63 44.85 45.10 46.50 44.69
48.00 44.50 45.71 44.30 44.19 48.19 47.78 45.59 43.85 43.18 43.49 44.41
43.70 43.15 43.54 41.58 39.40 38.51 34.00 34.89 36.10 37.92 40.31 41.50

122.00 123.08 125.50 120.08 125.00 123.00 32.30 31.10 31.50 31.00 27.16 25.50
23.50 22.95 25.00 26.00 28.50 31.00 26.50 30.00 32.00 31.00 31.03 32.50
33.33 33.60 34.00 34.00 34.03 38.00 38.40 30.43 36.06 36.97 33.01 30.19

143.00 133.50 128.65 123.50 139.60 132.00 123.94 113.40 101.27 93.00 100.32 105.00
91.00 98.21 101.34 95.00 95.89 109.36 123.52 129.40 127.40 132.20 141.00 146.26
159.74 167.50 174.50 173.00 174.02 163.00 169.47 173.45 179.91 190.38 193.49 205.05

72.02 71.90 88.40 80.40 81.45 73.99 42.20 49.00 46.21 35.90 34.90 35.10
23.31 20.06 22.84 25.03 26.40 28.55 29.54 34.45 32.45 31.47 31.38 35.00
35.04 38.80 36.69 34.17 34.61 35.39 32.80 29.60 26.86 27.85 26.06 27.50
```

20.91 19.60 20.15 19.05 18.90 17.47 17.35 15.29 10.67 11.95 13.83 13.10
11.25 8.56 8.81 8.69 6.59 7.77 8.85 9.70 8.15 9.12 10.12 10.25
8.94 8.74 9.51 9.99 9.81 10.52 10.20 9.75 9.28 11.49 10.80 11.81

59.20 56.90 56.95 46.73 52.80 48.80 38.20 39.30 32.99 33.15 29.95 30.94
25.09 24.63 22.59 20.80 20.70 23.10 27.09 30.30 28.98 29.55 30.75 35.25
34.56 36.50 36.38 31.33 31.15 32.80 31.81 30.40 28.11 32.25 30.46 33.92

A.2 Problema 2

13.15 13.00 11.75 10.13 10.75 10.30 10.95 11.75 10.17 10.31 9.53 9.52
7.20 8.25 8.01 7.32 7.07 7.97 9.38 8.97 8.72 9.57 9.68 10.72
10.01 11.21 11.85 11.23 11.56 12.18 11.99 11.15 10.71 13.08 12.30 12.92

272.00 181.80 163.00 165.00 196.00 196.80 162.32 175.00 158.00 138.00 114.40 81.02
56.00 46.74 53.04 57.72 51.86 59.00 64.40 79.00 83.22 88.56 88.60 93.42
92.00 102.40 98.78 99.20 102.64 121.84 112.22 91.80 96.22 98.96 88.06 22.70

54.55 52.15 50.75 44.77 49.86 50.79 53.59 54.36 47.70 40.85 38.00 32.71
25.40 24.51 23.89 28.01 26.83 26.75 30.48 32.50 29.85 32.40 32.90 36.60
39.45 42.37 42.20 42.14 44.45 49.55 46.53 42.88 37.91 42.67 41.43 46.30

51.10 50.51 42.65 37.24 35.48 33.80 38.18 35.25 30.32 32.00 32.32 36.00
22.99 15.50 19.09 17.28 16.15 12.70 14.94 16.80 20.20 22.90 24.05 23.00
25.25 29.94 38.21 34.47 33.49 30.40 30.80 24.28 25.00 29.41 34.26 38.31

143.00 133.50 128.65 123.50 139.60 132.00 123.94 113.40 101.27 93.00 100.32 105.00
91.00 98.21 101.34 95.00 95.89 109.36 123.52 129.40 127.40 132.20 141.00 146.26
159.74 167.50 174.50 173.00 174.02 163.00 169.47 173.45 179.91 190.38 193.49 205.05

26.85 25.00 23.93 21.90 24.50 26.80 25.00 29.49 29.60 30.40 29.50 28.30
26.25 27.00 25.89 26.16 26.20 25.92 28.80 26.84 28.65 28.70 27.55 27.60
24.91 29.00 36.34 40.04 23.40 26.20 24.70 22.65 23.86 22.80 21.37 21.56

49.40 49.50 45.50 39.53 42.60 40.01 47.43 50.46 32.60 33.50 30.75 31.90
23.30 26.64 26.10 23.50 22.20 25.65 30.40 31.90 31.10 33.60 31.98 35.70
33.50 37.95 38.69 36.13 36.50 39.00 37.70 34.70 32.50 39.59 37.82 40.47

20.30 17.20 15.55 13.35 15.40 13.05 11.90 13.05 10.50 11.90 10.33 8.35
5.96 6.40 6.27 6.39 5.65 6.49 9.00 9.42 9.18 10.50 11.25 11.96
11.60 14.50 15.53 13.23 13.00 13.25 12.93 12.40 13.08 15.00 14.79 15.48

72.02 71.90 88.40 80.40 81.45 73.99 42.20 49.00 46.21 35.90 34.90 35.10
23.31 20.06 22.84 25.03 26.40 28.55 29.54 34.45 32.45 31.47 31.38 35.00
35.04 38.80 36.69 34.17 34.61 35.39 32.80 29.60 26.86 27.85 26.06 27.50

38.00 37.30 34.15 45.40 42.21 46.60 38.00 42.40 40.19 37.20 37.95 32.50
28.20 33.51 32.21 28.70 29.65 32.50 34.12 35.11 31.30 28.89 30.20 33.50
33.70 36.90 37.10 33.60 31.30 31.51 26.24 29.77 27.00 25.75 23.90 24.20

A.3 Problema 3

42.99 42.00 36.00 27.60 33.57 33.10 39.20 40.50 31.88 31.39 29.50 23.50
15.70 13.28 15.67 14.15 14.10 14.25 20.00 21.72 21.62 27.40 29.90 31.06
31.00 38.25 39.30 35.98 39.02 40.80 43.01 43.50 49.00 58.96 56.70 57.99

47.70 45.93 43.76 33.95 29.28 26.09 26.34 25.22 17.92 15.96 14.66 11.95
9.38 8.72 9.91 10.25 9.43 6.65 7.53 8.55 10.89 13.73 17.35 18.36
18.25 25.01 26.08 23.10 24.55 22.30 23.10 20.30 21.38 24.50 22.70 25.95

51.10 50.51 42.65 37.24 35.48 33.80 38.18 35.25 30.32 32.00 32.32 36.00
22.99 15.50 19.09 17.28 16.15 12.70 14.94 16.80 20.20 22.90 24.05 23.00
25.25 29.94 38.21 34.47 33.49 30.40 30.80 24.28 25.00 29.41 34.26 38.31

54.00 50.90 51.96 45.31 55.26 54.10 65.51 81.88 38.40 34.30 30.70 21.21
14.00 14.55 15.06 15.05 12.66 12.90 15.70 20.61 20.50 21.93 22.06 23.75
26.30 27.10 29.14 25.34 26.45 28.90 28.45 25.00 23.59 25.40 23.36 22.70

20.91 19.60 20.15 19.05 18.90 17.47 17.35 15.29 10.67 11.95 13.83 13.10
11.25 8.56 8.81 8.69 6.59 7.77 8.85 9.70 8.15 9.12 10.12 10.25
8.94 8.74 9.51 9.99 9.81 10.52 10.20 9.75 9.28 11.49 10.80 11.81

59.20 56.90 56.95 46.73 52.80 48.80 38.20 39.30 32.99 33.15 29.95 30.94
25.09 24.63 22.59 20.80 20.70 23.10 27.09 30.30 28.98 29.55 30.75 35.25
34.56 36.50 36.38 31.33 31.15 32.80 31.81 30.40 28.11 32.25 30.46 33.92

122.00 123.08 125.50 120.08 125.00 123.00 32.30 31.10 31.50 31.00 27.16 25.50
23.50 22.95 25.00 26.00 28.50 31.00 26.50 30.00 32.00 31.00 31.03 32.50
33.33 33.60 34.00 34.00 34.03 38.00 38.40 30.43 36.06 36.97 33.01 30.19

49.40 49.50 45.50 39.53 42.60 40.01 47.43 50.46 32.60 33.50 30.75 31.90
23.30 26.64 26.10 23.50 22.20 25.65 30.40 31.90 31.10 33.60 31.98 35.70
33.50 37.95 38.69 36.13 36.50 39.00 37.70 34.70 32.50 39.59 37.82 40.47

30.92 33.45 33.19 29.70 33.05 29.01 36.49 36.35 27.66 27.00 23.34 24.10
15.26 8.76 10.49 11.70 9.91 11.65 18.98 17.80 16.39 23.80 27.31 26.68
26.14 29.04 28.24 24.37 13.70 12.29 11.84 11.00 10.80 13.20 12.10 13.05

36.00 32.64 28.80 25.50 26.00 29.19 32.40 35.00 30.99 29.20 29.03 24.88
24.00 26.56 25.70 26.25 25.00 28.05 27.55 28.76 30.15 27.75 25.80 27.25
26.15 26.80 29.00 26.35 26.30 32.90 28.70 27.50 25.50 26.56 24.00 26.25

A.4 Problema 4

47.70 45.93 43.76 33.95 29.28 26.09 26.34 25.22 17.92 15.96 14.66 11.95
9.38 8.72 9.91 10.25 9.43 6.65 7.53 8.55 10.89 13.73 17.35 18.36
18.25 25.01 26.08 23.10 24.55 22.30 23.10 20.30 21.38 24.50 22.70 25.95

40.84 42.56 37.64 40.60 40.64 40.92 48.40 45.04 41.16 36.08 33.68 31.16
23.56 29.00 28.24 33.00 38.70 30.55 35.27 39.90 37.32 42.50 43.25 44.70
43.20 52.53 54.48 53.49 49.10 48.30 46.20 49.97 46.40 47.05 42.20 46.20

11.99 9.55 8.40 8.59 9.35 7.26 7.65 6.74 5.98 5.44 6.05 7.00
5.50 6.90 4.91 6.75 6.54 5.90 6.56 7.66 6.70 5.53 5.45 5.87
5.50 6.36 7.15 7.00 7.35 7.00 6.10 7.01 7.30 6.95 6.99 7.26

93.89 81.00 71.00 62.51 71.00 59.00 55.10 70.43 58.69 59.00 58.15 45.94
27.80 24.01 23.80 24.85 21.36 21.57 34.36 41.28 37.11 43.80 48.62 49.80
51.00 50.50 47.80 38.00 37.50 38.40 38.37 30.60 30.08 33.80 28.00 31.30

54.00 50.90 51.96 45.31 55.26 54.10 65.51 81.88 38.40 34.30 30.70 21.21
14.00 14.55 15.06 15.05 12.66 12.90 15.70 20.61 20.50 21.93 22.06 23.75
26.30 27.10 29.14 25.34 26.45 28.90 28.45 25.00 23.59 25.40 23.36 22.70

20.91 19.60 20.15 19.05 18.90 17.47 17.35 15.29 10.67 11.95 13.83 13.10
11.25 8.56 8.81 8.69 6.59 7.77 8.85 9.70 8.15 9.12 10.12 10.25
8.94 8.74 9.51 9.99 9.81 10.52 10.20 9.75 9.28 11.49 10.80 11.81

20.30 17.20 15.55 13.35 15.40 13.05 11.90 13.05 10.50 11.90 10.33 8.35
5.96 6.40 6.27 6.39 5.65 6.49 9.00 9.42 9.18 10.50 11.25 11.96

11.60 14.50 15.53 13.23 13.00 13.25 12.93 12.40 13.08 15.00 14.79 15.48

122.00 123.08 125.50 120.08 125.00 123.00 32.30 31.10 31.50 31.00 27.16 25.50
23.50 22.95 25.00 26.00 28.50 31.00 26.50 30.00 32.00 31.00 31.03 32.50
33.33 33.60 34.00 34.00 34.03 38.00 38.40 30.43 36.06 36.97 33.01 30.19

42.99 42.00 36.00 27.60 33.57 33.10 39.20 40.50 31.88 31.39 29.50 23.50
15.70 13.28 15.67 14.15 14.10 14.25 20.00 21.72 21.62 27.40 29.90 31.06
31.00 38.25 39.30 35.98 39.02 40.80 43.01 43.50 49.00 58.96 56.70 57.99

0.90 0.79 0.75 0.61 0.64 0.60 0.61 0.66 0.59 0.62 0.60 0.48
0.31 0.35 0.34 0.30 0.29 0.36 0.37 0.38 0.38 0.44 0.46 0.40
0.49 0.50 0.48 0.52 0.52 0.50 4.20 3.80 4.03 3.57 3.60 3.34

A.5 Problema 5

36.00 32.64 28.80 25.50 26.00 29.19 32.40 35.00 30.99 29.20 29.03 24.88
24.00 26.56 25.70 26.25 25.00 28.05 27.55 28.76 30.15 27.75 25.80 27.25
26.15 26.80 29.00 26.35 26.30 32.90 28.70 27.50 25.50 26.56 24.00 26.25

17.24 17.50 18.25 17.65 18.78 19.30 20.00 19.87 17.28 18.64 16.76 15.74
12.58 14.80 13.64 11.88 11.69 12.90 13.53 14.00 13.05 13.40 13.65 15.68
14.99 16.99 16.75 13.09 12.42 11.40 11.16 11.89 12.14 11.60 11.11 11.15

54.55 52.15 50.75 44.77 49.86 50.79 53.59 54.36 47.70 40.85 38.00 32.71
25.40 24.51 23.89 28.01 26.83 26.75 30.48 32.50 29.85 32.40 32.90 36.60
39.45 42.37 42.20 42.14 44.45 49.55 46.53 42.88 37.91 42.67 41.43 46.30

0.90 0.79 0.75 0.61 0.64 0.60 0.61 0.66 0.59 0.62 0.60 0.48
0.31 0.35 0.34 0.30 0.29 0.36 0.37 0.38 0.38 0.44 0.46 0.40
0.49 0.50 0.48 0.52 0.52 0.50 4.20 3.80 4.03 3.57 3.60 3.34

59.45 47.50 45.20 48.88 48.32 44.50 43.81 45.63 44.85 45.10 46.50 44.69
48.00 44.50 45.71 44.30 44.19 48.19 47.78 45.59 43.85 43.18 43.49 44.41
43.70 43.15 43.54 41.58 39.40 38.51 34.00 34.89 36.10 37.92 40.31 41.50

13.15 13.00 11.75 10.13 10.75 10.30 10.95 11.75 10.17 10.31 9.53 9.52
7.20 8.25 8.01 7.32 7.07 7.97 9.38 8.97 8.72 9.57 9.68 10.72
10.01 11.21 11.85 11.23 11.56 12.18 11.99 11.15 10.71 13.08 12.30 12.92

32.50 33.90 34.30 25.70 26.50 23.65 24.90 26.80 19.90 19.79 12.44 8.50
5.79 5.15 6.02 6.69 5.95 7.03 8.87 11.38 11.78 12.06 11.62 13.06
11.40 11.80 12.25 12.73 11.84 12.04 11.50 12.13 11.60 12.95 12.80 14.15

31.99 31.30 27.50 28.35 28.85 25.89 32.00 32.71 31.69 31.86 29.82 25.00
21.40 22.63 23.63 22.85 23.18 20.85 26.33 30.89 31.43 30.50 31.05 30.31
34.80 37.85 39.89 40.22 38.15 39.49 40.24 37.15 37.38 40.06 40.30 43.65

272.00 181.80 163.00 165.00 196.00 196.80 162.32 175.00 158.00 138.00 114.40 81.02
56.00 46.74 53.04 57.72 51.86 59.00 64.40 79.00 83.22 88.56 88.60 93.42
92.00 102.40 98.78 99.20 102.64 121.84 112.22 91.80 96.22 98.96 88.06 22.70

93.89 81.00 71.00 62.51 71.00 59.00 55.10 70.43 58.69 59.00 58.15 45.94
27.80 24.01 23.80 24.85 21.36 21.57 34.36 41.28 37.11 43.80 48.62 49.80
51.00 50.50 47.80 38.00 37.50 38.40 38.37 30.60 30.08 33.80 28.00 31.30

Referências Bibliográficas

Abadie(1967) J. Abadie. On the kuhn-tucker theorem. Em *Nonlinear Programming*, páginas 21–36, New York, NY. John Wiley. Citado na pág. 15

Andreani *et al.***(2005)** R. Andreani, J. M. Martínez, e M. L. Schuverdt. On the relation between the constant positive linear dependence condition and quasinormality constraint qualification. *Journal of Optimization Theory and Applications*, 125:473–485. Citado na pág. 22

Baraniuk(2007) R. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, 24:118–121. Citado na pág. 43, 46

Benson e Morin(2000) H. P. Benson e T. L. Morin. The vector maximization problem: Proper efficiency and stability. *Mathematical Methods of Operations Research*, 51:479–498. Citado na pág. 28

Berge e Gouila-Houri(1965) C. Berge e A. Gouila-Houri. *Programming, Games, and Transportation Networks*. Wiley, New York. Citado na pág. 28

Bertsekas *et al.***(2003)** D. P. Bertsekas, A. Nedić, e A. E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, Belmont, Massachusetts. Citado na pág. 4, 70

Bigi(1999) G. Bigi. *Optimality and Lagrangian Regularity in Vector Optimization*. Tese de Doutorado, Dipartimento di Matematica, Università di Pisa, Italy. Citado na pág. 1, 11, 14, 21

Borwein e Zhuang(1993) J. M. Borwein e D. Zhuang. Super efficiency in vector optimization. *Transactions of the American Mathematical Society*, 338:105–122. Citado na pág. 8

Candès(2006) E. Candès. Compressive sampling. *International Congress of Mathematics*, 3: 1433–1452. Citado na pág. 43, 46

Candès *et al.***(2006)** E. Candès, J. Romberg, e T. Tao. Robust uncertainty principles: Exact recovery from highly incomplete fourier information. *IEEE Transactions on Information Theory*, 52:489–509. Citado na pág. 43, 45

Chankong e Haimes(1982) V. Chankong e Y. Y. Haimes. On the characterization of noninferior solutions of the vector optimization problem. *Automatica*, 18:697–707. Citado na pág. 1

Chen *et al.***(2001)** S. S. Chen, D. L. Donoho, e M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43:129–159. Citado na pág. 43, 46

Cohon(1978) J. L. Cohon. *Multiobejctive Programing and Planning*. Academic Press, New York. Citado na pág. 1, 3, 23, 24, 30

Cohon *et al.***(1979)** J. L. Cohon, R. L. Church, e D. P. Sheer. Generating multiobjective trade-offs: An algorithm for bicriterion problems. *Water Resources Research*, 15:1001–1010. Citado na pág. 23, 30

Cornuéjols e Tütüncu(2007) G. Cornuéjols e R. Tütüncu. *Optimization Methods in Finance*. Cambridge University Press, United Kingdom. Citado na pág. 66

Das e Dennis(1997) I. Das e J. E. Dennis. A closer look at drawbacks at minimizing weighted sums of objectives for pareto set generation in multicriteria optimization problems. *Structural and Multidisciplinary Optimization*, 14:63–69. Citado na pág. 29

Fliege e Svaiter(2000) J. Fliege e B. F. Svaiter. Steepest descent methods for multicriteria optmization. *Mathematical Methods of Operations Research*, 51:479–494. Citado na pág. 1, 23, 38, 40

Fliege *et al.***(2009)** J. Fliege, L. M. Graña Drummond, e B. F. Svaiter. Newton's method for multiobjective optimization. *SIAM Journal on Optimization*, 20:602–626. Citado na pág. 23

Geoffrion(1968) A. M. Geoffrion. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22:618–630. Citado na pág. 3, 7, 27, 35

Giorgi e Guerraggio(1992) G. Giorgi e A. Guerraggio. On the notion of tangent cone in mathematical programming. *Optimization*, 25:11–23. Citado na pág. 18

Haimes *et al.***(1971)** Y. Y. Haimes, L. S. Lasdon, e D. A. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-1:296–297. Citado na pág. 23, 32

Isermann(1974) H. Isermann. Proper efficiency and the linear vector maximum problem. *Operations research*, 22:189–191. Citado na pág. 8

Jiménez(2002) B. Jiménez. Strict efficiency in vector optimization. *Journal of Mathematical Analysis and Applications*, 265:264–284. Citado na pág. 8

Jorion(2003) P. Jorion. Portfolio optimization with tracking error constraints. *Financial Analysis Journal*, 59:70–82. Citado na pág. 49

Kraus e Litzenberger(1976) A. Kraus e R. Litzenberger. Skewness preference and the valuation of risky assets. *Journal of Finance*, 21:1085–1094. Citado na pág. 67

Kuhn e Tucker(1951) H. W. Kuhn e A. W. Tucker. Nonlinear programming. Em *Proceedings of the Second Berkeley Simposium on Mathematical Statistics and Probability*, páginas 481–492, Berkeley, California. Univeristy of California Press. Citado na pág. 3, 6, 7

Lai *et al.***(2006)** K. K. Lai, L. Yu, e S. Wang. Mean-variance-skewness-kurtosis-based portfolio optimization. Em *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences - Volume 2*, IMSCCS '06, páginas 292–297, Washington, DC, USA. IEEE Computer Society. Citado na pág. 67

Lai(1991) T. Y. Lai. Portfolio selection with skewness: A multiple-objective approach. *Review of Quantitative Finance and Accounting*, 1:293–305. Citado na pág. 66

Luc(1989) D. T. Luc. *Theory of Vector Optimization*. Springer, Berlin. Citado na pág. 1, 3, 11

Maciel *et al.***(2009)** M. C. Maciel, S. A. Santos, e G. N. Sottosanto. Regularity conditions in differentiable vector optimization revisited. *Journal of Optimization Theory and Applications*, 142:385–398. Citado na pág. 22

Maeda(1994) T. Maeda. Constraint qualifications in multiobjective optimization problems: Differentiable case. *Journal of Optimization Theory and Applications*, 80:483–500. Citado na pág. 1, 11, 16, 17, 18, 20, 21

Mangasarian(1969) O. L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, New York. Citado na pág. 14, 19, 34

Marglin(1967) S. Marglin. *Public Investment Criteria*. MIT Press, Cambridge, Massachusetts. Citado na pág. 32

Markowitz(1952) H. Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91. Citado na pág. 46

Markowitz(1956) H. Markowitz. The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, 3:111–133. Citado na pág.

Messac *et al.***(2000a)** A. Messac, C. P. Sukam, e E. Melachrinoudis. Aggregate objective functions and pareto frontiers: required relationships and pratical implications. *Optical Engineering*, 1:171–188. Citado na pág. 29

Messac *et al.***(2000b)** A. Messac, G. J. Sundararaj, R. V. Tappeta, e J. E. Renaud. Ability of objective functions to generate points on nonconvex pareto frontiers. *American Institute of Aeronautics and Astronautics Journal*, 38:1084–1091. Citado na pág. 29

Miettinen(1999) K. M. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston. Citado na pág. 1

Pappalardo e Bigi(1999) M. Pappalardo e G. Bigi. Regularity conditions in vector optimization. *Journal of Optimization Theory and Applications*, 102:83–96. Citado na pág. 1, 11, 16, 20

Qi e Wei(2000) L. Qi e Z. Wei. On the constant positive linear dependence condition and its application to sqp methods. *SIAM Journal on Optimization*, 10:963–981. Citado na pág. 22

Rockafellar(1997) R. T. Rockafellar. *Convex Analysis*. Princenton University Press, Princenton, NJ. Citado na pág. 4

Sawaragi *et al.***(1985)** Y. Sawaragi, H. Nakayama, e T. Tanino. *Theory of Multiobjective Optimization*. Academic, Orlando. Citado na pág. 3

van den Berg e Friedlander(2008) E. van den Berg e M. P. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31:890–912. Citado na pág. 43, 46

Wang e Xia(2002) S. Wang e Y. Xia. *Portfolio Selection and Asset Pricing*. Spring, Berlin. Citado na pág. 66

Wolf(1959) P. Wolf. The simplex method for quadratic programming. *Econometrica*, 27:382–398. Citado na pág. 48, 52

Zadeh(1963) L. A. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions On Automatic Control*, AC-8:59–60. Citado na pág. 3, 10, 23, 24

Zeleny(1973) M. Zeleny. Compromise programming. Em J. Cochrane e M. Zeleny, editors, *Multiple Criteria Decision Making*, páginas 262–301. University of South Carolina Press, Columbia. Citado na pág.

Zeleny(1974) M. Zeleny. *Linear Multiobjective Programming*. Springer-Verlag, Berlin and New York. Citado na pág. 23

Zoutendijk(1960) G. Zoutendijk. *Methods of feasible directions*. Elsevier, Amsterdam. Citado na pág. 40

Índice Remissivo

K-esparso, 44

análise
  de contexto, 23
  de decisão multicritério, 1

compressível, 44
compressão, 43
compressed sensing, 2, 43
condição de qualificação
  de Abadie, 15, 52, 58, 70
  de Guignard, 17
  de Kuhn-Tucker, 7
  de Mangasarian-Fromovitz, 19
condição de regularidade, 1, 16
  de Abadie (Maeda), 16
  de Abadie generalizada (Maeda), 17
  de Guignard (Maeda), 17
  de Guignard (Pappalardo e Bigi), 20
  de Mangasarian-Fromovitz (Maeda), 19
  de Mangasarian-Fromovitz (Pappalardo e Bigi), 21
  total, 1, 16
condições
  de Fritz John para POM, 1, 15
  KKT para o problema da ϵ-restrição, 35
  KKT para o problema dos pesos, 28
cone, 4
  convexo, 4
    pontudo, 3–5
  linearizado, 7, 13, 17
    fraco, 13
  pontudo, 4
  pseudotangente, 17
  tangente, 12, 13, 17, 18
crítica, 38, 40

eficiência, 1, 5, 6, 8, 20
  imprópria, 6
  própria, 6–8, 27, 35
eficiente, 5
  fraca
    local, 6
  fracamente, 6
    local, 11, 38, 40, 58
  fronteira, *veja* fronteira eficiente
  impropriamente, 8
  local, 6
  propriamente, 8, 27, 28, 33
    local, 8
erro máximo
  permitido, 31
  possível, 31
escalarização, 23
espaço
  de decisão, 9
  objetivo, 9

fronteira
  de Pareto, 9
  eficiente, 9, 23, 24, 30, 33, 46, 51
    risco-retorno, 50
função multiobjetivo, 3, 5, 9, 38

ineficiente, 5
inferior, 10

método
  da restrição, 32
  das direções viáveis, 40
  do gradiente, 38
  dos pesos, 24
  NISE, 30
métodos de geração, 24
Markowitz, 46, 48, 66
modelo
  da Média-Variância, 48
  da Média-Variância-Assimetria, 66
  de Markowitz, *veja* Markowitz

não-adaptativo, 44
não-dominância, 5
não-inferioridade, 5

ordem
  parcial, 3, 4
    linear, 4, 5
  total, 4
ortante não-negativo, 1, 5

otimização de portfolio, 2, 46, 48, 50, 52, 53

Pareto-otimalidade, 5
ponto minimal, 5
    ideal, 5
portfolio tangencial, 51
pseudonormalidade, 70

razão de Sharpe, 51
regular, 16
    fraco, 16, 29, 36, 37, 41
    totalmente, 16, 21
        fraco, 16, 21
relação binária, 3, 4
    anti-simétrica, 3
    conectada, 4
    linear, 4
    reflexiva, 3
    transitiva, 3
restricted isometry property, 45

Shannon-Nyquist, *veja* teorema de amostragem de Shannon-Nyquist
superior, 10

teorema
    da alternativa de Motzkin, 14, 15, 21
    de amostragem de Shannon-Nyquist, 43
tracking-error, 49
tradeoff, 10