

Este exemplar encaminhado à redação final da tese
defendida por Menotti Erasmo da Silva Machado

aprovada pela Comissão
Julgadora em 13 / 12 / 1991.

Paulo Augusto Valente Ferreira
Orientador

Universidade Estadual de Campinas

Faculdade de Engenharia Elétrica

Departamento de Telemática

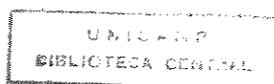
Projeção em Programação Multiobjetivo: Análise Algorítmica e Experiências Numéricas

Autor : Menotti Erasmo da Silva Machado

Orientador : Prof. Dr. Paulo Augusto Valente Ferreira

Dissertação apresentada à Faculdade de Engenharia Elétrica, da Universidade Estadual de Campinas, como parte dos requisitos exigidos para a obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA.

Dezembro de 1991



Aos meus pais Ana e Erasmo. por Tudo.

A Ana Beatriz. ...♡.

Agradecimentos

Ao Prof. Paulo Valente pela elaboração do tema, oportunidade oferecida, confiança depositada, constante incentivo, grande amizade e marcante orientação.

Aos professores Akebo, Basílio e Geromel pelos conhecimentos proporcionados, e a Pedro Peres pelas dicas no $\text{\LaTeX}(\mu)$. A Ana Beatriz pelas correções ortográficas finais. A Adailton, Cláudio, Eliane, ... e depois Plínio que através de gestos simples deram grande mostra de sua amizade. Aos colegas da FEE, com os quais compartilhei a alegria das festas, a ousadia das lutas e o prazer das conquistas, em especial à "galera" da Telemática pela importante troca de informações e pelo agradável ambiente de trabalho. Aos meus irmãos Almiro e Viviane que tanto torceram por mim.

Aos professores e funcionários da FEE e às pessoas que de alguma forma contribuíram para a elaboração deste trabalho.

Ao CNPq pelo suporte financeiro.

Resumo

O trabalho apresenta as bases teóricas de uma abordagem interativa recentemente proposta para tratar problemas de decisão envolvendo múltiplos objetivos. Como resultado deste desenvolvimento, obtem-se um *Algoritmo Básico*, composto de um nível de análise e de um nível de decisão, ao qual várias técnicas específicas podem ser agregadas. O trabalho discute a implementação do nível de decisão, inteiramente formulado no espaço dos objetivos, através de métodos multiobjetivos de reconhecida eficiência, incluindo-se métodos baseados em Teoria de Conjuntos Nebulosos. As implementações realizadas com base em um problema de controle de poluição em rios evidenciam uma série de simplificações introduzidas pela abordagem, e permitem a síntese de algoritmos eficientes tanto do ponto de vista computacional quanto do ponto de vista dos processos de decisão.

Abstract

The work presents the theoretical basis of a recently proposed interactive approach for solving decision problems with multiple objectives. The development originates a *Basic Algorithm* composed of an analysis level and of a decision level, in which a number techniques can be associated with. The work discusses the implementation of the decision level, entirely formulated in the objective space, through well-known multiobjective methods, including those based on Fuzzy Sets Theory. The implementations carried out on the basis in a water pollution control problem highlight a number of simplifications introduced by the approach, allowing the synthesis of very efficient algorithms both from the computational and the decision making point of views.

Simbologia

CONJUNTOS

\mathcal{R}	- conjunto de todos os números reais
\mathcal{R}^n	- conjunto de todos os vetores reais de dimensão n
$\mathcal{R}^n \times \mathcal{R}^n$	- produto cartesiano
$\mathcal{R}^{n \times m}$	- conjunto de todas matrizes de dimensão $n \times m$
\mathcal{A}	- espaço das decisões
\mathcal{F}	- espaço dos objetivos
$\mathcal{Y}^n \subset \mathcal{R}^n$	- subconjunto do \mathcal{R}^n
\tilde{A}	- subconjunto nebuloso
$\neg A$	- complemento do subconjunto A
$y \in \mathcal{Y}$	- elemento de \mathcal{Y}

VETORES E ESCALARES

y'	- transposição de y
$\langle y, x \rangle = y'x$	- produto escalar entre y e x
$\ y\ = \langle y, y \rangle^{1/2}$	- norma euclidiana de y
$\max_{1 \leq i \leq n} \{y_i\}$	- o maior valor real do conjunto $\{y_1 \dots y_n\}$
$ a $	- valor absoluto de a
\underline{y}	- limite inferior de y .
\bar{y}	- limite superior de y .
y^*	- valor ótimo de y
$\ y\ _\infty = \max_{1 \leq i \leq n} \{\alpha_i, y_i \}$	- norma Tchebyshev de y
$\dim(y)$	- dimensão de y

MATRIZES

I_r	- matriz Identidade de dimensão r
$S = \{s_{ij}\}$	- matriz de elementos s_{ij}
S'	- matriz transposta de S
$S \geq 0$	- matriz simétrica semi-definida positiva

Conteúdo

Introdução Geral	1
1 Otimização Multiobjetivo : Aspectos Básicos	5
1.1 Introdução	5
1.2 Definições e Teoremas Básicos	5
1.3 Geração de Soluções Eficientes - Técnicas Escalares	7
1.3.1 Método da Ponderação - P_w	7
1.3.2 Método das ϵ -Restrições - P_ϵ	7
1.4 Condições de Existência de Soluções Eficientes	8
1.5 Classificação dos Métodos Multiobjetivos	9
1.5.1 Métodos com Indicação À-Priori de Preferências	11
1.5.2 Métodos com Indicação Progressiva de Preferências	14
1.6 Conclusão	23
2 Otimização no Espaço dos Objetivos	24
2.1 Introdução	24
2.2 Descrição da Abordagem	25
2.3 Análise do Caso Bi-objetivo	29
2.4 Análise do Caso Multiobjetivo	30
2.5 Condições Fracas para \mathcal{V} Convexo	33
2.6 Solução do Problema Min-Max	35
2.7 Conclusão	40

3	O Problema de Poluição em Rios	41
3.1	Introdução	41
3.2	O Modelamento Físico	42
3.3	Modelo de Otimização	45
3.4	Modelo Multiobjetivo	49
3.5	Conclusão	52
4	Implementação do Método de Geoffrion no Espaço dos Ob- jetivos	54
4.1	Introdução	54
4.2	Descrição do Método - Nível de Decisão	54
4.3	Estimativa de TMS's através do Método de Dyer	57
4.4	Aspectos de Implementação do Método	60
4.5	Desempenho Numérico do Método	63
4.6	Comentários	68
4.7	Conclusão	69
5	Algoritmos Baseados em Conjuntos Nebulosos	70
5.1	Introdução	70
5.2	Conjuntos Nebulosos	71
5.3	Algoritmos Nebulosos	73
5.4	Algoritmos Nebulosos em Programação Matemática	74
5.5	Algoritmos de Baptistella e Ollero	76
5.6	Implementação no Espaço dos Objetivos	80
5.7	Redefinição do Problema Escalar	85
5.8	Aspectos de Implementação	87
5.9	Resultados Numéricos e Comentários	94
5.10	Conclusão	97
	Conclusão Geral	99
	Referências Bibliográficas	101

A Conjuntos Nebulosos	105
A.1 Operações Básicas	105
A.2 Termos Linguísticos	113
A.3 Delimitadores Linguísticos	114
A.4 Algoritmos Nebulosos	118

Introdução Geral

Nas últimas décadas a Programação Multiobjetivo vem se desenvolvendo rapidamente e atualmente representa uma das principais linhas de abordagem de processos de decisão. Com este enfoque surgiram vários métodos multiobjetivos, sendo que os mais eficientes apresentam algum tipo de procedimento interativo.

A característica fundamental dos métodos multiobjetivos é operar sobre conjuntos parcialmente ordenados [Zadeh, 1963], o que exclui em geral, a obtenção de soluções ótimas no sentido usual adotado em programação matemática. Isto se deve ao fato de problemas multiobjetivos geralmente apresentarem conflitos e incomensurabilidades entre os objetivos. Nestes casos, o conceito de solução ótima é substituído pelo conceito de *solução de compromisso*.

Visando a obtenção da solução de compromisso, os métodos multiobjetivos agregam aspectos relacionados às técnicas de otimização, o que se convencionou chamar de *Análise*, com aspectos essencialmente subjetivos, que correspondem ao nível de *Decisão* do método. A introdução de subjetividade é necessária para estabelecer uma relação de ordem entre as alternativas de solução do problema.

Uma classificação dos métodos multiobjetivos quanto a forma de participação do Decisor no processo de tomada de decisão pode ser encontrada em [Hwang e Masud, 1979].

Atualmente a programação multiobjetivo encontra aplicação em diversas áreas tais como, planejamento de produção de sistemas de manufatura, planejamento energético, gestão e alocação de recursos hidráulicos, planejamento acadêmico, planejamento de transportes e gerenciamento de tráfego, administração pública em geral, etc [Hwang e Masud, 1979].

A partir de meados da década de 70, a Teoria de Conjuntos Nebulosos [Zadeh, 1965] começou a ser aplicada em programação multiobjetivo, facilitando em muito o desenvolvimento de esquemas interativos, pois do ponto de vista do Decisor, o fornecimento de informações na forma linguística é mais simples e natural do que na forma numérica.

Neste trabalho utiliza-se um método multiobjetivo interativo recentemente proposto [Ferreira e Geromel, 1990], através do qual o processo de tomada de decisão se desenvolve inteiramente no espaço dos objetivos com substanciais vantagens tanto do ponto de vista de análise quanto de decisão.

De acordo com esta abordagem, pretende-se enfocar o nível de decisão e para tanto serão implementados os métodos de Geoffrion et al (1972) e Baptistella e Ollero (1980), inteiramente no espaço dos objetivos. Estes procedimentos serão aplicados a um clássico problema de controle de poluição em rios. Fazendo-se uso de propriedades importantes derivadas do emprego deste tipo de abordagem tais como linearidade, robustez, convergência infinita e homogeneidade, será possível abordar o problema multiobjetivo através de implementações mais elaboradas sem o acréscimo de complexidade para sua resolução e com simplificações relevantes.

Este desenvolvimento encontra-se organizado em 5 Capítulos como descrito a seguir:

Capítulo 1 - *Otimização Multiobjetivo : Aspectos Básicos*

Neste capítulo apresentam-se as definições básicas necessárias para a formulação e resolução de problemas multiobjetivos. Dentro da filosofia de decomposição de problemas multiobjetivos em uma estrutura de dois níveis, abordam-se alguns métodos importantes, dando-se ênfase àqueles essenciais para o desenvolvimento deste trabalho. Também é apresentada uma classificação destes métodos quanto a forma de participação do Decisor na resolução de problemas multiobjetivos.

Capítulo 2 - *Otimização no Espaço dos Objetivos*

Discutem-se neste Capítulo, aspectos relevantes para a formulação de problemas multicritérios no Espaço dos Objetivos. Apresentam-se as vantagens do emprego deste tipo de formulação quando aplicada a problemas

bi-critérios. Em seguida analisa-se o caso multicritério e comenta-se as vantagens proporcionadas pelo método em função de importantes propriedades derivadas. Também é feito um estudo sobre condições fracas de convexidade que possibilitam a convergência do *Algoritmo Básico* resultante deste tipo de formulação. Apresentam-se técnicas adequadas para se solucionar o problema Min-Max do *Algoritmo Básico*.

Capítulo 3 - O Problema de Poluição em Rios

Neste Capítulo apresenta-se um problema clássico de Planejamento de Recursos Hidráulicos, ao qual são aplicados os métodos de Geoffrion et al (1972) e Baptistella e Ollero (1980) descritos no Espaço dos Objetivos . Apresenta-se o modelo físico do problema de Controle de Poluição em Rios que a algum tempo vem sendo utilizado por diversos estudiosos com a finalidade de testar a eficiência de vários métodos [Haimes et al., 1975]. Apresenta-se a formulação via técnicas de otimização escalar, e também o modelo multiobjetivo utilizado neste trabalho e em [Haimes et al., 1975] cujos resultados servem como base para comparação.

Capítulo 4 - Implementação do Método de Geoffrion no Espaço dos Objetivos

Neste Capítulo discute-se a implementação do método de Geoffrion no Espaço dos Objetivos [Ferreira e Machado, 1990b]. [Ferreira e Machado, 1990c]. Importantes simplificações em relação ao procedimento original são testadas. Para a estimação das *Taxas Marginais de Substituição* utiliza-se o método de Dyer (1973), o qual baseia-se em comparações ordinais. São apresentados aspectos de implementação do método tanto do ponto de vista de análise quanto de decisão. Os resultados obtidos a partir da aplicação deste método ao problema de controle de poluição em rios, referido no Capítulo 3, possibilitam fazer uma análise de seu desempenho numérico.

Desta forma, é possível fazer comparações entre as versões implementadas (automática e interativa) e os resultados obtidos na literatura.

Capítulo 5 - Algoritmos Baseados em Conjuntos Nebulosos

Neste Capítulo discute-se a agregação de algoritmos baseados em Teoria de Conjuntos e Algoritmos Nebulosos ao Algoritmo Básico. Comenta-se a aplicação desta teoria em Programação Matemática, referenciando-se alguns importantes trabalhos encontrados na literatura. Enfatizam-se os algoritmos nebulosos propostos por Baptistella e Ollero, dos quais dois são implementados neste trabalho. Propõe-se a redefinição do problema escalar associado aos algoritmos nebulosos em termos de *Norma Infinito Generalizada*. São discutidos aspectos de implementação dos algoritmos tais como, construção dos termos linguísticos nebulosos e dos protocolos. Comentam-se as dificuldades encontradas para a definição precisa dos protocolos e dos universos de discurso e apresenta-se os resultados obtidos para a execução dos algoritmos nebulosos não-interativo e interativo, aplicados ao problema de controle de poluição em rios.

Apêndice A - Conjuntos e Algoritmos Nebulosos

Apresentam-se as operações básicas sobre conjuntos nebulosos e os conceitos necessários para a implementação de algoritmos baseados nesta teoria [Zadeh, 1965]. [Zadeh, 1973]. [Ferreira e Machado, 1990a].

Capítulo 1

Otimização Multiobjetivo : Aspectos Básicos

1.1 Introdução

A Programação Multiobjetivo apresenta atualmente uma ampla variedade de resultados e métodos destinados à resolução de problemas que envolvam múltiplos objetivos. Entre os métodos para a resolução de PMO's, os que apresentam resultados mais satisfatórios são aqueles que decompõem o problema em dois níveis: *Análise* e *Decisão*. O *Nível de Análise* ou *Inferior* está relacionado com os aspectos numéricos necessários para se obter uma solução ou um conjunto de soluções para o problema. O *Nível de Decisão* ou *Superior* está relacionado com os aspectos subjetivos do problema. Neste nível, um *Decisor* concentra-se na tarefa de induzir o nível de análise a gerar soluções que venham ao encontro de sua estrutura de preferências.

Neste capítulo apresentam-se os conceitos e definições básicas ligadas a estes dois níveis, juntamente com uma breve classificação de métodos multiobjetivos baseada na forma de participação do Decisor na resolução do problema.

1.2 Definições e Teoremas Básicos

A Programação Multiobjetivo está relacionada com o estudo de proble-

mas de otimização vetorial, isto é, que envolvem mais de uma função objetivo. No caso de minimização, a formulação geral desta classe de problemas é como segue:

$$\min_{x \in \mathcal{X}} f(x) \quad (1.1)$$

onde

$$f(x) = (f_1(x), f_2(x), \dots, f_m(x)) \quad (1.2)$$

$\mathcal{X} \subseteq R^n$ representa o *Espaço de Decisões* e $f_i(\cdot) : \mathcal{X} \rightarrow R$, $i = 1, \dots, m$. Como se pode observar, a aplicação $f(\cdot) : \mathcal{X} \rightarrow R^m$, indica que a característica fundamental do problema é o ordenamento parcial das soluções. Por outro lado, este ordenamento parcial é utilizado para caracterizar as possíveis soluções do problema.

Definição 1.1 Solução Eficiente

Uma solução factível $x^ \in \mathcal{X}$ é eficiente se não existe qualquer outra solução $x \in \mathcal{X}$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$. O conjunto de todas as soluções eficientes de (1.1) será representado por \mathcal{X}^* .*

A consequência de cada decisão factível $x \in \mathcal{X}$ em termos de função objetivo é $f(x) \in R^m$. O mapeamento completo de \mathcal{X} no R^m é representado pelo conjunto

$$\mathcal{F} = f[\mathcal{X}] \triangleq \{y \in R^m : y = f(x), x \in \mathcal{X}\}. \quad (1.3)$$

que será denominado neste trabalho como *Espaço dos Objetivos*. Analogamente, o conjunto de soluções eficientes no espaço dos objetivos será representado por $\Gamma = f[\mathcal{X}^*]$.

A caracterização analítica de soluções eficientes pode ser conseguida através da solução de problemas de otimização escalares.

1.3 Geração de Soluções Eficientes - Técnicas Escalares

Duas formas básicas de se gerar soluções eficientes através da escalari-zação do problema são discutidas nesta seção.

1.3.1 Método da Ponderação - P_w

As soluções eficientes são encontradas resolvendo-se parametricamente

$$\min_{x \in \mathcal{X}} \sum_{i=1}^m w_i f_i(x) \quad (P_w)$$

onde,

$$w \in \mathcal{W} \triangleq \{w \in R^m : w_i \geq 0 \text{ e } \sum_{i=1}^m w_i = 1\} \quad (1.4)$$

Os pesos w geralmente não refletem a importância relativa dos objetivos no sentido proporcional, sendo apenas parâmetros para a geração de soluções eficientes.

1.3.2 Método das ϵ -Restrições - P_ϵ

Neste método, a solução eficiente é gerada a partir do seguinte problema.

$$\begin{aligned} \min_{x \in \mathcal{X}} f_i(x) & \quad (P_\epsilon) \\ \text{s.a.} \quad f_j(x) & \leq \epsilon_j, \quad j = 1, \dots, m, \quad (i \neq j). \end{aligned} \quad (1.5)$$

O conjunto de soluções eficientes do problema é gerado a partir da variação dos ϵ_j 's. $j = 1, \dots, m$. ($j \neq i$).

Para um dado ponto $x^* \in \mathcal{X}$ define-se P_{ϵ^*} como o problema P_ϵ onde $\epsilon_j = \epsilon_j^* = f_j(x^*)$, $j \neq i = 1, \dots, m$.

Além destas duas formas apresentadas existe uma terceira, *Método do Lagrangeano* [Chankong e Haimes, 1982], que não será apresentada por não ser utilizada neste trabalho.

1.4 Condições de Existência de Soluções Eficientes

As condições necessárias e suficientes para a existência de soluções eficientes são análogas às clássicas *Condições de Kuhn-Tucker* (KT) de otimalidade para um problema de otimização escalar [Chankong e Haimes, 1982]. Assumindo-se que o espaço de decisões é representado por $\mathcal{X} = \{x \in R^n : g_i(x) \leq 0, i = 1, \dots, p\}$, estas condições podem ser estabelecidas como segue.

Definição 1.2 *Condições de Kuhn-Tucker*

Diz-se que x^ satisfaz as Condições de Kuhn-Tucker para Eficiência (KTE), se:*

1. todos os f_j 's e g_i 's são diferenciáveis;
2. existem multiplicadores $\lambda_j \geq 0, j = 1, \dots, m$, com desigualdade estrita valendo para no mínimo um j , e $\mu_i \geq 0, i = 1, \dots, p$, tais que

$$g_i(x^*) \leq 0, \quad \mu_i g_i(x^*) = 0, \quad i = 1, \dots, p \quad (1.6)$$

e

$$\sum_{j=1}^m \lambda_j \nabla f_j(x^*) + \sum_{i=1}^p \mu_i \nabla g_i(x^*) = 0 \quad (1.7)$$

Observando-se a Definição 1.2, pode-se notar que as equações (1.6) e (1.7) são quase equivalentes às condições de Kuhn-Tucker para a otimalidade de P_{ϵ^*} .

O *Diagrama de Implicações* (Fig. 1.1) apresentado a seguir resume os principais resultados e relações entre cada método de escalarização e as condições de Kuhn-Tucker para a eficiência [Chankong e Haimes, 1982]. Os

símbolos seguintes representam hipóteses e convenções adotadas neste Diagrama.

Hipóteses

- C Todas as funções envolvidas são convexas sobre \mathcal{X} convexo;
- D Todas as funções envolvidas são diferenciáveis;
- R x^* é um ponto regular das restrições. isto é, os gradientes de todas as restrições ativas em $x^* \in \mathcal{X}$ são linearmente independentes;
- U x^* é solução única do problema indicado;
- P Os pesos do problema ponderado são estritamente positivos;
- T x^* é a solução de $P_k(\epsilon^*)$ para todo k .

Convenções

- A $\xrightarrow{\alpha}$ B Aplicando-se a hipótese α , então A implica B.
- A $\xrightarrow{\alpha/\beta}$ B Aplicando-se as hipóteses α ou β , então A implica B.
- A $\xrightarrow{(\alpha)}$ B A implica α e B.

Além dos aspectos de geração de soluções eficientes, os métodos multiobjetivos agregam aspectos relacionados à tomada de decisão.

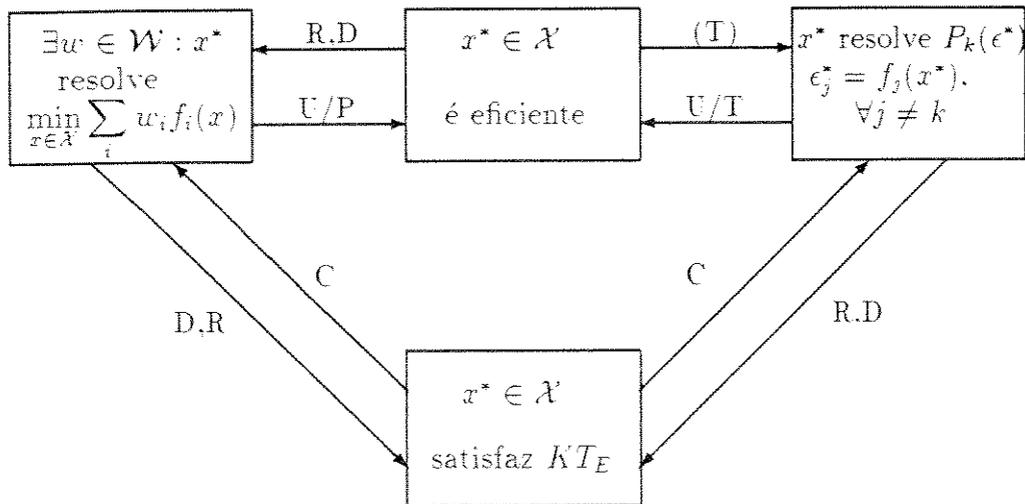


Figura 1.1: Diagrama de Implicações

1.5 Classificação dos Métodos Multiobjetivos

Os métodos multiobjetivos são geralmente classificados em termos da forma de participação do Decisor na escolha da solução preferida.

Para o entendimento de como as estruturas de preferências levam à solução do problema, apresenta-se a seguir uma breve descrição de alguns destes métodos.

1.5.1 Métodos com Indicação À-Priori de Preferências

Nessa classe de métodos, o Decisor fornece informações sobre suas preferências a um analista, o qual é responsável pela resolução do PMO, antes que o problema seja efetivamente resolvido. A solução preferida será a solução obtida pelo analista.

Abordagem Via Função Utilidade

Procura-se modelar matematicamente as atitudes do Decisor objetivando inferir uma aproximação de sua *Função Utilidade*.

Definição 1.3 *Função Utilidade*

Uma função $\mathcal{U}(\cdot)$ que associa um número real $\mathcal{U}(f)$ a cada ponto $f \in \mathcal{F}$, é denominada Função Utilidade representativa das preferências do Decisor.

Este tipo de método converte o PMO em um problema do tipo.

$$\min_{x \in \mathcal{X}} \mathcal{U}(f(x)), \quad (1.8)$$

onde $f(\cdot) : \mathcal{X} \rightarrow \mathcal{F}$ e $\mathcal{U}(\cdot) : \mathcal{F} \rightarrow R$ é a Função Utilidade ou Valor que agrega as preferências do Decisor em relação aos m objetivos. Assume-se que $\mathcal{U}(\cdot)$ é uma função não-decrescente em relação a cada objetivo. Com esta hipótese é fácil mostrar que no mínimo uma solução de (1.8) é eficiente [Geoffrion, 1967]. A estruturação das preferências do Decisor pode ser obtida através das curvas

de nível da sua função utilidade, as quais são também conhecidas como *curvas de indiferença ou isopreferência*. A solução do problema (1.8) será um ponto onde o conjunto das soluções eficientes for tangente às curvas de indiferença.

Suponha que x^* resolve $\min_{x \in \mathcal{X}} \mathcal{U}(f(x))$. Então sob hipóteses de convexidade, existe $w^* \in \mathcal{W}$ tal que $x^* = \arg \min_{x \in \mathcal{X}} \langle w^*, f(x) \rangle$. O método apresentado no Capítulo 2 explora esta propriedade buscando determinar w^* através de técnicas de decomposição [Ferreira e Geromel, 1990].

Método do Ordenamento Lexicográfico

Neste método, o Decisor ordena seus objetivos segundo suas prioridades. Seja $a = (a_1(x), \dots, a_m(x))$, o conjunto dos objetivos ordenados pelo Decisor, onde $a_i(x) = f_j(x)$, $i, j = 1 \dots m$. O problema fica formulado da seguinte maneira:

$$\min_{x \in \mathcal{X}_{i-1}} a_i(x), \quad (1.9)$$

onde

$$\mathcal{X}_i \triangleq \{x \in \mathcal{X}_{i-1} : x = \arg \min_{x \in \mathcal{X}_{i-1}} a_i(x)\}, \quad \mathcal{X}_0 = \mathcal{X} \quad (1.10)$$

Para se chegar a solução final, resolve-se (1.9) sequencialmente para $i = 1 \dots m$.

Este método apresenta a desvantagem da solução preferida x^* ser muito sensível ao ordenamento, o que é indesejável se os objetivos tiverem aproximadamente a mesma importância. Com o propósito de reduzir esta sensibilidade Waltz (1967) propôs que os critérios pudessem variar dentro de certas faixas percentuais do valor ótimo a_l^* encontrado nos passos anteriores. Deste modo, o problema (1.9) é reformulado como,

$$\min_{x \in \mathcal{X}} a_i(x) \quad (1.11)$$

$$\text{s.a. } a_l(x) \leq a_l^* + \delta_l, \quad l = 1, \dots, i-1.$$

onde $\delta_l > 0$, $l = 1, \dots, i-1$ são tolerâncias determinadas pelo Decisor.

Programação Alvo

A *Programação Alvo* (PA) foi inicialmente proposta para resolver problemas lineares [Charnes e Cooper, 1961], e alcançou um grande desenvolvimento nas últimas décadas.

A formulação do problema (1.1) através da Programação Alvo é apresentada a seguir:

$$\min_{x \in X} \sum_{i=1}^m [(d_i^- + d_i^+)^p]^{\frac{1}{p}}, \quad p \geq 1 \quad (1.12)$$

$$\text{s.a. } f_i(x) + d_i^- - d_i^+ = b_i, \quad i = 1, \dots, m \quad (1.13)$$

$$d_i^-, d_i^+ \geq 0 \quad (1.14)$$

$$d_i^- \cdot d_i^+ = 0, \quad (1.15)$$

onde b_i , $i = 1, \dots, m$ são as metas estabelecidas pelo Decisor para os objetivos $f_i(x)$, $i = 1, \dots, m$. d_i^+ representa quanto o objetivo $f_i(x)$ excedeu a meta estipulada b_i , e d_i^- representa quanto o objetivo $f_i(x)$ ficou abaixo da meta estipulada b_i . O valor de p depende da forma com que os desvios devem ser ponderados. Em geral, $p = 1$.

A solução preferida é definida como a que apresenta o *menor desvio total* em relação aos alvos ou metas desejados.

Uma variante para a formulação de PA pode ser conseguida incorporando-se ordenamento lexicográfico aos desvios. Assim, a formulação de PA para o problema (1.1) passa a ser

$$\min_{\substack{x \in X \\ d^+, d^-}} (a_1(d^+, d^-), \dots, a_m(d^+, d^-)) \quad (1.16)$$

$$\text{s.a.} \quad (1.13) - (1.15),$$

cujas soluções são conseguidas resolvendo-se inicialmente

$$\min_{\substack{x \in X \\ d^+, d^-}} a_1(d^+, d^-) \quad (1.17)$$

$$\text{s.a.} \quad (1.13) - (1.15),$$

obtendo-se a_i^* . Em seguida resolve-se para $i = 2, \dots, m$,

$$\min_{\substack{x \in X \\ d^+, d^- \geq 0}} a_i(d^+, d^-) \quad (1.18)$$

$$\begin{aligned} \text{s.a.} \quad & (1.13) - (1.15), \\ & a_j(d^+, d^-) \leq a_j^*, \quad j = 1, \dots, i. \end{aligned} \quad (1.19)$$

A grande vantagem de Programação Alvo é que o Decisor não precisa atribuir pesos numéricos aos objetivos. Por outro lado, apresenta como desvantagem sensibilidade à ordenação dos desvios e às metas estipuladas pelo decisor.

Uma outra variante da formulação de PA é o *Método da Realização das Metas* [Gembicki e Haimes, 1975]. O problema a ser resolvido é o seguinte

$$\min_{x \in X} \sigma \quad (1.20)$$

$$\text{s.a.} \quad f_i(x) - \sigma w_i \leq b_i, \quad i = 1, \dots, m,$$

onde $w \in W$, σ é um escalar e b_i , $i = 1, \dots, m$ são metas propostas pelo Decisor.

1.5.2 Métodos com Indicação Progressiva de Preferências

Esta classe de métodos, geralmente referida como de métodos interativos, tem tido nos últimos anos uma grande aceitação por parte dos pesquisadores e usuários. A maioria destes métodos requer que o Decisor forneça *trade-off's locais* (implícitos, como níveis de aspiração ou explícitos como taxas marginais de substituição) na vizinhança de uma alternativa viável. Estes

trade-off's assumem então a função de coeficientes variáveis para algum problema de otimização escalar. Apresenta-se a seguir alguns métodos desta classe.

Programação Alvo Interativa

Este método é uma extensão da Programação Alvo com estratégias interativas usando o algoritmo de Frank-Wolfe [Dyer, 1972].

Sob a hipótese de não decrescimento das preferências do Decisor em relação a cada objetivo, pode-se reformular o problema (1.16) como,

$$\begin{aligned} \min_{x \in X} \quad & (a_1(d^-), \dots, a_m(d^-)) & (1.21) \\ \text{s.a.} \quad & f_i(x) + d_i^- - d_i^+ = b_i, \quad i = 1, \dots, m \\ & d_i^-, d_i^+ \geq 0, \quad i = 1, \dots, m \\ & d_i^-, d_i^+ = 0, \quad i = 1, \dots, m \end{aligned}$$

Substituindo-se os níveis de prioridades por pesos relativos à importância de cada objetivo, pode-se reescrever (1.21) como

$$\begin{aligned} \min_{x \in X} \quad & \sum_{i=1}^m w_i d_i^- & (1.22) \\ \text{s.a.} \quad & f_i(x) + d_i^- - d_i^+ = b_i, \quad i = 1, \dots, m \\ & d_i^-, d_i^+ \geq 0, \quad i = 1, \dots, m \\ & d_i^-, d_i^+ = 0, \quad i = 1, \dots, m. \end{aligned}$$

onde w_i é o peso para o i -ésimo objetivo, cuja determinação será apresentada posteriormente quando apresentado o método de Geoffrion et al (1972).

Método SWT ("Surrogate Worth Trade-off Method")

Para a resolução de PMO's, este método emprega duas fases [Haimes et al., 1975]: identificação e geração de soluções eficientes, e escolha da solução preferida entre as soluções eficientes geradas.

Os trade-off's podem ser encontrados através dos valores das variáveis duais associados às restrições do problema (P_ϵ) , onde ϵ é definido como

$$\begin{aligned} \epsilon_j &= \underline{f}_j + \delta_j, \quad j = 1, \dots, m \quad (j \neq i), \quad e \\ \delta_j &> 0, \quad j = 1, \dots, m \quad (j \neq i) \end{aligned} \quad (1.23)$$

onde \underline{f}_j , $j = 1, \dots, m$, ($j \neq i$) representam as soluções ideais dos problemas,

$$\min_{x \in X} f_j(x), \quad j = 1, \dots, m \quad (j \neq i) \quad (1.24)$$

e δ_j são os desvios dos valores ideais dos objetivos, que tornam o problema ϵ -restrito factível.

O Lagrangeano associado a (P_ϵ) é

$$\mathcal{L}(x, \mu, \lambda) = f_i(x) + \sum_{i=1}^p \mu_i g_i(x) + \sum_{\substack{j=1 \\ j \neq i}}^m \lambda_{ij} (f_j(x) - \epsilon_j), \quad (1.25)$$

onde μ_i , $i = 1, \dots, p \in \lambda_{ij}$, $i, j = 1, \dots, m$ ($j \neq i$) são multiplicadores de Kuhn-Tucker.

De acordo com as condições de Kuhn-Tucker e supondo-se não degenerescência [Luenberger, 1970], os λ_{ij} 's correspondentes às restrições ativas indicam o custo marginal do objetivo f_i devido a uma variação de uma unidade em ϵ_j .

Se (x^*, μ^*, λ^*) minimiza (1.25) então

$$\lambda_{ij}^* = -\frac{\partial f_i}{\partial \epsilon_j}$$

e supondo-se que as ϵ -restrições estão ativas ($f_i(x^*) = \epsilon_j$)

$$\lambda_{ij}^* = -\frac{\partial f_i}{\partial f_j}.$$

A *Função de Valor Substituto* é definida da seguinte maneira:

$$\tau_{ij}(f_i^*) = \begin{cases} > 0, & \text{se } \lambda_{ij} \text{ unidades de } f_i \text{ são preferíveis a uma unidade de } f_j \\ < 0, & \text{se } \lambda_{ij} \text{ unidades de } f_i \text{ não são preferíveis a uma unidade} \\ & \text{de } f_j \\ = 0, & \text{se o Decisor é indiferente.} \end{cases}$$

O método então busca determinar f_j^* 's tais que $\tau_{ij}(f_j^*) = 0, \forall j \neq i$. Para tanto, utiliza-se um procedimento de interpolação baseado em avaliações do Decisor sobre as taxas λ_{ij} e os valores correspondentes dos objetivos $f_i^*, \forall j \neq i$. Uma vez determinados $f_j^*, \forall j \neq i$ tais que $\tau_{ij}(f_j^*) = 0$ resolve-se (P_i) com $e_j = f_j^*, \forall j \neq i$ para se obter a solução correspondente no espaço de decisões.

Este método apresenta como vantagens, a aplicação tanto a problemas estáticos quanto a problemas dinâmicos e a possibilidade de considerar preferências de vários Decisores. Como desvantagens, tem-se que o número de comparações duas a duas entre os objetivos pode se tornar excessivo se m é grande, assim como o número de soluções eficientes necessárias para encontrar f_j^* 's tais que $\tau_{ij}(f_j^*) = 0, \forall j \neq i$. A escolha arbitrária inicial dos e_j 's pode ainda excluir alguma solução eficiente que poderia ser uma solução preferida do Decisor.

Método STEM ("Step Method")

Este método foi proposto para a resolução de problemas multiobjetivos lineares [Benayoun et al., 1971], e permite ao Decisor reconhecer boas soluções de compromisso e a importância relativa dos objetivos. Esta técnica baseia-se na premissa de que a melhor solução de compromisso tem um desvio mínimo em relação a solução ideal.

A formulação do problema para uma iteração genérica l é dada a seguir:

$$\min_{\sigma, x \in \mathcal{X}^l} \sigma \quad (1.26)$$

$$\text{s.a. } \pi_i(f_i(x) - \underline{f}_i) \leq \sigma, \quad i = 1, \dots, m, \quad (1.27)$$

onde \mathcal{X}^l inclui $\mathcal{X} \triangleq \{x \in R^n : x \geq 0, Ax \leq b\}$, $A \in R^{p \times n}$ ($p \leq n$), $b \in R^p$ e $f_i(x) = \langle c^i, x \rangle$, $c_i \in R^n$, $i = 1, \dots, m$. O valor de π_i reflete a importância relativa da distância de $f_i(\cdot)$ à solução ideal \underline{f}_i , e pode ser calculado através da expressão,

$$\pi_i = \frac{\alpha_i}{\sum_{j=1}^m \alpha_j}, \quad i = 1, \dots, m. \quad (1.28)$$

onde

$$\alpha_i = \frac{\bar{f}_i - \underline{f}_i}{\underline{f}_i \|c^i\|}, \quad \text{se } \underline{f}_i > 0 \quad (1.29)$$

$$\alpha_i = \frac{\underline{f}_i - \bar{f}_i}{\underline{f}_i \|c^i\|}, \quad \text{se } \underline{f}_i \leq 0 \quad (1.30)$$

e \bar{f}_i é o valor pessimista de f_i , isto é,

$$\bar{f}_i = \max_{1 \leq j \leq m} \{f_i(x^j)\}$$

onde $x^j = \arg \min_{x \in \mathcal{X}} f_j(x)$, $j = 1, \dots, m$.

Pode-se notar que o peso π_i dado a cada diferença depende basicamente da variação $(\bar{f}_i - \underline{f}_i)$ de cada objetivo. Se esta variação for pequena, π_i também será e o objetivo $f_i(\cdot)$ não será muito sensível às variações de π_i .

A solução x^l de (1.26) é apresentada ao Decisor que compara subjetivamente $f_i(x^l)$ com \underline{f}_i , e classifica seu valor como satisfatório ou não.

Para algum i tal que f_i^l é satisfatório, o Decisor deverá fornecer uma quantidade $\delta_i > 0$ a qual está disposto a sacrificar, em troca de uma melhora em algum objetivo não satisfatório.

Deste modo, o conjunto de restrições para a iteração $l + 1$ fica sendo.

$$\mathcal{X}^{l+1} = \{x \in \mathcal{X}^l : f_j(x) \leq f_j^l + \delta_j; f_i(x) \leq f_i^l, \forall i \neq j\} \quad (1.31)$$

Faz-se $\pi_i = 0$, $l = l + 1$ e recalcula-se (1.26)-(1.27). A solução final é encontrada quando todos os objetivos forem satisfatórios.

A grande vantagem deste método é que a parte de cálculos pode ser desempenhada através de uma rotina de PL, um *simplex multiobjetivo* por exemplo, o qual permite otimizações sucessivas de várias funções objetivos.

Método de Geoffrion, Dyer e Feinberg

Este método será abordado em maiores detalhes pelo fato de ser utilizado no escopo deste trabalho. O estudo desta técnica demonstra que um algoritmo de gradiente pode ser utilizado para resolver um PMO se o Decisor for capaz de fornecer avaliações locais sobre sua função utilidade $\mathcal{U}(\cdot)$ [Geoffrion, 1967], [Geoffrion et al. 1972]. O método assume que o problema está formulado como segue:

$$\min_{x \in \mathcal{X}} \mathcal{U}(f_1(x), \dots, f_m(x)). \quad (1.32)$$

Para permitir a aplicação do algoritmo de Frank-Wolfe à resolução do problema assume-se que a função $\mathcal{U}(\cdot)$ e os objetivos $f_i(\cdot)$, $i = 1, \dots, m$ são diferenciáveis e convexos sobre um subconjunto convexo $\mathcal{X} \subset R^n$, com $\mathcal{U}(\cdot)$ não-decrescente com cada objetivo. Os passos deste algoritmo são apresentados a seguir.

Passo 1 : Determine uma solução inicial factível $x^0 \in \mathcal{X}$ e faça $k=0$.

Passo 2 : Obtenha uma solução ótima ξ^k para o problema de otimização

$$\min_{\xi \in \mathcal{X}} \langle \nabla_x \mathcal{U}(f(x^k)), \xi \rangle \quad (1.33)$$

Faça $d^k = \xi^k - x^k$.

Passo 3 : *Determine a solução ótima do problema de busca unidimensional*

$$\min_{t \in [0,1]} \mathcal{U}(f(x^k + td^k)) \quad (1.34)$$

Passo 4 : *Faça $x^{k+1} = x^k + t^k d^k$. Se $\|x^{k+1} - x^k\| < \epsilon$, para $\epsilon > 0$ arbitrariamente escolhido, pare: a solução x^{k+1} resolve (1.32). Caso contrário faça $k=k+1$ e retorne ao Passo 2.*

Devido ao fato da função utilidade $\mathcal{U}(\cdot)$ geralmente não ser explicitamente conhecida, requer-se o auxílio do Decisor para execução dos Passos 2, 3 e 4. Assume-se que o Decisor é capaz de estimar os trade-off's entre dois objetivos quaisquer em uma dada solução do problema. Estes trade-off's podem ser utilizados para encontrar a direção do gradiente de $\mathcal{U}(\cdot)$ naquela solução.

O gradiente da função utilidade pode ser expresso a partir da aplicação da regra da cadeia como

$$\nabla_x \mathcal{U}(f(x)) = \sum_{i=1}^m \frac{\partial \mathcal{U}}{\partial f_i} \nabla f_i(x). \quad (1.35)$$

Como a solução do problema (1.32) não será afetada pela multiplicação da função objetivo por um escalar, pode-se reescrevê-lo como,

$$\min_{\xi \in X} \left\langle \sum_{i=1}^m \tau_i^k \nabla f_i(x^k), \xi \right\rangle. \quad (1.36)$$

onde

$$\tau_i^k = (\partial \mathcal{U} / \partial f_i^k) / (\partial \mathcal{U} / \partial f_1^k), \quad i = 2, \dots, m \quad (1.37)$$

define a *Taxa Marginal de Substituição* do Decisor entre o critério i e o critério 1, arbitrariamente escolhido como critério de referência.

Uma forma de se obter τ_i é determinar uma mudança infinitesimal Δf_1 no objetivo de referência que é compensada *exatamente* por uma mudança Δf_i no i -ésimo objetivo, enquanto todos os outros objetivos são mantidos constantes. Pode-se então aproximar (1.37) por,

$$\tau_i^k = -\frac{\Delta f_i^k}{\Delta f_1^k} \quad i = 1, \dots, m. \quad (1.38)$$

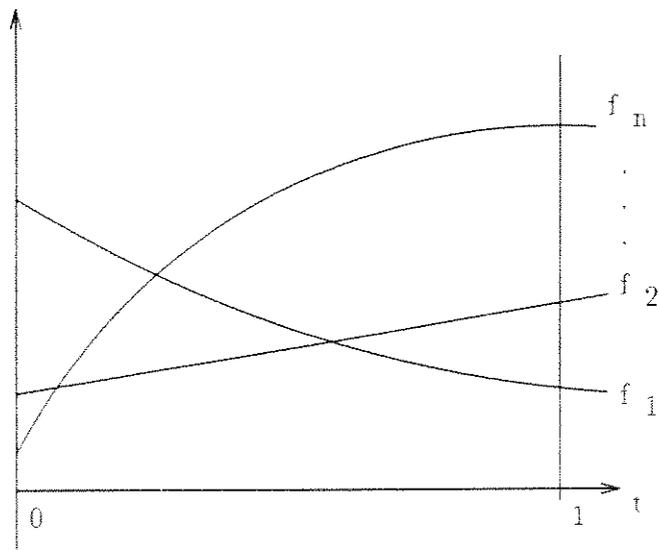
Devido a subjetividade na estimação dos trade-off's, o Decisor pode ter dificuldade em encontrar valores expressivos. Com o propósito de reduzir esta dificuldade, foi proposto um esquema interativo baseado em comparações ordinais que permite a determinação das taxas marginais de substituição com razoável precisão [Dyer, 1972]. Para a execução do Passo 3 resolve-se um problema de busca unidimensional com o auxílio do Decisor. A escolha do passo t pode ser conseguida plotando-se todos os m objetivos para $t \in [0, 1]$. Desta forma o Decisor poderá determinar qual o melhor passo $t^k \in [0, 1]$ (Fig. 1.2).

A condição teórica de parada do algoritmo (Passo 4) será satisfeita quando as soluções x^k e x^{k+1} forem iguais, o que nem sempre é possível. Em [Hwang e Masud, 1979] apresenta-se um critério de parada alternativo baseado na variação da função utilidade.

O método de Geoffrion apresenta-se como uma ferramenta eficiente para a resolução de problemas multiobjetivos, devido a simplicidade computacional e as fortes propriedades de convergência do algoritmo de Frank-Wolfe. [Geoffrion et al, 1972]. [Dyer, 1973]. Para analisar a robustez da convergência do algoritmo de Frank-Wolfe [Dyer, 1974]. [Ferreira, 1986], considera-se a introdução de erros no cálculo do gradiente da função utilidade, reescrevendo-se (1.33) como.

$$\min_{\xi \in X} \langle \nabla_{\mathcal{U}}(f(x^k)) + \eta^k, \xi \rangle. \quad (1.39)$$

Deste modo, pode-se mostrar que a sequência $\{\xi^l\}_0^\infty$ gerada pelo algoritmo converge para uma solução ótima se $\lim_{k \rightarrow \infty} \eta^k \rightarrow 0$. Analisando-se este fato no contexto multiobjetivo, isto indica que o erro na estimação da taxa marginal deve convergir para 0 quando $k \rightarrow \infty$, o que pode ser perfeitamente admitido supondo-se que o Decisor melhora progressivamente suas estimativas. Assumido estas condições, o algoritmo sempre convergirá.

Figura 1.2: Escolha do passo t

1.6 Conclusão

Aspectos teóricos básicos relativos a Otimização Multiobjetivo foram apresentados neste capítulo. Apresentou-se definições e teoremas básicos, juntamente com duas técnicas escalares para a geração de soluções eficientes, e um diagrama de implicações relacionando-as com a existência de soluções eficientes. Em seguida estudou-se as condições necessárias para a existência de soluções eficientes (KT_E) em analogia com o caso escalar (condições de Kuhn-Tucker). Finalmente apresentou-se uma classificação simplificada dos métodos multiobjetivos segundo a forma de participação do *Decisor*, juntamente com uma descrição sucinta de alguns destes métodos. Especial ênfase foi dada ao método de Geoffrion et al (1972) devido a sua futura implementação no Espaço dos Objetivos, a ser desenvolvida no Capítulo 4.

Capítulo 2

Otimização no Espaço dos Objetivos

2.1 Introdução

Neste capítulo, aspectos básicos referentes à Otimização Multiobjetivo no *Espaço dos Objetivos* são abordados. Para tanto, apresenta-se uma metodologia que decompõe o PMO em uma estrutura de dois níveis, permitindo assim a definição de um problema multiobjetivo mestre manipulado diretamente no Espaço dos Objetivos. Um subproblema testa a factibilidade da solução encontrada pelo problema mestre. A abordagem de PMO's no Espaço dos Objetivos oferece condições ideais para a implementação de vários métodos multiobjetivos.

Faz-se uma análise do caso bi-objetivo empregando-se o método descrito, e em seguida analisam-se problemas multiobjetivos gerais. O estudo de algumas propriedades como linearidade, robustez, homogeneidade e convergência, também será apresentado. Discute-se a aplicabilidade do método sob condições ligeiramente mais fracas do que as propostas por Ferreira e Geromel (1990) e faz-se um estudo sobre a técnica de resolução do problema *Min-Max*.

2.2 Descrição da Abordagem

A estratégia básica da abordagem proposta por Ferreira e Geromel (1990) é decompor o problema original (1.1) em um problema escalar do tipo *Min-Max*, e em um problema mestre formulado no Espaço dos Objetivos. A partir deste desenvolvimento, obtém-se um *Algoritmo Básico* ao qual várias técnicas específicas podem ser associadas.

Como comentado anteriormente, $\Gamma = f[\mathcal{X}^*]$ representa o conjunto de todas as soluções eficientes no Espaço dos Objetivos \mathcal{F} . É fácil mostrar que $\Gamma \subset \partial\mathcal{F}$, onde $\partial\mathcal{F}$ representa a fronteira de \mathcal{F} [Reid e Citron, 1971]. Sob hipóteses de convexidade pode-se obter importantes propriedades relacionadas a Γ .

A partir dos resultados obtidos no capítulo anterior, e supondo-se convexidade de $f_i(\cdot)$, $i = 1, \dots, m$, sobre $\mathcal{X} \subset R^n$ convexo, obter $x^* \in \mathcal{X}$ que resolve (1.1) é equivalente a determinar o vetor $w = w^*$ tal que $x^* = x(w^*)$, onde,

$$x(w^*) = \arg \min_{x \in \mathcal{X}} \langle w^*, f(x) \rangle . \quad (2.1)$$

O método explora esta equivalência considerando o problema no Espaço dos Objetivos

$$\min_{y \in \mathcal{F}} y \quad (2.2)$$

que pode ser colocado no sentido de (1.8) como

$$\min_{y \in \mathcal{F}} \mathcal{U}(y) \quad (2.3)$$

Como geralmente \mathcal{F} não é conhecido explicitamente, o problema (2.3) não pode ser diretamente solucionado. Uma maneira conveniente para representar a região factível \mathcal{F} é através de manipulações como projeção e relaxação. O desenvolvimento proposto em [Ferreira e Geromel, 1990], baseia-se no conceito de *Solução Satisfatória*, definido a seguir.

Definição 2.1 *Solução Satisfatória*

Um vetor de trade-off's implícitos $y \in R^m$ é dito ser satisfatório se $\exists x \in \mathcal{X}$ tal que $f(x) \leq y$.

Assumindo-se que a busca pela solução de (2.3) será restrita a decisões que têm imagem em Γ , o qual geralmente possui um número infinito de vetores eficientes, é interessante limitar Γ [Ferreira e Geromel, 1990]. Para tanto considera-se o conjunto compacto de trade-off's implícitos da seguinte forma,

$$\mathcal{Y} \triangleq \{y \in R^m : \underline{y} \leq y \leq \bar{y}\}, \quad (2.4)$$

onde \underline{y} e \bar{y} representam respectivamente os valores mínimos e máximos para as funções objetivos na região factível ou mais especificamente no conjunto de soluções eficientes Γ .

O conjunto \mathcal{Y} pode ser fornecido pelo Decisor através da análise das relações físicas que regem o problema, ou por meio de um procedimento mais sistemático. Note que os limites inferiores são dados por $\underline{y}_i = f_i(x^i)$, onde $x^i = \arg \min_{x \in \mathcal{X}} f_i(x)$. Já a obtenção dos limitantes superiores é menos evidente e maneiras de encontrá-los são propostas em Ferreira e Geromel (1990) e Geromel e Ferreira (1991).

A projeção das restrições $f(x) \leq y$, $x \in \mathcal{X}$, no espaço dos trade-off's implícitos é representada pelo conjunto convexo

$$\mathcal{V} \triangleq \{y \in R^m : f(x) \leq y, \text{ para algum } x \in \mathcal{X}\}, \quad (2.5)$$

o qual representa o conjunto de todas as soluções satisfatórias do problema.

Considerando este desenvolvimento, o problema (2.3) pode ser reformulado em termos de \mathcal{V} e \mathcal{Y} como

$$\min_{y \in \mathcal{Y} \cap \mathcal{V}} \mathcal{U}(y) \quad (2.6)$$

onde, por hipótese, $\mathcal{U}(\cdot)$ é crescente em relação a cada y_i , $i = 1, \dots, m$.

Teorema 2.1 *Teorema da Equivalência*

Se y^* é a solução ótima do problema (2.6), então y^* resolve o problema multiobjetivo (2.3).

Prova: Se $y^* \in \mathcal{V} \cap \mathcal{Y}$ resolve (2.6), então existe $x^* \in \mathcal{X}$ tal que $y^* = f(x^*)$. Caso contrário, se $y^* \geq f(x^*)$ e $y^* \neq f(x^*)$, então $y^0 = f(x^*)$ contradiz a otimalidade de y^* pois $y^0 \in \mathcal{Y} \cap \mathcal{V}$ e $\mathcal{U}(y^0) < \mathcal{U}(y^*)$. Então $y^* = y^0 = f(x^*)$ e, obviamente, $y^* \in \Gamma$. Note que $\mathcal{F} \subset \mathcal{V} \cap \mathcal{Y}$, permitindo concluir que $y^* \in \mathcal{F}$ também resolve (2.3). \square

Pelo teorema anterior, os problemas (2.3) e (2.6) são equivalentes, sendo que ao contrário da formulação (2.3), o conjunto de restrições de (2.6) pode ser caracterizado por propriedades fundamentais de conjuntos convexos. Em particular, o conjunto de soluções satisfatórias pode ser representado através dos semi-espacos suportes a \mathcal{V} .

Teorema 2.2 *Propriedade Fundamental dos Conjuntos Convexos*

Seja \mathcal{V} definido por (2.5). Então $y \in \mathcal{V}$ se e somente se y satisfaz o seguinte sistema linear infinitamente restrito.

$$\min_{x \in \mathcal{X}} \langle w, f(x) - y \rangle \leq 0, \quad \forall w \in \mathcal{W} \quad (2.7)$$

Prova: [Lasdon, 1970]. \square

A partir do Teorema 2.2, o problema (2.6) pode ser escrito na forma

$$\min_{y \in \mathcal{Y}} \mathcal{U}(y) \quad (2.8)$$

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

Observando-se a estrutura do problema (2.8), é conveniente empregar alguma estratégia de *relaxação* para resolvê-lo. As técnicas de relaxação baseiam-se em dois passos principais: i) determinar se a solução corrente de (2.8) satisfaz todas as restrições ignoradas, e caso contrário; ii) gerar e adicionar as restrições violadas à prévia aproximação da região factível. Esta tarefa pode ser realizada através da definição da função escalar $\theta(\cdot) : R^m \rightarrow R$,

$$\theta(y) \triangleq \max_{w \in \mathcal{W}} \min_{x \in \mathcal{X}} \langle w, f(x) - y \rangle. \quad (2.9)$$

Deste modo, nota-se que $y \in \mathcal{V}$ se e somente se $\theta(y) \leq 0$ e contrariamente $y \notin \mathcal{V}$ se e somente se $\theta(y) > 0$ [Ferreira e Geromel, 1990].

Com respeito à aplicação do Teorema 2.2, nota-se que se $\theta(y) > 0$, a solução ótima do problema Min-Max estará associada à restrição mais violada de (2.8). Propriedades adicionais originadas da aplicação deste teorema ajudarão a compreender os mecanismos do método.

Teorema 2.3 : Eficiência da Solução Ótima

Seja $(w^*, x(w^*))$ a solução ótima de (2.9) para um $y^* \in R^m$ fixo e considere também o semi-espço do R^m definido como

$$\mathcal{H}_w \triangleq \{y \in R^m : \langle w^*, y - f(x(w^*)) \rangle \geq 0\}. \quad (2.10)$$

então $\xi = f(x(w^*)) \in \Gamma \subset \mathcal{F} \subset \mathcal{V} \subset \mathcal{H}_w$.

Prova: [Ferreira e Geromel, 1990]. \square

A partir destes resultados é possível formular o seguinte problema de otimização a ser resolvido pelo Decisor

$$\min_y \mathcal{U}(y) \quad (2.11)$$

$$\text{s.a. } y \in \mathcal{Y} \cap \mathcal{H}_w, \quad \forall w \in \mathcal{W}$$

A solução ótima de (2.11) é obtida através do seguinte procedimento iterativo, que envolve apenas a adição da restrição mais violada com relação à solução atual. Este procedimento é referenciado aqui como *Algoritmo Básico*.

Passo 1 (*Inicialização*): Faça $\mathcal{Y}^0 = \mathcal{Y}$ e $k = 0$.

Passo 2 (*Decisão*): Resolva o problema relaxado

$$\min_{y \in \mathcal{Y}^k} \mathcal{U}(y) \quad (2.12)$$

Seja y^k a solução.

Passo 3 (*Análise*): Resolva o problema Min-Max

$$\theta(y^k) = \max_{w \in \mathcal{W}} \min_{x \in \mathcal{X}} \langle w, f(x) - y^k \rangle \quad (2.13)$$

Seja $(u^k, x(u^k))$ a solução. Se $\theta(y^k) \leq 0$, então $x^k = x(u^k)$ resolve (2.3) e $y^k = f(x^k)$ resolve (2.6). Caso contrário adicione a restrição mais violada à \mathcal{Y}^k

$$\mathcal{Y}^{k+1} = \{y \in \mathcal{Y}^k : \langle u^k, y \rangle \geq \langle u^k, f(x(u^k)) \rangle\} \quad (2.14)$$

faça $k=k+1$ e volte ao **Passo 2**.

O **Passo 2** para $k = 0$ pode ser prontamente executado pois supondo-se coerência do Decisor, obviamente, $y^0 = \underline{y}$.

2.3 Análise do Caso Bi-objetivo

O método apresentado adequa-se perfeitamente à resolução de problemas bicritérios. Do ponto de vista de análise, estes problemas levam a interessantes simplificações para o cálculo de $\theta(y^k)$. Neste caso, como $m=2$, o problema Min-Max (2.9) pode ser colocado como

$$\begin{aligned}
\theta(y^k) &= \max_{w_1 \in [0,1]} \min_{x \in \mathcal{X}} \{w_1(f_1(x) - y_1^k) + (1 - w_1)(f_2(x) - y_2^k)\} \\
&= \max_{w_1 \in [0,1]} \phi(w_1)
\end{aligned} \tag{2.15}$$

e para se obter $\theta(y^k)$, é suficiente empregar um simples procedimento de busca unidimensional para obter o valor ótimo de $w_1 \in [0, 1]$, com ou sem informações sobre as derivadas de $\phi(w_1)$ [Ferreira, 1986], e algum método apropriado (Primal ou Dual) para resolver o problema em $x \in \mathcal{X}$ para cada w_1 fixo.

Em termos de Decisão, problemas bicritérios oferecem vantagens ainda maiores. Neste tipo de problema, o procedimento de relaxação pode ser completamente representado através de dispositivos gráficos apropriados. Deste modo, a participação do Decisor na busca da melhor solução de compromisso torna-se mais real e direta. Um estudo mais detalhado sobre a solução de problemas bi-objetivos através da metodologia descrita neste capítulo pode ser encontrado em Ferreira e Geromel (1990).

Extensões dos resultados da análise do caso bi-objetivo podem ser conseguidas para o caso multiobjetivo.

2.4 Análise do Caso Multiobjetivo

Para problemas multiobjetivos onde $m \geq 3$, torna-se praticamente inviável a utilização de recursos gráficos, como no caso bi-objetivo. Entretanto, pode-se fazer uso de algumas propriedades derivadas da decomposição proposta em [Ferreira e Geromel, 1990].

A importante propriedade de *linearidade* encontra-se presente neste esquema, pois os objetivos são lineares, de um tipo muito especial e as restrições são também lineares. Em termos de implicação prática, qualquer método multiobjetivo linear, tal como STEM (seção 1.5), pode ser utilizado para resolver problemas multiobjetivos não-lineares. A estrutura do método apresenta também características de *robustez*, pois independentemente do erro cometido pelo Decisor em estimar subjetivamente a solução, o nível de

análise (resolução do problema Min-Max) sempre gerará uma solução eficiente (Fig. 2.1), o que por exemplo não ocorre com o método de Geoffrion et al (1972) que eventualmente poderá gerar uma solução não-eficiente.

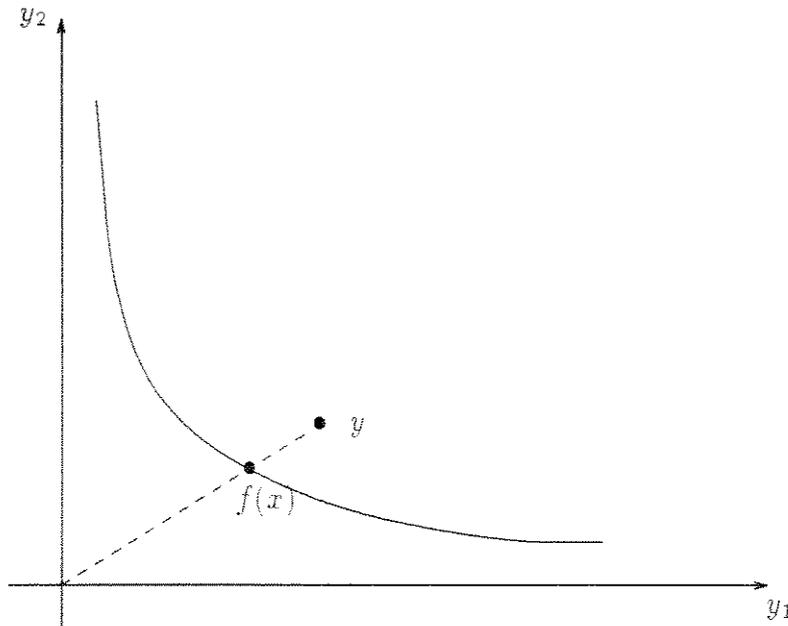


Figura 2.1: Ilustra a Robustez do Método

O problema multiobjetivo decomposto segundo esta abordagem sempre apresentará uma estrutura do tipo do problema (2.12), onde o conjunto factível \mathcal{Y}^k é um poliedro. Devido a esta simplificação, torna-se fácil a implementação de Sistemas de Suporte à Decisão baseados em Programação Multiobjetivo, pois problemas convexos (ou quase-convexos como se verá em seguida) podem ser sempre colocados nesta forma.

Supondo-se que a função utilidade $U(\cdot)$ seja perfeitamente conhecida o teorema a seguir garante a convergência infinita do Algoritmo Básico.

Teorema 2.4 *Convergência Infinita*

Suponha que o Algoritmo Básico gera a sequência $\{y^k\}$. Qualquer ponto limite desta sequência resolve (2.3).

Prova: [Ferreira e Geromel, 1990]. \square

Do teorema acima, fica claro que a convergência do Algoritmo Básico dependerá da precisão na determinação da solução do problema relaxado, quando suposto que $\mathcal{U}(\cdot)$ não é perfeitamente conhecida. Neste caso a determinação de y^k depende das propriedades de convergência do método multiobjetivo utilizado para resolver o problema relaxado.

A utilização do método multiobjetivo proposto em [Geoffrion et al, 1972] (método utilizado neste trabalho para solucionar o problema relaxado (2.11)), possibilita a garantia de convergência infinita para o Algoritmo Básico. Isto pode ser constatado observando-se a Figura 2.2, onde a região Γ de interesse para Decisor torna-se melhor aproximada à medida que o número de iterações cresce.

É interessante ressaltar que a convexidade do conjunto \mathcal{V} é de fundamental importância para o desenvolvimento do Algoritmo Básico.

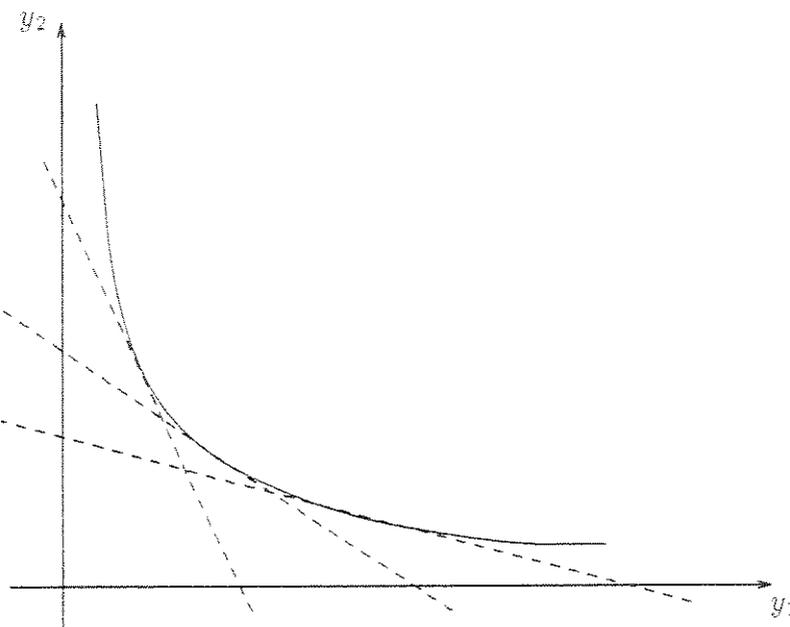


Figura 2.2: Procedimento de Linearização Externa

2.5 Condições Fracas para \mathcal{V} Convexo

Existem vários tipos de funções que são semelhantes às funções convexas ou côncavas, mas que mantêm apenas algumas de suas propriedades. Sabe-se que em algumas aplicações não se necessita da suposição restritiva de convexidade, mas apenas das suposições menos restritivas de *quase-convexidade* e *pseudo-convexidade*. O desenvolvimento a seguir baseia-se em [Bazaraa e Shetty, 1979], onde as demonstrações de Teoremas omitidas podem ser encontradas.

Definição 2.2 Quase-convexidade

Seja $f : \mathcal{X} \rightarrow R$, onde \mathcal{X} é um conjunto convexo não vazio do R^n . A função f é dita ser quase-convexa se, para todo x^1 e $x^2 \in \mathcal{X}$, a desigualdade

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

é verdadeira para todo $\lambda \in (0, 1)$.

Diz-se que a função f é quase-côncava se $-f$ for quase-convexa. Uma função quase-convexa é caracterizada pela convexidade do seu conjunto de *platôs*, o que é mostrado no teorema seguinte.

Teorema 2.5 Função Quase-convexa

Seja $f : \mathcal{X} \rightarrow R$, onde \mathcal{X} é um conjunto convexo não vazio do R^n . A função f é quase-convexa se e somente se o conjunto $\mathcal{X}_\alpha \triangleq \{x \in \mathcal{X} : f(x) \leq \alpha\}$ é convexo para todo número real α .

Uma variante de quase-convexidade é o conceito de quase-convexidade estrita, que assegura que um mínimo local de f sobre um conjunto convexo é também um mínimo global.

Definição 2.3 Quase-convexidade Estrita

Seja $f : \mathcal{X} \rightarrow R$, onde \mathcal{X} é um conjunto convexo não-vazio no R^n . A função f é dita ser estritamente quase-convexa se $\forall x^1, x^2 \in \mathcal{X}$ com $f(x^1) \neq f(x^2)$, implica em

$$f[\lambda x^1 + (1 - \lambda)x^2] < \max\{f(x^1), f(x^2)\}$$

para todo $\lambda \in (0, 1)$

Definição 2.4 Pseudo-Convergência

Seja \mathcal{X} um conjunto aberto não-vazio no R^n e seja $f : \mathcal{X} \rightarrow R$ diferenciável sobre \mathcal{X} . A função f é dita ser pseudo-convexa se $\forall x^1, x^2 \in \mathcal{X}$ com $\nabla f(x^1)'(x^2 - x^1) \geq 0$ tem-se $f(x^2) \geq f(x^1)$, ou equivalentemente, se $f(x^2) < f(x^1)$ então $\nabla f(x^1)'(x^2 - x^1) < 0$.

Diz-se que a função f é estritamente pseudo-convexa se $\forall x^1, x^2 \in \mathcal{X}$ distintos, satisfazendo $\nabla f(x^1)'(x^2 - x^1) \geq 0$ tem-se $f(x^2) > f(x^1)$, ou equivalentemente, se $\forall x^1, x^2 \in \mathcal{X}$ distintos, $f(x^2) \leq f(x^1)$ implica que $\nabla f(x^1)'(x^2 - x^1) < 0$.

O teorema a seguir mostra que uma função pseudo-convexa é também estritamente quase-convexa e quase-convexa.

Teorema 2.6 Função Pseudo-convexa

Seja \mathcal{X} um conjunto convexo aberto não-vazio no R^n , e seja $f : \mathcal{X} \rightarrow R$ uma função pseudo-convexa diferenciável sobre \mathcal{X} . Então f é tanto estritamente quase-convexa quanto quase-convexa.

Uma importante classe de funções pseudo-convexas é a representada pelas funções fracionárias.

Lema 2.1 Funções Fracionárias

Seja $f(x) = (\langle p, x \rangle + \alpha) / (\langle q, x \rangle + \beta)$, onde $p, q \in R^n$, e α e β são escalares, e seja \mathcal{X} um conjunto convexo tal que $\langle q, x \rangle + \beta \neq 0$ sobre \mathcal{X} . Então f é tanto pseudo-convexa quanto pseudo-côncava sobre \mathcal{X} .

Para um problema de minimização do tipo

$$\min_{x \in \mathcal{X}} f(x), \quad (2.16)$$

onde $f(x)$ é uma função fracionária, pode-se chegar a algumas conclusões a partir dos resultados obtidos nesta seção.

Como a função objetivo $f(\cdot)$ é tanto pseudo-convexa quanto pseudo-côncava sobre \mathcal{X} , então também será quase-convexa, estritamente quase-convexa, quase-côncava e estritamente quase-côncava. Assim, um ponto que satisfaça as condições de Kuhn-Tucker de (2.16) será um ponto de mínimo local sobre a região factível \mathcal{X} . Como a função objetivo é estritamente quase-convexa um ponto de mínimo local é também um ponto de mínimo global sobre a região factível [Bazarara e Shetty, 1979].

Note que a quase-convexidade da função

$$G(x, y) = f(x) - y, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y} \quad (2.17)$$

garante, pelo teorema 2.5, que o conjunto \mathcal{V} será convexo, e portanto, qualquer ponto da sua fronteira admite um semi-espço suporte. Se for possível obter o mínimo de $\langle w, f(x) - y \rangle$ sobre \mathcal{X} , qualquer solução eficiente poderá ser determinada. Note ainda, que isso não quer dizer que a função $\langle w, f(x) \rangle$ seja necessariamente convexa.

Apresenta-se a seguir um breve estudo sobre os aspectos relacionados à resolução do problema *Min-Max*.

2.6 Solução do Problema Min-Max

Nesta seção apresenta-se técnicas adequadas para solucionar o problema Min-Max do Passo 3 do Algoritmo Básico, que equivale ao nível de análise do método:

$$\theta(y) = \max_{w \in \mathcal{W}} \min_{x \in \mathcal{X}} \langle w, f(x) - y \rangle \leq 0 \quad (2.18)$$

$$= \max_{w \in \mathcal{W}} \phi(w) \quad (2.19)$$

Como a resolução deste problema não envolve aspectos subjetivos, pode-se utilizar técnicas convencionais de programação matemática. O problema *dual* (2.19) pode ser resolvido através de métodos baseados em cálculos de gradientes e de suas projeções. Para tanto é necessário garantir que $\nabla \phi(w^0)$

exista num determinado ponto w^0 do seu domínio \mathcal{W} . Se $\nabla\phi(w)$ existe para todo $w \in \mathcal{W}$, então é possível adotar processos iterativos do tipo

$$\begin{aligned} w^{l+1} &= w^l - \alpha^l \text{Proj}_{\mathcal{W}} \nabla\phi(w^l) \\ \phi(w^{l+1}) &> \phi(w^l), \quad l = 0, 1, \dots \end{aligned} \quad (2.20)$$

onde α^l é um tamanho de passo.

Mesmo admitindo-se que $\phi(w)$ seja diferenciável em \mathcal{W} , o uso de procedimentos com este enfoque torna-se oneroso em termos de esforço computacional.

Como um procedimento alternativo para a resolução de (2.18)-(2.19) pode-se utilizar a técnica de *Aproximação Tangencial* como descrito abaixo.

Sendo $\mathcal{X} \subset R^n$ compacto, demonstra-se em [Lasdon, 1970] que se

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

então

$$\xi = f(x(w)) - y$$

representa um elemento do conjunto de *subgradients* de $\phi(\cdot)$ no ponto w , que devido à concavidade de $\phi(\cdot)$ satisfaz a equação

$$\phi(w^1) \leq \phi(w^0) + \langle \xi, w^1 - w^0 \rangle, \quad \forall w^0, w^1 \in \mathcal{W}.$$

Supondo-se a existência de q subgradients $\xi^i = f(x(w^i)) - y$, $i = 1, \dots, q$ então

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

representa uma aproximação de ordem q para $\phi(\cdot)$. Deste modo pode-se aproximar (2.18)-(2.19) por

$$\max_{w \in \mathcal{W}} \min_{1 \leq i \leq q} \langle w, f(x(w^i)) - y \rangle \quad (2.21)$$

ou ainda

$$\max_{w \in W} \sigma \quad (2.22)$$

$$\text{s.a. } \sigma \leq \langle w, f(x(w^i)) - y \rangle, \quad i = 1, \dots, q \quad (2.23)$$

O problema (2.22)-(2.23) pode ser resolvido sem a necessidade de procedimentos de busca unidimensional. A solução do problema Min-Max é encontrada quando $(\sigma^q - \phi(w^q)) < \epsilon$, para um dado $\epsilon > 0$ arbitrariamente pequeno.

Para o cálculo de $\phi(\cdot)$, o método do *Gradiente Modificado*, proposto por Geromel e Ferreira (1986) pode ser empregado. Este método é voltado à resolução de problemas de otimização com função objetivo não-linear e restrições lineares. Genericamente resolve-se um problema do tipo

$$\min_{x \in X} h(x) \quad (2.24)$$

$$\text{s.a. } Ax = b \quad (2.25)$$

$$\underline{x} \leq x \leq \bar{x} \quad (2.26)$$

onde $x \in R^n$, $A \in R^{m \times n}$ com $\text{rank}(A) = m < n$ e a função objetivo $h(\cdot) : R^n \rightarrow R$ continuamente diferenciável.

A direção de decrescimento da função $h(\cdot)$ em x , $s(x) = \text{Proj}_X \nabla h(x)$ pode ser eficientemente determinada através da solução do problema abaixo:

$$s(x) \triangleq \arg \min_s \left\{ \frac{1}{2} \|\nabla h(x) - s\|^2 \text{ s.a. } As = 0, S(x)s \leq 0 \right\} \quad (2.27)$$

onde $S(x) \in R^{r \times n}$ é uma *matriz sinal* dada por

$$S(x) = \{S(x)_{ij}\} \triangleq \begin{cases} 1, & \text{se } x_j = \underline{x}_j \\ -1, & \text{se } x_j = \bar{x}_j \\ 0, & \text{caso contrário.} \end{cases}$$

Note que a dimensão r é igual ao número de restrições de bornes ativas e que as restrições não ativas geram linhas nulas em $S(x)$, as quais podem ser retiradas sem nenhum ônus para a solução do problema.

Definindo-se

$$\Omega(x) \triangleq \{s \in R^n \text{ s.a. } S(x)s \leq 0\}, \quad (2.28)$$

o problema (2.27) pode ser reescrito como

$$s(x) \triangleq \text{Proj}_{\mathcal{N}(A) \cap \Omega(x)} \nabla h(x) \quad (2.29)$$

Se $S(x)=0$ então $\Omega(x) \equiv R^n$ e tem-se

$$s(x) = \text{Proj}_{\mathcal{N}(A)} \nabla h(x) = P \nabla h(x) \quad (2.30)$$

onde $P = I - A'(AA')^{-1}A = P' \geq 0$ é o operador de projeção ortogonal sobre o espaço nulo de A , $\mathcal{N}(A)$.

Definição 2.5 Gradiente Modificado

Um vetor $\nabla_m h \in R^n$ é chamado gradiente modificado de $h(\cdot)$ com relação a $\mathcal{B} \subset R^n$, se satisfaz

$$\text{Proj}_{\mathcal{A}} \nabla_m h = \text{Proj}_{\mathcal{A} \cap \mathcal{B}} \nabla h \quad (2.31)$$

Aplicando-se a definição acima para $\mathcal{A} \triangleq \mathcal{N}(A)$ e $\mathcal{B} \triangleq \Omega(x)$ pode-se reescrever (2.30) como

$$s(x) = P \nabla_m h(x) \quad (2.32)$$

Teorema 2.7 Cálculo do Gradiente Modificado

O gradiente modificado da função $h(\cdot)$ em qualquer ponto factível é dado por

$$\nabla_m h(x) = \nabla h(x) - S(x)' \mu(x). \quad (2.33)$$

onde $\mu(x) \in R^r$ resolve o seguinte problema de programação quadrática,

$$\max_{\mu \geq 0} \left\{ -\frac{1}{2} \mu' Q(x) \mu + c(x)' \mu \right\}, \quad (2.34)$$

onde $Q(x) \triangleq S(x)PS(x)'$ e $c(x) \triangleq S(x)P\nabla f(x)$.

Prova: [Geromel e Ferreira, 1986]. \square

É importante observar que resolver (2.34) é muito mais simples que resolver o problema original (2.24), pois em geral $\dim(\mu) = r \ll \dim(x) = n$.

O procedimento numérico para resolver (2.24) pode ser definido como segue, supondo-se que uma solução inicial factível $x^0 \in \mathcal{X}$ tenha sido calculada.

Passo 1 Encontre $P, \nabla h(x^0) \in S(x^0)$. Faça $l = 0$.

Passo 2 Com $P, \nabla h(x^l) \in S(x^l)$ determine

$$\mu(x^l) = \arg \max_{\mu \geq 0} \frac{1}{2} \langle \mu, Q(x^l)\mu \rangle + \langle c(x^l), \mu \rangle \quad (2.35)$$

onde $Q(x^l) = -S(x^l)PS(x^l)'$ e $c(x^l) = S(x^l)P\nabla h(x^l)$.

Passo 3 Determine o gradiente modificado $\nabla h_m(x^l)$ e a direção de pesquisa s^l por

$$\nabla h_m(x^l) = \nabla h(x^l) - S(x^l)'\mu(x^l) \quad (2.36)$$

$$s^l = P\nabla_m h(x^l) \quad (2.37)$$

Passo 4 Se $\|s^l\| < \epsilon$, com $\epsilon > 0$ suficientemente pequeno, pare. A solução corrente x^l satisfaz as condições necessárias de Kuhn e Tucker para a existência de um mínimo local de $h(\cdot)$. Caso contrário, vá ao Passo 5.

Passo 5 Encontre o passo α^l do algoritmo resolvendo o problema de busca unidimensional

$$\min_{0 \leq \alpha \leq \bar{\alpha}} h(x^l - \alpha s^l)$$

atualize as variáveis primais $x^{l+1} = x^l - \alpha^l s^l$, faça $l = l + 1$ e volte ao Passo 1.

A determinação de $Q(x^l) \in R^{r \times r}$ e $c(x^l) \in R^r$ é simples porque é esperado que $r \ll n$. Estas representações são compostas pelas colunas de P e pelas linhas de $P\nabla h(x)$, com índices tais que $\{S(x)_{ij}\} \neq 0$. Portanto, envolvem apenas operações lógicas. É possível provar que se $s(x) = 0$, então x é uma solução ótima de (2.24) no sentido que satisfaz as condições de Kuhn e Tucker para a otimalidade [Geromel e Ferreira, 1986].

2.7 Conclusão

Neste capítulo apresentou-se uma metodologia baseada na projeção das variáveis de decisão sobre o espaço dos trade-off's implícitos que leva à formulação do problema multiobjetivo no espaço do objetivos. O problema é então resolvido através de uma técnica de relaxação. Esta abordagem de problemas multiobjetivos constitui um método efetivo para tratar processos que envolvam decisão.

Analisou-se o caso onde o Decisor interage com o espaço bi-objetivo de uma maneira simples e realística. Para a resolução de PMO's com mais de dois objetivos via o método proposto por Ferreira e Geromel (1990), apresentou-se um estudo sobre as propriedades derivadas da sua formulação, tais como *linearidade*, *robustez*, *convergência* e *homogeneidade*, o que possibilita colocar o problema numa forma suficientemente simples para permitir a implementação de qualquer método multiobjetivo.

Um estudo sobre a validade do esquema de decomposição apresentado sob a suposição de convexidade fraca do conjunto \mathcal{V} foi também realizado, o que permitiu estender os resultados prévios a funções objetivos pseudo-convexas.

Os aspectos numéricos relativos a resolução de problema *Min-Max* foram também abordados. O problema de maximização apresenta características que possibilitam a utilização do método de *Aproximação Tangencial*. Já para resolver o problema de minimização utilizou-se uma técnica, proposta por Geromel e Ferreira (1986) para a resolução de problemas de programação não-linear.

No próximo capítulo será feito o modelamento matemático do problema de controle de poluição em rios, ao qual os resultados descritos neste capítulo serão aplicados.

Capítulo 3

O Problema de Poluição em Rios

3.1 Introdução

Em um problema de planejamento de recursos hidráulicos se está interessado em otimizar tanto eficiência econômica, medida em termos de unidades monetárias, quanto qualidade ambiental, que é medida em unidades de concentração de poluentes. Tradicionalmente, para se resolver este tipo de problema considerava-se apenas um objetivo (eficiência econômica), com os demais objetivos sendo incluídos como restrições, ou através de comparações com o objetivo primário. Utilizavam-se técnicas de otimização clássicas para se resolver problemas multiobjetivos, o que tornava a solução bastante difícil e muitas vezes levava a soluções não satisfatórias.

Com o aumento da importância atribuída a objetivos secundários (difíceis de quantificar monetariamente) por parte da sociedade, a programação multiobjetivo tem encontrado uma grande aplicação em problemas desta natureza.

Geralmente, os projetos sobre a utilização de recursos hidráulicos, tais como, barragens, reservatórios, aqueodutos e sistemas de irrigação entre outros, são elaborados para servir a múltiplos objetivos. Por exemplo, um grande reservatório criado por uma barragem pode fornecer água para irrigação, suprir necessidades urbanas e industriais, servir à pesca e atividades de lazer, possibilitar a navegação e o controle de enchentes, servir à geração de energia elétrica, manter a qualidade da água aceitável em uma região, em resumo,

aumentar o desenvolvimento regional em termos de uma melhor economia e qualidade de vida.

Alguns resultados práticos sobre a aplicação da programação multi-objetivo a problemas de recursos hidráulicos podem ser encontrados em [Haimes et al., 1975] e [Ferreira, 1986] entre outros.

Na sequência deste Capítulo apresentam-se algumas formas de modelamento para o problema clássico de controle de poluição em rios, algumas formulações, mono e multiobjetivos, e os dados referentes ao problema prático de controle de poluição a ser tratado através dos algoritmos propostos neste trabalho.

3.2 O Modelamento Físico

O sistema físico em questão consiste de um reservatório multipropósito, como descrito anteriormente, formado a partir de uma barragem construída num determinado ponto de um rio. À jusante da barragem encontram-se n usuários, urbanos e industriais, que fazem uso da água liberada pelo reservatório e a devolvem poluída, sob a forma de detritos através de seus esgotos.

Os detritos poluentes são comumente classificados como degradáveis e não-degradáveis. Poluentes degradáveis tais como, lixo orgânico, bactérias e descargas térmicas utilizam a capacidade assimilativa da água para eliminar seus efeitos indesejáveis. Os poluentes não-degradáveis, tais como resíduos químicos inorgânicos e orgânicos sintéticos permanecem na água indefinidamente até que seus efeitos possam ser reduzidos através de diluição.

Muitos tipos de poluentes podem ser tratados em sua origem, embora em geral os métodos de tratamento sejam extremamente caros. Uma outra maneira de reduzir a poluição no rio é através do aumento do fluxo do reservatório, pois assim dilui-se tanto os poluentes degradáveis quanto os não-degradáveis, adiciona-se oxigênio na água para reduzir os efeitos de decomposição e alteram-se as características hidrológicas do *corpo de água*, melhorando sua capacidade assimilativa.

Uma forma de identificar o nível de poluição nos rios é através da quantidade de *oxigênio dissolvido*. Os detritos orgânicos servem como nutrientes para muitos organismos encontrados na água, que por sua vez aproveitam o

oxigênio da água no processo de utilização desses detritos. Quanto maior a quantidade de detritos orgânicos, maior será a quantidade de organismos, e por sua vez maior será a *demanda de oxigênio*.

Existem níveis mínimos específicos para a concentração de oxigênio dissolvido na água para cada rio ou seção de rio. Estes níveis são medidos a partir da concentração da demanda de oxigênio bioquímico (BOD - "Biochemical Oxygen Demand") [Loucks et al, 1967].

Duas equações diferenciais descrevendo os processos que determinam a concentração de oxigênio dissolvido (DO - "Dissolved Oxygen"), encontram-se desenvolvidas em [Loucks et al, 1967]. A primeira equação estabelece que a variação da concentração de BOD com o tempo é proporcional à concentração de BOD no instante presente, B , e à taxa de BOD adicional R , devido as chuvas e as correntezas.

$$\frac{dB}{dt} = -(k_1 + k_3)B + R, \quad (3.1)$$

onde k_1 e k_3 representam taxas constantes de desoxigenação e sedimentação, respectivamente.

A segunda equação determina que a variação no déficit de oxigênio dissolvido é proporcional à concentração de BOD atual B , ao déficit de oxigênio existente D e à taxa de produção ou redução de oxigênio A , e é dada por

$$\frac{dD}{dt} = k_1B - k_2D - A, \quad (3.2)$$

onde k_2 reflete a taxa em que o DO retorna ao rio através de *reaeração*.

Estas equações apesar de não descreverem exatamente os complexos fenômenos biológicos, químicos e físicos que ocorrem nos rios, são atualmente usadas por engenheiros projetistas e por órgãos oficiais para estabelecer os níveis de tratamento dos detritos lançados nos rios.

Integrando-se a equação (3.1) obtem-se,

$$B(t) = \left(B_0 - \frac{R}{k_1 + k_3} \right) \exp^{-(k_1 + k_3)t} + \frac{R}{k_1 + k_3} \quad (3.3)$$

onde $B(t)$ expressa a concentração de BOD no instante t em qualquer ponto do rio, abaixo do ponto de concentração inicial de BOD, B_0 .

Substituindo-se a equação (3.3) em (3.2) e integrando-se obtém-se

$$D(t) = \frac{k_1}{k_2 - (k_1 + k_3)} \left\{ \left(B_0 - \frac{R}{k_1 + k_3} \right) (\exp^{-(k_1 + k_3)t} - \exp^{-k_2 t}) \right\} + \frac{k_1}{k_2} \left\{ \left(\frac{R}{k_1 + k_3} - \frac{A}{k_1} \right) (1 - \exp^{-k_2 t}) \right\} + D_0 \exp^{-k_2 t} \quad (3.4)$$

onde $D(t)$ representa o déficit de DO no instante t , em qualquer ponto do rio, abaixo do ponto de concentração inicial de DO, D_0 .

Em qualquer instante t , a concentração de oxigênio dissolvido $C(t)$, é dada pela diferença entre a *concentração saturação* de oxigênio dissolvido CS e o déficit de oxigênio $D(t)$, como expresso abaixo.

$$C(t) = CS - D(t) \quad (3.5)$$

A equação (3.4) descreve a curva de decaimento do oxigênio dissolvido, como ilustrado na figura (3.1). O déficit crítico D_c e o instante crítico t_c ocorrem quando a concentração de DO está no seu menor valor C_c . O déficit crítico é dado pela diferença entre a concentração de saturação CS e a concentração no instante crítico t_c .

Portanto, para se determinar a concentração de DO em qualquer ponto de uma seção do rio, deve-se primeiro encontrar as concentrações de DO e BOD no início da referida seção [Loucks et al. 1967]. Conhecendo-se o tempo que a água leva para fluir do ponto inicial para o ponto onde se deseja determinar a concentração de DO desta seção, pode-se então fazer uso das equações (3.3) e (3.4), e obter as concentrações no ponto desejado.

A partir deste estudo várias formas de se modelar este tipo de sistema físico através de um problema de otimização podem ser sugeridas.

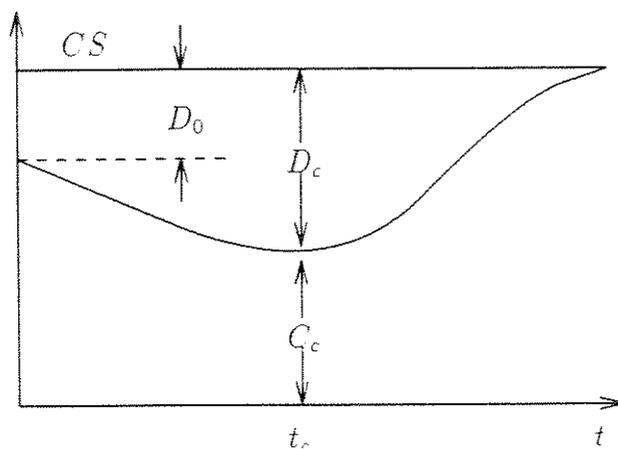


Figura 3.1: Concentração de Oxigênio Dissolvido

3.3 Modelo de Otimização

Para um dado rio e um conjunto padrões de DO para pontos ao longo de seu curso, desenvolve-se um conjunto de inequações lineares em termos das variáveis de tratamento e do acréscimo ao fluxo. Particionando o rio em n seções, onde no início de cada seção tem-se uma fonte de poluição associada (Figura 3.2), pode-se formular o seguinte problema de minimização de custo de tratamento dos resíduos lançados ao rio [Hass, 1970]:

$$\begin{array}{llllllll}
 \min & \phi_1(x_1) & + & \phi_2(x_2) & + & \dots & + & \phi_n(x_n) & + & \phi_0(x_{n+1}) \\
 \text{s.a.} & a_{11}w_1x_1 & + & & & \dots & + & b_1x_{n+1} & \geq & s_1 \\
 & a_{21}w_1x_1 & + & a_{22}w_2x_2 & + & & + & b_2x_{n+1} & \geq & s_2 \\
 & \vdots & & \vdots & & \dots & & \vdots & & \vdots \\
 & a_{n1}w_1x_1 & + & a_{n2}w_2x_2 & + & \dots & + & b_nx_{n+1} & \geq & s_n \\
 & x_i \geq 0 & & & & & & & & i = 1, \dots, n+1
 \end{array}$$

sendo que

$$s_i = \sum_{k=1}^i a_{ik} u_k,$$

e a seguinte notação é empregada,

n : número de seções do rio;

w_i : concentração de BOD no início da i -ésima seção;

x_i : percentual de w_i removido através do tratamento na i -ésima fonte poluidora;

a_{ij} : quantidade de oxigênio requerida na i -ésima seção para a decomposição de uma unidade de BOD descarregada no início da j -ésima seção; $a_{ij} = 0$, $j > i$;

y : acréscimo de vazão do reservatório para diluição dos detritos poluentes;

b_i : quantidade de oxigênio dissolvida, por unidade de vazão, disponível para o processo de decomposição na seção i ;

$\phi_i(x_i)$: custo de tratamento na fonte i ;

$\phi_0(x_{n+1})$: custo do acréscimo de vazão.

Os dados apresentados neste modelo são reais e obtidos de um rio americano (Miami River).

As restrições do problema expressam que a quantidade de oxigênio disponível para o processo de decomposição em cada seção deve ser igual ou maior que a demanda imposta pela carga de BOD descarregada naquela seção e em todas as seções precedentes.

A formulação com precisão das funções de custo locais $\phi_i(x_i)$ e do conjunto de restrições do problema anterior, parece ser um problema difícil, pois não depende apenas de técnicas de tratamento utilizadas pelos usuários, mas depende também das mudanças no processo de produção que têm resíduos poluentes (BOD) como sub-produtos.

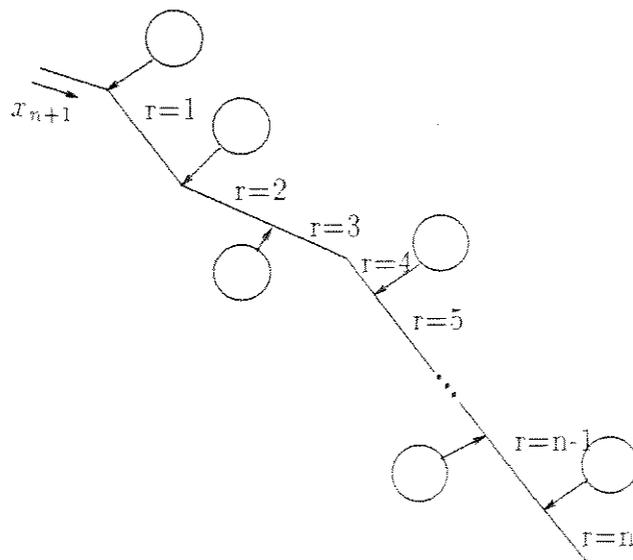


Figura 3.2: Bacia Hipotética do Rio

O trecho do rio que será modelado consta de um reservatório, abaixo do qual existem n usuários urbanos e industriais que descarregam seus resíduos poluentes no rio. Para o modelo, as variáveis de decisão x_i , $i = 1, \dots, n$ referem-se aos percentuais de carga BOD a serem removidos por cada um dos n usuários, e x_{n+1} é a constante de proporcionalidade a ser aplicada à vazão inicial, com o propósito de aumentar o fluxo liberado pelo reservatório e assim de diluir os resíduos. Admite-se que o tratamento primário mínimo dado aos resíduos poluentes liberados no rio remove cerca de 45% da carga de BOD, e que o máximo tratamento remove até 99% de BOD, devido às dificuldades tecnológicas. Entende-se por tratamento primário: clorificação, filtragem, sedimentação e remoção de sólidos volumosos. A variável de acréscimo de vazão x_{n+1} está limitada inferiormente por 0, quando nenhum acréscimo for necessário e superiormente por 3.47, que representa a quantidade de água disponível a ser liberada.

A função custo de tratamento empregada para cada usuário individual como apresentado em [Hass, 1970] é quadrática e dada por

$$\phi_i(x_i) = \alpha_i + \beta_i(x_i - 0.45)^2. \quad (3.6)$$

onde

$$\alpha_i = 160.8 + 26.7q_i, \quad (3.7)$$

$$\beta_i = 640.7 + 255.7q_i \quad (3.8)$$

e q_i , $i = 1, \dots, n$ é a quantidade total de resíduos gerada pelo i -ésimo usuário em milhões de galões por dia.

A função custo de acréscimo à vazão $\phi_0(x_{n+1})$ é dada por

$$\phi_0(x_{n+1}) = 726.6x_{n+1} \quad (3.9)$$

O modelo proposto por Hass resolve o problema multiobjetivo de controle de poluição em rios através de uma técnica de otimização escalar, colocando como função objetivo o somatório das funções de custos de tratamento dos resíduos somado à função de custo de acréscimo de vazão. As funções que

envolvem a qualidade da água e a quantidade de água disponível no reservatório são colocadas como restrições. Este tipo de formulação pode levar a resultados não satisfatórios, pois os objetivos deixados como restrições certamente serão colocados ao seu valor limitante. Na próxima seção apresenta-se o modelo multiobjetivo proposto por [Haimes et al., 1975].

3.4 Modelo Multiobjetivo

De modo diferente da abordagem de Hass (1970), o procedimento proposto por [Haimes et al., 1975] modela e resolve o problema de controle de poluição considerando que um *Decisor* estabelece os níveis de tratamento de cada usuário, e também, a quantidade de água liberada do reservatório com o propósito de diluir os resíduos poluentes lançados ao rio. Para estabelecer os níveis de tratamento para cada usuário, o Decisor controla a quantidade de BOD apenas num dado ponto do rio abaixo do último usuário.

Como já citado, o modelo proposto por Hass (1970) divide o rio em n seções e controla a concentração de BOD em cada uma destas seções. Uma relação entre os dois modelos pode ser obtida considerando-se que cada usuário representa o início de uma seção de rio (Fig. 3.3).

No modelo proposto em [Haimes et al., 1975] consideram-se três funções objetivos para a resolução do problema de poluição em rios.

O primeiro objetivo é o custo total de tratamento dos resíduos poluentes na região, que é representado pela soma dos custos de tratamento de cada usuário. Este objetivo é o mesmo apresentado pelo modelo de Hass

$$f_1(x) = \sum_{i=1}^n \phi_i(x_i) = \sum_{i=1}^n \alpha_i + \beta_i(x_i - 0.45)^2 \quad (3.10)$$

onde, α_i e β_i são dados por (3.7) e (3.8) respectivamente.

A quantidade de água disponível no reservatório para a diluição dos resíduos poluentes depende de fatores tais como, taxa de evaporação, quantidade de chuvas na região, vazão de entrada e da capacidade de armazenamento do reservatório que podem ser agrupados em um parâmetro $C_0 = 3.47$ (característica de cada reservatório). Portanto, a quantidade de água presen-

te no reservatório que será utilizada com outros propósitos que não a diluição dos resíduos é dada por

$$f_2(x) = C_0 - y \quad (3.11)$$

Sabe-se que o Decisor está interessado em manter a qualidade da água em um determinado ponto, abaixo do último usuário. Por exemplo, esta área pode ser usada para lazer ou pode estar próxima a regiões de fronteira, e portanto para se evitar problemas legais a qualidade da água deve ser mantida em níveis aceitáveis. Para a avaliação da qualidade da água utiliza-se a equação de Streeter-Phelps [Haimes et al., 1975] que relaciona a quantidade de BOD descarregada no rio, com o nível de DO em qualquer ponto do rio. Seguindo o modelo de Hass e associando a cada seção de rio um usuário, pode-se formular o último objetivo como

$$f_3(x) = \left[\sum_{i=1}^n a_i x_i + b x_{n+1} + c \right] / [d - \epsilon x_{n+1}] \quad (3.12)$$

onde a_i , b , c , d , e ϵ são constantes do modelo dadas. Este objetivo *substitui* o conjunto de restrições do modelo escalar, isto é, leva em conta as características de todas as seções. Em particular, a_i é diretamente proporcional à quantidade de BOD do *usuário* i .

As restrições para este modelo são as mesmas do modelo proposto por Hass,

$$\begin{aligned} 0.45 &\leq x_i \leq 0.99 & i = 1, \dots, n \\ 0 &\leq x_{n+1} \leq 3.47 \end{aligned}$$

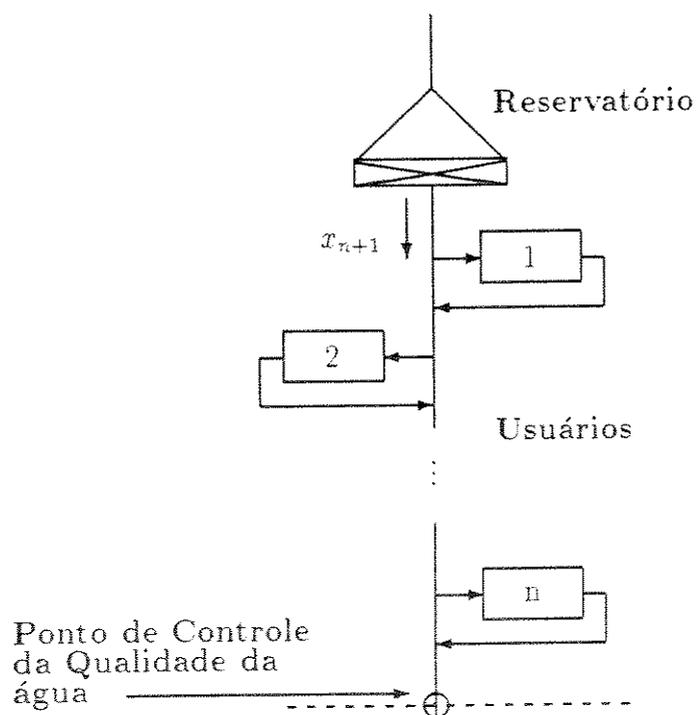


Figura 3.3: Um Reservatório e seus Usuários

O problema de poluição em rios é então formulado tendo em vista três objetivos:

- *Minimizar o custo total de tratamento*

$$\min_x f_1(x) = \sum_{i=1}^n a_i + \beta_i (x_i - 0.45)^2$$

- *Maximizar a quantidade de água no reservatório* favorecendo desta forma a realização de outros objetivos que dependem desta água (irrigação, navegação, lazer, pesca, etc).

$$\max_x f_2(x) = C_0 - x_{n+1}$$

- *Maximizar a quantidade de oxigênio dissolvido* num dado ponto do rio, após o último usuário

$$\max_x f_3(x) = [\sum_{i=1}^n a_i x_i + b x_{n+1} + c] / [d - \epsilon x_{n+1}]$$

A formulação multiobjetivo para este problema seria

$$\begin{aligned} \min_x \quad & f(x) = [f_1(x), f_2(x), f_3(x)] \\ \text{s.a.} \quad & 0.45 \leq x_i \leq 0.99 \quad i = 1, \dots, n \\ & 0 \leq x_{n+1} \leq 3.47 \end{aligned}$$

Em [Haimes et al., 1975] este problema foi resolvido através de método SWT (seção 1.5.2), para os valores numéricos apresentados na Tabela 3.1.

O desempenho do método SWT servirá de base para comparações com os métodos apresentados nos capítulos restantes deste trabalho.

3.5 Conclusão

Neste capítulo apresentou-se aspectos relativos ao modelamento do problema de Controle de Poluição em Rios. Apresentou-se como ocorre o processo de retirada de oxigênio da água por parte de microrganismos. Citou-se algumas formas de tratamento de resíduos na fonte poluidora e comentou-se que o aumento do fluxo de água pode diluir tanto resíduos orgânicos quanto inorgânicos. Os parâmetros que são usados para medir a qualidade da água em um determinado ponto do rio, concentração de BOD e concentração de DO, tiveram suas equações apresentadas.

Apresentou-se também dois modelos para resolução do problema, o modelo proposto por Hass (1970) e o modelo multiobjetivo proposto por Haimes et al (1975).

No próximo capítulo, a formulação multiobjetivo do problema de controle de poluição em rios será utilizada como exemplo de aplicação para a implementação do método de Geoffrion et al (1972) no Espaço dos Objetivos.

Usuário	a_i	q_i
1	3331.84	45.2
2	342.96	4.7
3	1539.69	4.2
4	886.02	3.6
5	73.83	0.5
6	172.06	1.2
7	189.40	0.8
8	433.07	0.6
9	199.94	0.5
10	1913.91	3.2
11	1741.59	8.4
12	722.94	2.7
13	238.59	0.6
14	3633.82	12.1
15	2266.56	8.4
	b=7650.31	c=5909.95
	d=2901.41	e=-971.27

Tabela 3.1: Tabela de Dados

Capítulo 4

Implementação do Método de Geoffrion no Espaço dos Objetivos

4.1 Introdução

O presente capítulo tem por objetivo discutir a implementação do método de Geoffrion no Espaço dos Objetivos a partir da abordagem descrita no Capítulo 2. O desempenho numérico do método será avaliado em termos do problema de poluição em rios descrito no Capítulo 3. Para a estimação das *Taxas Marginais de Substituição* (TMS's) por parte do Decisor, utiliza-se o método proposto por Dyer (1973), o qual se baseia em uma série de comparações ordinais. Os resultados obtidos são comparados com os dados existentes na literatura e importantes conclusões a respeito do desempenho global da implementação podem ser feitas.

4.2 Descrição do Método - Nível de Decisão

A partir dos resultados obtidos no Capítulo 2, torna-se possível implementar o método de Geoffrion no espaço dos objetivos. Com a implementação deste método, o problema relaxado (2.12) do Passo 2 do Algoritmo Básico, representado a seguir para uma iteração genérica k .

$$\min_{y \in \mathcal{Y}^k} \mathcal{U}(y),$$

poderá ser eficientemente resolvido, pois como será visto, a formulação no espaço dos objetivos leva a uma série de simplificações em relação ao procedimento original.

O algoritmo de Frank-Wolfe implementado segundo o método de Geoffrion no Espaço dos Objetivos pode ser descrito pela seguinte sequência de passos. Observe as várias simplificações decorrentes desta implementação.

Passo 1 *Determine uma solução inicial factível. Pode-se utilizar a última solução eficiente gerada pois $f(x^k) \in \Gamma \subset \mathcal{Y}^k$. Faça $l = 0$ e $y^0 = f(x^k)$*

Passo 2 *Obtenha uma solução ótima ξ^l para o problema de otimização*

$$\min_{\xi \in \mathcal{Y}^k} \langle \nabla_y \mathcal{U}(y^l), \xi \rangle \quad (4.1)$$

que pode ser reescrito como

$$\min_{\xi \in \mathcal{Y}^k} \langle \tau^l, \xi \rangle \quad (4.2)$$

onde

$$\tau_i^l = (\partial \mathcal{U} / \partial y_i) / (\partial \mathcal{U} / \partial y_1)$$

cujo valor pode ser aproximado por

$$\tau_i^l \approx \frac{y_1 - y_1^l}{y_i - y_i^l}, \quad i = 1, \dots, m \quad (4.3)$$

e que corresponde à obtenção da i -ésima Taxa Marginal de Substituição. Faça $d^i = \xi^l - y^l$

Passo 3 Determine a solução ótima do problema de busca unidimensional

$$\min_{t \in [0,1]} \mathcal{U}(y^l + td^l) \quad (4.4)$$

Passo 4 Faça $y^{l+1} = y^l + t^l d^l$. Se $\|y^{l+1} - y^l\| < \epsilon$, para $\epsilon > 0$ arbitrariamente pequeno pare: a solução y^{l+1} resolve 2.12. Caso contrário faça $l = l + 1$ e retorne ao Passo 2.

Assim como na versão para o Espaço das Decisões, a execução do Passo 3 do algoritmo no Espaço dos Objetivos envolve a resolução de um problema de busca unidimensional com a ajuda do Decisor. Só que neste caso, a estimativa do passo t^l pode ser facilmente obtida através do diagrama ilustrado na Figura 4.1 sem nenhum esforço adicional de cálculo. Note ainda que no Espaço dos Objetivos, os objetivos são sempre lineares, o que simplifica a obtenção de t^l . No Espaço das Decisões, a estimativa do passo t^l não é tão simples pois necessita-se calcular o valor de cada função $f_i(x)$, $i = 1, \dots, m$ no intervalo $t \in [0, 1]$.

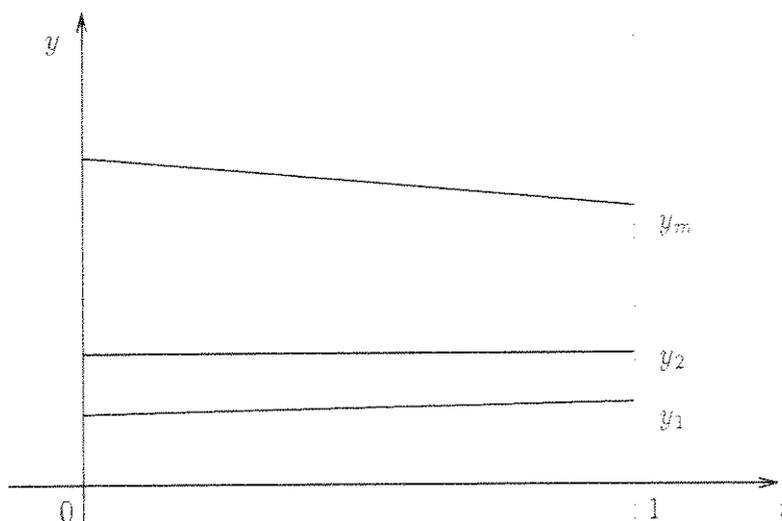


Figura 4.1: Escolha do Passo t

As Taxas Marginais de Substituição no Passo 2 do algoritmo podem ser estimadas de uma maneira sistemática como descrito na próxima seção.

4.3 Estimativa de TMS's através do Método de Dyer

Assumindo-se que o Decisor é capaz de estimar as suas TMS's entre os objetivos y_i , $i = 2, \dots, m$ e y_1 , sendo y_1 escolhido como objetivo de referência, a informação necessária para o seu cálculo pode ser obtida respondendo-se a seguinte questão: "Quanto seria necessário decrescer o objetivo y_i para compensar um acréscimo Δy_1 no objetivo y_1 ?"

$$\tau_i^k \approx \frac{\Delta y_1^k}{\Delta y_i^k}, \quad i = 1, \dots, m \quad (4.5)$$

Objetivando a estimativa das TMS's, Dyer [Dyer, 1973] propôs um esquema de comparações entre duas alternativas. O Decisor escolhe uma entre as alternativas ou mostra-se indiferente às mesmas. A informação fornecida pelo Decisor pode então ser utilizada para calcular o valor das TMS's.

Para a obtenção do trade-off τ_i^k para o i -ésimo objetivo ($i \neq 1$) compara-se o vetor y^k , representado por A com o vetor $(y_1^k + \Delta y_1^k, y_2^k, \dots, y_i^k - \Delta y_i^k, \dots, y_m^k)$, representado por B . Se o Decisor prefere B frente a A ($B \prec A$), aumenta-se Δy_i^k e repete-se a comparação e o acréscimo até que A seja preferido frente a B ($A \prec B$). Esta situação está ilustrada na Figura 4.2, onde a linha curva representa a curva de indiferença do Decisor relacionando os critérios 1 e i , passando através do ponto A .

Um algoritmo de bissecção ligeiramente modificado é aplicado ao valor de Δy_i^k com o objetivo de incrementá-lo se $A \prec B$, e decrementá-lo se $B \prec A$. Este procedimento é repetido até que o valor Δy_i^{k*} seja obtido, o que ocorre quando o Decisor mostra-se indiferente entre A e B ($A \sim B$), o que está ilustrada na Figura 4.3. Então, $\tau_i^k \approx \Delta y_1^k / \Delta y_i^{k*}$, e repete-se o procedimento até que todos os $m - 1$ trade-off's tenham sido calculados.

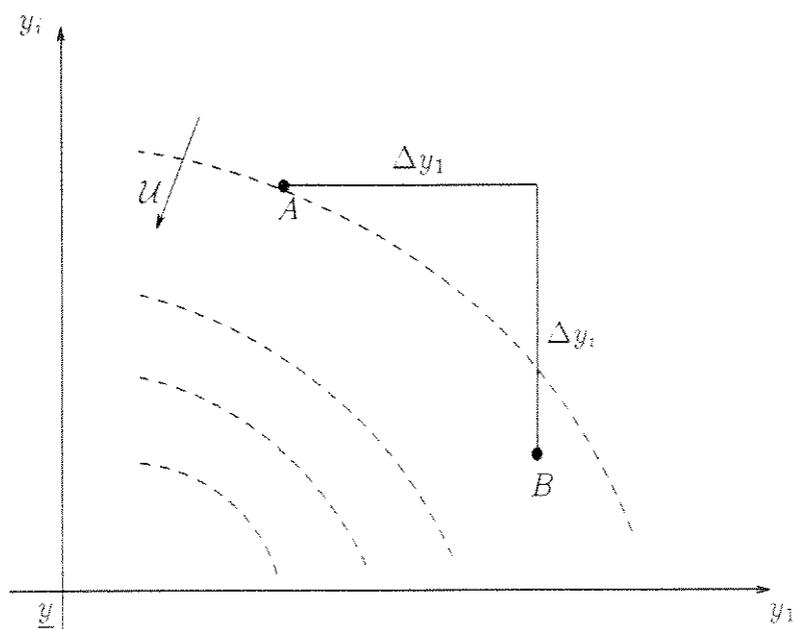


Figura 4.2: Escolha da Solução Preferida

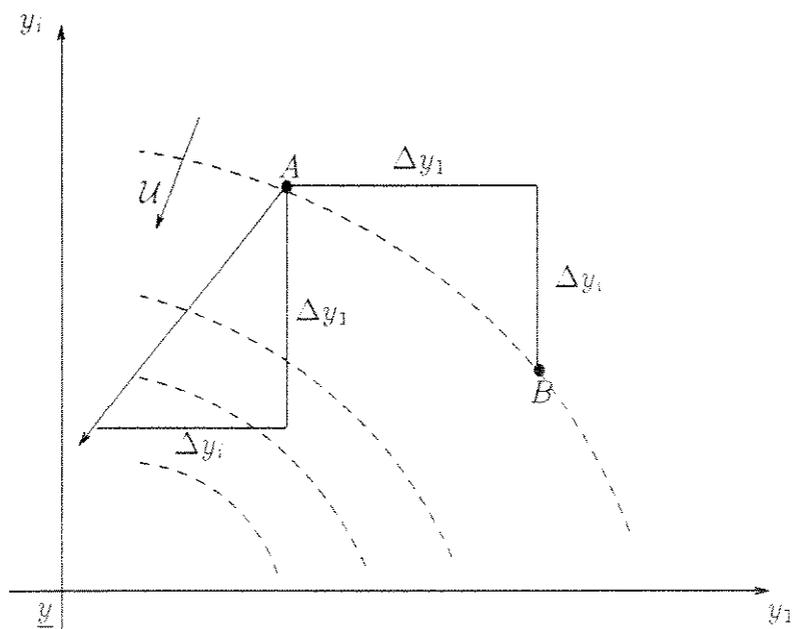


Figura 4.3: Indiferença entre Soluções

O algoritmo de Dyer para a estimativa dos trade-off's é apresentado a seguir.

Passo 1 Faça $A = y^k$ e $i = 2$.

Passo 2 Escolha Δy_1^k

Passo 3 Faça $j = 0$ e escolha Δy_i^j .

Passo 4 Faça $B = (y_1^k + \Delta y_1^k, \dots, y_i^k - \Delta y_i^j, \dots)$. Compare A com B .

Passo 5 Se $B \prec A$, faça $\Delta y_i^{j+1} = 2(\Delta y_i^j)$, $j = j + 1$, e vá para o Passo 4. Se $A \prec B$, faça $\theta^j = \Delta y_i^j$, e vá para o Passo 7. Se $A \sim B$ vá para o Passo 9.

Passo 6 Compare A e B . Se $A \prec B$ vá para o Passo 7. Se $B \prec A$ vá para o Passo 8. Se $A \sim B$ vá para o Passo 9.

Passo 7 Aumente a preferência por B fazendo, $\Delta y_i^{j+1} = \Delta y_i^j - \theta^j/2$. Faça $\theta^{j+1} = \theta^j/2$ e $j = j + 1$. Vá para o Passo 6.

Passo 8 Decresça a preferência por B fazendo, $\Delta y_i^{j+1} = \Delta y_i^j + \theta^j/2$. Faça $\theta^{j+1} = \theta^j/2$ e $j = j + 1$. Vá para o Passo 6.

Passo 9 Calcule $\tau_i^k = \Delta y_1^k / \Delta y_i^k$. Se $i \neq m$, faça $i = i + 1$ e vá para o Passo 3. Caso contrário termine o procedimento.

4.4 Aspectos de Implementação do Método

Como mencionado anteriormente, o nível de análise concentra a maior parte do esforço computacional do algoritmo básico devido a necessidade de resolução do problema min-max

$$\theta(y^k) = \max_{w \in W} \min_{x \in X} \langle w, f(x) - y^k \rangle$$

apresentado na seção 2.2.

Para a resolução do problema de minimização, utilizou-se o método do Gradiente Modificado [Geromel e Ferreira, 1986]. Como o problema em questão apresenta apenas restrições de bornes, importantes simplificações acerca do algoritmo deste método podem ser feitas. Neste caso, tem-se que $P = I$, $Q(x) = I$ e $c(x) = S(x)\nabla f(x)$.

Para a determinação do passo ótimo $\alpha \in R_+$, utilizou-se o procedimento de busca unidimensional de *Bisseção*. O número de iterações n necessário para a determinação do passo α deve satisfazer a seguinte equação.

$$0.5^n (b_1 - a_1) < \ell$$

onde $(b_1 - a_1)$ é o intervalo de incerteza inicial e ℓ é o intervalo de incerteza desejado. Utilizou-se o valor da norma euclidiana do gradiente modificado $\|\nabla_m f(x^k)\| < \epsilon_1$ como critério de parada para o algoritmo.

O problema de maximização baseado em Aproximação Tangencial foi solucionado através do método Primal Simplex. Como critério de parada para este procedimento utilizou-se a diferença entre os valores da função objetivo para os procedimentos de minimização e maximização: $(\sigma^l - \sigma^l) < \epsilon_2$.

Para testar a eficiência do método de Geoffrion no espaço dos objetivos foram desenvolvidas duas implementações básicas.

No primeiro algoritmo implementado (versão automática) utilizou-se uma função utilidade explícita do tipo norma l_2 , satisfazendo as condições de convexidade e não-decrescimento com cada objetivo que é apresentada a seguir:

$$U(y) = \sum_{i=1}^m \frac{1}{2} q_i^* (y_i - \underline{y}_i)^2 \quad (4.6)$$

onde $q_i^* = q_i / \sum_i q_i$; os valores de q_i , $i = 1, 2, 3$ foram gerados aleatoriamente através dos recursos do MATLAB. Em seguida, o problema de Decisão foi resolvido como um problema clássico de programação matemática.

O critério de parada utilizado para o nível de decisão foi o seguinte,

$$\frac{\|y^k - y^{k-1}\|}{\|y^k\| + 1} < \epsilon_3$$

Implementou-se também uma versão interativa onde o Decisor faz uso do método de Dyer (comparações binárias) para obter os trade-off's entre os objetivos. Os valores utilizados como incremento para os objetivos encontram-se na Tabela 4.2. Para a escolha do passo utilizou-se a discretização dos objetivos $f_i(x + td)$, $i = 1, \dots, m$ no intervalo $t \in [0, 1]$ dividido em 10 sub-intervalos. Nesta mesma versão simulou-se a participação do Decisor fazendo-se uso da função utilidade (4.6) para se indicar preferências entre as alternativas indicadas. Neste caso, o critério de parada utilizado para o nível de decisão foi o seguinte.

$$\frac{\langle w^i, y^i - y^{i-1} \rangle}{\langle w^1, y^1 - y^0 \rangle} < \epsilon_4$$

o mesmo utilizado por Hwang e Masud (1979). O objetivo de concentração de oxigênio dissolvido foi escolhido como referência, uma vez que conflita diretamente com os demais.

Para parar o algoritmo básico utilizou-se o seguinte critério:

$$\theta(y^k) < \epsilon_5.$$

As constantes utilizadas como tolerâncias para satisfazer as condições de convergência dos algoritmos implementados são apresentadas na Tabela 4.1.

Constantes					
ϵ_1	ϵ_2	ϵ_3	ϵ_4	ϵ_5	ℓ
0.01	0.01	0.01	0.01	0.01	0.01

Tabela 4.1: Tolerâncias para Convergência

Neste problema foi possível obter a faixa de variação dos objetivos a partir de análises de pior e de melhor caso. Especificamente,

$$\underline{y}_i = \min_{x \in X} f_i(x), \quad i = 1, \dots, m \quad (4.7)$$

e

$$\overline{y}_i = \max_{x \in X} f_i(x), \quad i = 1, \dots, m \quad (4.8)$$

Os valores obtidos são apresentados na Tabela 4.2.

i	y_i	\overline{y}_i	Δy_i
1	4.99	15.01	0.20
2	-3.47	0.00	0.20
3	-8.07	-4.78	0.20

Tabela 4.2: Faixa dos Objetivos e Incrementos

Todos os métodos referidos encontram-se implementados em linguagem C, tanto em máquinas compatíveis com IBM-PC quanto em estações de trabalho do tipo Sparc.

Na próxima seção, apresenta-se e analisa-se os resultados obtidos a partir da execução dos algoritmos referidos nesta seção.

4.5 Desempenho Numérico do Método

A partir dos resultados obtidos com a implementação dos algoritmos citados na seção anterior, é possível estabelecer comparações e avaliar o desempenho numérico do método em relação aos resultados disponíveis na literatura.

Utilizando-se o algoritmo com a função utilidade explícita, resolveu-se o problema de poluição em rios como um problema de programação matemática clássico, e chegou-se aos resultados expressos na Tabela 4.3, onde:

N_e : número de soluções eficientes geradas;

N_A : número de iterações do Algoritmo Básico.

q_1^*	q_2^*	q_3^*	w_1^*	w_2^*	w_3^*	f_1^*	f_2^*	f_3^*	N_ϵ	N_A
8.56	4.97	1.00	0.41	0.19	0.40	5.51	-3.18	-5.82	25	3
2.77	0.54	1.00	0.45	0.15	0.40	5.34	-2.22	-6.14	36	5
0.13	0.45	1.00	0.17	0.16	0.67	7.30	-2.45	-6.95	24	3
0.05	0.71	1.00	0.12	0.10	0.78	9.12	-3.37	-7.17	32	4
0.08	0.07	1.00	0.15	0.08	0.77	6.92	-0.84	-7.13	22	3

Tabela 4.3: Resultados para a Versão Automática

Na Tabela 4.3 as sequências de fatores de ponderação q_i^* , $i = 1, 2, 3$ encontram-se normalizados em relação ao terceiro objetivo. Nesta Tabela, ilustra-se apenas 5 casos de um total de 50 executados. Os resultados médios obtidos são apresentados na Tabela 4.4.

N_ϵ	\bar{N}_ϵ	μ_{N_ϵ}	σ_{N_ϵ}	N_A	\bar{N}_A	μ_{N_A}	σ_{N_A}
19	42	26.80	5.35	3	6	3.56	0.72

Tabela 4.4: Resultados Médios para a Versão Automática

Resolvendo-se o problema através dos métodos de Geoffrion e Dyer (versão interativa) e simulando-se a função utilidade do Decisor através da função (4.6), chegou-se aos resultados apresentados na Tabela 4.5, onde as sequências de fatores de ponderação utilizadas são as mesmas apresentadas na Tabela 4.3.

q_1^*	q_2^*	q_3^*	w_1^*	w_2^*	w_3^*	f_1^*	f_2^*	f_3^*	N_ϵ	N_A
8.56	4.97	1.00	0.36	0.21	0.43	5.53	-3.37	-5.71	37	5
2.77	0.54	1.00	0.33	0.17	0.50	5.47	-2.19	-6.27	36	5
0.13	0.45	1.00	0.21	0.13	0.67	7.59	-2.62	-6.99	22	3
0.05	0.71	1.00	1.00	0.00	0.00	8.79	-3.37	-7.11	27	5
0.08	0.07	1.00	0.16	0.11	0.73	8.16	-2.40	-7.17	26	3

Tabela 4.5: Resultados para a Versão Interativa

Os resultados médios obtidos para os mesmos 50 casos executados são apresentados na Tabela 4.6.

N_e	\bar{N}_e	μ_{N_e}	σ_{N_e}	N_A	\bar{N}_A	μ_{N_A}	σ_{N_A}
13	55	27.82	7.85	3	7	3.8	0.89

Tabela 4.6: Resultados Médios para a Versão Interativa

A Figura 4.4 ilustra a convergência do algoritmo de Aproximação Tangencial utilizado para a resolução do problema Min-Max, na segunda iteração do algoritmo básico, usando-se a sequência que mais se aproxima dos valores médios.

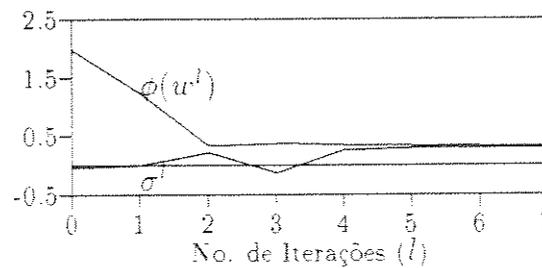
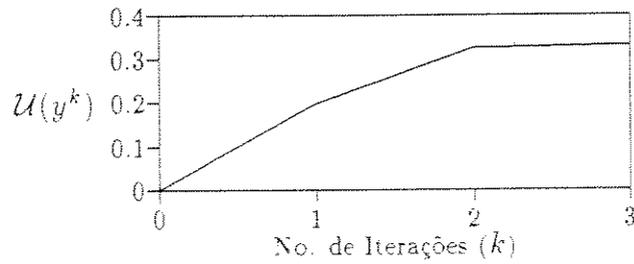
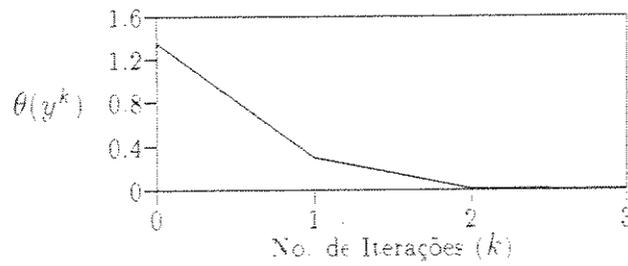


Figura 4.4: Convergência da Aproximação Tangencial

Para ilustrar a convergência do Algoritmo Básico, apresenta-se nas Figuras 4.5 e 4.6 o comportamento da função utilidade e da função θ , respectivamente, para a função utilidade que mais se aproxima dos valores médios.

Na Tabela 4.7, apresenta-se os valores das variáveis de decisão e dos objetivos obtidos para os 5 casos apresentados na Tabela 4.3, através da execução da versão automática do método. Apresenta-se também os resultados obtidos por Haimes et al (1975).

Figura 4.5: Convergência da Função Utilidade $\mathcal{U}(y)$ Figura 4.6: Convergência da Função $\theta(y)$

Var \ Caso	Q_{02}	Q_{23}	Q_{14}	Q_{34}	Q_{38}	Haimes
x_1	0.55	0.55	0.57	0.70	0.56	0.47
x_2	0.55	0.55	0.55	0.62	0.55	0.50
x_3	0.60	0.55	0.85	0.89	0.80	0.70
x_4	0.55	0.55	0.70	0.89	0.67	0.61
x_5	0.55	0.55	0.55	0.56	0.55	0.48
x_6	0.55	0.55	0.55	0.63	0.55	0.50
x_7	0.55	0.55	0.55	0.67	0.55	0.51
x_8	0.55	0.55	0.71	0.89	0.66	0.60
x_9	0.55	0.55	0.57	0.71	0.55	0.52
x_{10}	0.67	0.58	0.89	0.89	0.89	0.81
x_{11}	0.56	0.55	0.73	0.89	0.70	0.62
x_{12}	0.55	0.55	0.69	0.89	0.66	0.60
x_{13}	0.55	0.55	0.59	0.74	0.56	0.53
x_{14}	0.61	0.55	0.88	0.89	0.84	0.72
x_{15}	0.59	0.55	0.81	0.89	0.78	0.67
x_{16}	0.29	1.25	1.02	0.10	2.63	0.14
f_1	5.51	5.34	7.30	9.13	6.92	5.94
f_2	-3.18	-2.22	-2.45	-3.37	-0.84	-3.33
f_3	-5.82	-6.14	-6.95	-7.17	-7.13	-6.06

Tabela 4.7: Variáveis de Decisão e Objetivos

4.6 Comentários

Observando-se os resultados expressos nas tabelas 4.3, 4.4, 4.5 e 4.7, algumas conclusões importantes podem ser feitas.

- (a) Quanto ao número de soluções eficientes.

O número de soluções eficientes geradas é um dos critérios sugeridos por Cohon e Marks (1973) como medida de performance de um método multiobjetivo.

No trabalho de Haimes et al (1975), foram geradas 30 soluções eficientes a partir da aplicação do método SWT. Deve-se ressaltar que no referido trabalho não foi utilizada uma função utilidade explícita, o que em princípio dificulta comparações.

No presente trabalho, os números de soluções eficientes geradas pela versão interativa estão próximos à média e dentro da faixa do desvio padrão da versão automática. Comparando-se os valores mínimo, médio e máximo de soluções eficientes geradas com os resultados apresentados por Haimes et al (1975), nota-se que são da mesma ordem de grandeza, o que do ponto de vista de soluções eficientes viabiliza o método como uma forma simples e alternativa para tratar problemas multiobjetivos.

- (b) Quanto ao número de iterações do Algoritmo Básico.

O número de iterações médio foi de 3.56 com desvio padrão de 0.72. Este resultado indica claramente que o número de iterações do algoritmo básico praticamente independe das preferências do Decisor. Este resultado pode ser atribuído à capacidade do método em rapidamente identificar a região da superfície de trade-off de interesse para o Decisor, e assim localizar sua solução de compromisso.

- (c) Quanto à implementação interativa.

Na versão interativa foi empregada uma função utilidade como forma de se estabelecer comparações entre duas alternativas. Isto significa tão somente que o Decisor é capaz de *ordenar* as alternativas que lhe são apresentadas. Observe o pequeno aumento no número de iterações do algoritmo básico. Este aumento pode ser atribuído às imprecisões introduzidas no processo de decisão

que tem como consequência básica a geração de soluções que não produzem os cortes mais profundos do conjunto \mathcal{Y}^k .

- (d) Quanto à escolha dos incrementos.

A escolha dos incrementos Δy_i , $i = 1, 2, 3$ deve ser feita de forma tal que sejam pequenos em relação aos valores dos objetivos y_i , $i = 1, 2, 3$ mas grandes o suficiente para que o Decisor perceba as alterações nos valores dos objetivos. Isto representa uma das principais desvantagens do procedimento de Dyer, isto é, a necessidade do Decisor fornecer informações numéricas. Para solucionar este problema, implementa-se no próximo capítulo algoritmos com informações linguísticas.

- (e) Quanto ao valores das variáveis de decisão.

Observe que coerentemente com as sequências de ponderação apresentadas na Tabela 4.3, as variáveis de decisão da Tabela 4.7 refletem maior ou menor importância atribuída aos objetivos.

4.7 Conclusão

Neste capítulo abordou-se aspectos relativos à implementação do algoritmo de Geoffrion no espaço dos objetivos. Foram apresentadas as simplificações decorrentes desta formulação. O desempenho numérico do método foi analisado a partir dos resultados obtidos através da execução de suas versões automática e de Geoffrion/Dyer, interativa.

Os resultados obtidos demonstram a viabilidade computacional de implementações de métodos multiobjetivos no espaço dos objetivos, ao mesmo tempo em que evidenciam uma série de simplificações decorrentes da linearidade induzida no problema mestre pelas estratégias de relaxação e linearização externa.

A única dificuldade associada às implementações deste capítulo reside no cálculo das taxas marginais de substituição e principalmente na escolha dos incrementos para o método de Dyer. No próximo capítulo, algoritmos baseados em Teoria de Conjuntos Nebulosos são implementados com o objetivo de substituir as informações numéricas a serem fornecidas pelo Decisor, por informações linguísticas, mais naturais e facilmente manipuladas pelo Decisor.

Capítulo 5

Algoritmos Baseados em Conjuntos Nebulosos

5.1 Introdução

Como mencionado anteriormente, entre os métodos de Programação Multiobjetivo, os interativos são os que levam a resultados mais satisfatórios. A eficiência destes métodos depende basicamente da precisão das informações fornecidas pelo Decisor, que frequentemente são erráticas, inconsistentes e sujeitas a erros, devido a natureza imprecisa dos processos de decisão.

A *Teoria de Conjuntos e Algoritmos Nebulosos* introduzida por Zadeh [Zadeh, 1965] [Bellman e Zadeh, 1970] vem tendo aplicações em Programação Multiobjetivo desde o final da década de 70. O emprego deste tipo de abordagem é motivado principalmente pela sua capacidade de representação de conjuntos com limites imprecisos, frequentemente associados a julgamentos e avaliações subjetivas do Decisor.

Neste capítulo apresenta-se os conceitos básicos sobre conjuntos nebulosos e alguns algoritmos derivados da aplicação desta teoria [Ferreira e Machado, 1990a]. É dada ênfase aos algoritmos nebulosos propostos em [Baptistella e Ollero, 1980]. Propõe-se uma redefinição destes métodos em termos de norma infinito, e de acordo com esta nova formulação, algoritmos nebulosos interativo e não-interativo são implementados. Em seguida, compara-se os resultados obtidos com a solução nebulosa aos obtidos através da implementação do método de Geoffrion discutida no

capítulo anterior.

5.2 Conjuntos Nebulosos

Um subconjunto nebuloso A pode ser definido como uma classe de objetos com um contínuo grau de associação ou pertinência. O conjunto no qual cada elemento do subconjunto nebuloso A está definido é chamado de *Universo de Discurso* U . Um subconjunto nebuloso é caracterizado pela sua *Função Associação* $\mu_A(\cdot)$ que mapeia cada ponto de U num número real no intervalo $[0, 1]$. O valor de $\mu_A(\cdot)$ em x define o grau de associação de x em A .

Em conjuntos binários usuais, a Função Associação assume apenas dois valores, 0 ou 1, que representam as conhecidas relações de pertinência (\in) e não pertinência (\notin).

O *Conjunto Suporte* de A é o conjunto de pontos em U em que $\mu_A(\cdot) > 0$. Um *Ponto de Cruzamento* de um subconjunto nebuloso A é qualquer ponto do universo de discurso U em que $\mu_A(\cdot) = 0.5$. Um *Singleton Nebuloso* é um subconjunto nebuloso cujo suporte é um único ponto em U . Sua representação é dada por

$$A = \mu_A(x)x \quad (5.1)$$

ou

$$A = \mu_A(x)/x. \quad (5.2)$$

onde o símbolo “/” é utilizado para separar e identificar as componentes μ e x do termo μx . Um singleton de um conjunto não nebuloso pode ser representado por $1/x$.

Um conjunto nebuloso A em U pode ser representado pela união de seus singletons constituintes e denotado por

$$A = \int_U \mu_A(x)/x. \quad (5.3)$$

onde o sinal de integral representa união. Quando o suporte de A é um conjunto finito, é conveniente representar A como

$$\mu_1(x)/x_1 + \dots + \mu_m(x)/x_m \quad (5.4)$$

ou

$$\sum_{i=1}^m \mu_i(x)/x_i. \quad (5.5)$$

onde x_i , $i = 1, \dots, m$ representa o i -ésimo elemento do conjunto suporte de A e $\mu_i(x)$, $i = 1, \dots, m$ representa o seu respectivo grau de associação em A . O sinal '+' simboliza a união dos singletons de A . A representação de um conjunto nebuloso através de seu conjunto suporte pode ser entendida como a decomposição deste conjunto em seus singletons constituintes.

A partir destas definições é possível estabelecer as operações básicas sobre conjuntos nebulosos, tais como união, interseção, fuzzificação, etc. que servem de base para a construção de sentenças e algoritmos nebulosos. Por conveniência, estas operações são apresentadas no Apêndice A deste trabalho.

Baseando-se na premissa de que os elementos chaves do raciocínio humano não são números e sim classes de objetos em que a transição de pertinência para não-pertinência é gradual [Zadeh, 1973], pode-se fazer uso da teoria de conjuntos nebulosos em processos que envolvam avaliação subjetiva.

A capacidade humana de resumir informações encontra sua maior manifestação no uso de linguagens naturais. Cada termo x em uma linguagem natural L pode ser visto como a descrição resumida de um subconjunto nebuloso $M(x)$ de um universo de discurso U , com $M(x)$ representando o significado de x [Zadeh, 1973].

A construção de algoritmos nebulosos envolve definições sobre semântica nebulosa tais como, variáveis linguísticas, termos atômicos, delimitadores linguísticos e sentenças nebulosas, que também encontram-se no Apêndice A.

Existem ações que não são fáceis de representar através das operações sobre conjuntos nebulosos apresentadas ou referenciadas nesta seção. Em geral, estas ações são melhor representadas através do uso de algoritmos nebulosos.

5.3 Algoritmos Nebulosos

Um algoritmo nebuloso é definido como uma sequência de instruções ordenadas que podem conter atribuições e/ou declarações condicionais nebulosas.

Uma *Relação Nebulosa* é definida como um subconjunto nebuloso do produto cartesiano $\mathcal{X} \times \mathcal{Y}$, onde \mathcal{X} e \mathcal{Y} são dois conjuntos quaisquer.

$$R = \int_{\mathcal{X} \times \mathcal{Y}} \mu_R(x, y) / (x, y), \forall x \in \mathcal{X} \text{ e } y \in \mathcal{Y}$$

Declarações condicionais nebulosas como “se A então B”, onde A e B são termos com significado nebuloso, podem ser caracterizadas desde que sejam precisamente definidas como subconjuntos nebulosos específicos do universo de discurso a que pertencem. Uma *Declaração Condicional Nebulosa* pode ser interpretada como uma relação nebulosa entre as variáveis que a compõem.

$$\text{Se } A \text{ então } B \text{ senão } C = (A \times B) + (\neg A \times C), \quad (5.6)$$

onde A é um subconjunto nebuloso de um universo de discurso \mathcal{X} . “ \times ” entre os subconjuntos representa o produto cartesiano nebuloso (relação nebulosa) e “+” a união das relações nebulosas. Declarações deste tipo envolvem relações simples entre variáveis nebulosas. No caso de relações mais complexas de dependência de x com y , o uso de algoritmos nebulosos é mais apropriado.

Uma função nebulosa $f(x)$ pode ser definida através de um conjunto de declarações condicionais nebulosas, de uma maneira análoga a uma função não nebulosa, ou seja, através de tabelas de pares $(x, f(x))$.

Regra Composicional de Inferência : Se R é uma relação nebulosa de X para V , e x um subconjunto nebuloso de X , então o subconjunto y de V que é induzido por x , é dado pela composição de R e x .

$$y = x \circ R.$$

Os algoritmos nebulosos fornecem uma representação aproximada de uma grande variedade de fenômenos complexos devido à sua capacidade de tratar

imprecisões. Deste modo, eles podem fornecer um significado efetivo para descrições aproximadas de objetivos, restrições, desempenho de sistemas, estratégias, etc.

A execução de uma instrução nebulosa apresenta como resultado um conjunto nebuloso, e se for necessário a escolha de um único elemento deste conjunto como resposta, pode-se adotar várias estratégias. A estratégia mais comum é escolher o elemento de maior grau de associação no conjunto e no caso de ser mais de um, pode-se escolher um elemento de maior grau qualquer, ou utilizar um outro critério externo. Esta alternativa recebe o nome de *Regra da Alternativa Preponderante*.

5.4 Algoritmos Nebulosos em Programação Matemática

A teoria de conjuntos nebulosos vem sendo aplicada em otimização de sistemas desde 1975, usualmente associada a imprecisões quanto às restrições, recursos, parâmetros e variáveis de decisão do problema. Slowinski [Slowinski, 1985] apresenta um quadro comparativo de diferentes técnicas de otimização que empregam elementos nebulosos, propostas no período de 1975 a 1983.

Mais recentemente, em [Fabian e Stoica, 1984] apresenta-se uma formulação inteira determinística para o problema de Programação Inteira Nebulosa e um procedimento para resolução deste problema. Duas formulações para problemas de Programação 0-1 Linear Nebulosa são propostas em [Zimmermann e Pollastschek, 1978]. Em [Esogbue e Bellman, 1984] discute-se aplicações de Programação Dinâmica Nebulosa à otimização multicritério, sistemas de energia e processos de alocação de recursos, entre outras.

Aparentemente, o primeiro trabalho em programação multicritério envolvendo teoria de conjuntos nebulosos foi publicado por Zimmermann (1978) e trata apenas problemas lineares. Mostra-se que as soluções obtidas através de programação linear nebulosa são sempre eficientes.

Em [Baptistella e Ollero, 1980] propõe-se métodos iterativos que fazem um uso misto de técnicas, que têm comprovada eficiência em otimização dinâmica de grande porte com técnicas baseadas na teoria dos conjuntos nebulosos. São estudados dois diferentes tipos de procedimentos baseados

na hipótese do conhecimento local da função utilidade ser impreciso. No primeiro procedimento, assume-se que o Decisor pode estimar de um modo numérico aproximado os trade-off's locais entre os critérios. No segundo caso, assume-se que os trade-off's locais são estimados de um modo qualitativo através de termos linguísticos.

Como exemplo de algoritmo nebuloso, apresenta-se a seguir um dos procedimentos propostos em [Baptistella e Ollero, 1980]. Seja o problema escalar

$$\min_{x \in X} \alpha_1 f_1(x) + \dots + \alpha_m f_m(x) \quad (5.7)$$

$$\text{s.a. } Ax = b$$

$$\alpha \in \mathcal{A} \triangleq \left\{ \alpha_i \mid \alpha_i \geq 0, \sqrt{\sum_i (\alpha_i)^2} = 1 \right\}$$

Suponha que o Decisor pode estimar de um modo numérico aproximado o trade-off local τ_{ij} entre os critérios i e j . Por exemplo, " τ_{ij} deve estar na vizinhança de $\bar{\tau}_{ij}$ ", onde *vizinhança* é representada pela função

$$\mu_G(\tau_{ij}) = \left[1 + \left[\frac{\tau_{ij} - \bar{\tau}_{ij}}{\gamma_j} \right]^2 \right]^{-1}.$$

γ_j é um parâmetro de espalhamento que pode ser escolhido para representar as preferências do Decisor na vizinhança de $\bar{\tau}_{ij}$, e G é a meta nebulosa.

Seja f^k a solução do problema escalar equivalente e $\alpha_{1j}^k = \alpha_j^k / \alpha_1^k$, $j = 1, \dots, m$ os pesos relativos usados na sua determinação, onde α_1 é arbitrariamente escolhido como referência.

Determine o grau no qual a proporção de mudança "marginal ideal" entre os critérios 1 e j é correspondente à proporção ideal $\bar{\tau}_{1j}^k$, isto é,

$$\mu_{G_{1j}^k}(\alpha_{1j}^k) = \left[1 + \left[\frac{\alpha_{1j}^k - \bar{\tau}_{1j}^k}{\gamma_j} \right]^2 \right]^{-1}, \quad j = 1, \dots, m$$

Levando-se em conta que as metas nebulosas estão conectadas entre si por conectivos “e”, o efeito combinado total pode ser representado pela interseção dos subconjuntos nebulosos G_j^k , $G^k = G_2^k \cap G_3^k \cap \dots \cap G_m^k$.

$$\mu_{G^k} = \min [\mu_{G_2^k}(\alpha_{12}^k), \dots, \mu_{G_m^k}(\alpha_{1m}^k)].$$

O valor de μ_{G^k} representa o grau de satisfação do Decisor por f^k , que pode ser otimizado através de um algoritmo adequado. Na próxima seção, apresenta-se de forma mais detalhada os algoritmos nebulosos de Baptistella e Ollero, que utilizam estimativa qualitativa dos trade-off's locais.

5.5 Algoritmos de Baptistella e Ollero

Os métodos multiobjetivos interativos desenvolvidos por Baptistella e Ollero (1980) resolvem problemas do tipo

$$\min_{x \in \mathcal{X}} \mathcal{U}(f_1(x), \dots, f_m(x)) \quad (5.8)$$

$$\text{s.a. } Ax = b$$

$$\mathcal{X} \triangleq \{x \mid \underline{x}_i \leq x_i \leq \bar{x}_i, \quad i = 1, \dots, n\}. \quad (5.9)$$

Se o problema (5.8) apresenta propriedades de convexidade, ($f_i(x)$, $i = 1, \dots, m$, e \mathcal{X} convexos), pode-se utilizar a formulação equivalente,

$$\min_{x \in \mathcal{X}} \alpha_1 f_1(x) + \dots + \alpha_m f_m(x) \quad (5.10)$$

$$\text{s.a. } Ax = b$$

$$\alpha \in \mathcal{A} \triangleq \left\{ \alpha_i \mid \alpha_i \geq 0, \sqrt{\sum_i (\alpha_i)^2} = 1 \right\}$$

onde o Decisor fixa os pesos α_i , $i = 1, \dots, m$, durante o processo interativo.

Como comentado anteriormente, procedimentos interativos geralmente são imprecisos, e para tratar tais imprecisões Baptistella e Ollero propuseram quatro algoritmos que utilizam teoria de conjuntos nebulosos. O primeiro trata as informações aproximadas do Decisor de um modo numérico e foi apresentado resumidamente na seção anterior. O segundo procedimento utiliza resultados derivados de Álgebra Nebulosa [Dubois e Prade, 1978] para definir um algoritmo interativo. O terceiro e o quarto procedimentos são baseados na aplicação de algoritmos relacionais nebulosos utilizados quando o Decisor não está apto a fornecer estimativas numéricas mas apenas linguísticas de suas preferências. Sua essência é a hipótese de que o Decisor pode fornecer um *protocolo* com termos linguísticos qualitativos para modificação sequencial do problema escalar equivalente (5.10). Estes procedimentos são descritos a seguir.

Algoritmo Linguístico Não-Interativo

Este algoritmo baseia-se na síntese linguística de *Controladores Nebulosos* [Mandani e Procyk, 1978]. Seja x^k uma solução do problema escalar equivalente (5.10). Considera-se como entradas do controlador as funções objetivo $f_i(x^k)$, $i = 1, \dots, m$, e como saída os *Fatores de Mudança* c_i representando os pesos relativos entre os objetivos i e 1, $i = 2, \dots, m$, sendo o objetivo $f_1(x)$ escolhido como referência. Assume-se que o Decisor pode definir subconjuntos nebulosos dos universos de discurso de cada objetivo descrevendo por exemplo, os seguintes termos linguísticos básicos: *bom*, *aceitável*, *pobre* e *inaceitável*.

Os termos linguísticos c_i , $i = 2, \dots, m$ podem ser definidos heurística-mente como subconjuntos nebulosos dos seus valores prováveis (geralmente em percentuais de fatores de mudança). Por exemplo: *grande acréscimo*, *pequeno acréscimo*, *acrécimo desprezível*, *decrécimo desprezível*, *pequeno decréscimo* e *grande decréscimo*.

O protocolo consiste de um conjunto de regras heurísticas, objetivando inferir cada c_i , a partir de f_1 e f_i , $i = 2, \dots, m$. Estes protocolos podem ser representados através de tabelas.

Para cada $i = 2, \dots, m$, deve-se definir um protocolo contendo todas as r_i regras resultantes da combinação dos termos linguísticos básicos que definem

completamente um fator de mudança c_i . Essas regras são conectadas por “ou”, como segue:

Se F_1^1 e F_i^1 então c_i^1
 ou
 Se F_1^2 e F_i^2 então c_i^2
 ou
 \vdots
 ou
 Se $F_1^{r_i}$ e $F_i^{r_i}$ então $c_i^{r_i}$

onde F_1^j , F_i^j e c_i^j , $j = 1, \dots, r_i$, $i = 2, \dots, m$, são termos linguísticos associados aos valores de f_1 , f_i e c_i respectivamente.

Cada regra define uma relação nebulosa entre os valores imprecisos de f_1 e f_i , inferindo c_i^j , $j = 1, \dots, r_i$. Devido à presença do conectivo “ou”, o fator de mudança resultante será um conjunto nebuloso $c_i = c_i^1 \cup c_i^2 \cup \dots \cup c_i^{r_i}$. Como o fator de mudança requerido pelo algoritmo é determinístico, necessita-se de um procedimento para selecionar o valor mais apropriado dentro do conjunto nebuloso c_i . Neste trabalho, utiliza-se o valor que apresenta função associação máxima ou a média aritmética dos valores máximos, caso existam “picos” ou “platôs”. Outras formas de se transformar um número nebuloso em um escalar são sugeridas em [Zadeh, 1968].

O algoritmo nebuloso não interativo é descrito a seguir.

Passo 1 Encontre $\underline{f}_i = f_i(x^i)$, $i = 1, \dots, m$ onde x^i resolve

$$\min_{x \in X} f_i(x), \quad i = 1, \dots, m,$$

e se possível o valor máximo de cada objetivo, \overline{f}_i , $i = 1, \dots, m$.

Passo 2 Interaja com o Decisor para obter: subconjuntos nebulosos do universo de discurso para cada critério, $i = 1, \dots, m$, que têm associados termos linguísticos descritivos; todas as regras para se definir completamente a política de decisão (protocolos), inferindo o fator de mudança c_i ; e o suporte dos subconjuntos nebulosos c_i que definem as magnitudes dessas mudanças.

Passo 3 Escolha um vetor de ponderação inicial $\alpha^0 = (1, \alpha_2^0, \dots, \alpha_m^0)$ com $\alpha_i^0 > 0$, $i = 2, \dots, m$, e normalize-o. Faça $k = 0$.

Passo 4 Resolva o problema escalar equivalente (5.10) com α^k e obtenha a solução x^k e $f_i(x^k)$, $i = 1, \dots, m$.

Passo 5 Para cada critério, $i = 2, \dots, m$, determine o conjunto nebuloso associado com o fator de mudança c_i^j , $j = 1, \dots, r_i$, através dos protocolos definidos no Passo 4. Como mais de uma regra pode contribuir para o cálculo da decisão, faça $c_i = c_i^1 \cup c_i^2 \cup \dots \cup c_i^{r_i}$. Selecione o valor determinístico de c_i .

Passo 6 Teste o critério de parada ($c_i = 1$, $i = 2, \dots, m$). Caso não satisfeito, faça $\alpha^{k+1} = (1, \alpha_2^k c_2, \dots, \alpha_m^k c_m)$, normalize-o, faça $k = k + 1$ e volte ao Passo 4.

As regras que inferem os fatores de mudança nos protocolos têm que refletir *coerência* do Decisor, uma característica essencial para que o procedimento acima convirja, conforme discutido em [Baptistella e Ollero, 1980].

Procedimentos que utilizam esquemas de aprendizagem adaptativa podem ser acrescentados a este algoritmo através de interação com o Decisor a cada passo.

Algoritmo Linguístico Interativo

Neste algoritmo assume-se que o Decisor é acionado em cada passo da sequência anterior para formular um diagnóstico sobre os valores de $f_i(x^k)$, $i = 1, \dots, m$, através do uso de *delimitadores linguísticos* sobre os termos básicos dos protocolos, por exemplo, “ f_1 é muito bom” e “ f_1 é mais ou menos aceitável”. A Regra Composicional de Inferência de Zadeh permite encontrar o subconjunto nebuloso c_i , $i = 2, \dots, m$ correspondente a esta regra.

Ao contrário do algoritmo não interativo, os termos básicos usados nos protocolos podem ser subconjuntos nebulosos com um suporte finito no intervalo $[0,1]$ ao invés do intervalo $[\underline{f}_i, \overline{f}_i]$, $i = 1, \dots, m$. O procedimento interativo é descrito nos seguintes passos.

Passo 1 Defina os termos linguísticos básicos, os delimitadores linguísticos e os protocolos para inferir c_i , $i = 2, \dots, m$.

Passo 2 Escolha um vetor inicial $\alpha^0 = (1, \alpha_2^0, \dots, \alpha_m^0)$ e normalize-o. Faça $k = 0$.

Passo 3 Resolva o problema escalar equivalente (5.10) com α^k e obtenha x^k e $f_i(x^k)$, $i = 1, \dots, m$.

Passo 4 Apresente $f_i(x^k)$ ao Decisor que realiza então o diagnóstico linguístico nebuloso para cada $i = 1, \dots, m$, através do uso combinado de delimitadores linguísticos e termos linguísticos básicos.

Passo 5 Encontre os subconjuntos nebulosos que descrevem o diagnóstico linguístico através do uso da regra composicional de Zadeh, infira os subconjuntos nebulosos c_2, \dots, c_m e seus valores determinísticos associados.

Passo 6 Teste o critério de parada. Se não for satisfeito, faça $\alpha^{k+1} = (1, \alpha_2^k c_2, \dots, \alpha_m^k c_m)$, normalize-o, faça $k = k + 1$ e volte ao Passo 2.

Este algoritmo apresenta uma característica adaptativa ou de auto-aprendizagem, pois o Decisor pode modificar o diagnóstico linguístico dos resultados obtidos nas interações precedentes.

5.6 Implementação no Espaço dos Objetivos

Segundo a estrutura de dois níveis utilizada neste trabalho para a resolução do PMO, é possível interagir com o Decisor no espaço dos objetivos, evitando desta forma todas as referências às variáveis de decisão do problema. Como esta estrutura permite o acoplamento de várias formas de participação do Decisor, é possível utilizar os métodos linguísticos propostos em [Baptistella e Ollero, 1980], apresentados na seção anterior. Deste modo, a complexidade da interação com o Decisor será reduzida, pois estimar seus trade-off's de forma qualitativa (linguística) é bem mais simples do que de forma quantitativa (numérica).

Foram implementados os dois algoritmos nebulosos linguísticos (não iterativo e iterativo) propostos por Baptistella e Ollero.

A versão do algoritmo linguístico não-iterativo para o espaço dos objetivos, em uma iteração genérica k do algoritmo básico consiste dos seguintes passos:

Passo 1 Inicialize $y^0 = f(x^k)$, $\alpha_i^0 = w_i^k/w_r^k$, $i = 1, \dots, m$ ($w_r^k \neq 0$), onde r é o índice de um objetivo de referência. Faça $l = 0$.

Passo 2 Para cada critério, y_i^l , $i = 1, \dots, m$, use os protocolos definidos previamente e determine os conjuntos nebulosos associados aos fatores de mudança c_j^l , $j = 1, \dots, r_i$. Como mais do que uma regra pode contribuir para o cálculo da decisão, faça $c_i = c_1^l \cup c_2^l \cup \dots \cup c_{r_i}^l$. Selecione o valor determinístico de c_i .

Passo 3 Teste o critério de parada. Se não for satisfeito, faça $\alpha^{l+1} = [\alpha_1^l c_1, \alpha_2^l c_2, \dots, \alpha_{r-1}^l c_{r-1}, 1, \alpha_{r+1}^l c_{r+1}, \dots, \alpha_m^l c_m]$, e normalize-o. Faça $l = l + 1$.

Passo 4 Resolva o problema linear

$$\min_{y \in Y^k} \langle \alpha^l, y \rangle \quad (5.11)$$

obtendo y^l e vá para o Passo 2.

No caso do algoritmo nebuloso linguístico iterativo, os passos seriam os seguintes:

Passo 1 Inicialize $y^0 = f(x^k)$, $\alpha_i^0 = w_i^k/w_r^k$, $i = 1, \dots, m$ ($w_r \neq 0$), onde r é o índice de um objetivo de referência. Faça $l = 0$.

Passo 2 Apresente y^l ao Decisor que fará o diagnóstico linguístico para cada $i = 1, \dots, m$ combinando os termos linguísticos básicos e os delimitadores linguísticos.

Passo 3 Encontre os subconjuntos nebulosos que descrevem o diagnóstico linguístico através do uso da Regra Composicional de Zadeh. Infira os subconjuntos nebulosos c_1^l, \dots, c_m^l , $i \neq r$ e seus valores determinísticos associados.

Passo 4 Teste o critério de parada. Se não for satisfeito, faça $\alpha^{l+1} = [\alpha_1^l c_1, \alpha_2^l c_2, \dots, \alpha_{r-1}^l c_{r-1}, 1, \alpha_{r+1}^l c_{r+1}, \dots, \alpha_m^l c_m]$, e normalize-o. Faça $l = l + 1$.

Passo 5 Resolva o problema linear

$$\min_{y \in Y^k} \langle \alpha^l, y \rangle$$

obtendo y^l e vá para o Passo 2.

A Figura 5.1 ilustra a relação entre os subconjuntos nebulosos associados ao universo de discurso de cada objetivo, no espaço dos objetivos.

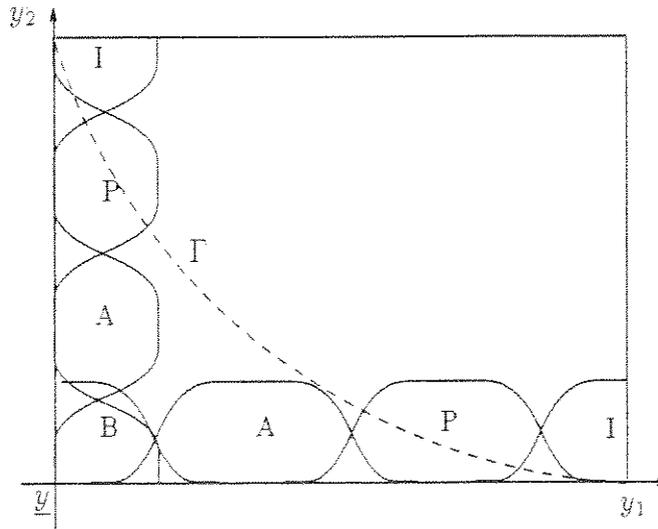


Figura 5.1: Subconjuntos dos Universos de Discurso

É claro que numa primeira iteração será selecionado o ponto $y^0 = \underline{y}$, uma vez que este ponto fornece o maior grau de associação com o conjunto *bom* em ambos os objetivos. Como em geral $y^0 = \underline{y}$ não é factível, o nível de análise fornecerá como solução (u^1, x^1) tal que $y^1 = f(x^1) \in \Gamma$. Conforme mencionado em [Ferreira e Geromel, 1990], esta solução tem como propriedade minimizar a norma infinito (Tchebycheff) entre \underline{y} e a curva de trade-off Γ (Fig. 5.2).

Do ponto de vista de análise no espaço dos objetivos, o que os métodos de Baptistella e Ollero fazem é rotacionar o vetor de ponderações (Fig. 5.3) através dos fatores de mudança c_i , de modo a refletir preferências do Decisor em relação aos subconjuntos nebulosos associados aos universos de discurso.

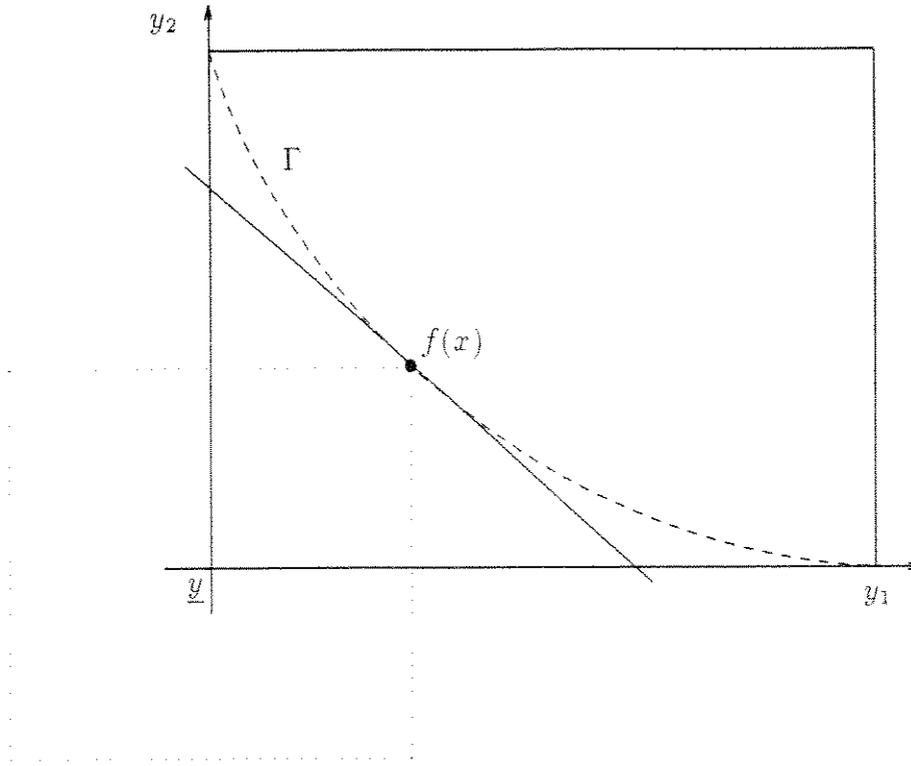


Figura 5.2: Primeira iteração do Algoritmo Básico

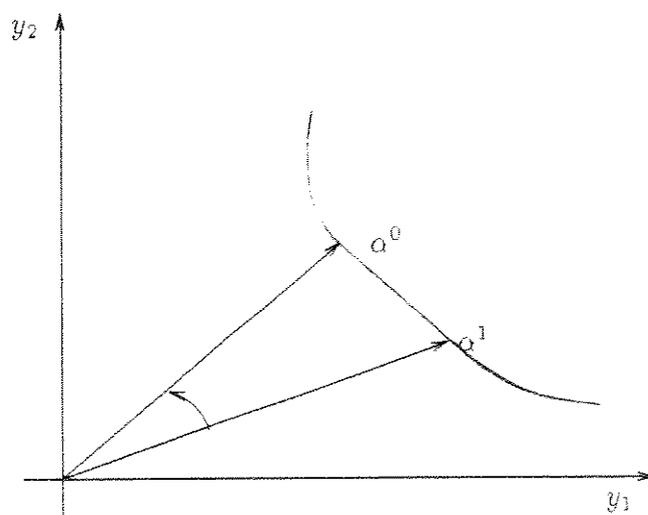


Figura 5.3: Efeito do Fator de Mudança

Entretanto, a implementação dos algoritmos nebulosos de Baptistella e Ollero no espaço dos objetivos não produziu bons resultados, por apresentar grandes oscilações em torno dos fatores de mudança, lenta convergência, e em alguns casos, não convergência no nível de decisão. Estes fatos, já observados por Sakawa (1984) em implementações no espaço das variáveis de decisão podem ser associados a situações em que a solução do problema multiobjetivo se encontra numa região de Γ em que exista uma lenta variação de trade-off's, o que exigiria grande precisão no cálculo dos fatores de mudança para produzir exatamente o vetor normal α . Esta situação é particularmente frequente em problemas multiobjetivos lineares, como as implementações no espaço dos objetivos.

A partir destas observações, buscou-se uma redefinição do problema escalar associado aos algoritmos de Baptistella e Ollero que é discutida na próxima seção.

5.7 Redefinição do Problema Escalar

Devido a lenta convergência apresentada pelos métodos linguísticos de Baptistella e Ollero, propõe-se nesta seção uma reformulação do problema escalar

$$\min_{y \in \mathcal{Y}^k} \langle \alpha, y \rangle$$

em termos de Norma Infinito (Tchebycheff) Generalizada [Bowman, 1975]. A norma infinito generalizada de um vetor $x \in R^n$ é definida como

$$\|x\|_\alpha = \max_{1 \leq i \leq n} \{\alpha_i |x_i|\}, \quad \alpha \in R^n, \quad \alpha > 0$$

O problema escalar associado fica então sendo

$$\min_{y \in \mathcal{Y}^k} \max_{1 \leq i \leq m} \{\alpha_i (y_i - \underline{y}_i)\} \quad (5.12)$$

cuja interpretação geométrica é fornecida na Figura 5.4.

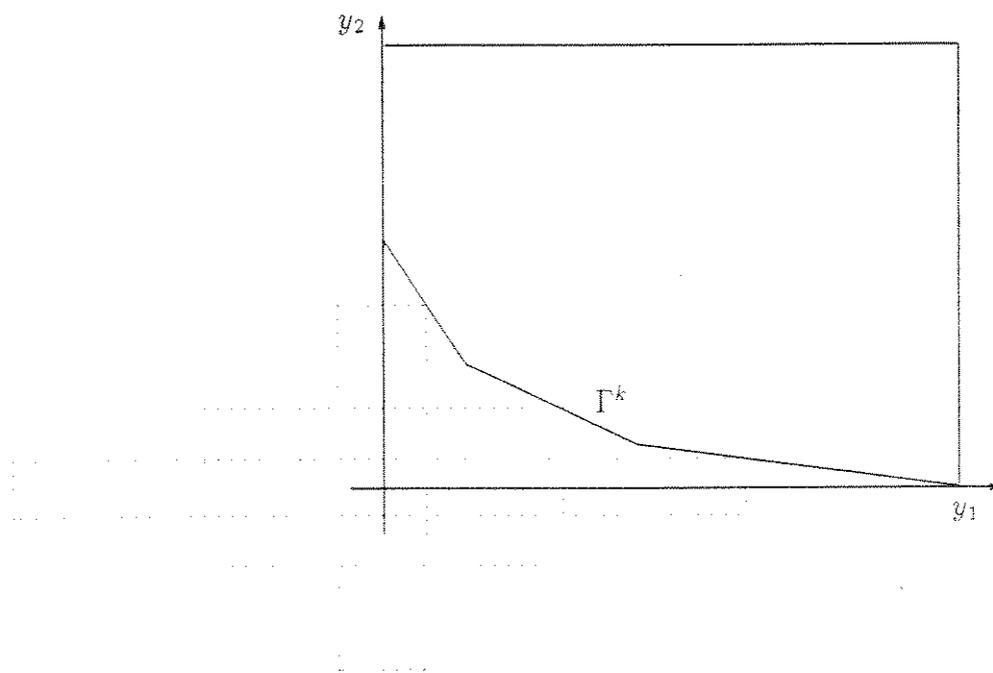


Figura 5.4: Redefinição em termos de Norma Infinito

Assim sendo, qualquer ponto associado à curva de soluções eficientes no espaço dos objetivos pode ser facilmente determinado através de uma escolha adequada de ponderações α_i , $i = 1, \dots, m$. Note ainda que o problema anterior pode ser visto como

$$\begin{aligned} \min_{y \in \mathcal{Y}^k, \sigma} \quad & \sigma & (5.13) \\ \text{s.a.} \quad & \sigma \geq \alpha_i (y_i - \underline{y}_i), \quad i = 1, \dots, m \end{aligned}$$

e portanto, o problema escalar permanece linear. O mesmo não ocorre, em geral, se a implementação através da norma infinito generalizada fosse feita no espaço das variáveis de decisão. Observe finalmente que a redefinição do problema escalar através de *qualquer* norma é imediata e deixa o problema muito simples no espaço dos objetivos, o que mais uma vez não ocorre, em geral, no espaço das decisões.

5.8 Aspectos de Implementação

Os algoritmos nebulosos de Baptistella e Ollero (1980), implementados no espaço dos objetivos de acordo com a reformulação em termos de norma infinito generalizada, foram aplicados ao problema multiobjetivo de controle de poluição de rios discutido nos Capítulos 3 e 4. Mais uma vez, o objetivo 3 foi adotado como objetivo de referência.

Por uma questão de simplicidade todos os subconjuntos nebulosos utilizados foram representados através de funções triangulares e/ou trapezoidais (Fig. 5.5).

No *Algoritmo Linguístico Não-Iterativo*, os termos linguísticos básicos utilizados pelo Decisor para cada critério foram os seguintes: Bom (B), Aceitável (A), Pobre (P) e Inaceitável (I).

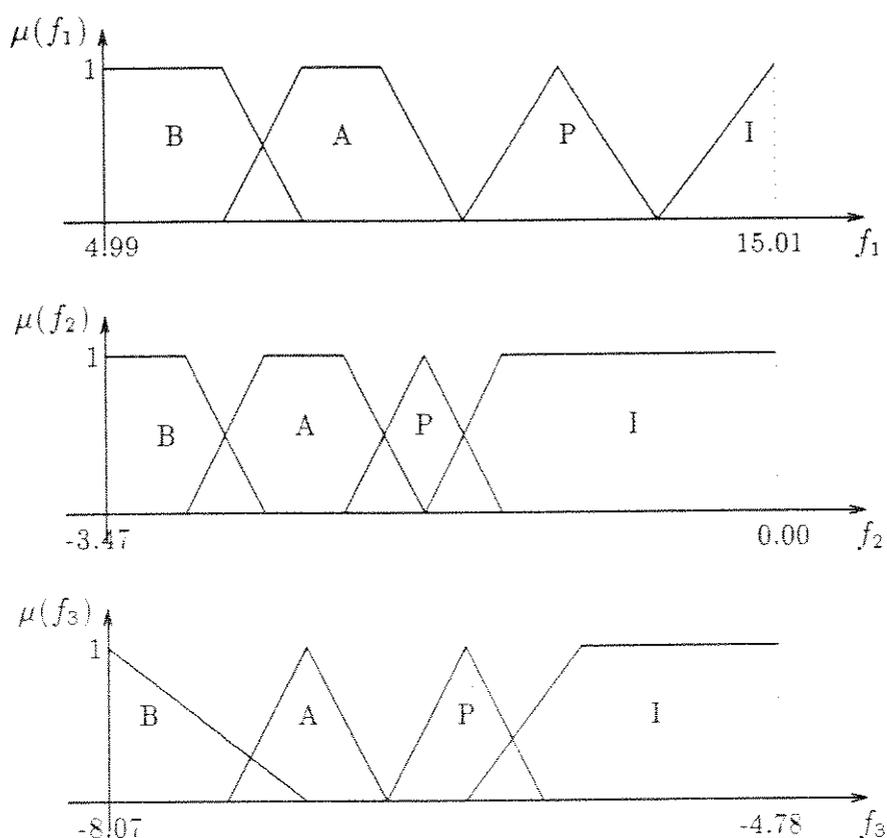


Figura 5.5: Termos Linguísticos Básicos

Os protocolos utilizados para inferir os fatores de mudança a partir dos termos linguísticos básicos estão apresentados na Figura 5.6.

Os termos linguísticos associados aos c_i 's foram os seguintes subconjuntos nebulosos: Decréscimo Desprezível (DD), Acréscimo Desprezível (AD), Pequeno Decréscimo (PD), Pequeno Acréscimo (PA), Grande Decréscimo Desprezível (GD) e Grande Acréscimo (GA), apresentados na Figura 5.7.

		f_1			
		B	A	P	I
f_3	B	AD	AD	PA	GA
	A	PD	DD	PA	PA
	P	PD	PD	PD	PA
	I	GD	GD	PD	PD

		f_2			
		B	A	P	I
f_3	B	AD	AD	PA	GA
	A	PD	DD	PA	PA
	P	PD	PD	PD	PA
	I	GD	GD	PD	PD

Figura 5.6: Protocolos

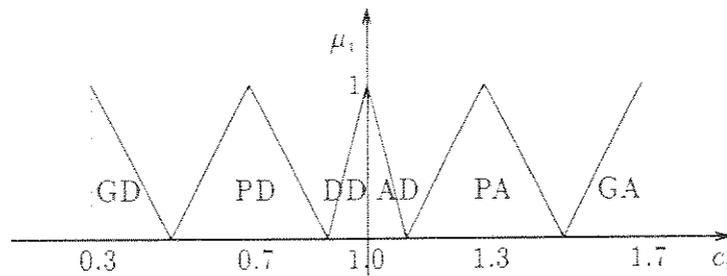


Figura 5.7: Fatores de Mudança

No caso do *Algoritmo Linguístico Iterativo*, os termos linguísticos básicos são os mesmos utilizados no algoritmo não-iterativo, mas com o suporte dos subconjuntos nebulosos representados no intervalo $[0, 1]$.

Os delimitadores linguísticos utilizados foram os seguintes. Não (NA), Muito (MT), Mais ou Menos (MM), Altamente (AL) e Não Muito (NM).

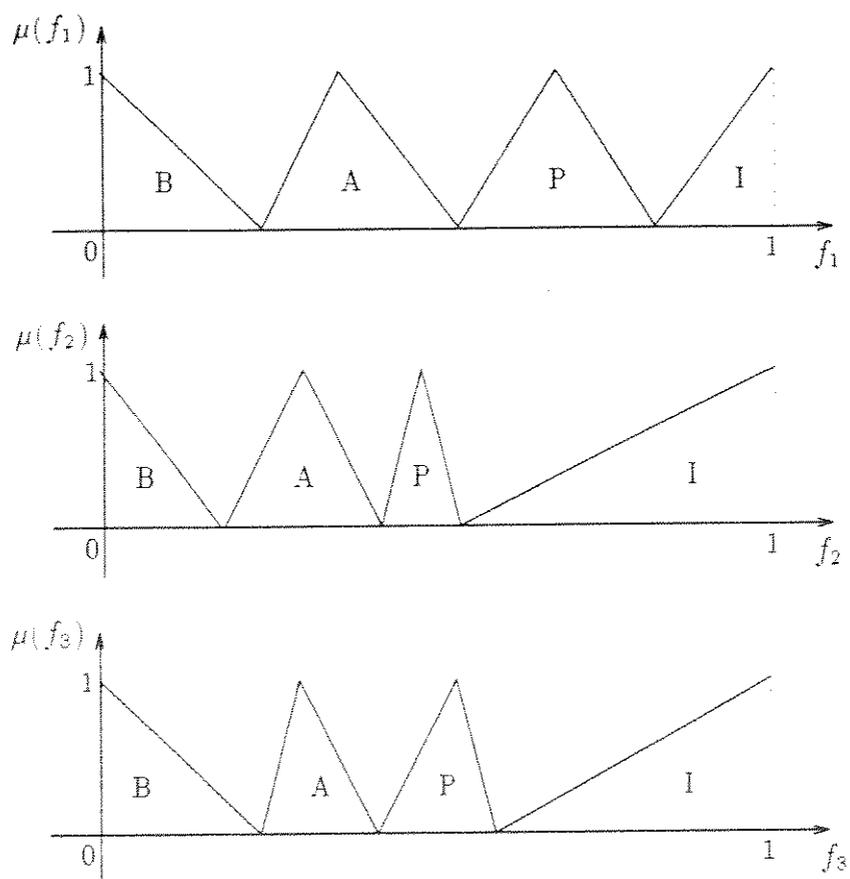


Figura 5.8: Termos Linguísticos Normalizados

Os protocolos e o suporte para os subconjuntos nebulosos c_i 's são os mesmos utilizados no algoritmo não-iterativo.

O critério de parada no nível de Decisão para os dois algoritmos nebulosos incorporados foi o seguinte:

$$c_i = 1.0, \quad i = 1, 2.$$

Como exemplo de execução de alguns passos do algoritmo nebuloso não-iterativo, suponha que numa iteração genérica l obtenha-se

$$y^l = \begin{bmatrix} 7.60 \\ -2.20 \\ -6.00 \end{bmatrix} \quad (5.14)$$

como solução fornecida pelo Analista ao Decisor. Utilizando-se os termos linguísticos definidos pelo Decisor (Fig. 5.8) tem-se:

$$\begin{aligned} y_1 &: (B) \text{ e } (A); \\ y_2 &: (A); \\ y_3 &: (P) \text{ e } (I). \end{aligned}$$

Consultando-se os protocolos (Figura 5.6) pode-se fazer a seguinte inferência:

• Protocolo 1:

1. y_1 é (B) e y_3 é (P) $\rightarrow c_1^1$ sofre um PD;
2. y_1 é (A) e y_3 é (P) $\rightarrow c_1^2$ sofre um PD;
3. y_1 é (B) e y_3 é (I) $\rightarrow c_1^3$ sofre um GD;
4. y_1 é (A) e y_3 é (I) $\rightarrow c_1^4$ sofre um GD.

Como se pode observar, 4 regras estão ativas e portanto o conjunto nebuloso resultante c_1 será a união dos conjuntos c_1^j , $j = 1, \dots, 4$, representada pela operação máximo. O conjunto nebuloso resultante apresenta dois picos. Utilizou-se a média aritmética dos pontos associados aos picos para inferir o fator de mudança resultante, fornecendo $c_1=1.5$.

• Protocolo 2:

1. y_2 é (A) e y_3 é (P) $\rightarrow c_2^1$ sofre um PD;
2. y_2 é (A) e y_3 é (I) $\rightarrow c_2^2$ sofre um GD;

Como se pode observar, 2 regras estão ativas e portanto o conjunto nebuloso resultante c_2 será a união dos conjuntos c_2^j , $j = 1, 2$ representada pela operação máximo. O conjunto nebuloso resultante apresenta dois picos. Utilizou-se a média aritmética dos pontos associados aos picos para inferir o fator de mudança resultante, fornecendo $c_2=1.5$. Como $c_1, c_2 \neq 1$, o critério de parada não é satisfeito, e portanto tem-se $\alpha^{l+1} = (\alpha_1^l c_1, \alpha_2^l c_2, 1)$. faz-se $l = l + 1$, normaliza-se α^l , resolve-se o problema escalar (5.12) e apresenta-se novamente a solução encontrada ao Decisor.

No caso do algoritmo linguístico iterativo considerando-se o mesmo vetor y^l , os passos são os seguintes:

Diagnóstico do Decisor:

- y_1 : MT.A (Muito Aceitável);
 y_2 : MM.B (Mais ou Menos Bom);
 y_3 : MT.P (Muito Pobre).

A partir dos termos básicos do diagnóstico e dos protocolos, tem-se as seguintes regras de inferência:

1. y_1 é (A) e y_3 é (P) então c_1 sofre um (PD);
2. y_2 é (B) e y_3 é (P) então c_2 sofre um (PD);

Pode-se montar as matrizes relacionais a partir do produto cartesiano entre os subconjuntos nebulosos envolvidos em cada regra.

$$R_1^1 = (A) \times (PD) \quad (5.15)$$

$$R_1^2 = (P) \times (PD) \quad (5.16)$$

$$R_2^1 = (B) \times (PD) \quad (5.17)$$

$$R_2^2 = (P) \times (PD) \quad (5.18)$$

Os subconjuntos nebulosos dos fatores de mudança, c_1 e c_2 , podem ser inferidos através da composição entre os conjuntos nebulosos resultantes do diagnóstico e as matrizes relacionais formadas.

$$c_1 = [(MT.A) \circ R_1^1] \wedge [(MT.P) \circ R_1^2] \quad (5.19)$$

$$c_2 = [(MM.B) \circ R_2^1] \wedge [(MT.P) \circ R_2^2] \quad (5.20)$$

Os valores determinísticos dos fatores de mudança podem ser obtidos pelo elemento de máximo grau de associação no conjunto nebuloso, no caso, $c_1 = 0.7$, $c_2 = 0.7$. Como $c_1, c_2 \neq 1$, o critério de parada não é satisfeito, e portanto tem-se. $\alpha^{l+1} = (\alpha_1^l c_1, \alpha_2^l c_2, 1)$. faz-se $l = l + 1$. normaliza-se α^l , resolve-se o problema escalar (5.12) e apresenta-se novamente a solução encontrada ao Decisor.

Para a resolução do problema em questão, inicialmente utilizou-se os protocolos definidos em [Baptistella e Ollero, 1980]. Como não ocorreu a convergência dos fatores de mudança, alterou-se gradativamente, estes protocolos até se atingir a convergência desejada. Todo este trabalho se faz necessário, pois protocolos muito restritivos podem levar a situações não factíveis. Note que a condição de término dos algoritmos ocorre quando a aplicação das regras dos protocolos fornecem termos linguísticos associados aos fatores de mudança do tipo *Acréscimo Desprezível*, *Decréscimo Desprezível*. Obviamente se o Decisor associar estes termos apenas quando os objetivos forem bons, por exemplo, então possivelmente não ocorrerá convergência, uma vez que soluções deste tipo podem ser infactíveis.

Isto demonstra que ao definir seus protocolos, o Decisor deve ter claro o que pode ser considerado Bom, Aceitável, Pobre e Inaceitável para um critério em relação aos demais. A não convergência do procedimento é uma evidência de que o Decisor deve relaxar seus protocolos, em função da impossibilidade estrutural de atendê-los na forma pré-estabelecida.

Este problema está presente nos dois algoritmos nebulosos implementados, sendo que para o algoritmo interativo é facilmente contornado, pois a cada interação o Decisor pode notar que a convergência não ocorre e *relaxar* um pouco sua decisão, permitindo assim a convergência, o que não ocorre no caso não-interativo.

No caso do algoritmo iterativo, simulou-se o diagnóstico do Decisor através da definição de Delimitadores Linguísticos (AL, MT, MM, NM e NA) em valores percentuais sobre cada um dos subconjuntos nebulosos (B, A, P e I) definidos para os objetivos, com a finalidade de testar a eficiência deste algoritmo sem incorrer em incoerências do Decisor.

Na próxima seção compara-se os resultados apresentados pelos algoritmos nebulosos frente aos não-nebulosos.

5.9 Resultados Numéricos e Comentários

A partir dos resultados obtidos com a implementação dos algoritmos citados na seção anterior é possível estabelecer comparações e avaliar desempenhos numéricos.

As Tabelas 5.1 e 5.2 apresentam os resultados para a execução dos algoritmos nebulosos não-iterativo e iterativo, respectivamente.

No caso iterativo foram utilizados os seguintes valores percentuais para simular o diagnóstico:

Altamente: entre 0 e 5% do subconjunto considerado:

Muito: entre 5 e 20% do subconjunto considerado:

Mais ou Menos: entre 20 e 80% do subconjunto considerado:

Não Muito: entre 80 e 95% do subconjunto considerado:

Não: entre 95 e 100% do subconjunto considerado.

k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
0	4.99	-3.47	-8.07	0.23	0.16	0.61	6.41	-2.28	-6.71
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.37	0.26	1.00	6.41	-2.28	-6.71	0.77	1.00	1.00
1	0.27	0.25	0.93	7.31	-0.93	-7.40	0.88	1.70	1.00
2	0.21	0.38	0.90	7.89	-1.82	-7.38	0.94	1.15	1.00
3	0.18	0.39	0.90	8.15	-2.00	-7.43	0.94	1.15	1.00
4	0.15	0.40	0.90	8.41	-2.17	-7.48	0.94	1.15	1.00
5	0.13	0.42	0.90	8.66	-2.32	-7.54	0.94	1.00	1.00
6	0.11	0.38	0.92	8.83	-2.34	-7.60	1.00	1.00	1.00
k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
1	8.83	-2.34	-7.60	0.12	0.09	0.79	8.70	-2.04	-7.33
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.16	0.11	1.00	8.70	-2.04	-7.33	0.94	1.15	1.00
1	0.15	0.13	0.98	8.58	0.00	-7.54	0.94	1.70	1.00
2	0.13	0.21	0.97	9.06	-0.89	-7.51	1.00	1.70	1.00
3	0.12	0.33	0.93	9.47	-1.80	-7.47	1.00	1.15	1.00
4	0.12	0.36	0.93	9.67	-1.95	-7.49	1.00	1.15	1.00
5	0.11	0.38	0.92	9.88	-2.09	-7.51	1.00	1.15	1.00
6	0.10	0.40	0.91	10.11	-2.21	-7.53	1.00	1.00	1.00
k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
2	10.11	-2.21	-7.53	0.09	0.10	0.82	9.67	-2.20	-7.43
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.11	0.12	1.00	9.67	-2.20	-7.43	1.00	1.00	1.00

Tabela 5.1: Algoritmo Não-Iterativo

k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
0	4.99	-3.47	-8.07	0.23	0.16	0.61	6.41	-2.28	-6.71
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.37	0.26	1.00	6.41	-2.28	-6.71	0.70	1.00	1.00
1	0.21	0.25	0.94	7.67	-1.19	-7.47	1.00	1.70	1.00
2	0.19	0.38	0.90	8.06	-1.93	-7.42	1.00	1.20	1.00
3	0.17	0.41	0.90	8.26	-2.11	-7.44	1.00	1.00	1.00
k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
1	8.26	-2.11	-7.44	0.15	0.11	0.74	7.81	-1.59	-7.23
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.20	0.15	1.00	7.81	-1.59	-7.23	0.70	1.30	1.00
1	0.14	0.19	0.97	8.75	-0.71	-7.54	1.00	1.70	1.00
2	0.13	0.30	0.95	9.21	-1.65	-7.50	1.00	1.30	1.00
3	0.12	0.36	0.93	9.45	-1.99	-7.50	1.00	1.00	1.00
k	y_1	y_2	y_3	w_1	w_2	w_3	f_1	f_2	f_3
2	9.45	-1.99	-7.50	0.10	0.09	0.82	9.55	-2.42	-7.39
l	α_1	α_2	α_3	y_1	y_2	y_3	c_1	c_2	c_3
0	0.12	0.11	1.00	9.55	-2.42	-7.39	1.00	1.00	1.00

Tabela 5.2: Algoritmo Iterativo

Observando-se as Tabelas 5.1 e 5.2 constata-se que como esperado, o número de iterações do algoritmo iterativo é menor que o número de iterações do algoritmo não-iterativo. Isto se deve ao fato do algoritmo iterativo levar em conta informações melhor qualificadas pelo Decisor a respeito dos valores das alternativas propostas.

O número de soluções eficientes geradas por estes algoritmos foi ligeiramente menor em alguns casos ($N_e = 20$ no caso não iterativo, $N_e = 25$ no caso iterativo) mas em geral da mesma ordem de grandeza daqueles apresentados pelos algoritmos não-nebulosos.

Comparando-se as implementações de algoritmos nebulosos e não-nebulosos do ponto de vista do Decisor, as implementações baseadas em variáveis linguísticas fornecem um mecanismo adequado para que o Decisor expresse preferências e oriente o processo de obtenção da solução de compromisso de uma maneira natural. No caso dos algoritmos não-nebulosos, além de realizar comparações ordinais que importam numa atribuição subjetiva de utilidade para cada alternativa, o Decisor deve ser capaz de indicar incrementos em relação aos objetivos de forma a se calcular taxas marginais de substituição e, em seguida, definir tamanhos de passos e avaliar a convergência do algoritmo.

Em qualquer caso, exige-se do Decisor um certo conhecimento a respeito do problema, de maneira a se obter protocolos e subconjuntos nebulosos associados aos termos linguísticos num caso e taxas de substituição, etc, em outro, embora, mais uma vez, estas informações sejam mais simples e naturais no caso dos algoritmos nebulosos.

5.10 Conclusão

Neste Capítulo apresentaram-se os conceitos básicos sobre a Teoria de conjuntos e Algoritmos Nebulosos, referenciando-se alguns trabalhos envolvendo a aplicação destes conceitos em programação matemática. No tocante a otimização multiobjetivo, estudaram-se os algoritmos linguísticos propostos por Baptistella e Ollero, os quais foram implementados no espaço dos objetivos. Posteriormente constatou-se um mau desempenho destes métodos ao tratar problemas com curva de trade-off linear ou achatada, e propôs-se sua reformulação em termos de norma infinito generalizada. Comentou-se também, o desempenho numérico destes algoritmos e discutiu-se suas vanta-

gens em relação aos algoritmos não-nebulosos implementados.

Conclusão Geral

Neste trabalho empregou-se uma forma de abordagem para a resolução de problemas multiobjetivos, que permite tratá-los inteiramente no espaço dos objetivos. Seguindo esta filosofia, implementou-se os métodos de Geoffrion et al (1972) e Baptistalla e Ollero (1980), aplicados a um problema clássico de controle de poluição em rios, obtendo-se resultados compatíveis com aqueles encontrados na literatura.

Com o emprego deste tipo de estrutura deriva-se propriedades importantes tais como, linearidade, robustez, convergência infinita e homogeneidade, as quais permitem a implementação de métodos multiobjetivos com simplificações relevantes. Em função destas facilidades, torna-se possível abordar o problema multiobjetivo através de implementações mais elaboradas sem o acréscimo de complexidade para a sua resolução.

Do ponto de vista de esforço computacional, os resultados obtidos neste trabalho, embora restritos a um único problema, mostram que as implementações no espaço dos objetivos são no mínimo tão eficientes quanto as implementações no espaço das decisões. Embora a resolução de um problema do tipo Min-Max seja em princípio uma tarefa dispendiosa, o fato de se aproximar a curva de soluções eficientes através de linearização externa e assim rapidamente identificar a região de interesse para o Decisor parece compensar plenamente este esforço.

Como possíveis extensões para este trabalho sugere-se:

- Tratar o problema mestre através de um método Simplex Multiobjetivo [Yu e Zeleny, 1975] pois o número de *bases eficientes* é, em geral, muito menor no espaço dos objetivos, do que no espaço das decisões. Com um pequeno número de bases, a tarefa de se gerar *faces eficientes* é

simples, o que também simplifica em muito os esquemas de enumeração presentes quando se trata o problema no espaço das decisões.

- Tratar as restrições como conjuntos nebulosos de forma a orientar com maior precisão e significado as ações do Decisor. Se for possível representar os cortes do espaço dos objetivos em termos de interseções de subconjuntos nebulosos, o diálogo Decisor-Analista torna-se muito mais natural e o Decisor terá como interpretar e orientar suas decisões no nível estrutural de maneira também linguística.
- Implementação de Sistemas de Suporte a Decisão baseados em Programação Multiobjetivo. Como o problema possui uma estrutura homogênea no Espaço dos Objetivos, objetivos lineares, restrições lineares e pequeno número de variáveis, a implementação de qualquer método multiobjetivo torna-se muito simples. Obviamente, a existência de um grande elenco de métodos integrados num sistema que ofereça facilidades de interação, inclusive para usuários sem conhecimento sobre programação matemática, é essencial para a abordagem de problemas genéricos de programação multiobjetivo.

Referências Bibliográficas

- [Baptistella e Ollero, 1980] Baptistella, L.F.B. & A. Ollero, Fuzzy Methodologies Multicriteria Optimization, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-10, N^o 7.
- [Bazaraa e Shetty, 1979] Bazaraa, M.S. & C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Jonh Wiley & Sons, New York.
- [Bellman e Zadeh, 1970] Bellman, R.E. & L.A. Zadeh, Decision-Making in a Fuzzy Environment, *Management Science*, Vol. 17, N^o 4.
- [Benayoun et all, 1971] Linear Programming with Multiple Objective Functions: Step Method (STEM), *Mathematical Programming*, 1.
- [Bowman, 1975] Bowman, V.J.J., On Relationship of Tchebycheff Norm and the Efficient Frontier of Multiple-Criteria Objectives, in Thiriez, H. & Zionts (Editors) *Multiple Criteria Decision Making*, Springer.
- [Chankong e Haimes, 1982] Chankong, V. & Y.Y. Haimes, On the Characterization of Noninferior Solutions of the Vector Optimization Problem, *Automatica*, Vol. 18, N^o 6.
- [Charnes e Cooper, 1961] Charnes, A. & W.W. Cooper, *Management Models and Industrial Applications of Linear Programming*, Vol. I, Wiley, New York.
- [Cohon e Marks, 1973] Cohon, J.L. & D.H. Marks, Multiobjctive Screening Models and Water Investment, *Water Resources Research*, Vol. 5, N^o 4.

- [Dubois e Prade, 1978] Dubois, D. & H. Prade, Operations on Fuzzy Numbers, *Int. J. Systems SCI*, Vol. 9, N^o 6.
- [Dyer, 1972] Dyer, J.S., Interactive Goal Programming, *Management Science*, Vol. 19, N^o 1.
- [Dyer, 1973] Dyer, J.S., A Time-Sharing Computer Program for the Solution of the Multiple Criteria Problem, *Management Science*, Vol. 19, N^o 12.
- [Dyer, 1974] Dyer, J.S., The Effects of Errors in the Estimation of the Gradient on the Frank-Wolfe Algorithm, with Implications for Interactive Programming, *Operations Research*, Vol. 22.
- [Esogbue e Bellman, 1984] Esogbue, A.O. & R.E. Bellman, Fuzzy Dynamic Programming and Its Extensions, *Management Science* 20.
- [Fabian e Stoica, 1984] Fabian, C. & M. Stoica, Fuzzy Integer Programming, *Management Science* 20.
- [Ferreira, 1986] Ferreira, P.A.V., Progamação Matemática Multiobjetivo: Desenvolvimento de Método Interativo Baseado em Projecção, *Tese de Doutorado FEE/UNICAMP*.
- [Ferreira e Geromel, 1990] Ferreira, P.A.V. & J.C. Geromel, An Interactive Projection Method for Multicriteria Optimization Problems, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.20 N^o 3.
- [Ferreira e Machado, 1990a] Ferreira, P.A.V. & M.E.S. Machado, Teoria de Conjuntos e Algoritmos Nebulosos, *Nota Interna, FEE/UNICAMP*, N^o 30/90.
- [Ferreira e Machado, 1990b] Ferreira, P.A.V. & M.E.S. Machado, Otimização Multiobjetivo: Algoritmo e Aplicação em Controle de Poluição de Rios, *8o. Congresso Brasileiro de Automática*, Belém - Pa (artigo selecionado para o 4o. Congresso Latino-Americano de Controle Automático - cidade do México - México).
- [Ferreira e Machado, 1990c] Ferreira, P.A.V. & M.E.S. Machado, A Relaxation Procedure for Multiobjective Decision Making, *29th Conference of Decision and Control*, Honolulu, Hawaii.

- [Gembicki e Haimes, 1975] Gembicki, F.W. & Y.Y. Haimes, Approach to Performance and Sensitivity Multiobjective Optimization: The Goal Attainment Method, *IEEE Transactions on Automatic Control*, Vol. AC-20, N^o 6.
- [Geoffrion, 1967] Geoffrion, A.M., Solving Bicriteria Mathematical Programs, *Operations Research*, Vol.15, N^o 1.
- [Geoffrion et al, 1972] An Interactive Approach for Multi-criterion Optimization, with an Application to the Operation of an Academic Department, *Management Science*, Vol. 19, N^o 4.
- [Geromel e Ferreira, 1986] Geromel, J.C. & P.A.V. Ferreira, A Dual-Projection Method for Large Scale Systems Optimization, *Control-Theory and Advanced Technology*, Vol. 2, N^o 2.
- [Geromel e Ferreira, 1991] Geromel, J.C. & P.A.V. Ferreira, Un Upper Bound on Properly Efficient Solutions in Multiobjective Optimization. *Operations Research Letters*, Vol. 10, N^o 3.
- [Haimes et al., 1975] Haymes, Y.Y., W.A. Hall and H.T. Freedman, *Multi-objective Optimization in Water Resources Systems*, Elsevier Scientific Publishing Co., Amsterdam.
- [Hass, 1970] Hass, J.E., Optimal Taxing for the Water Pollution. *Waster Resources Research*, Vol.6, N^o 2.
- [Hwang e Masud, 1979] Hang, C. & A.S.M. Masud, Multiple Objective Decision Making Methods and Applications, *Lecture Notes in Economics and Mathematical Systems Series*. Springer, New York..
- [Lasdon, 1970] Lasdon, L.S., *Otimization for Large Systems*. The Mcmillan co.. London.
- [Loucks et al. 1967] Loucks, D.P., C.S. Revelle and W.R. Lynn, Linear Programming Models for Water Pollution Control, *Management Science*, Vol 14, N^o 4.
- [Luenberger, 1970] Lunberger, D.G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Massachusetts.

- [Mandani e Procyk, 1978] Mandani, E.H. & T.J. Procyk, Application on Fuzzy Logic to Controller Design Based on Linguistic Protocol, *Progress in Cybernetics and System Research*, Vol. III, Wiley ed., London.
- [Reid e Citron, 1971] Reid, R.W. & S.J. Citron, On Noninferior Performance Index Vectors, *Journal of Optimization Theory and Applications*, Vol. 7, N^o 1.
- [Sakawa, 1984] Sakawa, M., Interactive Multiobjective Decision-Making by the Fuzzy Sequential Proxy Optimization Technique - FSPOT, *Management Sciences*, Vol. 20.
- [Slowinski. 1985] Slowinski, R., A Multicriteria Fuzzy Linear Programming Method for Water Supply System Development Planning, *Fuzzy Sets and Systems*.
- [Waltz. 1967] , Waltz, F.M., An Engineering Approach: Hierarchical Optimization Criteria. *IEEE Transactions on Automatic Control*, Vol. AC-12.
- [Yu e Zeleny, 1975] Yu, P.L. & M. Zeleny, The Set of all Nondominated Solutions in Linear Cases and a Multicriteria Simplex Method. *J. of Mathematical Analysis and Applications*, Vol. 49.
- [Zadeh. 1963] Zadeh, L.A., Optimaly and Nonscalar Value Performance Criteria, *IEEE Transactions on Automatic Control*, Vol. AC-8, N^o 1.
- [Zadeh, 1965] Zadeh, L.A., Fuzzy Sets. *Information and Control* 8.
- [Zadeh. 1968] Zadeh, L.A., Fuzzy Algorithms. *Information and Control* 12.
- [Zadeh. 1973] Zadeh, L.A., Outline of New Approach to the Analisis of the Complex Systems and Decision Processes, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-3, N^o 1.
- [Zimmermann, 1978] Zimmermann, H.J., Fuzzy Programming and Linear Programming with Several Objective Functions, *Fuzzy Sets and Systems*, 1.
- [Zimmermann e Pollastschek, 1978] Zimmermann, H.J. & M.A. Pollastschek, Fuzzy 0-1 Linear Programs. *Management Science* 20.

Apêndice A

Conjuntos Nebulosos

A.1 Operações Básicas

Um *Subconjunto Nebuloso* A é definido como uma classe de objetos com um contínuo grau de associação ou pertinência.

O conjunto no qual cada elemento do subconjunto nebuloso A está definido é chamado de *Universo de Discurso* U .

Um subconjunto nebuloso A é caracterizado pela *Função Associação* $\mu_A(\cdot)$ que mapeia cada ponto de U em um número real no intervalo $[0, 1]$.

O *Conjunto Suporte* de A é o conjunto de pontos em U em que $\mu_A(\cdot) > 0$. Um *Ponto de Cruzamento* de um subconjunto nebuloso A é qualquer ponto do universo de discurso U em que $\mu_A(\cdot) = 0.5$.

Um *Singleton Nebuloso* é um subconjunto nebuloso cujo suporte é um único ponto em U .

Um conjunto nebuloso A em U é dito **Vazio** se e somente se :

$$\mu_A(x) = 0, \quad \forall x \in U \quad (\text{A.1})$$

Igualdade: dois conjuntos nebulosos A e B , definidos em U , são iguais se e somente se

$$\mu_A(x) = \mu_B(x), \quad \forall x \in U. \quad (\text{A.2})$$

Complementação : a complementação é uma operação que transforma um conjunto nebuloso definido em um universo de discurso, em outro con-

junto nebuloso definido no mesmo universo de discurso. O complemento de um conjunto nebuloso A é denotado por $\neg A$ e definido por

$$\mu_{\neg A}(x) = 1 - \mu_A(x), \quad \forall x \in U. \quad (\text{A.3})$$

Estar contido em : um conjunto nebuloso A , em U , é dito estar contido em um conjunto nebuloso B , em U , (i.e, A é um subconjunto de B), se e somente se

$$\mu_A(x) \leq \mu_B(x), \quad \forall x \in U. \quad (\text{A.4})$$

União : a união de dois conjuntos nebulosos A e B , em U , é um terceiro conjunto nebuloso C , em U , representado por

$$C = A \cup B \text{ ou } C = A + B, \quad (\text{A.5})$$

e é definida por

$$\mu_C(x) = \mu_A(x) \vee \mu_B(x) \quad (\text{A.6})$$

$$= \max(\mu_A(x), \mu_B(x)), \quad \forall x \in U. \quad (\text{A.7})$$

A união de A com B é o menor conjunto nebuloso que contém tanto A quanto B .

Interseção : a interseção de dois conjuntos nebulosos A e B , em U , resulta em um terceiro conjunto nebuloso C , em U , escrito como,

$$C = A \cap B. \quad \forall x \in U \quad (\text{A.8})$$

e é definida por

$$\mu_C(x) = \mu_A(x) \wedge \mu_B(x) \quad (\text{A.9})$$

$$= \min(\mu_A(x), \mu_B(x)), \quad \forall x \in U. \quad (\text{A.10})$$

Algumas propriedades de união, interseção e complementação de conjuntos nebulosos :

$$\begin{aligned}
 & \left. \begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C \\ A \cap (B \cap C) &= (A \cap B) \cap C \end{aligned} \right\} \text{Leis Associativas} \\
 & \left. \begin{aligned} \neg(A \cup B) &= \neg A \cap \neg B \\ \neg(A \cap B) &= \neg A \cup \neg B \end{aligned} \right\} \text{Leis de De Morgan} \\
 & \left. \begin{aligned} C \cap (A \cup B) &= (C \cap A) \cup (C \cap B) \\ C \cup (A \cap B) &= (C \cup A) \cap (C \cup B) \end{aligned} \right\} \text{Leis Distributivas}
 \end{aligned}$$

Produto : o produto entre dois conjuntos nebulosos, A e B , em U é representado por AB e definido como

$$\mu_{AB}(x) = \mu_A(x) \cdot \mu_B(x), \quad \forall x \in U. \quad (\text{A.11})$$

Exponenciação : a exponenciação de conjuntos nebulosos, representada por $A^\alpha = (A.A.\dots.A)$, α vezes, é definida como

$$\mu_{A^\alpha}(x) = (\mu_A(x))^\alpha, \quad \forall x \in U. \quad (\text{A.12})$$

onde α é um número real positivo.

Produto por escalar : o produto de um escalar real por um conjunto nebuloso é representado por αA , e definido como

$$\mu_{\alpha.A}(x) = \alpha \cdot \mu_A(x), \quad \forall x \in U. \quad (\text{A.13})$$

onde α é um número real não negativo.

Soma Algébrica: a soma algébrica de dois conjuntos nebulosos A e B , definidos em U , é representada por $A + B$ e definida como

$$\mu_{A+B}(x) = \mu_A(x) + \mu_B(x), \quad \forall x \in U. \quad (\text{A.14})$$

satisfazendo a restrição de que $\mu_{A+B}(x) \leq 1$, para assim tornar a soma algébrica consistente com as definições anteriores.

Relações nebulosas : a relação-nebulosa R de um conjunto U para um conjunto Y , é um subconjunto do produto cartesiano $U \times Y$ e é representada por uma função associação de duas variáveis

$$R = \int_{U \times Y} \mu_R(x, y) / (x, y), \quad \forall x \in U \text{ e } \forall y \in Y. \quad (\text{A.15})$$

Uma relação-nebulosa pode também ser representada como uma matriz de relação, cujo elemento (i, j) é o valor de $\mu(x, y)$ para o i -ésimo valor de x e o j -ésimo valor de y .

Estendendo-se o conceito, uma relação-nebulosa n -ária é um subconjunto nebuloso de $U_1 \times U_2 \times \dots \times U_n$, definido por

$$R = \int_{U_1 \times \dots \times U_n} \mu_R(x_1, \dots, x_n) / (x_1, \dots, x_n), \quad \forall x_i \in U_i, \quad i = 1, \dots, n. \quad (\text{A.16})$$

Composição de relações nebulosas: seja R uma relação-nebulosa de U para Y e S uma relação-nebulosa de Y para Z . Então a composição de R e S é uma terceira relação-nebulosa (matriz relação-nebulosa), representada por $R \circ S$, e definida como :

$$R \circ S = \int_{U \times Z} \sup_y \min[\mu_R(x, y), \mu_S(y, z)] / (x, z) \quad (\text{A.17})$$

ou

$$R \circ S = \int_{U \times Z} \bigvee_y [\mu_R(x, y) \wedge \mu_S(y, z)] / (x, z) \quad (\text{A.18})$$

$\forall x \in U, \quad y \in Y$ e $z \in Z$ e \bigvee_y representa o *supremum* sobre o domínio de Y .

Se U, Y e Z forem conjuntos finitos, então a matriz de relação $R \circ S$ é dada pelo Max-Min entre as matrizes de relação R e S .

Normalização: seja $\bar{\mu}_A(x)$ o supremum de uma função associação $\mu_A(x)$ sobre o universo de discurso. Um conjunto nebuloso A é dito *normal* se $\bar{\mu}_A(x) = 1$, caso contrário é *subnormal*. A normalização de um conjunto nebuloso A subnormal pode ser feita através da divisão de $\mu_A(x)$ por $\bar{\mu}_A(x)$ e pode ser expressa por,

$$NORM(A) = (\bar{\mu}_A(x))^{-1} \cdot \mu_A(x), \quad \mu_A(x) \neq 0, \quad \forall x \in U. \quad (\text{A.19})$$

Concentração: a ação da operação de concentração sobre um conjunto nebuloso A , em U , possui como resultado um subconjunto nebuloso de A , com uma redução relativamente pequena nos graus de associação dos elementos de U que tem um alto grau de associação em A , e uma redução relativamente grande nos graus de associação dos elementos de U que têm um baixo grau de associação em A . Este efeito pode ser conseguido elevando-se, por exemplo, $\mu_A(x)$ ao quadrado. Simbolizando a aplicação da operação de concentração sobre A por $CONC(A)$, podemos escrever

$$\mu_{CONC(A)}(x) = \mu_A^2(x), \quad \forall x \in U, \quad (\text{A.20})$$

e pela definição de A^α temos,

$$CONC(A) = A^2. \quad (\text{A.21})$$

A operação de concentração apresenta as propriedades distributiva sobre a união, sobre a interseção e sobre o produto entre conjuntos nebulosos e também, pode ser composta com ela própria,

$$CONC(CONC(A)) = A^4 \text{ ou } CONC^\alpha(A) = A^{2^\alpha}, \quad (\text{A.22})$$

onde α é um inteiro maior ou igual a dois.

Dilatação : a dilatação é uma operação que possui o efeito contrário ao da concentração, pois no caso da sua aplicação a um conjunto nebuloso A , em U , produz um subconjunto nebuloso denotado por $DIL(A)$, com um aumento relativamente pequeno nos graus de associação dos x que possuem um alto grau de associação em A e um aumento relativamente grande nos graus de associação dos x que possuem um baixo grau de associação em A . A dilatação é definida como

$$\mu_{DIL(A)}(x) = (\mu_A(x))^{0.5}, \forall x \in U \quad (\text{A.23})$$

ou

$$DIL(A) = A^{0.5} \quad (\text{A.24})$$

Intensificação de Contraste: esta operação quando aplicada a um conjunto nebuloso A , em U , origina um aumento nos graus de associação $\mu_A(x)$ que estão abaixo de 0.5 e um decréscimo nos graus de associação $\mu_A(x)$ que estão acima de 0.5. A intensificação de contraste é expressa por $INT(A)$ e definida como

$$\mu_{INT(A)}(x) = 2\mu_A^2(x), \text{ para } 0 \leq \mu_A(x) \leq 0.5 \quad (\text{A.25})$$

$$\mu_{INT(A)}(x) = 1 - 2[1 - \mu_A(x)]^2, \text{ para } 0.5 \leq \mu_A(x) \leq 1 \quad (\text{A.26})$$

É interessante notar que a intensificação de contraste reduz a nebulosidade do conjunto. Como no caso da concentração, a intensificação possui a propriedade distributiva para a união, interseção e produto entre dois conjuntos nebulosos.

A função que define a intensificação de contraste é usada também para representar funções associação de conjuntos nebulosos. Logo, é interessante definir uma função S da linha real para o intervalo $[0,1]$ dada por

$$S(\mu_A(x)) = 0, \quad \text{para } \mu_A(x) < 0 \quad (\text{A.27})$$

$$S(\mu_A(x)) = 2\mu_A^2(x), \quad \text{para } 0 \leq \mu_A(x) \leq 0.5 \quad (\text{A.28})$$

$$S(\mu_A(x)) = 1 - 2(1 - \mu_A(x))^2, \quad \text{para } 0.5 \leq \mu_A(x) \leq 1 \quad (\text{A.29})$$

$$S(\mu_A(x)) = 1, \quad \text{para } \mu_A(x) > 1 \quad (\text{A.30})$$

Combinação Convexa: esta é uma operação que representa n conjuntos nebulosos A_1, A_2, \dots, A_n quaisquer, em um único conjunto nebuloso A , através de uma combinação ponderada sobre as funções associação dos mesmos.

$$\mu_A(x) = w_1\mu_{A_1}(x) + w_2\mu_{A_2}(x) + \dots + w_n\mu_{A_n}(x). \quad (\text{A.31})$$

onde $w_i \geq 0$ e $\sum_{i=1}^n w_i = 1, \quad \forall x \in U$.

Fuzzificação: a operação de fuzzificação tem o efeito de transformar um conjunto não-nebuloso em um conjunto nebuloso, ou apenas aumentar o grau de nebulosidade de um conjunto nebuloso. Representa-se um conjunto A fuzzificado por \tilde{A} ou $F(A)$, onde “ \sim ” representa o fuzzificador.

Seja U o conjunto dos números reais e $x \in U$. A expressão “ $x \approx 3$ ” significa que “ x é um número aproximadamente igual a 3”, enquanto que “ $x \underset{\sim}{=} 3$ ” significa que “ x é um conjunto nebuloso de números que se aproximam de 3”. Enquanto “ $=$ ” representa igualdade, “ $\underset{\sim}{=}$ ” representa igualdade aproximada. Enquanto “ $>$ ” denota “maior que”, “ $\underset{\sim}{>}$ ” denota “aproximadamente maior que”. A operação de fuzzificação é essencial para a definição de delimitadores linguísticos como “mais ou menos”, “ligeiramente”, “muito”, etc., como será visto posteriormente.

Um fuzzificador é caracterizado pelo seu Kernel $K(y)$, o qual é um conjunto nebuloso resultante da aplicação do fuzzificador a um singleton $1/y$ em

U , ou seja, $K(y) = \underline{1}/\underline{y}$ o que mostra a dependência de \underline{y} sobre y . Geralmente, $K(y)$ é um intervalo nebuloso cuja função associação $\mu_{K(y)}(z)$ é uma função não crescente com a distância entre z e y .

A **Fuzzificação Pontual** é definida como um processo que transforma um singleton nebuloso $1/x$, definido em U , em um conjunto nebuloso \underline{x} concentrado em torno de x . O resultado da aplicação do fuzzificador F a um conjunto nebuloso qualquer A , em termos de Kernal é,

$$F(A, K) = \int_U [\mu_A(x)/x] K(x), \quad (\text{A.32})$$

onde $[\mu_A(x)/x]K(x)$ representa o produto do escalar $\mu_A(x)/x$ pelo conjunto nebuloso $K(x)$. Pode-se fazer uma analogia entre a equação (A.32) e a representação integral de um operador linear, tendo $K(x)$ como resposta ao impulso.

Seja A um conjunto nebuloso definido por

$$A = \mu_1(x)/x_1 + \dots + \mu_n(x)/x_n, \quad (\text{A.33})$$

onde $\mu_i(x)$ é o grau de associação de x_i , $i = 1, \dots, n$ em A . Se considerarmos a fuzzificação como uma transformação linear, então fuzzificando-se A tem-se

$$\underline{A} = \underline{\mu}_1(x)/\underline{x}_1 + \dots + \underline{\mu}_n(x)/\underline{x}_n, \quad (\text{A.34})$$

podendo-se então considerar dois casos especiais,

Caso 1 : Se $\mu_i(x)$, $i = 1, \dots, n$, são constantes, então

$$\underline{\mu}_i(x)/\underline{x}_i = \mu_i(x)/x_i, \quad i = 1, \dots, n \quad (\text{A.35})$$

Caso 2 : Se x_i , $i = 1, \dots, n$, são constantes, então

$$\underline{\mu}_i(x)/\underline{x}_i = \mu_i(x)/x \quad (\text{A.36})$$

Caso 1: $K(x_i)$ caracteriza a transformação linear que mapeia $1/x_i$ em \underline{x}_i . A fuzzificação que transforma A em \underline{A} é expressa por $SF(A; K)$ (s-fuzzificação ou fuzzificação suporte), e é definida como,

$$\underline{A} = SF(A; K) = \mu_1(x_1)K(x_1) + \dots + \mu_n(x_n)K(x_n). \quad (\text{A.37})$$

onde $\mu_i(x_i)K(x_i)$ representa o produto de um escalar por um conjunto nebuloso e “+” simboliza a união dos conjuntos nebulosos. No caso acima, A apresenta um conjunto suporte finito. No caso mais geral $SF(A; K)$ é definida como

$$SF(A; K) = \int_U [\mu_A(x)/x] K(x), \quad \forall x \in U. \quad (A.38)$$

Se A é um singleton $1/x$, então $SF(A; K) = K(x)$. A propósito, um singleton é análogo a uma função delta e $K(x)$ assume a função de resposta ao impulso. Uma s-fuzzificação pode ser empregada para produzir uma translação de um conjunto nebuloso dentro ou não do seu universo de discuso. A s-fuzzificação tem um efeito semelhante à operação de dilatação, mas diferindo quanto ao fato da dilatação de um conjunto não-nebuloso ser também um conjunto não nebuloso, enquanto que a s-fuzzificação de um conjunto não-nebuloso resulta, em geral em um conjunto nebuloso.

Caso 2 : Seja $K(\mu_i(x))$ o Kernel que define a fuzzificação pontual que transforma $\mu_i(x)$ em $\underline{\mu}_i(x)$. Denominando-se $GF(A; K)$ (fuzzificação do grau de associação ou g-fuzzificação), à fuzzificação que transforma A em \underline{A} , onde

$$\underline{A} = \underline{\mu}_1(x)/x_1 + \dots + \underline{\mu}_n(x)/x_n \quad (A.39)$$

pode-se escrever

$$GF(A; K) = \underline{A} = K(\mu_1(x))/x_1 + \dots + K(\mu_n(x))/x_n \quad (A.40)$$

Generalizando, se

$$A = \int_U \mu_A(x)/x. \quad (A.41)$$

então a g-fuzzificação de A , define um novo conjunto nebuloso

$$GF(A; K) = \int_U \underline{\mu}_A(x)/x, \quad (A.42)$$

em que o grau de associação de x em $GF(A; K)$ é o conjunto nebuloso $\mu_A(x) = K(x)$.

Convexidade : um conjunto nebuloso A é dito convexo se e somente se os conjuntos Γ definidos por,

$$\Gamma_\alpha = \{x | \mu_A(x) \geq \alpha\}, \quad (\text{A.43})$$

forem convexos para todo $\alpha \in [0, 1]$, ou igualmente se e somente se

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min[\mu_A(x_1), \mu_A(x_2)], \quad (\text{A.44})$$

para quaisquer x_1 e x_2 em U e $\forall \lambda \in [0, 1]$.

Limitação : Um conjunto nebuloso A é limitado se e somente se os conjuntos definidos por (A.44) forem limitados para $\forall \alpha > 0$ isto é, existe um raio finito $R(\alpha)$ tal que $\|x\| < R(\alpha)$, $\forall x \in \Gamma_\alpha$.

Projeção: A projeção de um conjunto nebuloso A no R^n , com função associação sobre um hiperplano H definido como

$$H = \{x | x_1 = 0\} \quad (\text{A.45})$$

é um conjunto nebuloso denotado por $S_H(A)$ em R^{n-1} com função associação dada por

$$\mu_{S_H(A)}(x_1, \dots, x_n) = \sup_x \mu_A(x_1, \dots, x_n) \quad (\text{A.46})$$

A.2 Termos Linguísticos

Uma linguagem L pode ser vista como a correspondência entre um conjunto de termos T e um universo de discurso U , e pode ser caracterizada por uma relação-nebulosa N que associa cada termo y em T a um objeto x em U , com grau de associação $\mu_N(y, x) \in [0, 1]$. Por exemplo, $y = \text{jovem}$, $x = 23$ anos então $\mu_N(\text{jovem}, 23)$ pode ser 0.9, onde $U = \{x/x \in Z, 0 \leq x \leq 100\}$.

Um termo pode ser *atômico*, por exemplo $y = \text{alto}$, ou *composto*, $y = \text{muito alto}$. Para um x fixo, a função $\mu_M(y, x)$ define um subconjunto nebuloso $M(y)$ de U , que expressa o *significado* de x . Sua função associação é definida por

$$\mu_{M(y)}(x) = \mu_N(y, x), \quad x \in U \text{ e } \forall y \in T. \quad (\text{A.47})$$

Uma *variável linguística* é definida como uma variável cujos valores são sentenças em uma linguagem natural ou artificial. Se por exemplo, *alto*, *muito alto*, *não alto*, etc., são valores de estatura, então estatura é uma variável linguística.

Os termos atômicos dividem-se em quatro categorias :

1. Termos primários que servem de rótulos para um conjunto nebuloso específico de um universo de discurso.
2. Negação (“não”) e conectivos (“e”, “ou”).
3. Delimitadores, por exemplo, muito, mais ou menos, ligeiramente, etc.
4. Marcadores,“()”(parênteses).

Com a composição de termos atômicos, surge um problema que é encontrar o significado do termo composto. Por exemplo, $x = hu$, onde h é um delimitador e u é um termo com significado específico.

A.3 Delimitadores Linguísticos

Um delimitador linguístico “ h ” pode ser definido como um operador que transforma o conjunto nebuloso $M(u)$ (significado de u) em um conjunto nebuloso $M(hu)$. Logo, tem a função de gerar um conjunto de valores para as variáveis linguísticas, a partir de uma coleção de termos primários. Por exemplo, o termo primário “alto”, as conjunções “e” e “não” e o delimitador “muito”, podem gerar : muito alto, muito muito alto, não muito alto, alto e não muito alto, etc. O operando sobre o qual um delimitador linguístico age deve ter um significado nebuloso.

O estudo dos delimitadores linguísticos poder ser feito dividindo-os em duas categorias nebulosas básicas.

Tipo I - Os delimitadores desta categoria podem ser representados como operadores agindo diretamente sobre um conjunto nebuloso. São exemplos : muito, mais ou menos, altamente, ligeiramente, etc..

Tipo II - Os delimitadores deste tipo têm uma caracterização muito mais complexa que os do tipo I, e agem sobre as componentes do conjunto nebuloso. A definição deste tipo de delimitador é efetuada sob a forma de algoritmos nebulosos envolvendo delimitadores do tipo I. O efeito deste delimitador pode ser aproximado a uma modificação nos coeficientes de ponderação de uma combinação convexa. São exemplos: essencialmente, tecnicamente, atualmente, estritamente, praticamente, regularmente, etc.

Delimitadores do Tipo I

Muito : este delimitador age como um intensificador do significado do seu operando. A operação que tem esta propriedade é a concentração. Seja A um conjunto nebuloso em U representando o significado de um termo y e A^* o conjunto nebuloso resultante da aplicação da operação de concentração sobre A , com significado y^* .

$$A^* = CONC(A) \quad (A.48)$$

$$A^* = A^2, \quad (A.49)$$

e substituindo-se o termo pelo seu significado tem-se,

$$y^* = muito\ y = y^2 \quad (A.50)$$

Se

$$y = \mu_1(x_1)/x_1 + \mu_2(x_2)/x_2 + \dots + \mu_n(x_n)/x_n, \quad (A.51)$$

$$\forall x \in U, \quad i = 1, \dots, n,$$

então

$$\text{muito } y = \mu_1^2(x_1)/x_1 + \mu_2^2(x_2)/x_2 + \dots + \mu_n^2(x_n)/x_n \quad (\text{A.52})$$

ou mais genericamente, se

$$y = \int_U \mu(x)/x, \quad (\text{A.53})$$

então

$$\text{muito } y = \int_U \mu^2(x)/x, \quad (\text{A.54})$$

O delimitador *muito* pode ser composto com ele próprio, exemplo

$$\text{muito muito } x = x^4.$$

Não é correto escrever, $x = \text{muito não exato}$, pois a frase pode se tornar incompreensível. Substitui-se então não exato por inexato, assim tem-se. $x = \text{muito inexato} = \text{inexato}^2 = (\text{não exato})^2$. *Muito* distribui-se sobre a operação de união. *Muito* aplicado sobre um termo não-nebuloso resulta no próprio termo não-nebuloso.

O termo *muito* é bastante utilizado em relações comparativas, ex.: muito mais bonito, muito menos semelhante, etc. Seja a relação “maior que”, representada por “>” e definida como,

$$> = \int_D 1/(v, w), \quad (\text{A.55})$$

onde “>” tem um significado de um conjunto não-nebuloso e $v, w \in U$ e D é o conjunto de todos os pontos em $U \times U$ em que $v > w$. O delimitador *muito* age como um s-fuzzificador, pois transforma o conjunto não-nebuloso “>” em um conjunto nebuloso simbolizado por “>>” e expresso como.

$$\begin{aligned} \gg &= \int_D [1 + (v - w)^{-2}]^{-1}/(v, w) \\ \gg &= \neg\text{SF}(\neg >; K). \end{aligned} \quad (\text{A.56})$$

onde K é o seu Kernel definido como,

$$K(v, w) = \int_D [1 + (r - v - s + w)^{-2}]^{-1} / (r, s), \quad (\text{A.57})$$

onde $r, s \in U$, e

$$D = \{(r, s) : r - v \geq s - w\}. \quad (\text{A.58})$$

Mais e menos: estes delimitadores podem ser chamados de acentuadores e desacentuadores, respectivamente, e fornecem graus mais suaves de concentração e dilatação que aqueles associados às operações *CONC* e *DIL*. São definidos como,

$$\text{mais } x = x^{1.25} \quad (\text{A.59})$$

$$\text{menos } x = x^{0.75}, \quad (\text{A.60})$$

Os expoentes são escolhidos com o propósito de fazer valer a identidade aproximada, mais mais $x \cong$ menos muito x . Uma importante aplicação dos delimitadores *mais e menos*, é na definição de delimitadores naturais, cujo significado difere ligeiramente de outros, por exemplo,

$$\text{altamente} = \text{mais muito} \quad (\text{A.61})$$

ou

$$\text{altamente} = \text{menos muito muito}. \quad (\text{A.62})$$

O delimitador *mais ou menos* pode ser utilizado como um s-fuzzificador. Quando seu operando tem significado nebuloso, seu efeito é alcançado através da dilatação ou da desacentuação.

$$\text{mais ou menos} = \text{SF}(U, K), \quad (\text{A.63})$$

onde K é o Kernel associado ao conjunto U .

Delimitadores Tipo II

Essencialmente: seja um termo x cujas componentes são expressas por x_1, x_2, \dots, x_n . Assume-se que o conjunto nebuloso x é uma combinação conexa de suas componentes:

$$\mu(x) = w_1\mu_1 + w_2\mu_2 + \dots + w_n\mu_n, \quad (\text{A.64})$$

onde $\sum_{i=1}^n w_i = 1$ e $w_i \geq 0$, $i = 1, \dots, n$, sendo que w_i mede a importância do i -ésimo atributo x_i e μ_i é o grau de associação de um y individual em x_i . O delimitador *essencialmente*, aumenta o peso dos atributos importantes e diminui o peso dos atributos que são relativamente sem importância. Para isto, normaliza-se os pesos w_i 's tal que w_1 seja o peso mais alto (valor unitário), em seguida divide-se o quadrado dos pesos normalizados pela somatória do quadrado dos pesos normalizados. Obtem-se assim, o efeito de *essencialmente* sobre seu operando.

O procedimento descrito acima pode ser usado para expressar o efeito de outros delimitadores linguísticos deste tipo, tais como: *tecnicamente*, *regularmente*. etc.

Até o momento considerou-se delimitadores simples cujo efeito pode ser caracterizado através de combinações ou composições das operações elementares descritas. O que não se considerou foram delimitadores cuja ação não é fácil representar, nem mesmo aproximadamente em termos das operações elementares descritas neste texto. Para caracterização estas ações necessita-se do uso de algoritmos nebulosos.

A.4 Algoritmos Nebulosos

Um algoritmo nebuloso pode ser representado por uma sequência de instruções ordenadas que fornecem uma representação aproximada de uma grande variedade de fenômenos complexos.

Em algoritmos nebulosos encontram-se instruções dos tipos :

1. Declarações de atribuição

- * x é aproximadamente igual a 5;
- * x é grande;
- * x é positivo;
- * x não é grande e x não é pequeno;
- * $x = 3$.

2. Declarações condicionais nebulosas

- * se x é pequeno, então y é grande, senão y não é grande;
- * se x é positivo, então decresça y ligeiramente;
- * se x é muito maior que 5, então pare;
- * se x é muito pequeno, então vá para 7.

3. Declarações incondicionais

- * multiplique x por x ;
- * decresça ligeiramente x ;
- * delete primeiro algumas ocorrências de 1;
- * vá para 7;
- * escreva x ;
- * pare.

A execução de uma instrução-nebulosa, apresenta como resultado um conjunto nebuloso, mas se for necessário a escolha de um único elemento deste conjunto como resposta, pode-se adotar várias estratégias para fazê-la. A estratégia mais comum é escolher o elemento de maior grau de associação no conjunto e no caso de ser mais de um, pode-se escolher um elemento aleatoriamente, ou utilizar um outro critério externo qualquer. Esta alternativa recebe o nome de *regra da alternativa preponderante*.

Os algoritmos nebulosos podem ser classificados de acordo com o tipo de aplicação. Podendo ser utilizados para a definição e geração de conjuntos nebulosos, para a descrição do comportamento de sistemas e de estratégias ou regras de decisão.