

Universidade Federal de Minas Gerais
Departamento de Matemática

Notas de Aula

Otimização Escalar e Vetorial

Volume 3: Otimização Vetorial

Professor: Ricardo H. C. Takahashi

Belo Horizonte, Janeiro de 2007

Conteúdo

I	Introdução e Conceitos Preliminares	6
1	Introdução	7
1.1	Otimização em Projeto Assistido por Computador	7
1.2	Sistemas de Projeto Assistido por Computador	9
1.3	Otimização em PAC	12
1.3.1	A Abordagem Escalar	12
1.3.2	A Abordagem Vetorial	16
1.4	Formulação do Problema de Otimização Vetorial	17
1.4.1	Etapa de Determinação das Soluções Eficientes	17
1.4.2	Etapa de Decisão	18
2	Definições de Referência	20
2.1	Espaços e Normas	20
2.2	Espaços Topológicos	24
2.3	Cones	27
2.4	Hiperplanos e Poliedros	28
3	Caracterização das Funções	29
3.1	Superfícies de Nível e Modalidade	30
3.1.1	Bacias de Atração	32
3.2	Continuidade e Diferenciabilidade	32
3.3	Convexidade e Quasi-Convexidade	33
3.4	Mínimos Locais e Mínimos Globais	36
3.5	Caracterização dos Mínimos Locais	37
4	Convergência de Algoritmos	42
4.1	Algoritmos	42

II	Otimização Escalar	46
5	Interpretação Geométrica	47
5.1	O Jogo da Otimização	47
5.1.1	Formulação do Problema de Otimização	48
5.1.2	As Regras do Jogo	54
5.2	Otimização Sem Restrições	57
5.2.1	Estratégias de Direção de Busca	62
5.2.2	Estratégias de Exclusão de Regiões	67
5.2.3	Estratégias de Populações	74
5.3	Otimização com Restrições de Desigualdade	80
5.3.1	Interpretação geométrica de uma restrição de desigualdade	82
5.3.2	Interpretação geométrica de várias restrições de desigualdade	84
5.3.3	Barreiras e Penalidades	85
5.3.4	Composição pelo Máximo	89
5.4	Otimização com Restrições de Igualdade	90
5.5	Otimização Linear	93
6	Direções de Busca	98
6.1	Estrutura Básica	99
6.2	Busca em Direções Aleatórias	100
6.3	Algoritmo do Gradiente	102
6.3.1	Cálculo do Gradiente	103
6.3.2	Otimização Unidimensional	104
6.3.3	Critérios de Parada	108
6.3.4	Convergência	112
6.4	Aproximações Quadráticas	115
6.4.1	Algoritmo de Newton	118
6.4.2	Método de Newton Modificado	119
6.4.3	Determinação Numérica da Hessiana	122
6.4.4	Construção da Hessiana	122
6.4.5	Correção de Posto 1	124
6.4.6	Métodos Quasi-Newton	129
6.5	Tratamento de Restrições	132
6.5.1	Método de Barreira	132
6.5.2	Método de Penalidades	133

6.6	Comportamento dos Métodos de Direção de Busca	135
6.6.1	Não-Diferenciabilidade	135
6.6.2	Não-Convexidade	137
6.6.3	Multimodalidade	138
7	Exclusão de Semi-Espaços	139
7.1	Formulação Geral	140
7.2	Métodos de Planos de Corte	141
7.2.1	Algoritmo de Planos de Corte de Kelley	144
7.3	Algoritmo Elipsoidal	144
7.3.1	Algoritmo Elipsoidal com “Deep Cut”	146
7.4	Tratamento de Restrições	147
7.5	Características de Comportamento	149
7.5.1	Descontinuidades e Não-Diferenciabilidade	149
7.5.2	Não-Convexidade	150
7.5.3	Multimodalidade	150
7.5.4	Velocidade de Convergência	150
7.6	Algoritmo Cone-Elipsoidal	151
7.7	Definição do Problema	152
7.8	Método Elipsoidal Convencional	152
7.8.1	Problemas Difíceis para o Método Convencional	153
7.9	Cones das Direções Factibilizantes	155
7.10	O Método Cone-Elipsoidal	157
7.10.1	Primeira Reformulação do Problema	158
7.10.2	Segunda Reformulação do Problema	160
7.11	O Algoritmo MCE	163
7.12	Não-Convexidade de Restrições de Igualdade	164
7.13	Conclusões	165
8	Otimização por Populações	167
8.1	Algoritmo Evolucionário Simples	169
8.2	Algoritmo de Simulated Annealing	170
8.3	Algoritmos Genéticos	173
8.3.1	Algoritmo Genético - Codificação Binária	174
8.3.2	Algoritmo Genético - Codificação Real - Polarizado	176
8.4	Sobre a Estrutura do AG-B e do AG-RP	181
8.4.1	Resultados para o AG-B	181
8.4.2	Resultados para o AG-RP	194

8.4.3	Teste das Propriedades de Convergência	198
8.5	Metodologia de Avaliação da Eficiência de AG's	204
8.5.1	Metodologia de Avaliação	206
8.6	Tratamento de Restrições	212
8.7	Características de Comportamento	212
8.7.1	Descontinuidades e Não-Diferenciabilidade	212
8.7.2	Multimodalidade	212
8.7.3	Velocidade de Convergência	213
9	Exercícios - Otimização Escalar	216
 III Otimização Vetorial		 222
10	Soluções de Pareto	223
10.1	O Problema de Otimização Vetorial	223
10.1.1	Notação	224
10.2	Ordenamento de Soluções	225
10.3	O Conjunto Pareto-Ótimo	229
10.3.1	Conjunto localmente Pareto-ótimo	236
10.3.2	Solução utópica	237
10.4	O Problema de Determinação das Soluções Eficientes	239
10.5	Condições de Kuhn-Tucker para Eficiência	240
11	Geração de Soluções Eficientes	244
11.1	Abordagem via Problema Ponderado	244
11.1.1	Interpretação geométrica	245
11.1.2	Algoritmos P_λ	252
11.2	Abordagem via Problema ϵ -Restrito	254
11.2.1	Algoritmos P_ϵ	259
11.3	Abordagem híbrida: Ponderando e Restringindo	261
11.4	Abordagem da Programação-Alvo	262
11.5	Abordagem P_χ	267
11.6	Teste de Eficiência	274
11.6.1	Algoritmos P^*	275

12 Propriedades de Grupo	276
12.1 Verificação versus Falseamento	277
12.2 Estrutura do Conjunto Pareto-Ótimo	279
12.3 Análise Multiobjetivo	282
12.3.1 Consistência	283
12.3.2 Ordenamento e Dominância	284
12.3.3 Extensão	284
12.3.4 Dados Extremos	286
12.4 Decisão e Síntese Multiobjetivo	286
12.5 Algoritmo Genético Multiobjetivo	288
12.5.1 Construção do Algoritmo Genético Multiobjetivo	288
12.5.2 AG-RPMO	290
12.6 Exemplo de Aplicação: Projeto de Controladores	296
12.6.1 Realimentação completa de estados	296
12.6.2 Realimentação estática de saídas	299
13 Exercícios - Otimização Vetorial	302
Exercícios Computacionais	307

Parte III
Otimização Vetorial

Capítulo 10

Caracterização Analítica e Formulação do Problema de Geração das Soluções de Pareto

Neste capítulo são apresentados os conceitos que conduzem à definição do *conjunto Pareto-ótimo* de um problema de otimização vetorial. Com tal definição estabelecida, formulam-se a seguir condições analíticas que caracterizam as soluções pertencentes a tal conjunto.

10.1 O Problema de Otimização Vetorial

Considere-se o problema de otimização vetorial (POV):

$$(POV) \begin{cases} \mathcal{X}^* = x^* \in \mathbb{R}^n \mid \\ x^* = \arg \min_x f(x) \\ \text{sujeito a: } x \in \mathcal{F}_x \end{cases} \quad (10.1)$$

sendo $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$ o vetor de objetivos do problema e $\mathcal{F}_x \subset \mathbb{R}^n$ a região factível. Os vetores $x \in \mathbb{R}^n$ são os chamados *vetores de parâmetros* do POV, formando o *espaço de parâmetros* \mathcal{X} . Os vetores $f(x) \in \mathbb{R}^m$ encontram-se

num espaço vetorial que será aqui denominado *espaço de objetivos*, sendo denotado por \mathcal{Y} .

Deseja-se, portanto, com o POV, realizar a determinação de um conjunto de pontos X^* pertencentes ao espaço de parâmetros do problema, tais que estes minimizem, em certo sentido, uma função vetorial. A primeira questão que se coloca, portanto, é a de definir *o que é essa minimização*.

10.1.1 Notação

Neste capítulo, um conjunto de conceitos novos será definido. A lista de notações associadas a tais conceitos é reunida aqui para facilitar a leitura deste capítulo e dos próximos.

\mathcal{X} : Espaço dos parâmetros de otimização. Neste trabalho, tal espaço será sempre identificado com um espaço \mathbb{R}^n .

\mathcal{X}^* : Conjunto de *soluções eficientes*, ou *conjunto Pareto-ótimo*, de um POV. O conjunto \mathcal{X}^* é um subconjunto do espaço de parâmetros \mathcal{X} .

x^* : Um ponto pertencente ao conjunto \mathcal{X}^* .

\mathcal{F}_x : O conjunto dos pontos factíveis, subconjunto do conjunto \mathcal{X} .

\mathcal{Y} : Espaço dos “objetivos”, isto é, o espaço no qual se representa a imagem da função $f(\cdot)$. Neste trabalho, tal espaço será sempre identificado com um espaço \mathbb{R}^m .

\mathcal{I} : O conjunto imagem da função $f(\cdot)$, tendo por domínio todo o espaço \mathcal{X} .

\mathcal{F}_y : O conjunto imagem da função $f(\cdot)$ restrita ao domínio \mathcal{F}_x . O conjunto \mathcal{F}_y é um subconjunto do espaço \mathcal{Y} .

\mathcal{Y}^* : O conjunto imagem da função $f(\cdot)$ com o domínio restrito ao conjunto \mathcal{X}^* . O conjunto \mathcal{Y}^* é um subconjunto de \mathcal{Y} , e é denominado *conjunto Pareto-ótimo no espaço de objetivos*, ou ainda *fronteira Pareto-ótima*.

$\partial(\cdot)$: A fronteira do conjunto-argumento.

$(\cdot)^\circ$: O interior do conjunto-argumento.

10.2 Ordenamento de Soluções

O problema de otimização vetorial (POV) se define a partir da análise do ordenamento das soluções, levando em conta os diversos objetivos. As soluções multiobjetivo, ou soluções de Pareto, são as melhores soluções entre as quais não existe um ordenamento (ou seja, não há como definir, a partir da avaliação dos funcionais objetivo, que uma solução é melhor que a outra).

Para fundamentar essa definição, é apresentada inicialmente a definição de *conjunto ordenado*.

Definição 10.1 (Conjuntos Ordenados) *Um conjunto C é ordenado de acordo com a relação de ordem “ \preceq ” se dados quaisquer dois elementos $x, y \in C$ é sempre verdade que $x \preceq y$ ou $y \preceq x$ e as seguintes propriedades com relação a “ \preceq ” são válidas:*

$$i. \quad x \preceq x \quad \text{(reflexividade)}$$

$$ii. \quad x \preceq y \text{ e } y \preceq z \Rightarrow x \preceq z \quad \text{(transitividade)}$$

$$iii. \quad x \preceq y \text{ e } y \preceq x \Rightarrow x = y \quad \text{(anti-simetria)}$$

□

A seguir, é apresentada a definição que será relevante no caso da otimização vetorial, que é a de *conjunto parcialmente ordenado*.

Definição 10.2 (Conjuntos Parcialmente Ordenados) *Diz-se ainda que C é parcialmente ordenado se valem as propriedades (i), (ii) e (iii) mas nem sempre $x \preceq y$ ou $y \preceq x$, isto é, nem sempre x e y são comparáveis. □*

A relação usual de \leq (menor ou igual) atende a todos os requisitos para estabelecer uma ordenação no conjunto dos números reais (o conjunto \mathbb{R}). Note-se que a relação de $<$ (menor, no sentido de “estritamente menor”) não atende à reflexividade nem a anti-simetria, não servindo portanto para estabelecer um ordenamento dos reais no sentido acima definido.

No caso do conjunto \mathbb{R}^n , com $n \geq 2$, é possível estabelecer generalizações das relações de “menor” e de “menor ou igual”, definindo-se relações \prec e \preceq por exemplo da seguinte maneira:

Definição 10.3 (Comparação de Vetores) *As operações de comparação entre vetores pertencentes ao espaço \mathbb{R}^n serão definidas da seguinte forma:*

$$x \preceq y \Rightarrow \{x_i \leq y_i, i = 1, \dots, n\}$$

$$x \prec y \Rightarrow \{x_i < y_i, i = 1, \dots, n\}$$

$$x = y \Rightarrow \{x_i = y_i, i = 1, \dots, n\}$$

Os operadores \succeq e \succ são definidos analogamente. Observe-se que o operador \neq é definido de outra forma:

$$x \neq y \Rightarrow \{\exists i \mid x_i \neq y_i\}$$

ou seja:

$$x \neq y \not\Rightarrow \{x_i \neq y_i \forall i = 1, \dots, n\}$$

Dessa forma, a condição de desigualdade (\neq) permanece sendo complementar à condição de igualdade ($=$). \square

Pode-se verificar que a relação \preceq assim definida estabelece um ordenamento parcial do conjunto \mathbb{R}^n , pois as propriedades de reflexividade, transitividade e anti-simetria de \preceq são facilmente verificáveis, enquanto que nem sempre será verdade que $a \preceq b$ ou $b \preceq a$, no caso de vetores a e b nesse conjunto.

EXEMPLO 10.1 *Considerem-se os vetores x , y e z , todos em \mathbb{R}^3 :*

$$x = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \quad y = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \quad z = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

Usando a Definição 10.3, são verdadeiros os resultados das seguintes operações de comparação desses vetores dois a dois:

$$y \prec x \Rightarrow y \preceq x$$

$$z \prec x \Rightarrow z \preceq x$$

Definindo agora a relação $\not\preceq$ como significando “a relação \preceq não ocorre”, obtém-se também:

$$z \not\preceq y$$

$$y \not\preceq z$$

Portanto, não há uma ordenação total dos vetores x , y e z . \diamond

É possível definir outras relações diferentes desta capazes de estabelecer ordenamentos parciais no conjunto \mathbb{R}^n . A definição e a proposição a seguir estabelecem esse fato, mostrando possibilidades de construção de ordenamentos parciais alternativos.

Definição 10.4 (Ordem Parcial Linear) *Considere-se o espaço vetorial \mathbb{R}^n sobre o corpo dos reais. Uma relação de ordenação parcial “ \triangleleft ” nesse espaço é dita linear se, além de satisfazer às relações de transitividade, reflexividade e anti-simetria, também satisfaz:*

$$i. \ x \triangleleft y \text{ e } t \geq 0 \Rightarrow tx \triangleleft ty$$

$$ii. \ x \triangleleft y \text{ e } z \triangleleft w \Rightarrow (x + z) \triangleleft (y + w)$$

para $x, y, z, w \in \mathbb{R}^n$ e $t \in \mathbb{R}$. □

Proposição 10.1 *Seja $\mathbf{0}$ a origem do espaço \mathbb{R}^n em um sistema de coordenadas. Suponha-se que “ \triangleleft ” seja uma relação de ordem parcial linear em \mathbb{R}^m . Então o conjunto C_0 definido por*

$$C_0 \triangleq \{y \in \mathbb{R}^n \mid \mathbf{0} \triangleleft y\}$$

é um cone convexo com ponta em $\mathbf{0}$. Equivalentemente, se $C \subseteq \mathbb{R}^m$ é um cone convexo com ponta, então a relação \triangleleft definida por

$$x \triangleleft y \Leftrightarrow (y - x) \in C$$

é uma ordem parcial linear em \mathbb{R}^n . □

Essa maneira de definir uma relação de ordem parcial “ \triangleleft ” é uma generalização da relação “ \preceq ” que se estabelece através da Definição 10.3. Esta última corresponde à situação em que o cone C é identificado com o cone não-negativo do \mathbb{R}^n .

EXEMPLO 10.2 *Considere-se o cone C definido da seguinte forma:*

$$x \in C \Leftrightarrow x = \alpha_1 c_1 + \alpha_2 c_2$$

sendo os vetores c_1 e c_2 dados por:

$$c_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad c_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

e os escalares α_1 e α_2 tais que $\alpha_1 \geq 0$ e $\alpha_2 \geq 0$. Considerem-se os vetores y, z, w dados por:

$$w = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad y = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad z = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Pode-se verificar que

$$w - y = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \times c_1 + 0 \times c_2 \Rightarrow (w - y) \in C \Rightarrow y \triangleleft w$$

Por outro lado, verifica-se também que

$$w - z = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 1 \times c_1 + (-1) \times c_2 \Rightarrow (w - z) \notin C \Rightarrow y \not\triangleleft z$$

$$z - w = \begin{bmatrix} 0 \\ -2 \end{bmatrix} = -1 \times c_1 + 1 \times c_2 \Rightarrow (z - w) \notin C \Rightarrow z \not\triangleleft y$$

◇

Retomando agora o problema de otimização vetorial conforme enunciado na expressão (10.1):

$$(POV) \begin{cases} \mathcal{X}^* = x^* \in \mathbb{R}^n \mid \\ x^* = \arg \min_x f(x) \\ \text{sujeito a: } x \in \mathcal{F}_x \end{cases} \quad (10.2)$$

Esse problema fica bem definido quando se define a operação de min como a determinação dos elementos de um conjunto que não possuam predecessores naquele conjunto, de acordo com uma relação de ordenação parcial “ \triangleleft ”. Essa definição é apresentada a seguir.

Definição 10.5 (Operação min) *Seja um conjunto Ω cujos elementos são dotados de uma ordem parcial \triangleleft . A operação min sobre esse conjunto, denotada por*

$$y = \min_{x \in \Omega} x$$

é definida de forma que $y \in \Omega$ não possua nenhum predecessor além de si próprio, ou seja, de forma que seja verdadeira a expressão

$$\nexists x \in \Omega \mid x \neq y \text{ e } x \triangleleft y.$$

□

Pode-se verificar que um caso particular da operação min definida dessa forma corresponde à minimização usual quando o conjunto Ω for um subconjunto dos reais e a relação de ordem for dada pelo operador \leq (menor ou igual) usual.

A nomenclatura predominantemente empregada hoje em dia faz referência a um *Problema de Otimização Vetorial* quando o conjunto $\Omega = \text{Imagem}(f(\mathcal{F}_x))$ é um subconjunto do espaço \mathbb{R}^n e a ordem parcial definida pelo operador \triangleleft é uma ordem parcial linear. Um caso particular deste, no qual o cone que induz a ordem parcial é o cone não-negativo do \mathbb{R}^n (ou seja, quando a ordem parcial é definida a partir do operador \preceq), é chamado de *Problema de Otimização Multiobjetivo*.

10.3 O Conjunto Pareto-Ótimo

Considere-se o problema de otimização vetorial definido por (10.1). A figura 10.1 mostra o que seria o mapeamento $f(\cdot)$ levando de um espaço \mathcal{X} a pontos de um espaço \mathcal{Y} , num problema sem restrições (ou seja, um problema em que $\mathcal{F}_x = \mathcal{X}$). A região \mathcal{I} corresponde à imagem da função $f(\cdot)$. O ponto $y_A \in \mathcal{Y}$ corresponde ao vetor de objetivos que ocorre como imagem do ponto x_A que minimiza o funcional $f_1(\cdot)$. O ponto $y_B \in \mathcal{Y}$ corresponde ao vetor de objetivos que ocorre como imagem do ponto x_B que minimiza o funcional $f_2(\cdot)$.

A figura 10.2 mostra o que seria o mapeamento $f(\cdot)$ levando de um espaço \mathcal{X} a pontos de um espaço \mathcal{Y} , num problema agora com restrições (ou seja, um problema em que $\mathcal{F}_x \subset \mathcal{X}$). A região $\mathcal{F}_y \subset \mathcal{I}$ corresponde à imagem da função $f(\cdot)$ restrita ao domínio \mathcal{F}_x . O ponto $y_A \in \mathcal{Y}$ corresponde ao vetor de objetivos que ocorre como imagem do ponto x_A que minimiza o funcional $f_1(\cdot)$. O ponto $y_B \in \mathcal{Y}$ corresponde ao vetor de objetivos que ocorre como imagem do ponto x_B que minimiza o funcional $f_2(\cdot)$. Observe-se que embora um ponto de mínimo de algum funcional possa ocorrer no interior da região \mathcal{F}_x (na figura 10.2, esse é o caso do ponto x_A), os pontos correspondentes às imagens de todos os mínimos no espaço \mathcal{Y} necessariamente se encontrarão na fronteira do conjunto \mathcal{F}_y (vide o ponto y_A).

O objeto fundamental da otimização multiobjetivo consiste em um conjunto de soluções \mathcal{X}^* , denominado *conjunto Pareto-ótimo*, que irá conter as possíveis soluções x^* do POV acima definido. Os elementos desse conjunto são definidos a seguir. Preliminarmente, define-se o conceito de *dominância*.

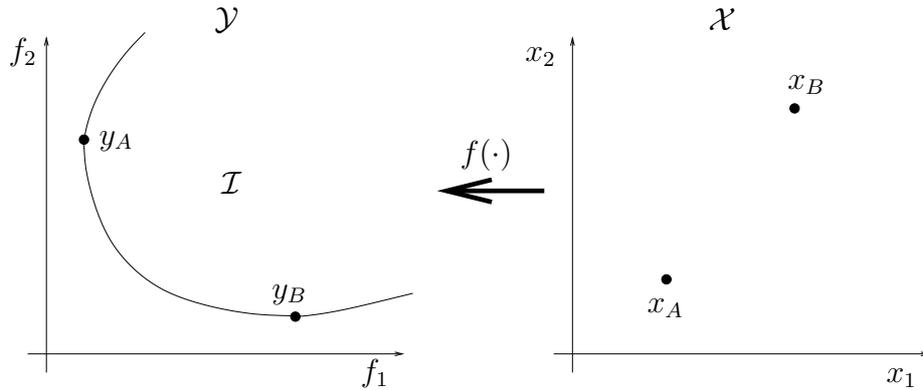


Figura 10.1: Mapeamento do espaço de parâmetros \mathcal{X} no espaço de objetivos \mathcal{Y} feito pela função $f(\cdot)$. O conjunto-imagem da função $f(\cdot)$ é designado por \mathcal{I} . O ponto de mínimo do funcional f_1 no espaço de parâmetros é designado por x_A , e sua imagem no espaço \mathcal{Y} é designada por y_A . O ponto de mínimo do funcional f_2 no espaço de parâmetros é designado por x_B , e sua imagem no espaço \mathcal{Y} é designada por y_B .

Definição 10.6 (Dominância) Diz-se que o ponto $x_1 \in \mathcal{X}$ domina o ponto $x_2 \in \mathcal{X}$ se $f(x_1) \leq f(x_2)$ e $f(x_1) \neq f(x_2)$. Equivalentemente, diz-se que $f(x_1) \in \mathcal{Y}$ domina $f(x_2) \in \mathcal{Y}$, nessas mesmas condições. \square

O conceito de *dominância* encontra-se ilustrado nas figuras 10.3 e 10.4.

Definição 10.7 (Solução Pareto-Ótima) Diz-se que $x^* \in \mathcal{F}_x$ é uma solução Pareto-Ótima do POV se não existe qualquer outra solução $x \in \mathcal{F}_x$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$, ou seja, se x^* não é dominado por nenhum outro ponto factível. \square

NOTA 10.1 Na literatura é encontrada uma terminologia diversificada para fazer referência ao conjunto de soluções Pareto-ótimas:

$$\text{Pareto-Ótima} = \text{eficiente} = \text{não-dominada} = \text{não-inferior}$$

\diamond

Os conjuntos de Pareto (ou seja, conjuntos das soluções Pareto-ótimas) associados às figuras 10.1 e 10.2 são mostrados respectivamente nas figuras 10.5 e 10.6. A figura 10.7 mostra um caso em que o conjunto Pareto-ótimo é não-conexo, embora o conjunto \mathcal{I} assim como sua fronteira sejam conexos.

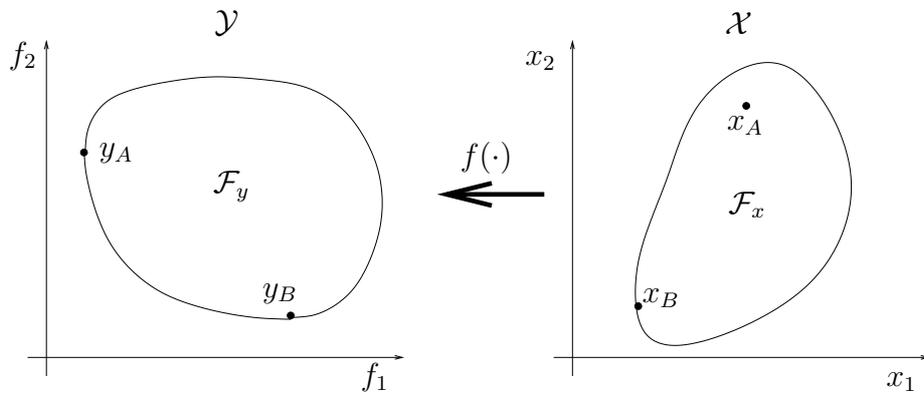


Figura 10.2: Mapeamento do espaço de parâmetros \mathcal{X} no espaço de objetivos \mathcal{Y} feito pela função $f(\cdot)$. A região factível no espaço de parâmetros é designada por \mathcal{F}_x , sendo o conjunto-imagem da função $f(\cdot)$ restrita à região factível designado por \mathcal{F}_y . O ponto de mínimo do funcional f_1 no espaço de parâmetros é designado por x_A , e sua imagem no espaço \mathcal{Y} é designada por y_A . O ponto de mínimo do funcional f_2 no espaço de parâmetros é designado por x_B , e sua imagem no espaço \mathcal{Y} é designada por y_B .

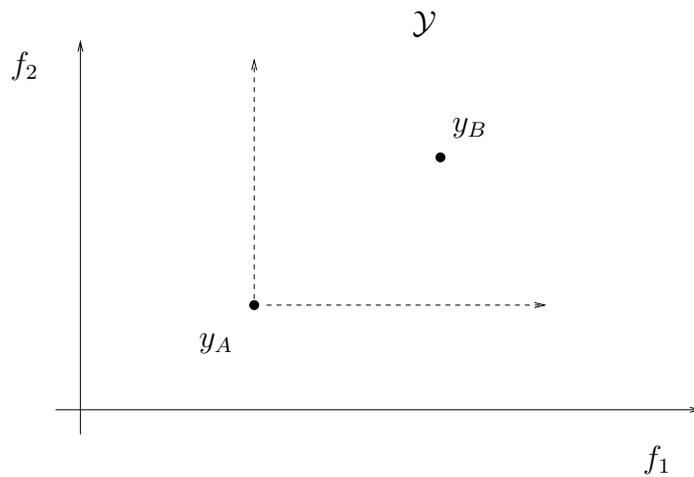


Figura 10.3: São mostrados os pontos y_A e y_B pertencentes ao espaço de objetivos \mathcal{Y} . Um cone paralelo aos eixos coordenados do espaço \mathcal{Y} é colocado com vértice no ponto y_A . Todos os pontos no interior desse cone são dominados por y_A , de forma que y_A domina y_B .

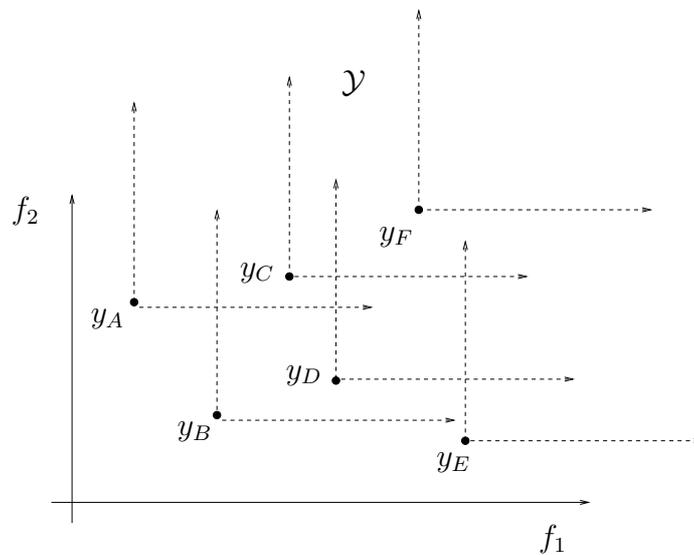


Figura 10.4: São mostrados os pontos y_A a y_F pertencentes ao espaço de objetivos \mathcal{Y} . Cones paralelos aos eixos coordenados do espaço \mathcal{Y} são colocados com vértices em cada um desses pontos. Todos os pontos no interior de cada cone são dominados pelo ponto que se localiza no seu vértice, de forma que: (i) entre y_A , y_B e y_E não há relação de dominância; (ii) y_A domina y_C e y_F ; (iii) y_B domina y_C , y_D e y_F ; (iv) entre y_C , y_D e y_E não há relação de dominância; (v) y_C e y_D dominam y_F ; (vi) y_E e y_F não dominam nenhum outro ponto mostrado na figura.

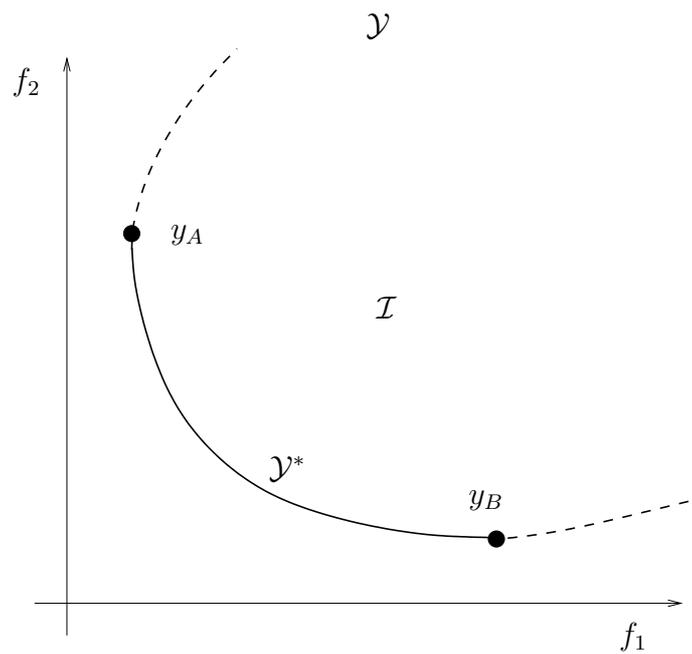


Figura 10.5: O conjunto de Pareto $\mathcal{Y}^* \in \mathcal{Y}$, associado ao conjunto-imagem \mathcal{I} da função $f(\cdot)$ encontra-se representado em linha contínua. O restante da fronteira do conjunto \mathcal{I} está representado em linha tracejada. O conjunto \mathcal{Y}^* possui extremos nos pontos y_A e y_B , correspondentes aos mínimos individuais dos funcionais f_1 e f_2 . Esta figura corresponde à figura 10.1.

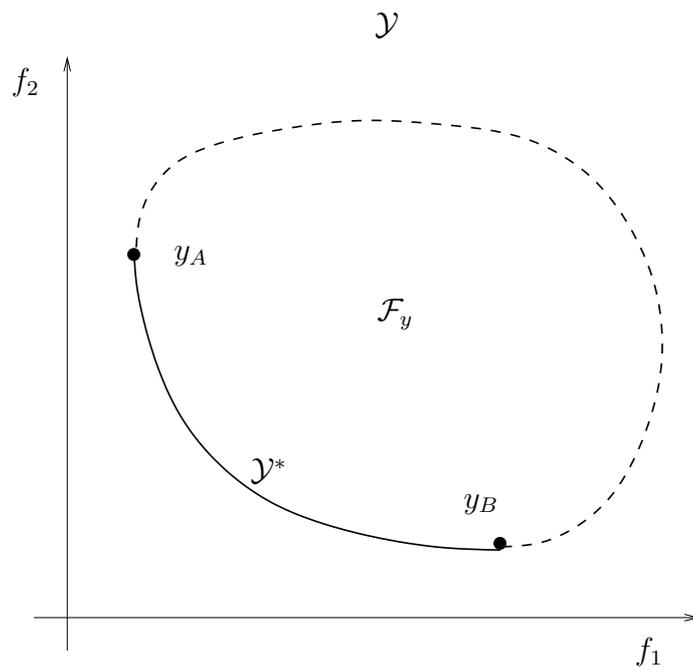


Figura 10.6: O conjunto de Pareto $\mathcal{Y}^* \in \mathcal{Y}$, associado ao conjunto-imagem \mathcal{F}_y da função $f(\cdot)$ encontra-se representado em linha contínua. O restante da fronteira do conjunto \mathcal{F}_y está representado em linha tracejada. O conjunto \mathcal{Y}^* possui extremos nos pontos y_A e y_B , correspondentes aos mínimos individuais dos funcionais f_1 e f_2 . Esta figura corresponde à figura 10.2.

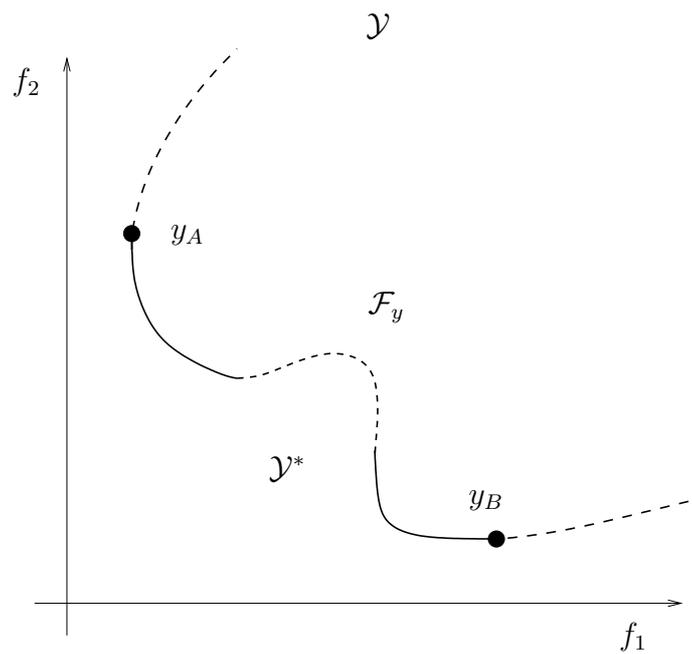


Figura 10.7: O conjunto de Pareto $\mathcal{Y}^* \in \mathcal{Y}$, associado ao conjunto-imagem \mathcal{F}_y da função $f(\cdot)$ encontra-se representado em linha contínua. O restante da fronteira do conjunto \mathcal{F}_y está representado em linha tracejada. O conjunto \mathcal{Y}^* possui extremos nos pontos y_A e y_B , correspondentes aos mínimos individuais dos funcionais f_1 e f_2 . Esta figura mostra que o conjunto Pareto-ótimo pode eventualmente ser não-conexo.

As seguintes relações podem ser facilmente estabelecidas:

$$\begin{aligned}\mathcal{X} \supset \mathcal{F}_x \supset \mathcal{X}^* \\ \mathcal{Y} \supset \mathcal{I} \supset \mathcal{F}_y \supset \mathcal{Y}^*\end{aligned}\tag{10.3}$$

sendo que:

$$\begin{aligned}\mathcal{I} &= f(\mathcal{X}) \\ \mathcal{F}_y &= f(\mathcal{F}_x) \\ \mathcal{Y}^* &= f(\mathcal{X}^*)\end{aligned}\tag{10.4}$$

A proposição a seguir estabelece outras relações úteis.

Proposição 10.2 *Todas as soluções eficientes do POV estão contidas em $\partial\mathcal{F}_y$ ou, equivalentemente, nenhuma solução eficiente está contida em \mathcal{F}_y° , ou seja:*

$$\begin{aligned}\mathcal{Y}^* &\subset \partial\mathcal{F}_y \\ \mathcal{Y}^* \cap \mathcal{F}_y^\circ &= \emptyset\end{aligned}\tag{10.5}$$

□

10.3.1 Conjunto localmente Pareto-ótimo

É possível estabelecer a definição de conjunto Pareto-ótimo num sentido local. Assim como no caso da otimização escalar, a otimização vetorial frequentemente irá trabalhar com soluções apenas locais, seja nos casos em que é difícil estabelecer a globalidade das soluções encontradas, seja nos casos em que não é necessária essa globalidade para propósitos práticos.

Definição 10.8 (Solução Localmente Pareto-Ótima) *Diz-se que $x \in \mathcal{F}_x$ é uma solução localmente Pareto-Ótima do POV numa dada vizinhança $\mathcal{N}(x^*, \delta)$ se existe $\delta > 0$ tal que não existe qualquer outra solução $x \in \mathcal{N}(x^*, \delta) \cap \mathcal{F}_x$ que faça $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$, ou seja, se x^* não é dominado por nenhum outro ponto naquela vizinhança.* □

O teorema a seguir conecta a propriedade de Pareto-otimalidade no sentido local com tal propriedade no sentido global, para problemas convexos.

Teorema 10.1 *Sejam $f_i(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \dots, m$ funções convexas definidas sobre um conjunto convexo $\mathcal{X} \subset \mathbb{R}^n$. Então toda solução localmente Pareto-Ótima é globalmente Pareto-Ótima.* \square

10.3.2 Solução utópica

Um conceito útil, relacionado ao conceito de conjunto Pareto-ótimo, é o de solução utópica.

Definição 10.9 (Solução Utópica) *A solução utópica y_* do POV é definida como:*

$$y_{*i} = f_i(x^i) \quad , \quad i = 1, \dots, m \quad (10.6)$$

onde:

$$x^i = \arg \min_{x \in \mathcal{F}_x} f_i(x)$$

\square

Note-se que se $y_* \in \mathcal{Y}$, isto é, se existe $x^* \in \mathcal{F}$ tal que $y_* = f(x^*)$ então o problema está resolvido: o ponto x^* é o ótimo correspondente a todas as funções-objetivo, e deve ser adotado como solução. Entretanto, via de regra, y_* não é uma solução existente, ou seja, não é um ponto pertencente à imagem da função $f(\cdot)$ restrita ao domínio \mathcal{F} . Essa é a situação que justifica um estudo da otimização vetorial segundo técnicas diferentes daquelas da otimização escalar, quando existe um número possivelmente ilimitado de soluções eficientes.

A solução utópica y_* é uma escolha bastante conveniente para ser tomada como “origem” do espaço de objetivos \mathcal{Y} , uma vez que tal escolha deixa todo o conjunto de Pareto incluído no primeiro quadrante, com interseções com os eixos coordenados correspondentes aos mínimos individuais de cada um dos funcionais que compõem o vetor de objetivos. De fato, o conjunto Pareto-ótimo encontra-se necessariamente contido no hiper-paralelepípedo definido no espaço de objetivos pela solução utópica e pelas soluções ótimas individuais de cada problema escalar que compõe o vetor de objetivos. Isso está ilustrado na figura 10.8.

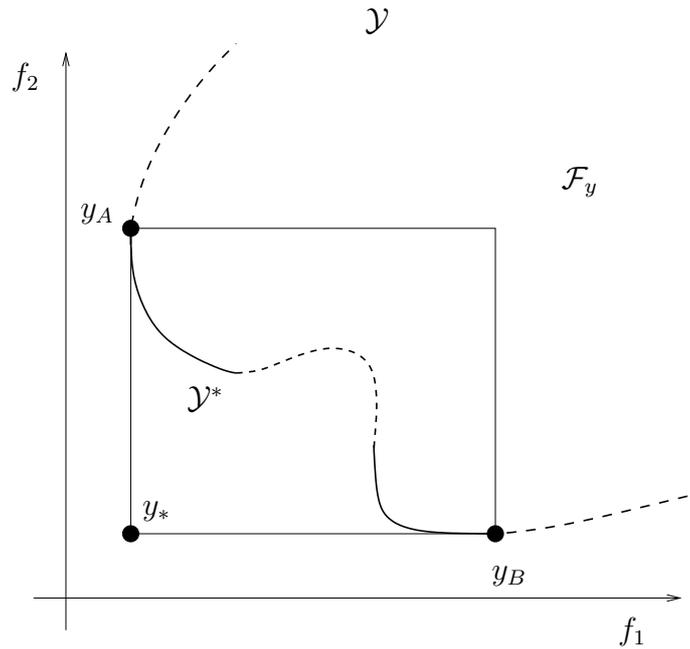


Figura 10.8: O conjunto de Pareto $\mathcal{Y}^* \in \mathcal{Y}$ (representado em linha contínua) encontra-se contido no hiper-paralelogramo definido pela solução utópica y_* e pelas soluções correspondentes aos mínimos individuais dos funcionais f_1 e f_2 (respectivamente os pontos y_A e y_B).

Interpretação em termos de dominância

Sejam os conjuntos definidos por:

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

$$C_{\geq}(x^*) = \{x : f(x) \geq f(x^*)\}$$

$$C_{\sim}(x^*) = \{x : f(x) \text{ não comparável a } f(x^*)\}$$

Propriedades:

- i. $\mathbb{R}^n = C_{<}(x^*) \cup C_{\geq}(x^*) \cup C_{\sim}(x^*)$
- ii. Se $\mathcal{F}_x \cap C_{<}(x^*) = \emptyset$ e $x^* \in \mathcal{F}_x$ então x^* é eficiente.

NOTA 10.2 Considere o caso em que as funções $f_i(\cdot)$ sejam lineares:

$$f_i(x) = c^i x \quad ; \quad c^i \in \mathbb{R}^{1 \times n}, \quad i = 1, \dots, m$$

Seja C o cone gerado pelos gradientes $\{ -c^1 \ -c^2 \ \dots \ -c^m \}$. Nesse caso:

$$C^* = C_{<}(x^*)$$

ou seja, o cone C sobre o ponto x^* coincide com o conjunto $C_{<}(x^*)$.

◇

10.4 O Problema de Determinação das Soluções Eficientes

A seguir é discutida a possibilidade da aplicação dos métodos mono-objetivo de otimização não-linear para a geração de soluções eficientes em problemas de otimização vetorial. As estruturas básicas de formulação dos problemas multiobjetivo: (P_λ) , (P_ϵ) , (P_χ) , (P^*) e (P_{KTE}) são apresentadas e desdobradas em algoritmos, segundo uma formulação conceitual apoiada apenas na existência abstrata de mecanismos de otimização.

Deve-se entender toda a discussão aqui apresentada tendo como objetivo a geração, sempre, de um subconjunto representativo do conjunto de soluções eficientes do Problema Multiobjetivo (PMO). Ou seja, deseja-se gerar um grande número de pontos que de certa forma descreva o conjunto de soluções eficientes, sem ainda levar em consideração a preferência do decisor quanto a tais pontos (isso significa que a “função de utilidade” não está sendo levada em consideração neste capítulo). Toda a estrutura conceitual aqui apresentada seria aplicável, com adaptações mínimas, para o problema da geração de um único ponto pertencente ao conjunto de soluções eficientes, a partir do qual se iniciaria um processo de interação com o decisor para a geração de novos pontos. Esse tema, entretanto, será assunto de capítulos posteriores, que irão constituir a última parte deste texto.

Para efeito de definir o problema que será referenciado ao longo do restante deste capítulo, considere-se o problema de otimização multiobjetivo com um vetor de objetivos $f(\cdot)$:

$$\mathcal{X}^* = \{x^* \in \mathcal{F}_x : \nexists x \in \mathcal{F}_x \mid f(x) \leq f(x^*) ; f(x) \neq f(x^*)\} \quad (10.7)$$

sendo que $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$ e $g(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^p$. O problema a ser tratado a partir de agora é o de determinar um conjunto de pontos que seja representativo do conjunto \mathcal{X}^* de soluções eficientes do problema. Esse problema se subdivide em: (i) localizar tais pontos e (ii) estabelecer condições que garantam que os mesmos de fato pertençam ao conjunto de soluções eficientes.

10.5 Condições de Kuhn-Tucker para Eficiência

Condições equivalentes às condições de Kuhn-Tucker podem ser estabelecidas a partir do seguinte resultado:

Teorema 10.2 *Se x^* é eficiente então x^* resolve os m problemas*

$$(P_i, i = 1, \dots, m) \begin{cases} \min_{x \in \mathcal{F}_x} f_i(x) \\ \text{sujeito a } f_j(x) \leq f_j(x^*) \quad \forall j \neq i \end{cases} \quad (10.8)$$

Reciprocamente, se x^ resolve $(P_i, i = 1, \dots, m)$ então x^* é eficiente. \square*

DEMONSTRAÇÃO:

(\Rightarrow)

1. x^* é eficiente
2. então: $\nexists x \in \mathcal{F}_x : f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$
3. logo x^* resolve $(P_i, i = 1, \dots, m)$

(\Leftarrow)

1. x^* resolve $(P_i, i = 1, \dots, m)$
2. se x^* não é eficiente, $\exists x \in \mathcal{F}_x$ e algum i tal que $f_i(x) < f_i(x^*)$
3. se fosse esse o caso, x^* não resolveria, portanto, o problema P_i .

■

Lembrando que $\mathcal{F}_x = \{x : g(x) \leq 0\}$, as condições de Kuhn-Tucker de cada problema (P_i) seriam:

$$(KT_i) \begin{cases} \lambda_j^* \geq 0 ; \mu_k^* \geq 0 \\ g_k(x^*) \leq 0 ; \mu_k^* g_k(x^*) = 0 \quad , \quad k = 1, \dots, p \\ \nabla f_i(x^*) + \sum_{j \neq i} \lambda_j^* \nabla f_j(x^*) + \sum_{k=1}^p \mu_k^* \nabla g_k(x^*) = 0 \end{cases} \quad (10.9)$$

Definição 10.10 (Condições de Kuhn-Tucker para Eficiência) *Uma solução factível x^* satisfaz as condições necessárias de Kuhn-Tucker para eficiência (KTE) se:*

- i. todos f_i e g_i são diferenciáveis e*
- ii. existem vetores multiplicadores $\mu^* \geq 0$, $\lambda^* \geq 0$, com pelo menos uma desigualdade estrita $\lambda_i^* > 0$, tais que:*

$$g_k(x^*) \leq 0 \quad ; \quad \mu_k^* g_k(x^*) = 0 \quad ; \quad k = 1, \dots, p$$

$$\sum_{j=1}^m \lambda_j^* \nabla f_j(x^*) + \sum_{k=1}^p \mu_k^* \nabla g_k(x^*) = 0 \tag{10.10}$$

□

NOTA 10.3 *Observe-se que:*

$$(KT_i, i = 1, \dots, m) \Rightarrow (KTE)$$

$$(KTE) \Rightarrow (KT_i, i = 1, \dots, m) \text{ se } \lambda_j^* > 0, j = 1, \dots, m$$

(10.11)

◇

EXEMPLO 10.3 (Caráter apenas necessário das KTE) *Considere o problema bi-objetivo:*

$$\min_x [f_1(x) \quad f_2(x)]$$

$$g(x) = -x \leq 0$$

(10.12)

onde as funções $f_1(\cdot)$ e $f_2(\cdot)$ são como descritas na figura 10.9.

Tem-se que:

- $f_1(\cdot)$ e $f_2(\cdot)$ são convexas;
- KTE é sempre satisfeita no intervalo (b, c) :

$$\lambda \geq 0 \quad ; \quad \mu \geq 0$$

$$g(x) \leq 0 \quad ; \quad \mu g(x) = 0$$

$$\lambda_1 \frac{d}{dx} f_1(x) + \lambda_2 \frac{d}{dx} f_2(x) - \mu = 0$$

Basta escolher:

$$\lambda_1 > 0 \quad ; \quad \lambda_2 = \mu = 0$$

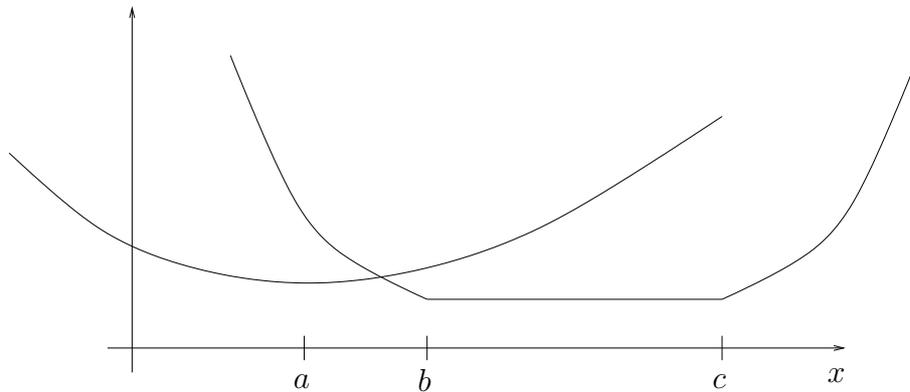


Figura 10.9: Exemplo de situação em que as condições KTE são satisfeitas em pontos não-eficientes.

Entretanto, as soluções no intervalo (b, c) não são eficientes. O leitor deverá notar que as soluções eficientes encontram-se no intervalo $[a, b]$. \diamond

Condições suficientes

Embora as KTE sejam condições necessárias, mas não suficientes para a Pareto-otimalidade de soluções, sob determinadas circunstâncias essas condições tornam-se suficientes.

Teorema 10.3 (Condição 1) *Se o POV atender às KTE em x^* com todos os multiplicadores λ_i^* , $i = 1, \dots, m$ estritamente positivos então x^* é uma solução eficiente.* \square

DEMONSTRAÇÃO: Esse resultado vem diretamente do fato de que agora KTE se reduz a $(KT_i, i = 1, \dots, m)$ e, portanto, x^* é eficiente. \blacksquare

Teorema 10.4 (Condição 2) *Se o POV atender à KTE em x^* e tiver todas as funções $f_i(\cdot)$, $i = 1, \dots, m$ e $g_k(\cdot)$, $k = 1, \dots, p$ estritamente convexas, então x^* é uma solução eficiente.* \square

DEMONSTRAÇÃO:

1. Como por hipótese x^* satisfaz KTE, existe no mínimo um i tal que $\lambda_i^* > 0$. Nesse caso:

$$\nabla f_i(x^*) + \sum_{j \neq i} \frac{\lambda_j^*}{\lambda_i^*} \nabla f_j(x^*) + \sum_{k=1}^p \frac{\mu_k^*}{\lambda_i^*} \nabla g_k(x^*) = 0$$

e portanto x^* resolve (P_i) .

2. Como $f_i(\cdot)$ é estritamente convexa, a solução de (P_i) é única na região convexa definida pelas restrições $g_k(\cdot) \leq 0$ e $f_j(\cdot) \leq f_j(x^*)$.
3. Portanto, x^* é eficiente.

■

NOTA 10.4 *Como interpretação geométrica das KTE, pode-se notar que, para pontos do conjunto de Pareto, o cone no qual as funções $f_i(\cdot)$ decrescem, a partir de x^* :*

$$-\sum_{j=1}^m \lambda_j^* \nabla f_j(x^*)$$

está estritamente fora da região factível, ou seja, não contém nenhuma direção factível. As direções factíveis são dadas pelo cone:

$$-\sum_{k=1}^p \mu_k^* \nabla g_k(x^*)$$

Note-se ainda que o próprio cone

$$\sum_{j=1}^m \lambda_j^* \nabla f_j(x^*)$$

pode ser nulo.

◇

Capítulo 11

Formulação do Problema de Geração do Conjunto de Soluções Eficientes

Neste capítulo são discutidas metodologias conceituais para geração dos pontos x^* que pertencem ao conjunto de soluções eficientes de um problema de otimização vetorial (POV). Dessa forma, o problema de geração de pontos pertencentes ao conjunto Pareto-ótimo fica reduzido a problemas de otimização mono-objetivo.

Por fim, coloca-se a questão de caracterizar computacionalmente as soluções que pertencem ao conjunto de soluções eficientes do POV. Um novo problema de otimização mono-objetivo é definido com tal propósito.

11.1 Abordagem via Problema Ponderado

Retomando a discussão sobre condições de Kuhn-Tucker, observe-se que as condições KTE são também condições necessárias (de Kuhn-Tucker mono-objetivo) para o problema escalar:

Problema P_λ

$$\min_{x \in \mathcal{F}_x} \sum_{i=1}^m \lambda_i f_i(x) \quad (11.1)$$

Isso sugere um método prático para a geração de soluções eficientes.

Teorema 11.1 *Seja x^* a solução de (P_λ) para $\lambda \geq 0$ dado. Então x^* é uma solução eficiente do PMO se:*

- i. x^* é a solução única de (P_λ) ou*
- ii. $\lambda_i > 0$; $i = 1, \dots, m$.*

□

Note que, sem perda de generalidade, é possível considerar

$$\lambda \in \Lambda = \left\{ \lambda : \lambda \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}$$

e, em seguida, resolver (P_λ) .

11.1.1 Interpretação geométrica

Uma interpretação geométrica do problema ponderado (P_λ) pode ser feita através de sua formulação no espaço dos objetivos.

$$\min_{y \in \mathcal{F}_y} \sum_{i=1}^m \lambda_i y_i \quad (11.2)$$

O problema pode ser visto como: determinar o ponto y^* pertencente ao conjunto \mathcal{F}_y que minimiza o valor do produto escalar $\lambda'y$. Se α^* é o valor mínimo do produto, então:

$$\lambda'y = \alpha^* \quad (11.3)$$

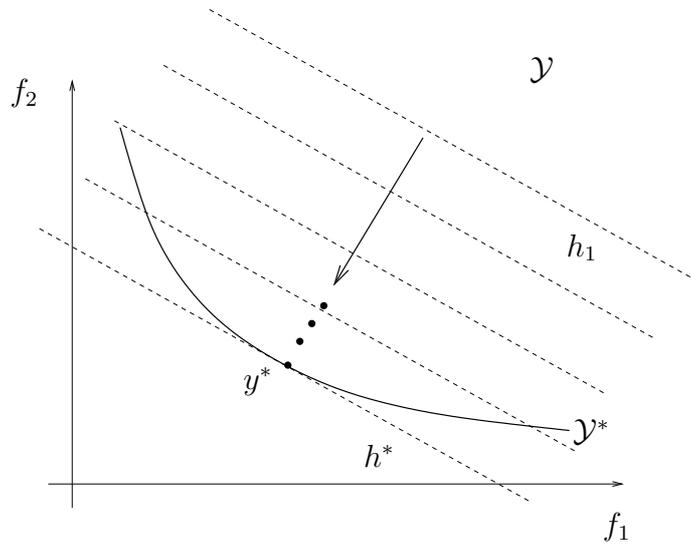


Figura 11.1: Interpretação do problema P_λ no espaço de objetivos: Partindo de um hiperplano inicial h_1 , cuja inclinação é definida pelo vetor de ponderações λ , um algoritmo de otimização mono-objetivo determina uma seqüência de hiperplanos paralelos ao hiperplano inicial, buscando a minimização da distância de cada hiperplano em relação à origem do espaço \mathcal{Y} , com a restrição de que o hiperplano contenha algum ponto de \mathcal{F}_y . O algoritmo converge para o hiperplano-suporte h^* , cujo único ponto de interseção com \mathcal{F}_x é o ponto y^* pertencente ao conjunto de Pareto \mathcal{Y}^* .

é a equação do hiperplano suporte a \mathcal{F}_y no ponto y^* . Essa interpretação é ilustrada na figura 11.1.

Apenas pontos do conjunto Pareto-ótimo \mathcal{Y}^* que admitam hiperplanos-suporte podem ser gerados através de (P_λ) . O raciocínio que conduz a esta conclusão pode ser elaborado desta maneira: um dado vetor de ponderações λ define a inclinação de um hiperplano. O processo de minimização descrito por (11.2) faz a busca do hiperplano com tal inclinação que se encontra a mínima distância da origem do espaço \mathcal{Y} . Esse hiperplano necessariamente será um hiperplano-suporte de algum ponto. Assim, mudando-se o vetor de ponderações λ , muda-se o hiperplano suporte que será encontrado, com seu respectivo ponto tangente, que é o ponto do conjunto de Pareto determinado por esse procedimento. Em casos gerais, apenas pontos pertencentes à fronteira da casca convexa do conjunto \mathcal{F}_y poderão ser gerados. Isso é ilustrado na figura 11.2.

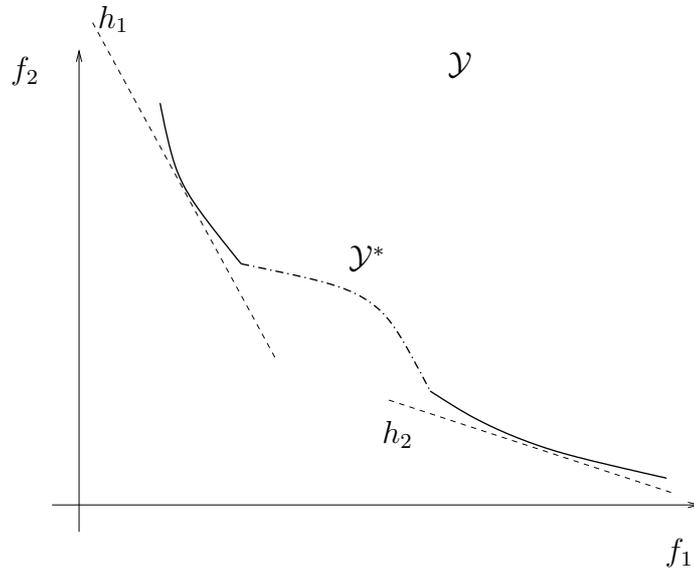


Figura 11.2: O conjunto de Pareto $\mathcal{Y}^* \in \mathcal{Y}$ possui trechos, representados em linha contínua, nos quais todos os pontos possuem hiperplano-suporte, e outros trechos, representados em linha tracejada, nos quais nenhum ponto admite hiperplano-suporte. Dois hiperplanos-suporte, h_1 e h_2 , são mostrados.

EXEMPLO 11.1 *Considerem-se as duas funções de uma única variável:*

$$f_1(x) = 1 - \exp\left(\frac{-(x-4)^2}{9}\right)$$

$$f_2(x) = 1 - \exp\left(\frac{-(x+4)^2}{9}\right)$$

Ambas as funções são quasi-convexas, e estão representadas no gráfico da figura 11.3 em traço contínuo. A função dada por:

$$F_1(x) = 0.7f_1(x) + 0.3f_2(x)$$

aparece representada no mesmo gráfico em traço-traço. A função:

$$F_2(x) = 0.4f_1(x) + 0.6f_2(x)$$

aparece em ponto-ponto. Pode-se perceber que, apesar dos pontos da superfície Pareto-ótima do vetor de funções $[f_1 \ f_2]'$ serem todos aqueles para $x \in [-4, 4]$, a solução do problema formulado como:

$$\min_x \alpha f_1(x) + (1 - \alpha)f_2(x)$$

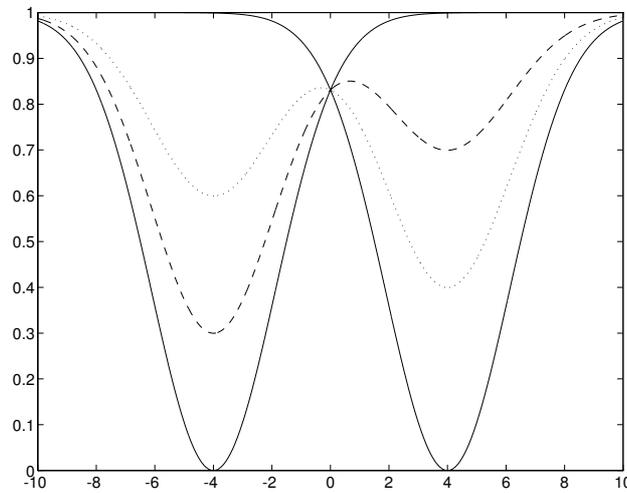


Figura 11.3: Representação das funções-objeto f_1 , f_2 e das funções ponderadas $0.7f_1 + 0.3f_2$ e $0.4f_1 + 0.6f_2$.

irá sempre produzir como resultados apenas os pontos $x = 4$ e $x = -4$. Esses são os únicos pontos da superfície de Pareto que pertencem também à casca convexa da região mostrada na figura 11.4. \diamond

EXEMPLO 11.2 Considerem-se as duas funções de uma única variável:

$$f_3(x) = \frac{(x - 5)^2}{100}$$

$$f_4(x) = |x + 5|$$

Ambas as funções são convexas. O problema ponderado (P_λ) , correspondente à minimização da função mono-objetivo:

$$F(x) = \alpha f_3(x) + (1 - \alpha)f_4(x)$$

é capaz de gerar todas as soluções eficientes deste problema, ou seja, todos os pontos no intervalo $[-5, 5]$. Isso pode ser visualizado na figura 11.5. No espaço de objetivos, tem-se agora uma região \mathcal{F} cuja fronteira Pareto-ótima é convexa, conforme pode ser verificado na figura 11.6. Observe-se que a região como um todo não é convexa. \diamond

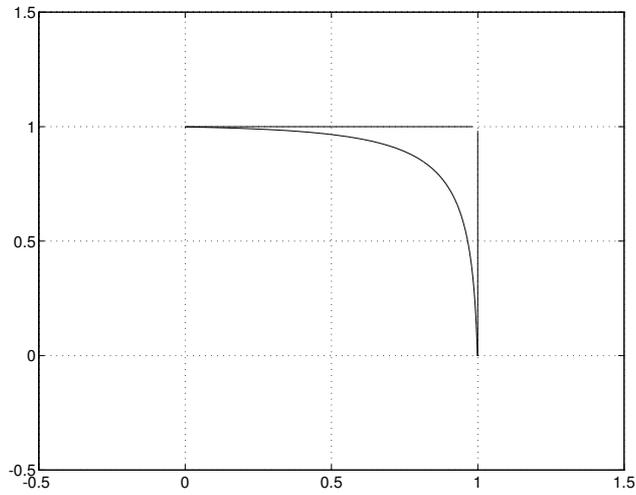


Figura 11.4: Representação da região \mathcal{F} (espaço de objetivos), para um vetor de objetivos composto de duas funções quasi-convexas não convexas.

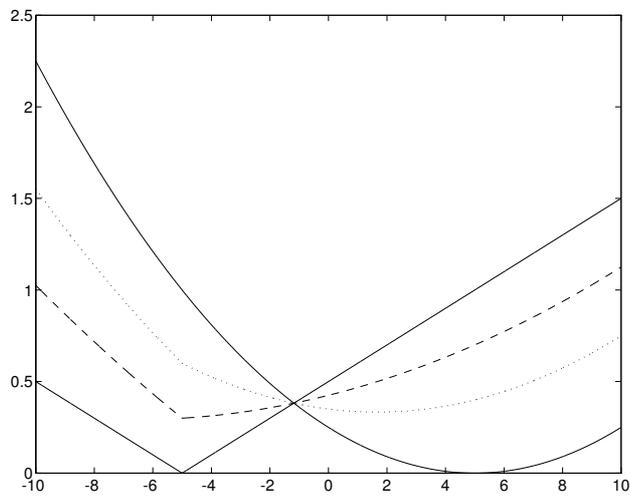


Figura 11.5: Representação das funções-objetivo f_3 , f_4 e das funções ponderadas $0.7f_3 + 0.3f_4$ e $0.4f_3 + 0.6f_4$.

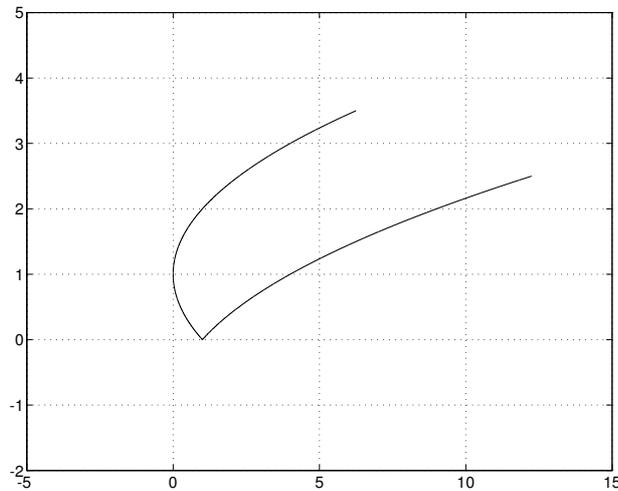


Figura 11.6: Representação da região \mathcal{F} (espaço de objetivos), para um vetor de objetivos composto de duas funções convexas.

EXEMPLO 11.3 *Considerem-se as duas funções de uma única variável:*

$$f_5(x) = \frac{(x - 5)^2}{100}$$

$$f_6(x) = 1,5 \left(1 - \exp\left(\frac{-(x+4)^2}{16}\right) \right)$$

A função f_5 é convexa, enquanto f_6 é não-convexa. O problema ponderado (P_λ) , correspondente à minimização da função mono-objetivo:

$$F(x) = \alpha f_5(x) + (1 - \alpha) f_6(x)$$

é capaz de gerar algumas soluções eficientes deste problema, ou seja, alguns pontos próximos dos mínimos individuais $x = -4$ e $x = 5$. Isso pode ser visualizado na figura 11.7. No espaço de objetivos, tem-se agora uma região \mathcal{F} cuja fronteira Pareto-ótima não é convexa, mas que possui regiões localmente convexas (pertencentes à casca convexa de \mathcal{F}), conforme pode ser verificado na figura 11.8. Essas regiões correspondem precisamente às soluções que podem ser geradas através do problema ponderado. \diamond

NOTA 11.1 *Foi observado nesses exemplos que é possível gerar desde todas as soluções eficientes (quando o problema é convexo) através do problema ponderado,*

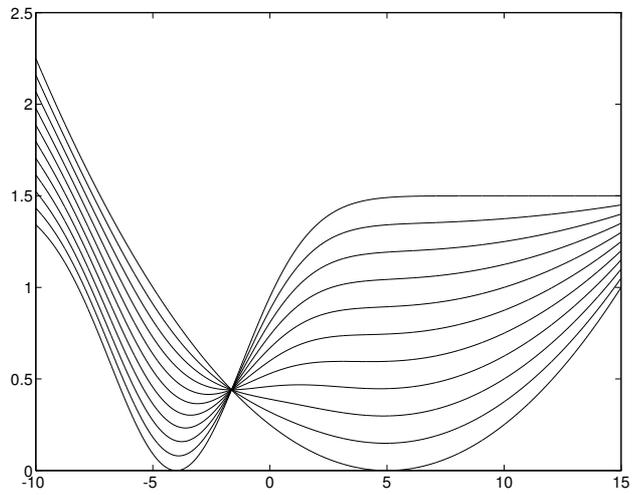


Figura 11.7: Representação das funções-objetivo f_5 , f_6 e das funções ponderadas com α variando de 0,1 a 0,9.

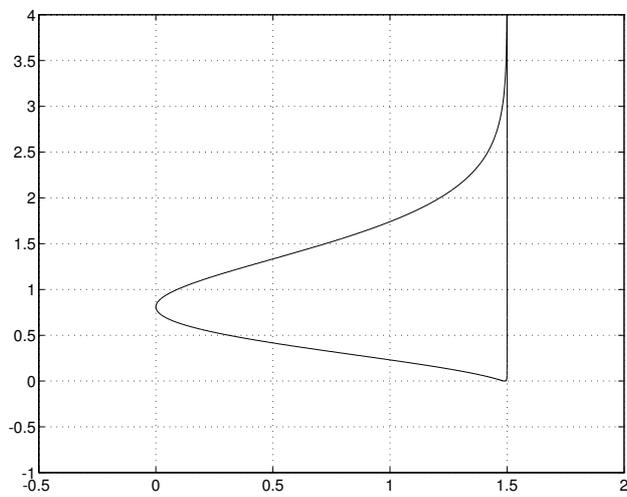


Figura 11.8: Representação da região \mathcal{F} (espaço de objetivos), para um vetor de objetivos composto de uma função convexa e uma não-convexa.

até um caso limite em que apenas as soluções mono-objetivo individuais são geradas via (P_λ) . Em geral, o problema ponderado é capaz de percorrer apenas parte do conjunto de soluções eficientes de um problema de otimização multiobjetivo, não gerando todo esse conjunto.

◇

11.1.2 Algoritmos P_λ

A estrutura do problema (P_λ) , conforme foi visto, funciona adequadamente caso todos os objetivos $f_i(\cdot)$ em (10.7) sejam funções convexas. Caso isso não se verifique, qualquer método baseado nessa estrutura somente terá acesso a parte do conjunto de soluções eficientes, deixando de mapear parcelas desse conjunto.

Define-se a função:

$$\bar{f}(x) = \sum_{i=1}^m \lambda_i f_i(x) \quad (11.4)$$

para $\lambda \geq 0 \in \mathbb{R}^m$. Tal função deverá ser otimizada empregando algum mecanismo de otimização, sendo o problema formulado como:

$$\mathcal{X}^* = \left\{ x^* \in \mathbb{R}^n \mid x^* = \arg \min_x \bar{f}(x) ; g(x^*) < 0 \right\} \quad (11.5)$$

Qualquer mecanismo de otimização que fosse capaz de resolver cada problema mono-objetivo individualmente será capaz de resolver essa formulação. Em particular, deve-se observar que:

- As restrições do problema modificado correspondem às restrições do problema original. Dessa forma, caso as restrições originalmente fossem, por exemplo, lineares, elas permaneceriam assim. Isso permitiria, nesse caso, que fossem utilizados, por exemplo, métodos projetivos de tratamento de restrições. Caso as restrições deixassem de ser lineares, tais métodos não se aplicariam (ao menos não diretamente).
- Algumas propriedades dos funcionais-objetivo também se preservariam. Por exemplo, caso fossem todos convexas, o funcional modificado também o seria. Métodos baseados na convexidade de funcionais seriam portanto transpostos diretamente.

- Uma propriedade particularmente relevante de ser preservada, tanto nos funcionais de restrição quanto nos de objetivos, é a diferenciabilidade. Tal propriedade, em particular, se preserva desde que ela esteja inicialmente presente.

Observe-se ainda que:

- Se se admitir que o problema seja convexo, é possível gerar dessa forma, variando o vetor λ , todo o conjunto de soluções eficientes do problema (10.7);
- se se adotar a desigualdade estrita $\lambda_i > 0$ para $i = 1, \dots, m$, tem-se ainda a garantia de que toda solução gerada será pertencente ao conjunto de soluções eficientes.

Dessa forma, para a geração de um conjunto de pontos representativo do conjunto de soluções eficientes, pode-se adotar por exemplo a seguinte heurística:

Passo 1: Utilizando um gerador de números aleatórios, gere um conjunto de v vetores λ diferentes, com distribuição uniforme.

Passo 2: Faça a otimização para cada um desses vetores.

Observe-se que está sendo suposto que nenhum componente desses vetores λ será menor que ou igual a zero. Essas premissas devem ser verificadas dentro do algoritmo, caso esteja sendo utilizado um gerador de números aleatórios que possa eventualmente falsificá-las.

NOTA 11.2 A utilização de um gerador de números aleatórios com distribuição uniforme é talvez a maneira mais simples de gerar uma grande quantidade de vetores, todos com grande probabilidade de serem dois a dois linearmente independentes (o que significa que não haverá cálculos a princípio redundantes), e varrendo todo o espaço de interesse de maneira aproximadamente uniforme.

◇

A seguir, são então sintetizadas as conclusões sobre a geração de soluções eficientes (conjunto \mathcal{X}^*) através do problema ponderado (P_λ):

- Computacionalmente muito simples. Equivale à solução de problemas mono-objetivos, um para cada ponto gerado. Cada problema mono-objetivo possui a mesma estrutura de restrições do problema original.
- Desvantagens:
 - Várias ponderações podem estar associadas à mesma solução eficiente;
 - Não se aplica a problemas não convexos, nos quais talvez não seja possível gerar várias soluções eficientes;
 - Nesse caso, combinações de pontos extremos gerarão aproximações de \mathcal{Y}^* .

O problema (P_λ) irá herdar todas as características de complexidade de cada um dos problemas mono-objetivo que caracterizariam a otimização de cada um dos objetivos individuais do problema. Se todos esses problemas forem, a princípio, simples, o problema (P_λ) resultante também o será. Os métodos de “direção de busca” tendem a ser aplicáveis nestas situações, a menos que haja uma particular complexidade preexistente nos problemas individuais. As outras famílias de métodos seriam provavelmente também aplicáveis, mas teriam provavelmente menor eficiência na determinação dos pontos do conjunto de Pareto.

11.2 Abordagem via Problema ϵ -Restrito

Uma técnica alternativa de geração de soluções eficientes pode ser desenvolvida a partir do seguinte resultado:

Teorema 11.2 *Se $x^* \in \mathcal{F}_x$ é eficiente então existem um inteiro $i \in \{1, 2, \dots, m\}$ e números reais ϵ_j , $j = 1, \dots, m$ ($j \neq i$) tais que x^* resolve:*

$$\begin{aligned}
 x^* &= \arg \min_{x \in \mathcal{F}_x} f_i(x) \\
 \text{sujeito a: } &\{f_j(x) \leq \epsilon_j, j = 1, \dots, m (j \neq i)\}
 \end{aligned}
 \tag{11.6}$$

□

DEMONSTRAÇÃO: Defina $f_i^* = f_i(x^*)$, $i = 1, 2, \dots, m$ e suponha que x^* não resolve o problema acima para nenhum $i \in [1, 2, \dots, m]$ e números reais ϵ_j , $j = 1, \dots, m$. Neste caso, para $i = 1$ e $\epsilon_j = f_j^*$, $j = 2, \dots, m$ deve existir $x^o \in \mathcal{X}$ tal que:

$$f_1(x^o) < f_1(x^*) \text{ e } f_j(x^o) \leq f_j^*, \quad j = 2, \dots, m$$

contradizendo a hipótese de que x^* é eficiente. ■

NOTA 11.3 Como implicação prática do teorema 11.2 conclui-se que, variando parametricamente ϵ_j , $\forall i$ e $\forall j \neq i$, é possível gerar completamente o conjunto \mathcal{X}^* .

◇

Problema P_ϵ

$$\begin{aligned} \min_{x \in \mathcal{F}_x} \quad & f_i(x) \\ \text{sujeito a:} \quad & \{f_j(x) \leq \epsilon_j, \quad j = 1, \dots, m, \quad j \neq i\} \end{aligned} \tag{11.7}$$

NOTA 11.4 Através de (P_ϵ) é possível gerar \mathcal{X}^* completamente, mesmo que o problema seja não-convexo.

◇

Nas figuras 11.3 e 11.4, pode-se verificar que o problema (P_ϵ) gera efetivamente todo o conjunto de soluções eficientes do problema.

O processo de determinação dos pontos pertencentes ao conjunto de soluções eficientes do problema é ilustrado na figura 11.9.

Em síntese, a geração do conjunto \mathcal{X}^* através de (P_ϵ) possui as seguintes características:

- Não é computacionalmente simples no caso de problemas com objetivos não-lineares (correspondente a problemas mono-objetivo com restrições não-lineares). Os problemas gerados passam a ter uma estrutura de restrições mais complexa que o problema multi-objetivo original.

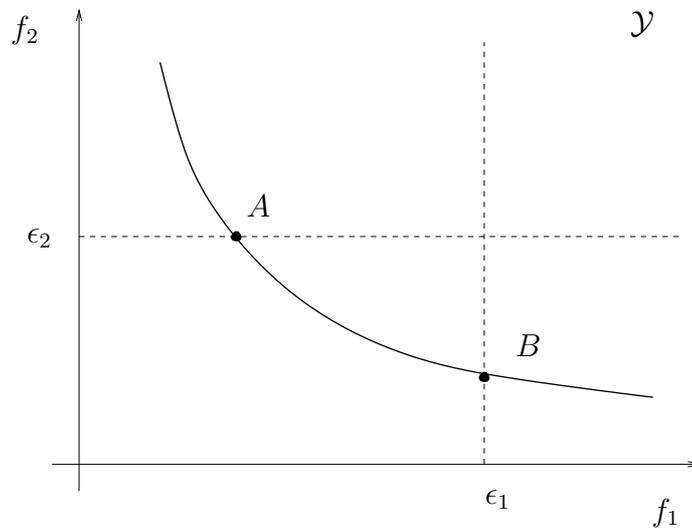


Figura 11.9: O conjunto Pareto-ótimo \mathcal{Y}^* encontra-se representado na figura pela linha em traço contínuo. Os pontos A e B pertencem ao conjunto Pareto-ótimo, sendo encontrados através do problema P_ϵ respectivamente: com a minimização de f_1 sujeita a $f_2 \leq \epsilon_2$, e com a minimização de f_2 sujeita a $f_1 \leq \epsilon_1$.

- Deficiência básica: a seleção dos reais ϵ_j que geram soluções eficientes não é tarefa trivial:
 - O ponto obtido pode não ser eficiente;
 - Os valores de ϵ_j podem tornar (P_ϵ) infactível.

A figura 11.10 ilustra a situação em que a formulação P_ϵ irá gerar alguns pontos não pertencentes ao conjunto de soluções eficientes de um problema.

A dificuldade da formulação P_ϵ para gerar soluções factíveis é ilustrada na figura 11.11.

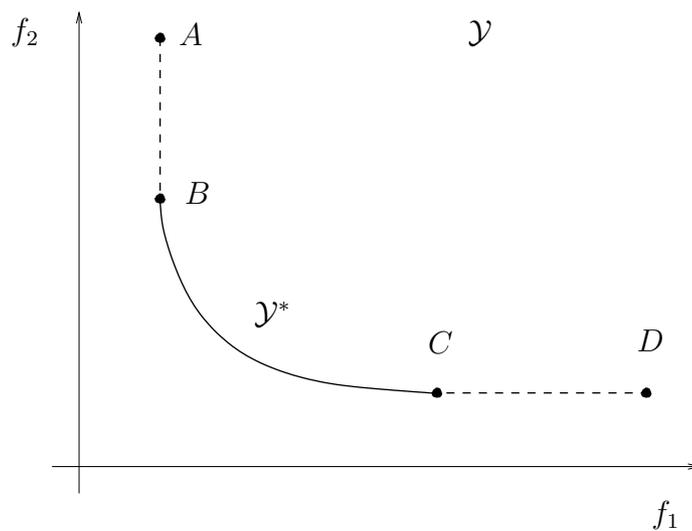


Figura 11.10: O conjunto-imagem da função $f(\cdot)$ possui trechos de sua fronteira não pertencentes ao conjunto de Pareto: os trechos AB e CD . O trecho BC corresponde ao conjunto de Pareto do problema. Como o valor da função f_1 em todo o trecho AB é igual ao seu valor no ponto B , e corresponde ao mínimo valor de f_1 , a minimização desta função com a restrição $f_2 < \epsilon_2$ pode gerar pontos pertencentes a AB e portanto não pertencentes ao conjunto de Pareto. Simetricamente, a minimização de f_2 com a restrição $f_1 < \epsilon_1$ pode gerar pontos pertencentes ao trecho CD , também não pertencentes ao conjunto de soluções eficientes.

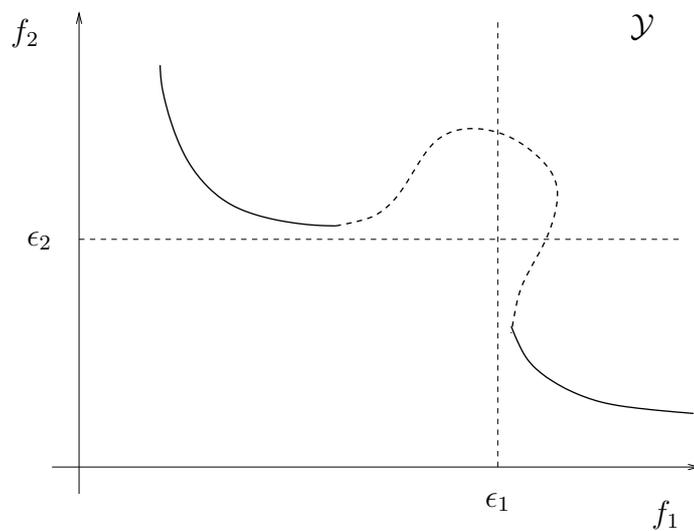


Figura 11.11: A fronteira do conjunto imagem da função $f(\cdot)$ encontra-se representada por uma linha contínua no trecho que corresponde ao conjunto Pareto-ótimo do problema, e por uma linha tracejada no trecho que não faz parte do conjunto Pareto-ótimo. Caso sejam estabelecidas as restrições $f_1 < \epsilon_1$ e $f_2 < \epsilon_2$ simultaneamente, não haverá solução factível, embora cada restrição isoladamente admita soluções.

11.2.1 Algoritmos P_ϵ

Considerando o mesmo problema (10.7) com um vetor de objetivos $f(\cdot)$, suponha-se, desta vez, que uma ou mais funções $f_i(\cdot)$ sejam não-convexas. Nesse caso, o método das ponderações não se aplica. Deve ser empregado então, por exemplo, o método do problema ϵ -restrito.

O problema é então redefinido em termos de m problemas com o seguinte formato:

$$x^* = \arg \min_x f_i(x)$$

$$\text{sujeito a: } \begin{cases} g(x) \leq 0 \\ f_j(x) \leq \epsilon_j \ ; \ j \neq i \end{cases} \quad (11.8)$$

sendo $x \in \mathbb{R}^n$, $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$ e $g(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^p$.

A respeito dessa formulação, observa-se que:

- Como boa parte dos problemas mono-objetivo assim gerados deverá ter região factível pequena, há uma grande chance de que o algoritmo desenvolva parte de sua busca em regiões infactíveis. Isto significa que o mecanismo de otimização que vier a ser empregado deve necessariamente ser adequado para realizar parte da busca dentro de regiões infactíveis. Isso exclui alguns tipos de algoritmos.
- A estrutura das restrições do problema original (10.7) fica bastante modificada nos problemas modificados (11.8). Em particular, algumas propriedades que poderiam estar presentes nas restrições originais, tais como a linearidade ou a convexidade, muito dificilmente serão preservadas. Isso significa que os métodos de otimização a serem empregados deverão fundamentalmente ser capazes de lidar com restrições de tipo geral.
- A maneira de transformar funcionais-objetivo em restrições também pode potencialmente definir aspectos importantes do problema mono-objetivo resultante. Alguns métodos de implementação dessa transformação causam, por exemplo, a perda das propriedades de diferenciabilidade e/ou de continuidade dos funcionais. Isso pode causar problemas em alguns tipos de métodos de otimização.

Feitas essas considerações, o algoritmo básico que implementa o problema (P_ϵ) é:

Método ϵ -restrito:

1. Preliminarmente, é preciso encontrar valores adequados para inicializar o vetor ϵ . Isso é feito resolvendo os m problemas mono-objetivo. Obtêm-se assim os valores f_i^* , que são os ótimos individuais de cada objetivo (os quais, agregados, compõem o vetor f^* , correspondente à solução utópica do problema).
2. Paralelamente, nessa mesma operação, são determinados os piores valores atingidos por cada objetivo quando um outro objetivo está em seu ótimo, formando o vetor f^o .
3. Cada problema (11.8) é resolvido para vetores ϵ resultantes de um gerador de números aleatórios com distribuição de probabilidade uniforme, atendendo à restrição:

$$f^* \leq \epsilon \leq f^o$$

NOTA 11.5 *Grande parte dos problemas gerados segundo esta metodologia será infactível, para problemas com número de objetivos $m \geq 3$. Isso pode tornar esse método ineficiente.*

◇

Ao contrário do problema (P_λ), o problema (P_ϵ) tende a apresentar características de complexidade que restringem a aplicabilidade dos métodos de otimização. Em geral:

- A região factível tende a se tornar pequena em relação à região infactível. Métodos do tipo “busca por populações” tendem a não encontrar pontos factíveis.
- Pela mesma razão, a utilização de métodos de “direção de busca” com tratamento de restrição do tipo “barreira” é inviável. Caso sejam empregados métodos de “direção de busca”, estes devem ser acoplados a tratamento de restrição por “penalidades”.
- A quantidade de restrições tende a aumentar. As restrições ficam, evidentemente, não-lineares, o que pode ainda significar um aumento de complexidade da estrutura de restrições. Isso pode significar uma

dificuldade para os métodos do tipo “direções de busca”. Outra dificuldade que pode surgir é a necessidade de utilizar “penalidades” que tornem o problema não-diferenciável (isso pode ocorrer pela dificuldade de se formular penalidades diferenciáveis e suficientemente precisas). Essas razões indicariam a utilização de métodos de “exclusão de semi-espacos” como formulação em geral mais adequada para os problemas com estrutura (P_ϵ) .

11.3 Abordagem híbrida: Ponderando e Restringindo

Soluções não-inferiores também podem ser caracterizadas em termos de soluções ótimas do seguinte problema $(P_{\lambda,\epsilon})$:

Problema $P_{\lambda,\epsilon}$

$$\min_{x \in \mathcal{F}_X} \sum_{i=1}^m \lambda_i f_i(x) \tag{11.9}$$

sujeito a: $\{f_j(x) \leq \epsilon_j ; j = 1, \dots, m\}$

O seguinte teorema se aplica:

Teorema 11.3 *O ponto x^* é uma solução eficiente se e somente se x^* for uma solução ótima de $P_{\lambda,\epsilon}$ para qualquer vetor de ponderações $\lambda \geq 0 \in \mathbb{R}^m$ e para algum vetor de restrições $\epsilon \in \mathbb{R}^m$ para o qual o problema $P_{\lambda,\epsilon}$ é factível.*

□

NOTA 11.6 *O resultado do teorema 11.3 tem a vantagem de permitir a geração de todos os pontos eficientes sem a necessidade de testar a eficiência dos pontos gerados a posteriori. Combinam-se as vantagens do método (P_λ) , que produz pontos que certamente são eficientes (quando λ é estritamente positivo), com a do*

método (P_ϵ) , que é capaz de gerar todos os pontos eficientes.

◇

11.4 Abordagem da Programação-Alvo

O método de programação-alvo (ou “goal programming”) foi originalmente formulado como:

$$\min_{x \in \mathcal{F}_x, d^+, d^-} \left(\sum_{i=1}^m (d_i^+ + d_i^-)^p \right)^{\frac{1}{p}} ; \quad 1 \leq p < \infty$$

$$\text{sujeito a: } \begin{cases} f_j(x) - d_j^+ + d_j^- = M_j, & j = 1, 2, \dots, m \\ d_j^+, d_j^- \geq 0 \\ d_j^+ \cdot d_j^- = 0 \quad \forall j \end{cases} \quad (11.10)$$

onde:

M_j : meta (ou alvo) estipulada para o j -ésimo objetivo.

d_j^+ : indica o quanto o j -ésimo objetivo excedeu a meta estipulada M_j .

d_j^- : indica o quanto o j -ésimo objetivo ficou abaixo da meta estipulada M_j .

O método como definido acima pode ser interpretado como a minimização de uma norma vetorial p em relação ao vetor de referência (ou alvo), veja-se a definição 2.5. Essa re-interpretação, que está mais de acordo com a forma de estruturar os problemas empregada na atualidade, é aqui designada como *Problema P_μ* :

Problema P_μ

$$\min_{x \in \mathcal{F}_x} \|f(x) - M\|_p$$

$$1 \leq p \leq \infty \quad (11.11)$$

sendo $x \in \mathbb{R}^n$, $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$ e $M \in \mathbb{R}^m$. Note-se que, pela interpretação de norma, é possível adotar a norma vetorial ∞ , sem prejuízo das propriedades desejadas da formulação¹. Definem-se as *soluções*, tanto no espaço de parâmetros quanto no espaço de objetivos, associadas a esse problema:

$$\mu_x^p(M) = \arg_{\mathcal{F}_x} \min_{x \in \mathcal{F}_x} \|f(x) - M\|_p \quad (11.12)$$

$$\mu_y^p(M) = \arg_{\mathcal{F}_y} \min_{y \in \mathcal{F}_y} \|y - M\|_p \quad (11.13)$$

Evidentemente, se o ponto M (a meta) for escolhido no interior do conjunto-imagem da região factível, \mathcal{F}_y , a solução do problema $\mu_x^p(M)$ irá igualar o ponto (ou os pontos, no caso de soluções não-únicas) cuja imagem é M , e este não será pertencente ao conjunto Pareto-ótimo (note-se que os pontos pertencentes ao Pareto-ótimo, no espaço de objetivos, não podem estar no *interior* de \mathcal{F}_y , pois encontram-se todos em sua *fronteira*). Dessa forma, para que P_μ forneça pontos pertencentes ao conjunto Pareto-ótimo \mathcal{X}^* , é necessário que M seja escolhido fora de \mathcal{F}_y ou em sua fronteira. As diferentes situações possíveis encontram-se representadas na figura 11.12.

Passa-se agora à análise do comportamento do problema P_μ sob o ponto de vista das soluções no espaço de objetivos \mathcal{Y} . As seguintes situações são previstas:

- A escolha de diferentes normas faz com que, para um mesmo ponto M situado fora de \mathcal{F}_y , sejam determinadas diferentes soluções, ou seja, $\mu_y^p(M) \neq \mu_y^q(M)$ para $p \neq q$. Isso é mostrado na figura 11.13.
- É possível que, dada uma certa meta M , dada uma certa norma vetorial p , haja mais de uma solução $\mu_y^p(M)$ para o problema P_μ (ou seja, não há unicidade garantida da solução no espaço de objetivos). Essa situação é mostrada na figura 11.14.
- É possível que para diferentes metas, $M_1 \neq M_2$, seja obtido como resultado do problema P_μ o mesmo ponto $\mu_y^p(M_1) = \mu_y^p(M_2)$. Essa possibilidade é ilustrada na figura 11.15.

¹Seria também possível a adoção de Q -normas, sem nenhum problema.

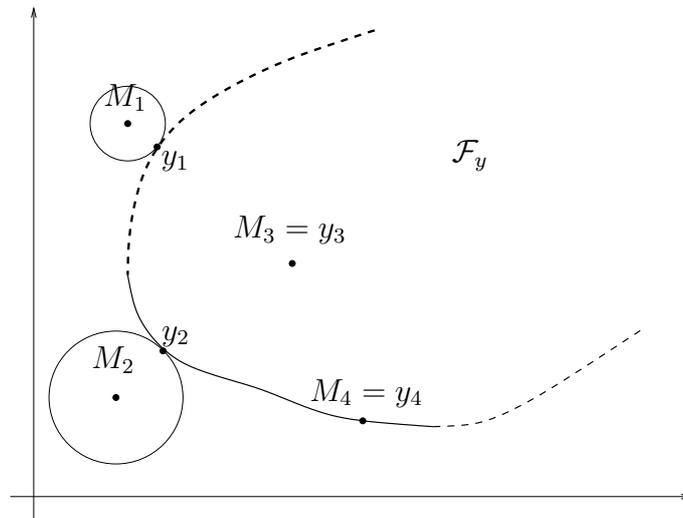


Figura 11.12: Interpretação do problema P_χ no espaço de objetivos. A imagem do conjunto factível, \mathcal{F}_y , encontra-se delimitada pelas linhas representadas na figura, sendo que as linhas tracejadas não constituem parte do conjunto Pareto-ótimo, enquanto a linha em traço contínuo representa o conjunto Pareto-ótimo, \mathcal{Y}^* . Escolhendo-se a meta M_1 , o resultado do problema P_μ é o ponto y_1 (o ponto mais próximo de M_1 dentro do conjunto \mathcal{F}_y). Nesse caso, o resultado não pertence a \mathcal{Y}^* . Da mesma forma, escolhendo-se a meta M_2 , o resultado do problema P_μ é o ponto y_2 , agora pertencente ao conjunto Pareto-ótimo \mathcal{Y}^* . Escolhendo-se metas já pertencentes a \mathcal{F}_y , o resultado do problema P_μ é o próprio ponto-meta. Esse é o caso das metas M_3 e M_4 , que resultam em $y_3 = M_3$ e em $y_4 = M_4$. O primeiro não pertence a \mathcal{Y}^* , enquanto o segundo pertence. Nesta figura, foi assumido que $p = 2$, ou seja, que a norma é a norma euclidiana de vetores.

- É finalmente possível que soluções $\mu_y^p(M)$ não pertençam ao conjunto Pareto-ótimo \mathcal{Y}^* para determinadas escolhas de $M \notin \mathcal{F}_y$. Isso é ilustrado nas figuras 11.12 e 11.14.

NOTA 11.7 *Todos esses pontos são inconvenientes da abordagem P_μ . Deve-se notar que a abordagem P_ϵ também apresenta as três últimas dificuldades (a primeira não se aplica). Deve-se notar que apenas os dois últimos pontos constituem de fato dificuldades que afetam negativamente a implementação computacional de métodos de otimização multiobjetivo baseados em P_ϵ e em P_μ .*

◇

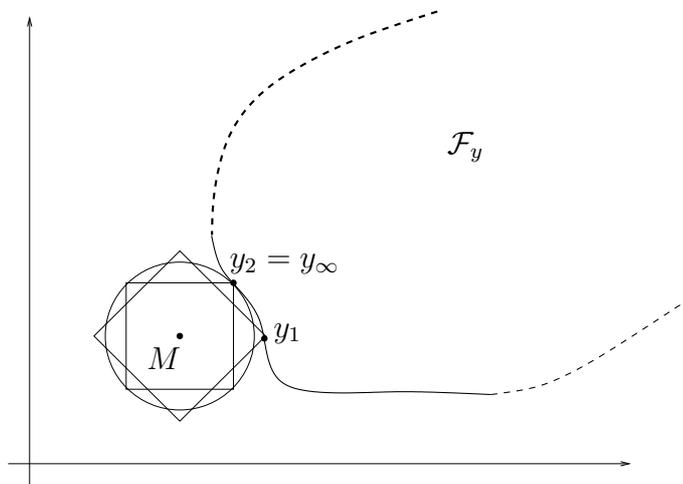


Figura 11.13: Esta figura ilustra a dependência do resultado do problema P_μ com a norma p adotada. A imagem da região factível, \mathcal{F}_y , encontra-se delimitada na figura. Com o uso de diferentes normas vetoriais, o problema P_μ resulta em diferentes soluções. A figura mostra a meta M fora de \mathcal{F}_y , que resulta na solução y_1 para $p = 1$, e na solução $y_2 = y_\infty$, para $p = 2$ e $p = \infty$.

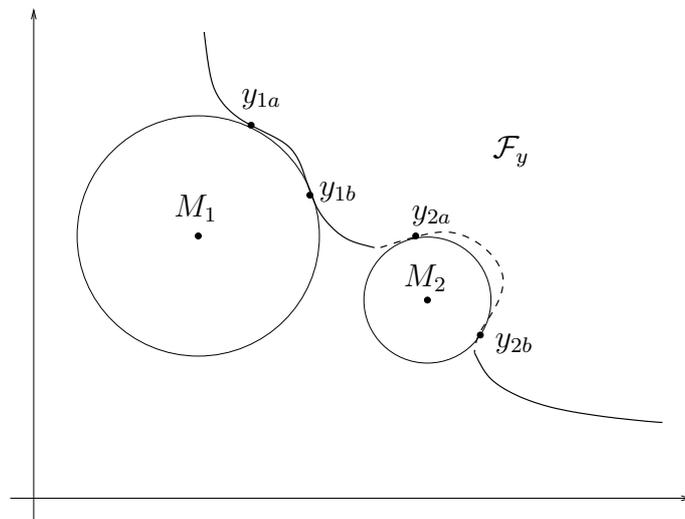


Figura 11.14: Esta figura mostra que o problema P_μ pode conduzir a múltiplos resultados, para uma mesma escolha da meta M , dada uma norma p fixa (no caso da figura, $p = 2$). Dada a meta M_1 , o problema P_μ fornece as soluções y_{1a} e y_{1b} , ambas pertencentes ao conjunto Pareto-ótimo do problema. Dada por sua vez a meta M_2 , o problema P_μ fornece as soluções y_{2a} e y_{2b} , ambas não pertencentes ao conjunto Pareto-ótimo do problema.

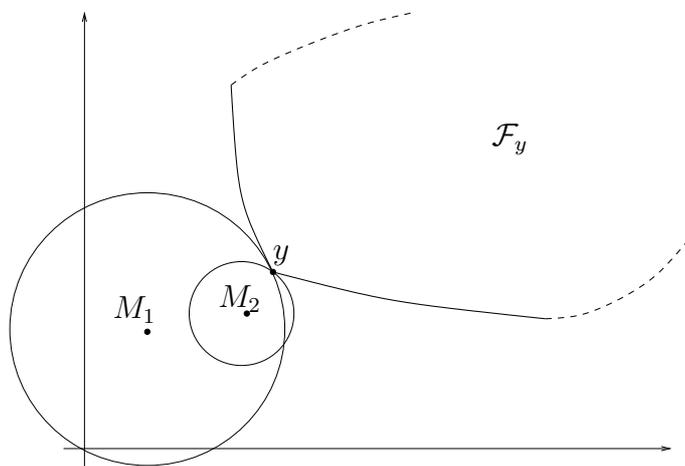


Figura 11.15: Esta figura mostra que o problema P_μ pode conduzir ao mesmo resultado y para metas diferentes M_1 e M_2 , para uma mesma norma p .

Os teoremas a seguir mostram os pontos fortes da abordagem P_μ .

Teorema 11.4 *Seja um POV com conjunto factível (no espaço de objetivos) \mathcal{F}_y não vazio. Então, para toda escolha de $M \in \mathcal{Y}$ o problema P_μ será factível, ou seja:*

$$\mu_y^p(M) \neq \emptyset \quad \forall M \in \mathcal{Y}, p \geq 1 \quad (11.14)$$

□

Teorema 11.5 *Seja um POV cujo conjunto Pareto-ótimo no espaço de objetivos \mathcal{Y}^* é não-vazio. Então:*

$$\exists M \in \mathcal{Y} \mid \mu_y^p(M) \supset y^* \quad \forall y^* \in \mathcal{Y}^*, p \geq 1 \quad (11.15)$$

ou seja, existe alguma meta M que, aplicada no problema P_μ (com uma escolha qualquer de norma p) resulta na solução y^ .* □

NOTA 11.8 *Em comparação com o problema P_ϵ , o problema P_μ apresenta como vantagem a propriedade garantida no teorema 11.4 da factibilidade de toda execução realizada sobre um POV factível. O teorema 11.5 mostra que a abordagem P_μ compartilha com a abordagem P_ϵ a capacidade de gerar todo o conjunto Pareto-ótimo.*

◇

11.5 Abordagem P_χ

Em (Gembiki & Haimes 1975) foi proposto o algoritmo de *goal attainment*. Para a finalidade de se fazer uma referência rápida, tal formulação será aqui denominada método (P_χ). Em (Takahashi, Peres & Ferreira 1997), tal algoritmo foi empregado, com uma formulação levemente expandida, para o tratamento de um problema de projeto de controladores. Essa é a formulação aqui apresentada.

Esse método contorna o problema básico de ineficiência apontado nas seções anteriores. Essa formulação, partindo do mesmo problema (10.7) com um vetor de objetivos $f(\cdot)$, também admite que uma ou mais funções $f_i(\cdot)$ sejam não-convexas.

Essa formulação é assim construída:

- Seja $f^* \in \mathbb{R}^m$ o vetor de objetivos correspondente à “solução utópica” do problema.

- Seja $f_i^* \in \mathbb{R}^m$; $i = 1, \dots, m$ o vetor de objetivos associado a cada mínimo individual do problema.
- Seja \mathcal{C} o cone gerado pelos vetores $(f_i^* - f^*)$, com origem em f^* .
- Seja $w \in \mathcal{C}$ um vetor construído segundo a fórmula:

$$w = f^* + \sum_{i=1}^m \gamma_i (f_i^* - f^*) \quad (11.16)$$

para $\gamma_i > 0$.

O problema é então redefinido em termos de um único problema com o seguinte formato:

Problema P_χ

$$\begin{aligned} & \min_{x, \eta} \eta \\ & \text{sujeito a: } \begin{cases} x \in \mathcal{F}_x \\ f(x) \leq f^* + \eta w \end{cases} \end{aligned} \quad (11.17)$$

Nessa expressão, $x \in \mathbb{R}^n$, $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$ e $g(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^p$. As correspondentes soluções no espaço de parâmetros \mathcal{X} e no espaço de objetivos \mathcal{Y} são também definidas:

Soluções no espaço de parâmetros:

$$\begin{aligned} \chi_x(w) &= \arg_x \min_{x, \eta} \eta \\ & \text{sujeito a: } \begin{cases} x \in \mathcal{F}_x \\ f(x) \leq f^* + \eta w \end{cases} \end{aligned} \quad (11.18)$$

Soluções no espaço de objetivos:

$$\chi_y(w) = \arg_y \min_{y,\eta} \eta$$

$$\text{sujeito a: } \begin{cases} y \in \mathcal{F}_y \\ y \leq f^* + \eta w \end{cases} \quad (11.19)$$

A interpretação gráfica da abordagem P_χ é fornecida na figura 11.16.

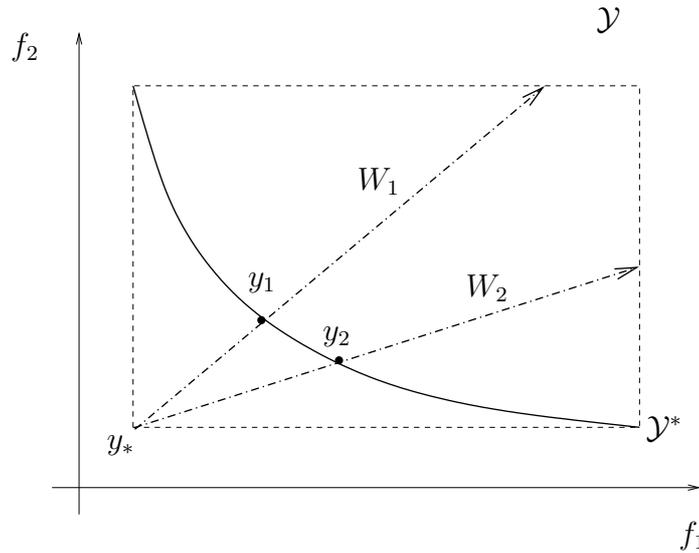


Figura 11.16: Interpretação do problema P_χ no espaço de objetivos: O conjunto Pareto-ótimo \mathcal{Y}^* encontra-se contido num hiper-paralelepípedo no espaço \mathcal{Y} , com um vértice sobre a solução utópica y_* . Tomam-se vetores, por exemplo W_1 e W_2 , com ponto inicial y_* , e ponto final na fronteira do paralelepípedo. Esses vetores são então multiplicados por um escalar η , que é minimizado com a restrição de que pontos dominados pelo ponto final do vetor continuem pertencendo à imagem factível \mathcal{F}_y . Dessa forma, são determinados os pontos y_1 e y_2 pertencentes ao conjunto \mathcal{Y}^* .

Os teoremas a seguir se aplicam ao problema P_χ :

Teorema 11.6 *Seja um POV com conjunto factível (no espaço de objetivos) \mathcal{F}_y não vazio. Então, dados os vetores w_1 e w_2 linearmente independentes, ambos pertencentes ao cone \mathcal{C} , tem-se que:*

$$\chi_y(w_1) \neq \chi_y(w_2) \quad (11.20)$$

□

Teorema 11.7 *Seja um POV cujo conjunto Pareto-ótimo no espaço de objetivos \mathcal{Y}^* é não-vazio, e seja o cone \mathcal{C} conforme definido nesta seção. Então:*

$$\exists w \in \mathcal{C} \mid \chi_y(w) \supset y^* \quad \forall y^* \in \mathcal{Y}^*, \quad p \geq 1 \quad (11.21)$$

ou seja, existe algum vetor w que, aplicado no problema P_χ resulta na solução y^ .* □

As seguintes considerações são feitas em relação à estrutura do novo problema que foi construído:

- Este método trabalha num espaço de busca com $n + 1$ variáveis e $m + n$ restrições. O problema P_ϵ corresponde a uma combinação de m problemas em n variáveis e $m + n - 1$ restrições). Já o problema P_μ corresponde a uma busca num espaço de n variáveis de otimização e m restrições.
- O teorema 11.6 afirma que o problema P_χ possibilita a obtenção de respostas espacialmente organizadas a partir de uma seqüência de escolhas de vetores w . Tal não é possível com o método P_μ , uma vez que este não permite a escolha prévia da orientação espacial (no espaço \mathcal{Y}) da solução a ser obtida.
- Embora não haja a garantia da factibilidade do problema P_χ , a ocorrência de infactibilidade no mesmo é muito menos freqüente que no problema P_ϵ . Por construção, é possível garantir que o problema de otimização seja factível e já inicie a busca em um ponto factível, desde que o vetor-suporte w esteja apontando para uma direção na qual exista algum ponto do conjunto de soluções factíveis.
- O funcional-objetivo do problema modificado é linear, sendo portanto trivial o cálculo de seu gradiente e hessiana.
- A estrutura de restrições do problema, no entanto, passa a conter, além das restrições do problema original, toda a complexidade do conjunto de funções-objetivo. Deve-se notar que propriedades que possivelmente estariam presentes nas restrições originais, tais como linearidade ou convexidade, provavelmente não irão se preservar no novo conjunto de restrições.

O problema P_χ é capaz de gerar, assim como o problema P_ϵ e o problema P_μ , todo o espaço de soluções Pareto-ótimas. Assim como eles, entretanto, o método P_χ também pode eventualmente gerar pontos não pertencentes ao conjunto de soluções eficientes (a situação, de fato, é análoga àquela que ocorre no caso do problema P_ϵ). Essa situação é ilustrada na figura 11.17.

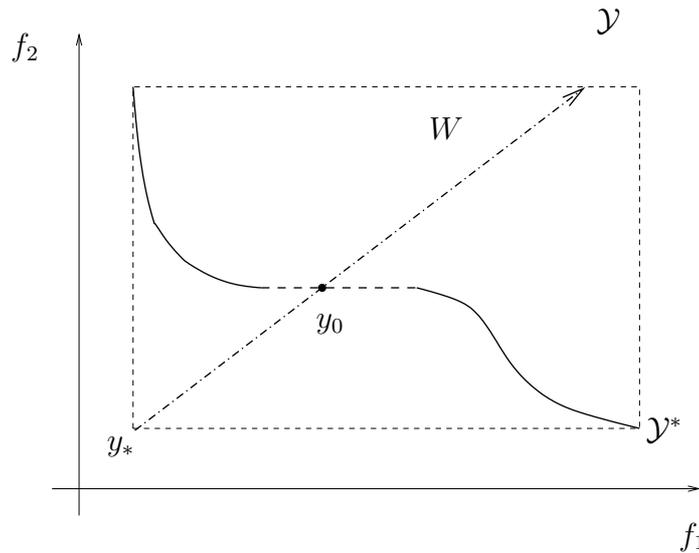


Figura 11.17: O problema P_χ pode gerar pontos não eficientes: O conjunto Pareto-ótimo \mathcal{Y}^* (representado em traço contínuo) encontra-se contido num hiper-paralelepípedo no espaço \mathcal{Y} , com um vértice sobre a solução utópica y_* . Há ainda um trecho da fronteira do conjunto imagem factível \mathcal{F}_y , representado por uma linha tracejada horizontal, que não está contido no conjunto Pareto-ótimo (observe-se que apenas o extremo esquerdo dessa linha faz parte do conjunto Pareto-ótimo). Toma-se o vetor W com ponto inicial y_* , e ponto final na fronteira do paralelepípedo. Esse vetor é multiplicado por um escalar η , que é minimizado com a restrição de que pontos dominados pelo ponto final do vetor ηW continuem pertencendo à imagem factível \mathcal{F}_y . Dessa forma, é determinado o ponto y_0 que, nesta figura, embora pertença à fronteira do conjunto imagem factível \mathcal{F}_y , não pertence ao conjunto de soluções eficientes \mathcal{Y}^* .

É importante mencionar um tipo de situação que pode ocorrer em alguns tipos de problemas: a solução do problema P_χ pode não se encontrar na direção do vetor-suporte, embora ainda assim ocorra em um ponto pertencente à fronteira Pareto-ótima. Essa situação é ilustrada na figura 11.18.

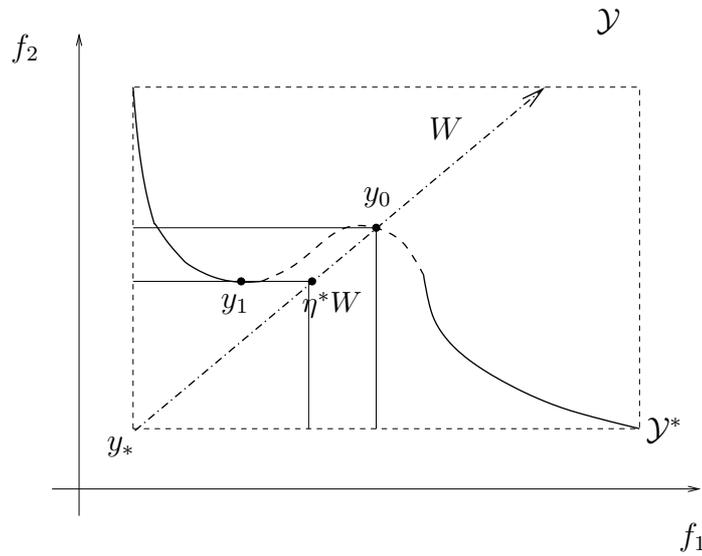


Figura 11.18: O problema P_x pode gerar pontos fora da direção do vetor-suporte W : O conjunto Pareto-ótimo \mathcal{Y}^* (representado em traço contínuo) encontra-se contido num hiper-paralelepípedo no espaço \mathcal{Y} , com um vértice sobre a solução utópica y_* . Há ainda um trecho da fronteira do conjunto imagem factível \mathcal{F}_y , representado por uma linha tracejada, que não está contido no conjunto Pareto-ótimo. Toma-se o vetor W com ponto inicial y_* , e ponto final na fronteira do paralelepípedo. Esse vetor é multiplicado por um escalar η , que é minimizado com a restrição de que um ponto dominado pelo ponto final de ηW continue pertencendo à imagem factível \mathcal{F}_y . Dessa forma, é determinado o ponto y_1 que pertence ao conjunto de soluções eficientes \mathcal{Y}^* , e é dominado por η^*W . Observe-se que o "último" ponto pertencente ao conjunto-imagem factível \mathcal{F}_y , denotado por y_0 , ainda possui pontos por ele dominados pertencentes a \mathcal{F}_y , e portanto não corresponde ao ponto de término do processo de otimização (ou seja, do processo de minimização de η). Observe-se ainda que η^*W não pertence ao conjunto imagem \mathcal{F}_y .

Tendo em vista essas considerações, a heurística a seguir pode ser empregada para gerar um conjunto representativo de pontos do espaço de soluções eficientes:

Passo 1: Determinar f^* e f_i^* ; $i = 1, \dots, m$.

Passo 2: Escolher aleatoriamente um conjunto W de vetores w , utilizando a equação (11.16) e vetores γ gerados através de um gerador de números aleatórios com distribuição uniforme.

Passo 3: Para cada vetor $w \in W$, escolher um número η_0 suficientemente grande (o quão grande deve ser esse número depende do problema em questão) para tornar (11.18) factível. Note-se que, uma vez que γ_i na equação (11.16) é um número estritamente positivo, sempre haverá algum valor de η que torna o problema factível, desde que o conjunto factível \mathcal{X} seja não-vazio.

Passo 4: Resolver cada problema assim construído. Se se utilizar o problema irrestrito do tipo barreira, será necessário iniciar com pontos factíveis conhecidos a priori. Se se empregar métodos de penalidade, é possível inicializar o problema com condições iniciais arbitrárias.

O problema $(P_{\mathcal{X}})$, embora envolva um acréscimo da complexidade da estrutura de restrições do problema, como no caso do problema (P_{ϵ}) , tem a peculiaridade de em geral conduzir a problemas não apenas factíveis, mas ainda em que é possível freqüentemente partir de pontos factíveis. Esse tipo de situação viabiliza a utilização de todas as categorias de métodos de otimização. Cada família de métodos terá suas vantagens específicas, que podem ser relevantes para problemas particulares:

- Os métodos de “direção de busca” podem ser bastante eficientes, caso não se detenham em descontinuidades.
- Os métodos de “exclusão de semi-espaços” simplificam sobremaneira o tratamento das restrições não-lineares que caracterizam o problema.
- Os métodos de “busca por populações” possivelmente permitiriam determinar os pontos pertencentes ao que seria o “conjunto Pareto-ótimo global”, que talvez não estivessem acessíveis aos outros dois grupos de métodos.

11.6 Teste de Eficiência

O último problema a ser tratado aqui é o de testar se um dado ponto x^* é ou não uma solução eficiente de um problema de otimização multi-objetivo. Para tal finalidade, escolhe-se um vetor $\alpha > 0 \in \mathbb{R}^m$. Formula-se a seguir o problema (P^*) :

$$\delta^* = \max_{x, \epsilon} \alpha^T \epsilon$$

$$(P^*) \quad \text{sujeito a: } \begin{cases} f(x) + \epsilon \leq f(x^*) \\ x \in \mathcal{F}_x \\ \epsilon \geq 0 \end{cases}$$

Uma das três alternativas se verifica:

1. (P^*) possui solução ótima limitada com $\delta^* = 0$;
2. (P^*) possui solução ótima limitada com $\delta^* > -\infty$ um valor finito ótimo; ou
3. (P^*) possui uma solução ótima ilimitada, sendo que o valor ótimo ilimitado de δ^* é tal que $\delta^* = -\infty$ (o máximo, no caso, se reduz a um supremo).

O caso 1 implica na eficiência de x^* . Por outro lado, os casos 2 e 3 implicam na não-eficiência do mesmo. O caso 2 implica na existência de algum ponto eficiente (diferente de x^*), enquanto o caso 3 pode indicar a não existência de soluções factíveis para o problema. Esses fatos são sumarizados no teorema a seguir.

Teorema 11.8 *Para qualquer $\alpha > 0 \in \mathbb{R}^m$ e $x \in \mathcal{F}_x$, seja δ^* o valor ótimo (ou supremo) definido pelo problema (P^*) . Nesse caso:*

- i. x^* é eficiente se e somente se $\delta^* = 0$;*
- ii. uma solução ótima x de (P^*) é eficiente se $-\infty < \delta^* < 0$;*

iii. se (P^*) possui uma solução ilimitada com $\delta^* = -\infty$ e se \mathcal{F}_x é um conjunto convexo, sendo também convexas as funções f_j , $j = 1, \dots, m$, então o problema de otimização multiobjetivo não possui soluções eficientes próprias.

□

11.6.1 Algoritmos P^*

Finalmente, o problema de construção de algoritmos (P^*) deve ser examinado, uma vez que o mesmo serve como teste para verificar a pertinência de um ponto ao conjunto de soluções eficientes, e uma vez que, exceto o problema (P_λ) , todos os demais podem conduzir (se não forem tomados os devidos cuidados) a pontos não eficientes. Além disso, deve-se notar que o problema (P^*) , caso verifique que determinado ponto x não é eficiente, ainda é capaz de gerar, sem esforço computacional adicional, um outro ponto que é eficiente. Basta observar a estrutura do problema:

$$\delta^* = \max_{x, \epsilon} \alpha^T \epsilon$$

$$(P^*) \quad \text{sujeito a: } \begin{cases} f(x) + \epsilon \leq f(x^*) \\ x \in \mathcal{X} \\ \epsilon \geq 0 \end{cases}$$

Toma-se então:

$$\bar{x} = \arg_x \delta^*(x, \epsilon) \quad (11.22)$$

Esse ponto \bar{x} possivelmente será eficiente, o que deve entretanto ser testado pelo próprio problema (P^*) . Caso não seja, é possível procurar outro ponto que possivelmente seja, e assim por diante.

Deve-se notar a similaridade estrutural deste problema com o problema (P_λ) . Desta forma, as observações aplicáveis ao problema (P_λ) , relativas à sua complexidade e aos métodos de otimização recomendados, poderão ser transpostas para o problema (P^*) .

Capítulo 12

Propriedades de Grupo do Conjunto Pareto-Ótimo

Problemas de projeto multiobjetivo dão origem a um objeto bem definido: o conjunto Pareto-ótimo. Existem condições analíticas que podem ser desdobradas em testes numéricos, para verificar se um determinado ponto pertence ao conjunto Pareto-ótimo de determinado problema. Um desses possíveis testes foi mostrado no capítulo 10, sendo denominado *Problema (P^*)*. Entretanto, tais condições se assentam na premissa de que é possível resolver globalmente determinados problemas de otimização de um tipo genérico (possivelmente não-lineares, não-convexos e multimodais), que se originam dessas formulações analíticas.

Na prática, problemas com tais características são resolvidos apenas de maneira heurística, isto é, não é possível na maioria dos casos obter provas formais de que determinada solução seja um ótimo global. Deste fato decorrem considerações filosóficas que conduzem o problema de construção do conjunto Pareto-ótimo a uma abordagem “refutacionista” (Popper 1974). Neste capítulo são desenvolvidas condições computacionalmente verificáveis (com tempo de execução e memória finitos), que são aplicáveis a conjuntos com um número finito de elementos, para *corroborar* ou *falsear* a hipótese de que os elementos desse conjunto constituam amostras do conjunto Pareto-ótimo de um PMO.

Essas condições levam a alguns critérios genéricos que podem ser empregados na avaliação de algoritmos enquanto mecanismos de otimização multiobjetivo. Um algoritmo multiobjetivo conceitual, baseado no princípio dos algoritmos genéticos, é proposto explorando as *propriedades de grupo* de

estimativas intermediárias do conjunto Pareto-ótimo, de forma a gerar uma estimativa final consistente de tal conjunto.

As idéias apresentadas neste capítulo foram inicialmente formuladas em (Takahashi, Palhares, Dutra & Gonçalves 2001).

12.1 Verificação versus Falseamento

Assim como as soluções ótimas de problemas de otimização mono-objetivo, os conjuntos Pareto-ótimos de problemas de otimização multiobjetivo constituem objetos conceituais que podem ser acessados a partir de algoritmos numéricos adequados. Embora a existência desses objetos, em um sentido analítico, possa ser provada, a determinação analítica desses objetos em geral não é factível.

Dessa forma, coloca-se a questão de “saber” se determinado objeto que foi encontrado computacionalmente corresponde ao objeto conceitual que estava sendo procurado. Essa questão pode ser estabelecida sob o ponto de vista de um mecanismo de “verificação” da hipótese de correspondência, ou sob o ponto de vista de um mecanismo de “corroboração e falseamento” dessa hipótese (Popper 1974).

Em princípio, demonstra-se que a solução do problema P_ϵ repetidas vezes, conduziria deterministicamente aos pontos constituintes do conjunto Pareto-ótimo de um problema multiobjetivo (PMO), sendo possivelmente encontrados também alguns pontos não pertencentes ao conjunto de Pareto (vide capítulo 10). Posteriormente, os pontos encontrados poderiam ser submetidos a testes de *verificação* da pertinência ao conjunto de Pareto, por exemplo o teste P^* (vide capítulo 10). Em termos de objetos e mecanismos conceitualmente definidos, não haveria problema com essa seqüência de procedimentos.

Entretanto, há alguns problemas associados com a natureza geral (possivelmente não-linear, não convexa e multimodal) dos funcionais que definem o PMO. Sabe-se que não há algoritmos de otimização mono-objetivo que sejam capazes de encontrar, de maneira determinística, o ótimo global desses problemas. Isso significa que: (i) todo ponto de ótimo que for encontrado pode se tratar de um mínimo apenas local, e (ii) mesmo que o ótimo global seja encontrado, não há maneira de saber se o ponto que foi determinado é de fato o ótimo global, ou seja, não é possível excluir a hipótese de que haja outro ponto de ótimo que venha a tornar o ponto anteriormente determinado um ótimo apenas local. As condições de otimalidade conhecidas (vide capítulo 3)

possuem caráter apenas local para funções de tipo geral. Dessa forma, como tanto o problema P_ϵ quanto o problema P^* são fundamentados na hipótese de que haja mecanismos capazes de solucionar globalmente problemas de otimização mono-objetivo, em ambos os casos é possível que não seja encontrada a solução do problema, para funções de tipo geral. Além disso, mesmo que essas soluções sejam encontradas, não é possível haver certeza quanto a esse fato. Assim:

- É possível que regiões do espaço de parâmetros que pertençam ao conjunto Pareto-ótimo venham a não ser amostradas pelos algoritmos de otimização utilizados para resolver o problema P_ϵ .
- É possível também que pontos que não pertençam ao espaço de parâmetros venham a ser fornecidos como “soluções” do problema P_ϵ e, simultaneamente, sejam confirmados como soluções do PMO pelo problema P^* .

Dessa forma, a possibilidade de *verificação* das soluções de um PMO parece estar descartada, no caso geral. Resta a possibilidade de utilização de um mecanismo de *corroboração e falseamento* para fundamentar a crença de que determinado conjunto de soluções trate-se de amostra representativa do conjunto Pareto-ótimo. A lógica da *corroboração e falseamento* consiste num esquema em que se procura a construção de dados que *falseiem* a hipótese de que determinados pontos pertençam ao conjunto Pareto-ótimo, ou seja, que mostrem que tais pontos *não podem pertencer* ao Pareto-ótimo. Caso se consigam dados que falseiem dados anteriormente obtidos, os novos dados passam a ser a melhor estimativa disponível até o momento de pontos que *podem pertencer* ao Pareto. Se tais dados não forem obtidos após tentativas sistemáticas, fica *corroborada* a hipótese de que os dados anteriormente disponíveis de fato pertençam ao Pareto. Note-se que essa *corroboração* não se trata de uma *verificação*, tendo o sentido de uma “hipótese provisoriamente acatada”.

Trivialmente, todo mecanismo de otimização mono-objetivo se apóia em um sistema de *corroboração e falseamento*, na medida em que o mecanismo computacional permanece operando na busca do ótimo enquanto for possível determinar novas soluções melhores que as anteriores, ou seja, soluções que *falseiem* a possibilidade de que a solução anterior seja a solução ótima. A partir do momento em que o método passar a não mais gerar soluções melhores apesar de tentativas sistematicamente realizadas, faz-se a *parada* do

método, que ocorre portanto quando há a *corroboração* da hipótese de que a solução disponível é a solução ótima.

No caso da otimização multi-objetivo, seria possível fazer os mecanismos de otimização simplesmente herdarem o sistema de *corroboração e falseamento* que já está presente nos mecanismos mono-objetivo. Esse seria o caso, por exemplo, da possível implementação computacional direta do problema P_ϵ seguida da verificação pelo problema P^* . Entretanto, é possível apoiar mecanismos de otimização multi-objetivo em sistemas mais sutis de *corroboração e falseamento*, construídos a partir de propriedades de grupo que estariam associadas a um hipotético conjunto Pareto-ótimo, e que podem ser verificadas sobre conjuntos de pontos candidatos a serem considerados “Pareto-ótimos”. Dessa forma, a hipótese de que tais pontos pertençam ao conjunto Pareto-ótimo pode ser *falseada* sendo que, caso isso não ocorra, tal hipótese fica *corroborada*.

Diante da argumentação desenvolvida, que mostra que os sistemas de “verificação” apresentam dificuldades de construção em situações gerais, restam os sistemas de “corroboração e falseamento”. Apesar de sua aparente fragilidade, eles se mostram adequados para o tratamento dessa questão de “saber” se há ou não correspondência entre objetos empiricamente determinados através de experimentos numéricos e objetos conceitualmente definidos, nas situações gerais em que o problema de otimização não possui particularidades que favoreçam o emprego de metodologias analíticas de “verificação”. Um argumento adicional em favor dos sistemas de “corroboração e falseamento” é o fato de que estes tornam explícito um grau de “crença” que se faz necessária no momento em que determinada solução é acatada. Os sistemas de “verificação”, ao contrário, se apóiam em conceitos não computacionalmente concretizáveis, o que pode obscurecer o fato de que existe incerteza associada aos seus resultados.

12.2 Estrutura do Conjunto Pareto-Ótimo

Considere-se um problema de projeto genérico:

Problema 1 [Problema de Projeto] *Considere-se o vetor de parâmetros de projeto $x \in \mathbb{R}^n$, o conjunto de soluções factíveis $\mathcal{X}_f \subset \mathbb{R}^n$, e o vetor de objetivos de projeto $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^m$. Considere-se a conotação de que $f_i(a) < f_i(b)$ implica que a solução a é melhor que a solução b no objetivo $f_i(\cdot)$.*

Encontrar um vetor solução $x = x^* \in \mathcal{X}_f$ tal que todo funcional $f_i(x^*)$, $i = 1, \dots, m$ atinja os menores valores possíveis. \square

O problema de encontrar soluções ótimas de projeto pode ser definido no contexto dos métodos de otimização mono-objetivo como:

Problema 2 [Problema P_ϵ] Dado o problema 1, encontrar $x^* \in \mathbb{R}^n$ tal que:

$$x^* = \arg \min f_i(x)$$

$$\text{sujeito a: } \begin{cases} x \in \mathcal{X}_f \\ f_j(x) \leq \gamma_j \quad , \quad (j = 1, \dots, m, j \neq i) \end{cases} \quad (12.1)$$

\square

O problema 2 é o problema P_ϵ , conforme definição apresentada no capítulo 10. Este é um dos dois esquemas para reduzir problemas com múltiplos objetivos a formulações mono-objetivo (o outro esquema básico é o do problema P_λ). Sabe-se que uma seqüência de execuções do problema 2 é capaz de produzir todas as soluções ditas *eficientes* do problema multiobjetivo, embora sejam eventualmente produzidas também soluções não-eficientes. Conceitualmente, o conjunto de todas as soluções eficientes do PMO é obtido pela seguinte operação de conjuntos:

Definição 12.1 ([Conjunto Pareto-Ótimo (1)]) Defina-se \mathcal{S}_i como o conjunto de todas as soluções do problema 2 quando o objetivo f_i é tomado como referência e o vetor $\gamma \in \mathbb{R}^{m-1}$ assume todos os valores possíveis. O conjunto \mathcal{P} definido por:

$$\mathcal{P} = \mathcal{S}_1 \cap \mathcal{S}_2 \cap \dots \cap \mathcal{S}_m \quad (12.2)$$

é chamado de conjunto de soluções eficientes, ou de conjunto Pareto-Ótimo do problema 1. \square

Esta definição é equivalente à que seria obtida com a aplicação do problema P_ϵ sobre apenas um objetivo de referência, seguida da realização do teste P^* .

Trabalhando-se explicitamente sob o ponto de vista da otimização multi-objetivo, passa a estar definido um *conjunto de soluções* que constitui o objeto da manipulação matemático-computacional a ser executada. Tal conjunto surge em contraponto ao objeto com o qual lida a otimização mono-objetivo, que é um único ponto-solução. Dessa forma, um objeto bem definido surge da definição do problema 1: o *Conjunto Pareto-Ótimo*. Deve-se admitir que este conjunto contém todas as “soluções razoáveis” para o problema 1.

NOTA 12.1 Esta forma de definir o conjunto Pareto-ótimo causa alguns inconvenientes de ordem operacional: O conjunto Pareto-ótimo \mathcal{P} está definido como uma interseção de conjuntos contínuos (que são subconjuntos de \mathbb{R}^n). Operacionalmente, está sendo tentada a determinação de uma amostra de \mathcal{P} , a partir de conjuntos de pontos que são soluções de problemas com formulação P_ϵ , sendo portanto amostras do que seriam os conjuntos \mathcal{S}_i (os quais também seriam conceitualmente contínuos). O teste de interseção de conjuntos fica, portanto, mal-definido, uma vez que os elementos a serem testados são amostras, e muito dificilmente haveria alguma interseção entre dois conjuntos de amostras de pontos tomadas de dois conjuntos contínuos, ainda que esses dois conjuntos possuíssem interseção. A interseção dos conjuntos de amostras só ocorreria caso fosse feita a amostragem de um mesmo ponto, pertencente ao conjunto interseção, duas vezes. Portanto, a definição 12.1 não possui valor operacional.

◇

Uma definição equivalente para o conjunto Pareto-ótimo pode ser estruturada segundo um esquema “relativo”:

Definição 12.2 ([Conjunto Pareto-Ótimo (2)]) *Considerem-se $\mathcal{X}_f \subset \mathbb{R}^n$ e $f : \mathbb{R}^n \mapsto \mathbb{R}^m$. Então:*

$$\mathcal{P} \triangleq \{x^* \in \mathcal{X}_f \mid \nexists x \in \mathcal{X}_f \text{ tal que } f(x) \leq f(x^*) \text{ e } f(x) \neq f(x^*)\} \quad (12.3)$$

define o conjunto \mathcal{P} , que é denominado conjunto Pareto-ótimo, ou conjunto de soluções eficientes do Problema 1. □

A equivalência das definições 12.1 e 12.2 pode ser demonstrada (Chankong & Haimes 1983). Embora ambas as definições dêem origem a um mesmo conjunto solução \mathcal{P} , a primeira tem suporte conceitual em uma série de soluções independentes advindas da aplicação de um mecanismo de otimização mono-objetivo a um conjunto de problemas oriundos da formulação P_ϵ . Esses problemas, uma vez definidos, passam a estar desconectados um do outro.

A segunda definição, ao contrário, liga qualquer solução pertencente ao conjunto \mathcal{P} a todas as demais soluções pertencentes ao mesmo conjunto. As *propriedades de grupo* que caracterizam os conjuntos Pareto-ótimos são enfatizadas com esta definição. Esta definição servirá de base, portanto, para o desenvolvimento de ferramentas de análise e síntese no presente capítulo. Com esta definição, o problema de projeto se torna:

Problema 3 [Projeto Multiobjetivo] *Dado o problema 1, encontrar o conjunto \mathcal{P} .* □

12.3 Análise Multiobjetivo

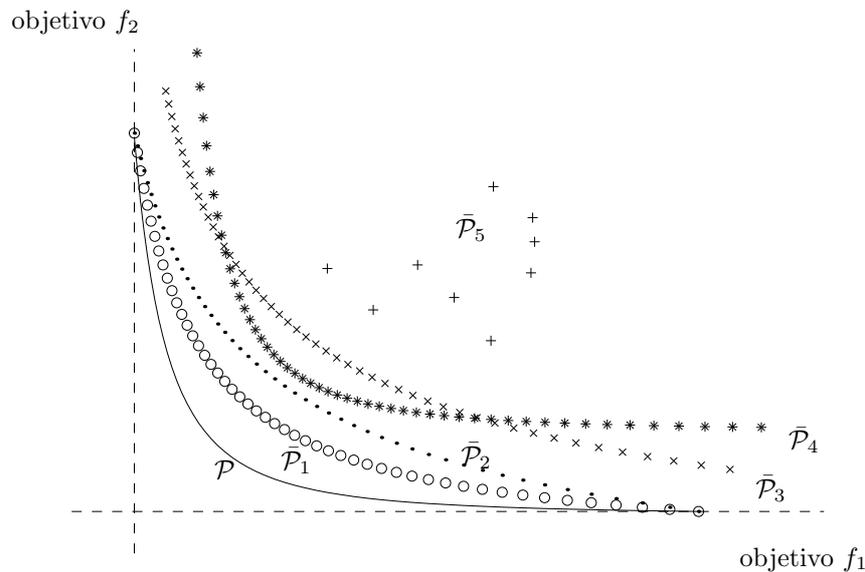


Figura 12.1: Estimativas computacionais típicas do conjunto Pareto-ótimo \mathcal{P} , no espaço de objetivos. Na figura: o conjunto Pareto-ótimo exato \mathcal{P} (linha contínua), conjunto estimado $\bar{\mathcal{P}}_1$ (o), conjunto estimado $\bar{\mathcal{P}}_2$ (\cdot), conjunto estimado $\bar{\mathcal{P}}_3$ (x), conjunto estimado $\bar{\mathcal{P}}_4$ (*), conjunto estimado $\bar{\mathcal{P}}_5$ (+). As linhas tracejadas denotam os valores ótimos dos objetivos f_1 e f_2 .

Para o propósito de discutir as propriedades do conjunto \mathcal{P} , Pareto-ótimo, que são relevantes aqui, a figura 12.1 mostra estruturas típicas que emergem

após a execução de qualquer cálculo realizado com o objetivo de estimar pontos pertencentes a esse conjunto. Um problema com apenas dois objetivos é considerado, com o propósito de facilitar a visualização. A análise a ser compreendida, no entanto, é válida para qualquer número de funcionais-objetivo.

O conjunto Pareto-ótimo exato, \mathcal{P} , é a curva contínua na figura 12.1. Esse conjunto exato, entretanto, não se encontra em princípio disponível, no caso de funcionais genéricos f_1 e f_2 , no sentido em que mesmo que se dispusesse de um conjunto de pontos que pertencessem a ele, não haveria maneira de provar tal fato. Para caracterizar as soluções que “provavelmente” pertençam ao conjunto \mathcal{P} , ou seja, as soluções *Pareto-candidatas*, uma outra relação é definida a seguir, utilizando apenas os pontos que se encontram disponíveis:

Definição 12.3 ([Conjunto Pareto-Candidato])

Sejam $f(\cdot)$ um vetor de objetivos e $\mathcal{K} \subset \text{Dom}(f(\cdot))$ um conjunto com um número finito de elementos: $\mathcal{K} = \{x_1, \dots, x_v\}$. O conjunto $\Psi(\mathcal{K})$, definido por:

$$\Psi(\mathcal{K}) \triangleq \{x_p \in \mathcal{K} \mid \nexists x_i \in \mathcal{K} \text{ tal que } f(x_i) \leq f(x_p) \text{ e } f(x_i) \neq f(x_p)\} \quad (12.4)$$

é denominado conjunto Pareto-candidato, associado ao “conjunto amostra” \mathcal{K} . □

De fato, devido à indisponibilidade do conjunto \mathcal{P} , a caracterização de conjuntos de soluções $\Psi(\cdot)$ como *conjuntos Pareto-candidatos* deve ser realizada a partir de procedimentos de *falseamento*, que podem mostrar que alguns conjuntos não são Pareto-candidatos, mas que não podem jamais mostrar que algum conjunto é de fato um subconjunto do conjunto Pareto-ótimo. Este é o papel do conceito de *conjunto Pareto-candidato*.

12.3.1 Consistência

Dado qualquer conjunto \mathcal{X} , este pode ser considerado como um *Pareto-candidato* apenas se $\Psi(\mathcal{X}) = \mathcal{X}$. Se isto ocorre, o conjunto é denominado *auto-consistente*. Caso contrário, a possibilidade de \mathcal{X} ser um subconjunto do conjunto Pareto-ótimo \mathcal{P} fica *falseada*.

12.3.2 Ordenamento e Dominância

Dados dois conjuntos, \mathcal{X}_1 e \mathcal{X}_2 , as seguintes relações de ordem entre esses conjuntos são definidas:

$$\begin{aligned}\mathcal{X}_1 \prec \mathcal{X}_2 &\Leftrightarrow \{\Psi(\mathcal{X}_1 \cup \mathcal{X}_2) \supset \mathcal{X}_1 \text{ e } \Psi(\mathcal{X}_1 \cup \mathcal{X}_2) \not\supset \mathcal{X}_2\} \\ \mathcal{X}_1 \preceq \mathcal{X}_2 &\Leftrightarrow \{\Psi(\mathcal{X}_1 \cup \mathcal{X}_2) \supset \mathcal{X}_1\}\end{aligned}\tag{12.5}$$

No caso em que $\mathcal{X}_1 \prec \mathcal{X}_2$, a possibilidade do conjunto \mathcal{X}_2 ser um subconjunto do conjunto Pareto-ótimo \mathcal{P} torna-se *falseada*, enquanto o conjunto \mathcal{X}_1 permanece sendo um *Pareto-candidato*. Neste caso, diz-se que \mathcal{X}_1 *domina* \mathcal{X}_2 . Existem duas possibilidades de não-dominância: se ambas as relações $\mathcal{X}_1 \preceq \mathcal{X}_2$ e $\mathcal{X}_2 \preceq \mathcal{X}_1$ não se verificam, então ambos os conjuntos se tornam *falseados* enquanto *Pareto-candidatos*; e se ambas as relações $\mathcal{X}_1 \preceq \mathcal{X}_2$ e $\mathcal{X}_2 \preceq \mathcal{X}_1$ funcionam, ambos os conjuntos seguem sendo *Pareto-candidatos*.

A partir desses conceitos, a figura 12.1 é analisada. O conjunto de estimativas $\bar{\mathcal{P}}_5$ não é *auto-consistente*, e portanto fica *falseado* como Pareto-candidato. Os conjuntos $\bar{\mathcal{P}}_1$ a $\bar{\mathcal{P}}_4$ são cada um auto-consistente e podem ser considerados, portanto, como *Pareto-candidatos* se apenas um deles estiver disponível. Existe entretanto uma relação de ordem entre esses conjuntos:

$$\mathcal{P} \prec \bar{\mathcal{P}}_1 \prec \bar{\mathcal{P}}_2 \prec \{\bar{\mathcal{P}}_3, \bar{\mathcal{P}}_4\} \prec \bar{\mathcal{P}}_5\tag{12.6}$$

Esse ordenamento corresponde a um ordenamento de dominância. Se todos esses conjuntos estiverem disponíveis, o único *Pareto-candidato* torna-se $\bar{\mathcal{P}}_1$, uma vez que:

$$\bar{\mathcal{P}}_1 = \Psi(\bar{\mathcal{P}}_1 \cup \bar{\mathcal{P}}_2 \cup \bar{\mathcal{P}}_3 \cup \bar{\mathcal{P}}_4 \cup \bar{\mathcal{P}}_5)\tag{12.7}$$

Os únicos conjuntos que não se encontram “ordenados” na figura 12.1 são $\bar{\mathcal{P}}_3$ e $\bar{\mathcal{P}}_4$. Se ambos estiverem disponíveis, ambos estariam *falseados* como *Pareto-candidatos*, sem necessidade de informação adicional.

12.3.3 Extensão

Um outro tipo de análise que é útil para a avaliação de uma estimativa do conjunto Pareto-ótimo está associado à extensão na qual essa estimativa cobre a superfície de Pareto. Para este propósito, algumas definições adicionais se fazem necessárias.

Considere-se o espaço \mathcal{Y} dos vetores-objetivo. Seja $\epsilon > 0$ um número inteiro real fixo, e $h \in \mathcal{Y}$ um ponto solução. O conjunto $\delta(\cdot, \cdot)$ é definido por:

$$\delta(\epsilon, h) = \{g \in \mathcal{Y} \text{ tal que } |g - h| \leq \epsilon\} \quad (12.8)$$

Tome-se o conjunto $\mathcal{X} = \{x_1, \dots, x_v\}$, $\mathcal{X} \subset \mathcal{Y}$.

Definição 12.4 ([ϵ -Extensão]) *O conjunto $\Theta(\epsilon, \mathcal{X})$ definido por*

$$\Theta(\epsilon, \mathcal{X}) = \bigcup_{i=1}^v \delta(\epsilon, x_i) \quad (12.9)$$

é a ϵ -extensão do conjunto \mathcal{X} . □

Para qualquer conjunto \mathcal{X} , $\Theta(\epsilon, \mathcal{X}) \supset \mathcal{X}$ trivialmente se verifica.

Considerem-se agora dois conjuntos \mathcal{X}_1 e \mathcal{X}_2 tais que $\mathcal{X}_1 \preceq \mathcal{X}_2$ e $\mathcal{X}_2 \preceq \mathcal{X}_1$, isto é, ambos os conjuntos são *Pareto-candidatos*, e seja dado um $\epsilon > 0$. As seguintes relações ficam definidas:

$$\begin{aligned} \Theta(\epsilon, \mathcal{X}_1) \supset \mathcal{X}_2 &\Leftrightarrow \mathcal{X}_1 \stackrel{\epsilon}{\supset} \mathcal{X}_2 \\ \Theta(\epsilon, \mathcal{X}_1) \not\supset \mathcal{X}_2 &\Leftrightarrow \mathcal{X}_1 \stackrel{\epsilon}{\not\supset} \mathcal{X}_2 \end{aligned} \quad (12.10)$$

As seguintes situações podem ocorrer:

$\mathcal{X}_1 \stackrel{\epsilon}{\supset} \mathcal{X}_2$ **and** $\mathcal{X}_2 \stackrel{\epsilon}{\supset} \mathcal{X}_1$: Neste caso, os conjuntos \mathcal{X}_1 e \mathcal{X}_2 são ditos *extenso-equivalentes*.

$\mathcal{X}_1 \stackrel{\epsilon}{\supset} \mathcal{X}_2$ **and** $\mathcal{X}_2 \stackrel{\epsilon}{\not\supset} \mathcal{X}_1$: Neste caso, o conjunto \mathcal{X}_2 é dito ser um *extenso-subconjunto* do conjunto \mathcal{X}_1 .

$\mathcal{X}_1 \stackrel{\epsilon}{\not\supset} \mathcal{X}_2$ **and** $\mathcal{X}_2 \stackrel{\epsilon}{\not\supset} \mathcal{X}_1$: Neste caso, os conjuntos são ditos *extenso-incomensuráveis*.

Se um conjunto é um *extenso-subconjunto* ou se é *extenso-incomensurável* quando comparado com outro conjunto, então ele se torna *falseado* enquanto *Pareto-candidato*.

12.3.4 Dados Extremos

Outra informação capaz de guiar uma análise de consistência de conjuntos Pareto-candidatos pode também estar disponível *a priori* em determinadas situações: trata-se dos ótimos individuais de algumas ou todas as funções-objetivo. Tal informação pode ficar disponível, por exemplo, pelo emprego de ferramentas analíticas, ou pelo conhecimento da física do sistema em questão, ou ainda pelo fato do mínimo de determinadas funções ser trivialmente conhecido (um volume ou uma massa, por exemplo, devem ser no mínimo zero). Isto significa que alguns pontos extremos do conjunto Pareto-ótimo são conhecidos.

Na figura 12.1, se este for o caso, os conjuntos $\bar{\mathcal{P}}_3$ e $\bar{\mathcal{P}}_4$ ficariam ambos *falseados* como *Pareto-candidatos*. Os conjuntos $\bar{\mathcal{P}}_1$ e $\bar{\mathcal{P}}_2$, tomados individualmente, permaneceriam sendo *candidatos*.

12.4 Decisão e Síntese Multiobjetivo

A análise multiobjetivo que foi apresentada pode ser utilizada de duas formas distintas:

I) Como um procedimento de validação para a análise dos resultados de algoritmos específicos. Neste caso, qualquer algoritmo que reivindicar que “resolva” um certo problema multiobjetivo deve produzir resultados que, sendo submetidos a todo o procedimento de análise, permaneçam sendo *Pareto-candidatos*. Sumarizando tal procedimento de análise, os seguintes critérios devem ser empregados para a avaliação de um algoritmo \mathbb{A} contra um algoritmo “de referência” \mathbb{B} :

Critério 1: Qualquer conjunto de pontos-resultado $\bar{\mathcal{P}}_a$, gerado por \mathbb{A} , deve ser *auto-consistente*.

Critério 2: Dado um conjunto de pontos-resultado $\bar{\mathcal{P}}_b$, gerado pelo algoritmo \mathbb{B} , a relação de ordem $\bar{\mathcal{P}}_a \preceq \bar{\mathcal{P}}_b$ deve prevalecer, qualquer que seja o algoritmo \mathbb{B} .

Critério 3: A relação $\bar{\mathcal{P}}_a \stackrel{\epsilon}{\supset} \bar{\mathcal{P}}_b$ deve prevalecer para todo algoritmo \mathbb{B} .

Critério 4: Se o problema possui ótimos individuais conhecidos *a priori*, então a relação $\Psi(\bar{\mathcal{P}}_a \cup \mathcal{O}) \supset \bar{\mathcal{P}}_a$ deve prevalecer, com \mathcal{O} denotando o conjunto de soluções individuais das funções objetivo.

II) Como um procedimento de agregação que tome um conjunto de pontos (que podem ter sido gerados por um algoritmo único ou por qualquer combinação de algoritmos diferentes) e produza um subconjunto que seja Pareto-candidato dada toda a informação disponível. Sejam $\bar{\mathcal{P}}_1, \dots, \bar{\mathcal{P}}_N$ conjuntos de pontos-solução resultantes da execução de N algoritmos diferentes. A estimativa do conjunto Pareto-ótimo, dados esses conjuntos, é: $\bar{\mathcal{P}}_e = \Psi(\bar{\mathcal{P}}_1 \cup \dots \cup \bar{\mathcal{P}}_N)$.

Na forma (I), um conjunto de pontos-solução que são obtidos a partir de um certo algoritmo é tomado como representativo das possibilidades de escrutínio desse algoritmo sobre o conjunto Pareto-ótimo do problema específico sob análise (o qual pode ainda ser considerado como um problema representativo de uma classe de problemas). Esse conjunto de pontos é analisado através dos critérios 1-4.

A análise a ser realizada irá comparar algoritmos, tendo como base conjuntos de soluções obtidas com os mesmos. Cada conjunto de pontos, originário de um único algoritmo, é visto como um único objeto. A comparação entre dois desses objetos é tomada como representativa da comparação entre os próprios algoritmos que originaram os conjuntos de pontos.

A utilização (II), ao contrário, irá tomar um conjunto de pontos-solução que podem ter sido obtidos de qualquer forma, e irá aplicar os critérios 1 e 4 para eliminar pontos não-consistentes. Ao serem fundidos todos os pontos em uma única massa de dados, os critérios 2 e 3 estarão automaticamente atendidos.

Na próxima seção, como exemplo desse tipo de uso dos critérios, será proposto um algoritmo genético multiobjetivo (AGM) que irá utilizar como condições iniciais os conjuntos de soluções encontradas por outros algoritmos multiobjetivo. Dessa forma, o AGM se tornará, por construção, “melhor ou igual” que qualquer desses algoritmos.

O algoritmo genético multiobjetivo proposto permite os seguintes ganhos de qualidade nos resultados da otimização:

- Os resultados de diversos algoritmos são agregados em um único conjunto de dados consistentes;
- São eventualmente produzidos mais dados melhores que os anteriores, devido às propriedades de “otimização global” associadas aos operadores genéticos;

- Caso o algoritmo genético não seja capaz de gerar dados “melhores” que os anteriores, fica *corroborada* a hipótese de que os dados anteriormente disponíveis já constituíssem parte do conjunto Pareto-ótimo.

12.5 Algoritmo Genético Multiobjetivo

Solucionar o problema da otimização de funcionais arbitrários tem sido, desde os primórdios do desenvolvimento da teoria de otimização, uma meta importante. Entretanto, cada método de otimização diferente que já foi construído se baseia em diversas premissas a respeito da estrutura do funcional a ser otimizado: linearidade, convexidade, diferenciabilidade, etc. A classe de métodos que atingiu a melhor aproximação para o problema da otimização de funcionais arbitrários é a família dos “métodos estocásticos de otimização”. Um grupo de métodos que atingiu grande aplicabilidade, dentro de tal classe, é a família dos “Algoritmos Genéticos”.

Devido às propriedades de “otimização global” dos Algoritmos Genéticos, eles se tornaram uma ferramenta natural para abordar problemas em que se desconhece a estrutura da função (ou funções) a ser otimizada. Uma outra razão potencial para esta adequabilidade é apontada aqui: uma vez que os algoritmos genéticos trabalham com populações de soluções candidatas, ao invés de trabalhar com uma única solução candidata (como seria o caso dos outros algoritmos usuais de otimização), eles são capazes de incorporar operadores que exploram as “propriedades de grupo” das estimativas do conjunto Pareto-ótimo.

12.5.1 Construção do Algoritmo Genético Multiobjetivo

Um algoritmo genético multiobjetivo (AGM) pode ser construído a partir de modificações introduzidas em qualquer algoritmo genético mono-objetivo.

Um “Algoritmo Genético” pode ser definido como a aplicação sucessiva das seguintes operações sobre um conjunto de soluções-candidatas do problema (esse conjunto sendo denominado “população”):

cruzamento: A população é dividida em pares, e cada par de soluções é substituído por um novo par, o qual é gerado empregando informação obtida do par original;

mutação: Algumas soluções (“indivíduos”) são aleatoriamente escolhidos para receber uma perturbação em seus parâmetros, sendo substituídos pelo indivíduo “perturbado”;

seleção: A população que surge após a aplicação das operações de cruzamento e mutação é modificada, com a exclusão de alguns indivíduos e a replicação de outros, sendo mantido o tamanho total da população. A probabilidade de um indivíduo ser replicado é tão maior quanto menor for o valor dos funcionais-objetivo a ele associados (considerando problemas de minimização);

elitismo: Alguns indivíduos (os “melhores”) são mantidos deterministicamente dentro da população.

Após algumas aplicações dessas operações, a “população” converge para soluções que, em algum sentido, são “boas aproximações” das soluções globais do problema.

Qualquer algoritmo genético mono-objetivo pode ser adaptado através das seguintes linhas, para a obtenção de um algoritmo genético multiobjetivo:

- Selecionar, de dentro da população inicial \mathcal{Q}_0 , o grupo de indivíduos que forma o máximo subconjunto consistente $\bar{\mathcal{P}}_0$. Esta operação é definida por: $\bar{\mathcal{P}}_0 = \Psi(\mathcal{Q}_0)$.
- A cada iteração, uma nova população \mathcal{Q}_i é gerada pela aplicação dos operadores genéticos. Re-calcular a estimativa do conjunto Pareto-ótimo, fazendo $\bar{\mathcal{P}}_i = \Psi(\mathcal{Q}_i)$, eventualmente excluindo alguns indivíduos e incluindo outros. O conjunto $\bar{\mathcal{P}}_i$ é empregado como o “conjunto-elite” na operação de elitismo.
- Uma técnica de “nicho” pode ser empregada, para evitar a inclusão de pontos que estiverem muito próximos um do outro no conjunto $\bar{\mathcal{P}}$. Desta forma, o conjunto-solução $\bar{\mathcal{P}}_{mga}$ sofre uma pressão para cobrir todo o conjunto \mathcal{P} .
- O funcional que guia o operador de seleção deve ser composto com os funcionais individuais que compõem o vetor de objetivos. Na implementação específica utilizada nos exemplos deste capítulo, eles são normalizados para a faixa de $[0, 1]$ e então agregados através do operador *max*.

Desta forma, ao invés de se procurar por soluções únicas, a totalidade do conjunto \mathcal{P} é perseguida, como um objeto com propriedades intrínsecas que auxiliam nessa busca. O procedimento de projeto começa com qualquer algoritmo *não-consistente* (ou *conservativo*), o qual fornece um conjunto de soluções iniciais $\bar{\mathcal{P}}_a$. Esse conjunto de soluções é a seguir refinado pelo algoritmo genético multiobjetivo (AGM).

Denote-se por $\bar{\mathcal{P}}_i$ o conjunto Pareto-candidato produzido pelo AGM na sua i -ésima iteração. Os operadores genéticos foram formados tais que:

nicho+seleção: Produz uma pressão que leva a estimativa do conjunto de Pareto $\bar{\mathcal{P}}_{mga}$ a aumentar sua “extensão”.

elitismo: Garante deterministicamente que: (i) $\bar{\mathcal{P}}_{i+1} \preceq \bar{\mathcal{P}}_i$ e (ii) $\bar{\mathcal{P}}_{i+1} \stackrel{\epsilon}{\supset} \bar{\mathcal{P}}_i$.

seleção: Produz a pressão por melhorias que permite que em algum momento possa ocorrer $\bar{\mathcal{P}}_{i+1} \prec \bar{\mathcal{P}}_i$ e $\bar{\mathcal{P}}_i \stackrel{\epsilon}{\not\supset} \bar{\mathcal{P}}_{i+1}$.

O algoritmo genético multiobjetivo, portanto, estende as soluções iniciais (obtidas de outros algoritmos) nos seguintes sentidos:

- A superfície de Pareto usualmente é movida “para baixo” no espaço de objetivos.
- Algumas “falhas” na estimativa de \mathcal{P} podem ainda ser “reparadas” pelo AGM. Tais falhas seriam “buracos” na superfície estimada, ou trechos não cobertos anteriormente.

12.5.2 Algoritmo Genético - Real Polarizado Multiobjetivo

Como instância do algoritmo conceitual AGM proposto acima, é aqui apresentado o chamado Algoritmo Genético Real Polarizado Multiobjetivo (AG-RPMO), que foi proposto pelo presente autor e que será empregado aqui como mecanismo de otimização para a execução de exemplos. Tal algoritmo é uma variação do Algoritmo Genético Real Polarizado (AGRP), apresentado no capítulo 8, tendo sido construído segundo as diretrizes estabelecidas na subseção anterior.

Há registros na literatura de outros métodos de utilização de algoritmos genéticos para otimização multiobjetivo, os quais utilizam metodologias bastante diversas daquela aqui apresentada (Viennet, Fonteix & Marc 1996)

Algoritmo Genético Real Polarizado Multiobjetivo (AG-RPMO)

- Cada parâmetro de projeto é descrito por uma variável real, sendo o conjunto de parâmetros armazenado em um vetor no espaço \mathbb{R}^n . Cada indivíduo corresponde a um vetor nesse espaço.
- Existe uma faixa admissível para cada um dos parâmetros (ou seja, para cada coordenada do vetor de parâmetros), dentro da qual estarão localizados os respectivos componentes de todos os indivíduos.
- O algoritmo se inicia com a geração aleatória de um número N par, de vetores (indivíduos) dentro das faixas admissíveis.
- São realizadas em seqüência as operações de: *cruzamento*, *mutação*, *avaliação*, *cálculo da função de ajuste* (“fitness function”), *elitização com nicho* e *seleção*. A *elitização* pode gerar um número variável de indivíduos, de forma que o total de indivíduos pode crescer ou diminuir de geração para geração, sendo sempre maior ou igual a N e sendo sempre par. Como subproduto da operação de elitização, vai sendo criado um conjunto que corresponde a uma estimativa do conjunto Pareto-ótimo, o qual vai sendo aperfeiçoado a cada iteração.
- O algoritmo termina seja atingindo determinada condição de término, seja excedendo o número máximo permitido de iterações.

As operações realizadas são definidas da seguinte forma:

Cruzamento: Divide-se a população em duas metades. Para cada par formado, verifica-se se vai ou não ocorrer cruzamento, com probabilidade de ocorrência de 0,6. Caso vá ocorrer cruzamento, são gerados dois

novos indivíduos segundo a lei:

$$x_g = \alpha x_1 + (1 - \alpha)x_2$$

$$-0,1 < \alpha < 1,1$$

sendo x_g o novo indivíduo gerado, x_1 e x_2 os indivíduos ancestrais. Deve-se observar neste caso a restrição de que:

$$B(x_2) < B(x_1)$$

sendo $B(\cdot)$ a função objetivo a ser minimizada (ver definição dessa função na equação (12.13)). Para a geração de α , verifica-se se o cruzamento será *polarizado* ou *não-polarizado*, sendo que a probabilidade de ser polarizado é de 0,3. Caso não seja polarizado, adota-se α com distribuição uniforme de probabilidade dentro do intervalo de valores possíveis para ambos os novos indivíduos gerados. Caso seja polarizado, para um dos novos indivíduos escolhe-se:

$$\alpha = 1,4\beta_1\beta_2 - 0,2 \quad (12.11)$$

sendo β_1 e β_2 escolhidas aleatoriamente e independentemente, com distribuição de probabilidade uniforme no intervalo $[0, 1]$. O outro indivíduo sempre será escolhido sem polarização. Sendo definida dessa forma a operação de cruzamento, torna-se possível que os indivíduos gerados estejam localizados fora das faixas admissíveis de parâmetros. Caso isso ocorra, ainda é realizada uma operação de *reflexão* do indivíduo para o interior da região admissível. Essa operação é definida como:

$$x_r = x_L + |x - x_L|$$

para reflexão no limite inferior, sendo x o indivíduo que violava a restrição, x_r o resultado da reflexão, e x_L o vetor de limites inferiores. Para a reflexão no limite superior, a operação é definida por:

$$x_r = x_U - |x_U - x|$$

para x_U o vetor de limites superiores, e as demais variáveis com mesmo significado que anteriormente.

Mutação: Determina-se para cada indivíduo se o mesmo sofrerá ou não mutação, com probabilidade igual a 0,02. Caso vá ocorrer mutação, é somado ao indivíduo x um vetor δ cujas componentes são dadas por:

$$\delta_i = 0,05\beta_i(x_R)_i$$

sendo β_i um número aleatório com distribuição gaussiana, média zero e variância um, e x_R o vetor de diferença entre os máximos e mínimos dos parâmetros:

$$x_R = (x_U - x_L)$$

Avaliação: Cada indivíduo é avaliado em cada uma das funções objetivo, sendo criado um vetor de objetivos associado a cada indivíduo.

Função de Ajuste: Cada função objetivo é injetada na *função de ajuste*, dada por (12.12), sendo obtido para cada indivíduo um vetor de funções de ajuste. Foi adotado um valor de $\gamma = 1,8$. De cada vetor é extraído o *mínimo* ou o *máximo* (o usuário define qual dos dois critérios é empregado em cada simulação). Considerando agora o conjunto de valores obtidos, cada um associado a um indivíduo da população, como sendo um valor de “função objetivo”, aplica-se ao conjunto novamente a *função de ajuste*.

Elitização com Nicho: O conjunto de elementos que são candidatos a Pareto-ótimos é montado, na primeira iteração, após a avaliação dos pontos, através de uma simples verificação de *não-dominância* associada à verificação de *nicho*. Esse conjunto é atualizado a cada iteração, considerando os pontos anteriormente incluídos e os novos pontos gerados com suas respectivas avaliações, sendo que cada ponto pode ser incluído ou não, e pode causar ou não a exclusão de pontos anteriormente incluídos, com a verificação de *não-dominância* e a verificação de *nicho*. Todo o conjunto de pontos resultantes dessa operação em uma iteração é incluído na população da iteração seguinte.

Seleção: É realizada uma seleção de pelo menos p indivíduos, sendo $p < N$, de forma a garantir que o total de indivíduos na nova população seja maior que ou igual a N . A probabilidade de um indivíduo ser selecionado a cada vez é igual ao valor da fração de sua função de ajuste em relação à soma das funções de ajuste de todos os indivíduos.

Para completar a caracterização do AG-RPMO:

Função de Ajuste: Seja J o vetor das avaliações de uma das funções-objetivo para os N indivíduos da população. A equação da função de ajuste (FT) é dada por:

$$\begin{aligned} \bar{J} &= \text{média}(J) \\ J_M &= \max(J) \\ J_m &= \min(J) \\ v &= \frac{(\gamma J_m - J_M)}{(\gamma - 1)} \\ J_m \geq v &\Rightarrow \begin{cases} \alpha = \bar{J} \frac{(\gamma - 1)}{(J_M - \bar{J})} \\ \beta = \bar{J} \frac{(J_M - \gamma \bar{J})}{(J_M - J)} \end{cases} \\ J_m < v &\Rightarrow \begin{cases} \alpha = \frac{\bar{J}}{(J_M - J_m)} \\ \beta = -\frac{\bar{J} J_m}{(J_M - J_m)} \end{cases} \end{aligned} \quad (12.12)$$

$$FT = \alpha J + \beta$$

Função Objetivo: A “função objetivo” a ser empregada como critério de seleção e de polarização é definida como:

$$B(x) = \max(FT(f_1(x)), FT(f_2(x)), \dots, FT(f_m(x))) \quad (12.13)$$

para um problema com m funções-objetivo.

Avaliação de Dominância: Seja um ponto cujo vetor de objetivos é $f(x) = [f_1(x) \ \dots \ f_m(x)]'$. Seja ainda

$$F = \begin{bmatrix} f_1(x_1) & \dots & f_1(x_p) \\ f_2(x_1) & \dots & f_2(x_p) \\ \vdots & \vdots & \vdots \\ f_m(x_1) & \dots & f_m(x_p) \end{bmatrix}$$

o conjunto das avaliações das funções-objetivo do conjunto de pontos previamente pertencentes à estimativa do conjunto Pareto-ótimo, (x_1, \dots, x_p) . Denotando-se as colunas de F por F^i , tem-se que:

$$f(x) \text{ não é dominado em } F \Leftrightarrow \nexists F^i \leq f(x) ; F^i \neq f(x) \quad (12.14)$$

e:

$$f(x) \text{ domina em } F \Leftrightarrow \exists F^i \geq f(x) ; F^i \neq f(x) \quad (12.15)$$

Se $f(x)$ domina em F , o ponto x é automaticamente introduzido no conjunto F , sendo excluídos os pontos dominados. Se $f(x)$ não domina mas também não é dominado em F , é feita a avaliação de nicho.

Avaliação de Nicho: Seja um ponto x , com vetor de objetivos $f(x)$, e seja um conjunto de pontos com avaliações previamente armazenadas na matriz F , que constitui a estimativa corrente do conjunto Pareto-ótimo. Caso o ponto x não domine nem seja dominado em F , sua avaliação será introduzida em F se:

$$d(f(x), F^i) \geq \delta \quad \forall i = 1, \dots, p \quad (12.16)$$

sendo $d(\cdot, \cdot)$ uma medida de distância (por exemplo a norma euclideana) e δ o tamanho do nicho.

NOTA 12.2 Deve-se observar que a escolha do critério de máximo para a definição da função objetivo agregada, no cálculo da função de ajuste, significa que serão privilegiados os pontos que não tiverem desempenho ruim em nenhum dos funcionais objetivo. Usualmente esses pontos estão longe dos ótimos individuais de cada um dos funcionais, onde em geral as funções que não estiverem otimizadas têm valores elevados. Por outro lado, a escolha do critério de mínimo privilegia a escolha dos pontos que tiverem desempenho bom em algum funcional, o que significa que os pontos tendem a se aproximar dos extremos das funções objetivo. Dessa forma, o chaveamento dos dois critérios seria uma forma de descrever em maior extensão e detalhe a curva Pareto-ótima. Com a adoção da técnica de nicho, entretanto, esse chaveamento torna-se desnecessário, sendo que ficam muito semelhantes os desempenhos para ambos os critérios (máximo e mínimo).

◇

12.6 Exemplo de Aplicação: Projeto de Controladores

Nesta seção, dois exemplos de aplicação das técnicas desenvolvidas neste capítulo são apresentados. Os exemplos se referem ao projeto de controladores para sistemas dinâmicos lineares, com os objetivos de minimizar a norma \mathcal{H}_2 e \mathcal{H}_∞ do sistema resultante em malha fechada. Este é um problema em aberto hoje (2001) na teoria de controle, sendo que não há algoritmos capazes de resolver de maneira exata o projeto multiobjetivo nesses dois objetivos.

Uma característica interessante desse problema “multiobjetivo $\mathcal{H}_2/\mathcal{H}_\infty$ ” é que o mesmo vem sendo abordado na literatura essencialmente por meio de métodos derivados do problema (P_ϵ), vide por exemplo (Chen, Cheng & Lee 1995, Scherer, Gahinet & Chilali 1997). Uma exceção a tal regra pode ser encontrada em (Takahashi et al. 1997). Os algoritmos que servirão de base de comparação com o algoritmo genético multiobjetivo são apresentados em (Khargonekar & Rotea 1991) (algoritmo LMI padrão) e em (Cao, Lam & Sun 1998, Shimomura & Fujii 2000) (algoritmos BMI), sendo todos eles baseados em metodologias do tipo (P_ϵ).

12.6.1 Realimentação completa de estados

Um sistema bastante simples é apresentado em primeiro lugar, para permitir a visualização do problema tanto no espaço de objetivos (o espaço das normas \mathcal{H}_2 e \mathcal{H}_∞ do sistema em malha fechada) quanto no espaço de parâmetros do controlador.

As equações do sistema são:

$$\dot{x} = \begin{bmatrix} -0.3868 & 0.0751 \\ 0 & -0.0352 \end{bmatrix} x + \begin{bmatrix} -0.6965 \\ 1.6961 \end{bmatrix} u + \begin{bmatrix} 0.0591 & 0 \\ 0 & 1.7971 \end{bmatrix} w$$

$$z = \begin{bmatrix} 0.0346 & 0.0535 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0.5297 \end{bmatrix} u$$

Este sistema é controlado com um controlador por realimentação estática de estados:

$$u = [K_1 \quad K_2] x$$

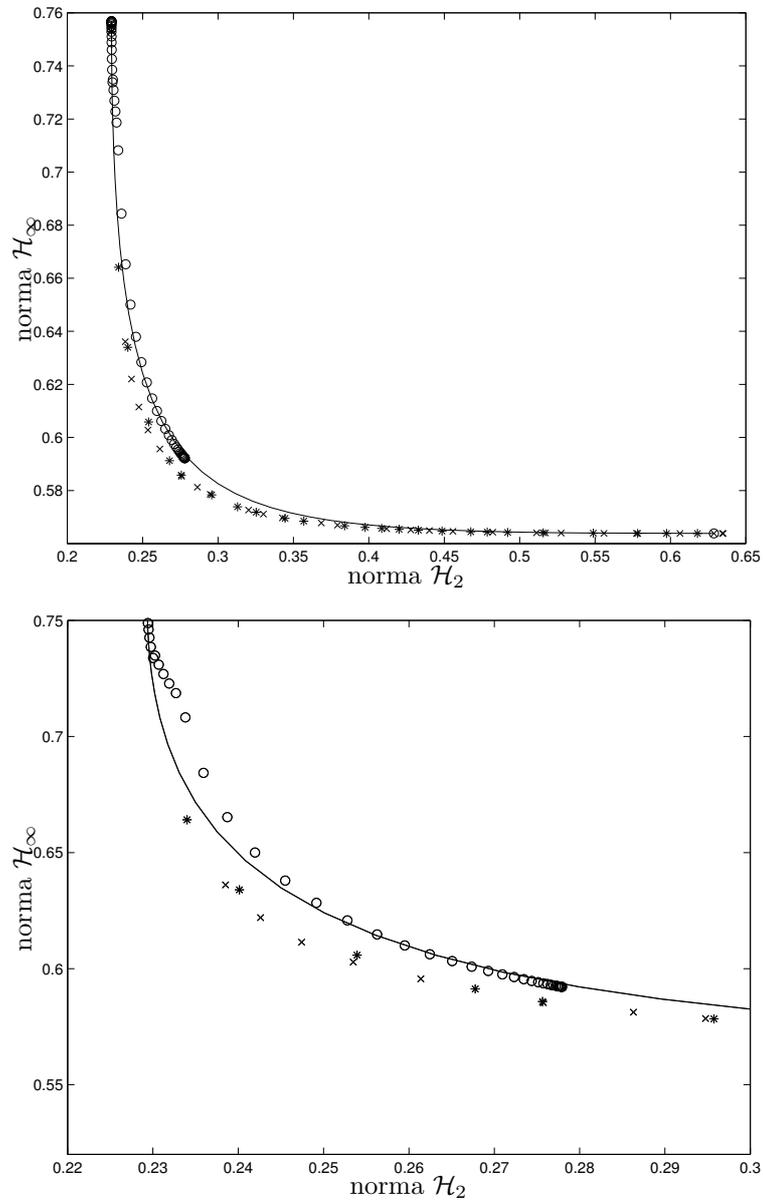


Figura 12.2: Estimativas do conjunto Pareto-ótimo, no espaço de objetivos, obtidas a partir de: (linha contínua)- formulação padrão LMI; (o)- formulação BMI; (x)- algoritmo genético multiobjetivo inicializado com as soluções LMI; (*)- algoritmo genético multiobjetivo inicializado com as soluções BMI. Abaixo, é mostrado um detalhe da figura.

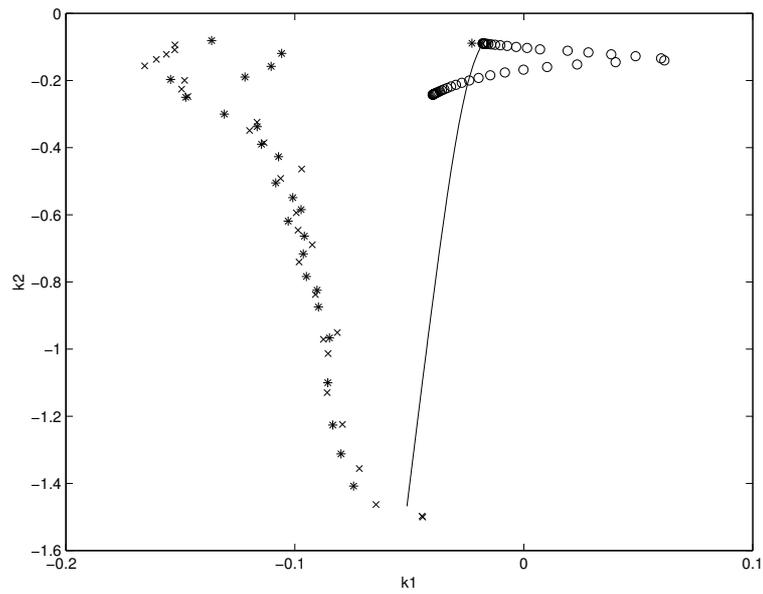


Figura 12.3: Estimativas do conjunto Pareto-ótimo, no espaço dos parâmetros do controlador, obtidas a partir de: (linha contínua)- formulação padrão LMI; (o)- formulação BMI; (x)- algoritmo genético multiobjetivo inicializado com as soluções LMI; (*)- algoritmo genético multiobjetivo inicializado com as soluções BMI.

O problema de projeto desse controlador é resolvido através de dois métodos aproximados: (i) um método padrão de LMI's (Linear Matrix Inequalities); (ii) um método de BMI's (Bilinear Matrix Inequalities); e (iii) o algoritmo genético multiobjetivo inicializado com as soluções obtidas dos métodos anteriores. As normas de malha fechada obtidas estão representadas na figura 12.2, e os parâmetros do controlador na figura 12.3.

Essas figuras mostram que, neste caso, a formulação BMI e a formulação LMI são mutuamente excludentes, isto é, $\mathcal{X}_{LMI} \not\subseteq \mathcal{X}_{BMI}$ e $\mathcal{X}_{BMI} \not\subseteq \mathcal{X}_{LMI}$. Por outro lado, $\mathcal{X}_{agm} \subseteq \mathcal{X}_{LMI}$ e $\mathcal{X}_{agm} \subseteq \mathcal{X}_{BMI}$. Também a questão da extensão revela a superioridade do AGM. O AGM ainda chega ao mesmo resultado, independentemente do conjunto de soluções iniciais utilizado como população inicial. Isso corrobora a hipótese de que o conjunto atingido corresponda ao conjunto Pareto-ótimo exato.

12.6.2 Realimentação estática de saídas

Agora é investigado o seguinte sistema instável de ordem três, com duas entradas de controle e duas saídas medidas:

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 1.9574 & -0.3398 & 1.1902 \\ 0.5045 & 0 & -1.1162 \\ 1.8645 & -0.2111 & 0.6353 \end{bmatrix} x + \begin{bmatrix} -0.6014 & 0 \\ 0.5512 & -2.0046 \\ -1.0998 & -0.4931 \end{bmatrix} u + \\ &+ \begin{bmatrix} 0.4620 & 0 & 0 \\ 0 & -0.3210 & 0 \\ 0 & 0 & 1.2366 \end{bmatrix} w \\ y &= \begin{bmatrix} 0 & 0.2311 & 0 \\ -2.3252 & 0 & 0 \end{bmatrix} x \\ z &= \begin{bmatrix} 0.1372 & 0.4374 & 0.7258 \\ 0.5216 & 0.4712 & 0 \\ 0.8952 & 0 & 0.3584 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.6264 & 0.9781 \\ 0.2412 & 0.6405 \end{bmatrix} u \end{aligned}$$

Emprega-se agora a estrutura de controle de realimentação estática de saídas:

$$u = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}$$

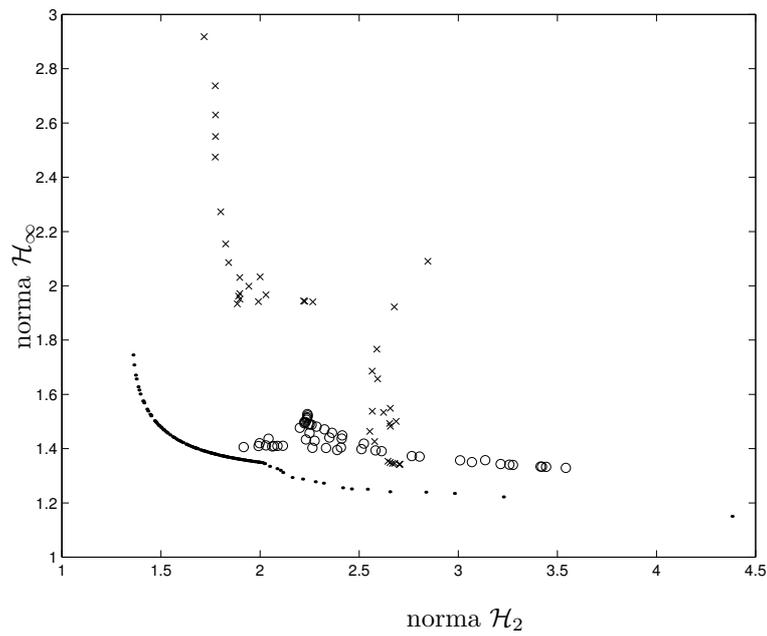


Figura 12.4: Estimativas do conjunto Pareto-ótimo, no espaço dos objetivos, obtida a partir de: (o) e (x) - duas formulações BMI; (·) algoritmo genético multiobjetivo inicializado com o conjunto das soluções BMI anteriores.

Dois algoritmos baseados em BMI's são executados, fornecendo soluções que estão representadas na figura 12.4. Nesta mesma figura, também estão representadas as soluções obtidas com o AGM, executado a partir das soluções anteriores.

Deve-se notar agora que as estimativas obtidas com os dois algoritmos de BMI's não são sequer auto-consistentes. Por outro lado, o algoritmo multiobjetivo genético produz uma estimativa do conjunto Pareto-ótimo que é: auto-consistente, dominante (em relação às estimativas anteriores) e possui maior extensão (também em relação às anteriores).

Capítulo 13

Exercícios - Otimização Vetorial

1. Mostre que, num problema convexo (ou seja, com todas as funções-objetivo e restrições convexas) não existe nenhuma solução eficiente que não seja solução de um problema ponderado (P_λ).
2. Construa um exemplo em que as condições de Kuhn-Tucker para eficiência estejam satisfeitas em um ponto não pertencente ao conjunto de soluções eficientes.
3. Compare a forma geométrica dos conjuntos de soluções possíveis para problemas lineares (isto é, problemas com função objetivo linear e restrições lineares) mono-objetivo e problemas lineares multi-objetivo. Faça a análise gráfica para o caso de duas variáveis de otimização.
4. Determine o conjunto de soluções eficientes do problema cujos três ob-

jetivos são f_1 , f_2 e f_3 definidos abaixo:

$$f_1(x) = x_1^2 + x_2^2$$

$$f_2(x) = (x_1 - 3)^2 + x_2^2$$

$$f_3(x) = x_1^2 + (x_2 - 2)^2$$

sendo $x \in \mathbb{R}^2$.

5. Considerando ainda as mesmas funções f_1 , f_2 e f_3 acima, suponha que é imposta a restrição:

$$f_3(x) < \epsilon$$

Sendo agora consideradas duas funções objetivo, $f_1(x)$ e $f_2(x)$, determine qual é o conjunto de soluções eficientes do problema, para ϵ variando de $\epsilon = 0.5$ até $\epsilon = 4.0$, com intervalo de variação de 0.5.

6. Considere ainda as mesmas funções. Sejam as restrições:

$$f_3(x) < 1$$

$$x_2 < 2$$

Seja considerado o problema bi-objetivo com os objetivos $f_1(x)$ e $f_2(x)$. Utilizando as condições de Kuhn-Tucker para eficiência, verifique se podem ser eficientes os pontos $[0 \ 2]$ e $[0 \ 0]$.

7. Mostre que é possível, em problemas de otimização multi-objetivo:
- (a) haver casos em que o problema ponderado (P_λ) não produz todas as soluções eficientes;

- (b) haver casos em que o problema ϵ -restrito produz resultados que não são soluções eficientes;
- (c) haver casos em que as condições de Kuhn-Tucker para eficiência são satisfeitas em pontos não pertencentes ao conjunto de soluções eficientes do problema.

8. Considere as funções-objetivo:

$$f_1(x) = x_1 + x_2$$

$$f_2(x) = x_1^2$$

Considere também as restrições:

$$g_1(x) = x_1^2 + x_2^2 \leq 4$$

$$g_2(x) = x_2 \leq 1$$

$$g_3(x) = x_1 - x_2 \leq 2$$

Determine o conjunto de soluções eficientes desse problema.

9. Considere as funções-objetivo:

$$f_1(x) = x_1^2 + x_2^2$$

$$f_2(x) = x_1^2 + (x_2 - 2)^2$$

$$f_3(x) = (x_1 - 2)^2 + (x_2 - 2)^2$$

$$f_4(x) = (x_1 - 2)^2 + x_2^2$$

Escreva as equações:

- (a) de dois problemas ponderados cujos resultados são (ambos) a solução eficiente $x = [1 \ 1]'$;
- (b) de dois problemas ϵ -restritos que também apresentem como resultado essa mesma solução eficiente.

10. Considere as funções:

$$f_1(x) = -x_1^2 - x_2^2$$

$$f_2(x) = (x_1 - 3)^2 + (x_2 - 3)^2$$

- (a) Determine o conjunto Pareto-ótimo do problema de minimização simultânea de f_1 e f_2 .
- (b) Avalie a adequação dos métodos P_λ e P_ϵ para a determinação dos pontos desse conjunto de Pareto.

11. Considere as funções:

$$f_1(x) = \max(x_1, x_2)$$

$$f_2(x) = (x_1 - 2)^2 + x_2^2$$

- (a) Determine o conjunto de soluções eficientes do problema de minimização simultânea de f_1 e f_2 .
- (b) Mostre que em parte do conjunto de soluções eficientes as condições de Kuhn-Tucker encontram-se satisfeitas. Explique por quê essas condições não se aplicam no restante do conjunto.

12. Considere as funções:

$$f_1(x) = \min(x_1, x_2)$$

$$f_2(x) = (x_1 - 2)^2 + x_2^2$$

- (a) Determine o conjunto de soluções eficientes do problema de minimização simultânea de f_1 e f_2 .
- (b) Mostre que em parte do conjunto de soluções eficientes as condições de Kuhn-Tucker encontram-se satisfeitas. Explique por quê essas condições não se aplicam no restante do conjunto.
- (c) Formule o problema (P_χ) que determina o conjunto de soluções eficientes. Discuta que tipo de algoritmo deve ser escolhido para a resolução numérica do problema assim formulado.

13. Sejam os vetores abaixo as avaliações de um conjunto de funções-objetivo em alguns pontos do espaço de parâmetros:

$$a = \begin{bmatrix} 10 \\ 15 \\ 23 \end{bmatrix} \quad b = \begin{bmatrix} 8 \\ 12 \\ 23 \end{bmatrix} \quad c = \begin{bmatrix} 11 \\ 9 \\ 21 \end{bmatrix}$$

$$d = \begin{bmatrix} 15 \\ 11 \\ 17 \end{bmatrix} \quad e = \begin{bmatrix} 9 \\ 11 \\ 22 \end{bmatrix} \quad f = \begin{bmatrix} 15 \\ 12 \\ 17 \end{bmatrix}$$

Tendo apenas estas informações:

- (a) Quais são os pontos que se pode afirmar, com certeza, que pertencem ao conjunto de soluções eficientes do problema de minimização das três funções-objetivo?
 - (b) Quais são os pontos que se pode afirmar, com certeza, que não pertencem ao conjunto de soluções eficientes desse problema?
14. Construa um problema de otimização multiobjetivo com duas variáveis de otimização no qual existam pontos que satisfazem às condições de Kuhn-Tucker sem pertencerem ao conjunto de soluções eficientes do problema. Mostre graficamente o conjunto das soluções eficientes do problema e também o conjunto das soluções de Kuhn-Tucker que não são eficientes.

Exercícios Computacionais

1. Considere as seguintes funções de duas variáveis:

$$f_1(x) = x'Q_1x$$

$$f_2(x) = x'Q_2x$$

$$Q_1(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1.2 \end{bmatrix} \quad Q_2(x) = \begin{bmatrix} 1 & 0 \\ 0 & 100 \end{bmatrix}$$

- (a) Faça a minimização dessas funções utilizando: (i) o algoritmo do gradiente, (ii) um algoritmo quasi-Newton, (iii) o algoritmo elipsoidal básico. Faça, para cada função e cada algoritmo, cinco testes, começando de pontos iniciais gerados aleatoriamente. Descreva as trajetórias das soluções sobre um mapa de curvas de nível das funções.
- (b) Acrescente agora as seguintes restrições às mesmas funções:

$$g_1(x) = (x - x_0)'Q_3(x - x_0) \leq 1$$

$$g_2(x) = B(x - x_0) \leq 0$$

$$x_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad Q_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad B = [1 \quad -1]$$

Faça, para cada função e cada algoritmo, cinco testes, começando de pontos iniciais gerados aleatoriamente. Descreva as trajetórias das soluções sobre uma figura que superponha um mapa de curvas de nível da função com o traçado da região factível. Analise os dados observados.

2. Implemente o algoritmo elipsoidal básico para minimização de funções no espaço \mathbb{R}^n . Com esse algoritmo, mais o algoritmo do gradiente e um algoritmo quasi-Newton anteriormente implementados, realize a minimização da função de duas variáveis:

$$f(x) = \max \{|x_1|, |x_2|\}$$

Faça, para cada método, cinco testes, começando de pontos iniciais gerados aleatoriamente. Descreva as trajetórias das soluções sobre um mapa de curvas de nível das funções. Analise os dados observados.

3. Faça a implementação computacional de um algoritmo da família de “direção de busca”, com tratamento de restrições. Construa problemas de otimização mono-objetivo de duas variáveis de otimização e com restrições, tais que:

- (a) o método implementado funcione;
- (b) o método implementado não convirja para a solução.

Mostre graficamente a seqüência de soluções nos dois casos.

4. Faça a implementação computacional de um algoritmo da família de “exclusão de semi-espaço”, com tratamento de restrições. Construa problemas de otimização mono-objetivo de duas variáveis de otimização e com restrições, tais que:

- (a) o método implementado funcione;
- (b) o método implementado não convirja para a solução.

Mostre graficamente a seqüência de soluções nos dois casos.

5. Construa um algoritmo que realize a operação de, dado um conjunto de vetores, determinar qual é o maior sub-conjunto não-dominado.

6. Faça a adaptação de um algoritmo de “direções de busca” e de um algoritmo de “exclusão de semi-espaços”, para a determinação de soluções

eficientes de problemas de otimização multiobjetivo, utilizando a estrutura P_ϵ ou P_χ . Considere as funções-objetivo:

$$f_1(x) = x_1 + x_2$$

$$f_2(x) = x_1^2$$

Considere também as restrições:

$$g_1(x) = x_1^2 + x_2^2 \leq 4$$

$$g_2(x) = x_2 \leq 1$$

$$g_3(x) = x_1 - x_2 \leq 2$$

Determine o conjunto de soluções eficientes desse problema, utilizando os algoritmos adaptados. Mostre graficamente as soluções obtidas.

7. Considere as funções-objetivo:

$$f_1(x) = x_1 + x_2$$

$$f_2(x) = x_1^2$$

e as restrições:

$$g_1(x) = x_1^2 + x_2^2 \leq 4$$

$$g_2(x) = x_2^2 \geq 1$$

$$g_3(x) = x_1 - x_2 \leq 2$$

Determine o conjunto de soluções eficientes desse problema, utilizando um algoritmo de direções de busca e um algoritmo de exclusão de semi-espacos. Mostre graficamente as soluções obtidas.

8. Construa situações, em problemas de otimização multiobjetivo com dois objetivos e duas variáveis de otimização, nas quais:
- (i) O método baseado em “direções de busca” convirja e o método baseado em “exclusão de semi-espacos” não convirja para pontos do conjunto de Pareto, ambos sendo inicializados no mesmo ponto.
 - (ii) O método baseado em “direções de busca” não convirja e o método baseado em “exclusão de semi-espacos” convirja para pontos do conjunto de Pareto, ambos sendo inicializados no mesmo ponto.

Mostre graficamente os resultados acima.

Bibliografia

- Ackley, D. H., Hinton, G. E. & Sejnowski, T. J. (1985). A learning algorithm for Boltzman Machines, *Cognitive Science* **9**: 147–169.
- Akgül, M. (1984). *Topics in Relaxation and Ellipsoidal Methods*, number 97 in *Research Notes in Mathematics*, Pitman Publishing Inc., London, UK.
- Belmont-Moreno, E. (2001). The role of mutation and population size in genetic algorithms applied to physics problems, *International Journal of Modern Physics C* **12**(9): 1345–1355.
- Bland, R. G., Goldfarb, D. & Todd, M. J. (1981). The ellipsoid method: a survey, *Operations Research* **29**(6): 1039–1091.
- Cao, Y.-Y., Lam, J. & Sun, Y.-X. (1998). Static output feedback stabilization: an ILMI approach, *Automatica* **34**(12): 1641–1645.
- Chankong, V. & Haimes, Y. Y. (1983). *Multiobjective Decision Making: Theory and Methodology*, North-Holland (Elsevier), New York.
- Chen, B. S., Cheng, Y. M. & Lee, C. H. (1995). A genetic approach to mixed $\mathcal{H}_\infty/\mathcal{H}_2$ optimal PID control, *IEEE Control Systems Magazine* **15**(5): 51–60.
- Chen, C. T. (1984). *Linear System Theory and Design*, Hartcourt Brace College Pub.
- Choi, D. H. & Oh, S. Y. (2000). A new mutation rule for evolutionary programming motivated from backpropagation learning, *IEEE Transactions on Evolutionary Computation* **4**(2): 188–190.

- Dias-Filho, W. (2003). *Algoritmos Cone-Elipsoidais para Geração de Soluções Eficientes: Construção e Aplicações*, PhD thesis, Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Dorigo, M., Maniezzo, V. & Colomi, A. (1996). Ant system: optimization by a colony of cooperating agents, *IEEE Trans. Sys. Man Cyb. - Part B* **26**(1): 29–41.
- Dziuban, S. T., Ecker, J. G. & Kupferschmid, M. (1985). Using deep cuts in an ellipsoidal algorithm for nonlinear programming, *Math. Program. Study* **25**: 93–107.
- Gembiki, F. W. & Haimes, Y. Y. (1975). Approach to performance and sensitivity multiobjective optimization: the goal attainment method, *IEEE Transactions on Automatic Control* **20**(6): 769–771.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley.
- J. C. Potts, T. D. G. & Yadav, S. B. (1994). The development and evaluation of an improved genetic algorithm based on migration and artificial selection, *IEEE Transactions on Systems, Man and Cybernetics* **24**(1): 73–86.
- Jain, A. & Zongker, D. (1997). Feature selection: evaluation, application, and small sample performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(2): 153–158.
- Johnson, J. M. & Ramat-Semii, Y. (1997). Genetic algorithms in engineering electromagnetics, *IEEE Antennas and Propagation Magazine* **39**(4): 7–25.
- K. F. Man, K. S. T. & Kwong, S. (1996). Genetic algorithms: concepts and applications, *IEEE Transactions on Industrial Electronics* **43**(5): 519–534.
- K. Rashid, J. A. R. & Freeman, E. M. (2000). A general approach for extracting sensitivity analysis from a neuro-fuzzy model, *IEEE Transactions on Magnetics* **36**(4): 1066–1070.

- Khargonekar, P. P. & Rotea, M. A. (1991). Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control: a convex optimization approach, *IEEE Transactions on Automatic Control* **36**(7): 824–837.
- Luenberger, D. G. (1984). *Linear and Nonlinear Programming*, Addison-Wesley.
- Meneguim, R. A. (1999). *Análise da sensibilidade de soluções em otimização através de elipsóides mínimos*, Master's thesis, Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Miranda, M. F. (2000). *Controle Multivariável na Presença de Incertezas*, PhD thesis, Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Naudts, B. & Kallel, L. (2000). A comparison of predictive measures of problem difficulty in evolutionary algorithms, *IEEE Transactions on Evolutionary Computation* **4**(1): 1–15.
- Popper, K. R. (1974). *A Lógica da Pesquisa Científica (trad.)*, Ed. Cultrix. (Logic der Forschung, 5a. ed. 1973; 1a. ed. 1934).
- Rudin, W. (1991). *Functional Analysis*, McGraw-Hill.
- Saldanha, R. R., Takahashi, R. H. C., Vasconcelos, J. A. & Ramirez, J. A. (1999). Adaptive deep-cut method in ellipsoidal optimization for electromagnetic design, *IEEE Transactions on Magnetics, Part I* **35**(3): 1746–1749.
- Sareni, B. & Krahenbuhl, L. (1998). Fitness sharing and niching methods revisited, *IEEE Transactions on Evolutionary Computation* **2**(3): 97–106.
- Schell, T. & Wegenkittl, S. (2001). Looking beyond selection probabilities: adaptation of the χ^2 measure for the performance analysis of selection methods in GA's, *Evolutionary Computation* **9**(2): 243–256.
- Scherer, C., Gahinet, P. & Chilali, M. (1997). Multiobjective output-feedback control via LMI optimization, *IEEE Transactions on Automatic Control* **42**(7): 896–911.

- Shah, S., Mitchell, J. E. & Kupferschmid, M. (2001). An ellipsoid algorithm for equality-constrained nonlinear programs, *Computers and Operations Research* **28**: 85–92.
- Shimomura, T. & Fujii, T. (2000). Multiobjective control design via successive over-bounding of quadratic terms, *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, pp. 2763–2768.
- Takahashi, R. H. C., Palhares, R. M., Dutra, D. A. & Gonçalves, L. P. S. (2001). Synthesis and characterization of Pareto-optimal solutions for the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control problem, *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, FL, USA, pp. 3997–4002.
- Takahashi, R. H. C., Peres, P. L. D. & Ferreira, P. A. V. (1997). Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$ guaranteed cost PID design, *IEEE Control Systems Magazine* **17**(5): 37–47.
- Takahashi, R. H. C., Ramirez, J. A., Vasconcelos, J. A. & Saldanha, R. R. (2001). Sensitivity analysis for optimization problems solved by stochastic methods, *IEEE Transactions on Magnetics* **37**(5): 3414–3417.
- Takahashi, R. H. C., Saldanha, R. R., Dias-Filho, W. & Ramirez, J. A. (2003). A new constrained ellipsoidal algorithm for nonlinear optimization with equality constraints, *IEEE Transactions on Magnetics* **39**(3): 1289–1292.
- Takahashi, R. H. C., Vasconcelos, J. A., Ramirez, J. A. & Krahenbuhl, L. (2003). A multiobjective methodology for evaluating genetic operators, *IEEE Transactions on Magnetics* **39**(3): 1321–1324.
- Tanomaru, J. (1995). Motivação, fundamentos e aplicações de algoritmos genéticos, *Anais do II Congr. Bras. de Redes Neurais*, Vol. 1, Curitiba, PR, Brasil.
- Vasconcelos, J. A., Ramirez, J. A., Takahashi, R. H. C. & Saldanha, R. R. (2001). Improvements in genetic algorithms, *IEEE Transactions on Magnetics* **37**(5): 3414–3417.
- Vidyasagar, M. (1993). *Nonlinear Systems Analysis*, Prentice-Hall.

- Viennet, R., Fonteix, C. & Marc, I. (1996). Multicriteria optimization using a genetic algorithm for determining a Pareto set, *Int. J. Sys. Sci.* **27**(2): 255–260.
- Wolpert, D. H. & Macready, W. G. (1997). No free lunch theorems for optimization, *IEEE Transactions on Evolutionary Computation* **1**(1): 67–82.
- Z. Michalewicz, K. Deb, M. S. & Stidsen, T. (2000). Test-case generator for nonlinear continuous parameter optimization techniques, *IEEE Transactions on Evolutionary Computation* **4**(3): 197–215.