



Controle Multicritério de Sistemas Dinâmicos:
Uma Abordagem Integrada

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de
Campinas, como parte dos requisitos exigidos para a obtenção do título de
Doutor em Engenharia Elétrica.

por

José Reginaldo Hughes Carvalho
Engenheiro Eletricista - UFBA
Mestre em Engenharia Elétrica - UNICAMP

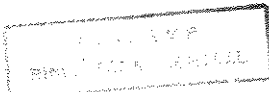
em 21 de março de 1997 perante a banca examinadora

- | | |
|--|---------------------------|
| Prof. Dr. Paulo Augusto Valente Ferreira | FEEC/UNICAMP (Orientador) |
| Prof. Dr. Pedro Luis Dias Peres | FEEC/UNICAMP |
| Prof. Dr. Rafael Santos Mendes | FEEC/UNICAMP |
| Prof. Dr. Jorge Roberto Brito de Souza | DEE/UFPa |
| Prof. Dr. Eugênio de Bona Castelan Neto | LCMI/EEL/UFSC |

9408549

Este exemplar corresponde à redação final da tese
defendida por José Reginaldo Hughes
Carvalho e aprovada pela Comissão
Jugadora em 21 03 97.

Paulo Augusto Valente Ferreira
Orientador



UNIDADE	BC
N.º CHAMADA	UNICAMP
	C253c
V.	Ex.
PREÇO B.	31202
PROG.	284197
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	25/07/97
N.º CPD	

CM-0009944 8-9

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

C253c Carvalho, José Reginaldo Hughes
Controle multicritério de sistemas dinâmicos: uma
abordagem integrada / José Reginaldo Hughes Carvalho.--
Campinas, SP: [s.n.], 1997.

Orientador: Paulo Augusto Valente Ferreira.
Tese (doutorado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Programação convexa. 2. Processo decisório por
critério múltiplo. 3. Controle automático. I. Ferreira, Paulo
Augusto Valente. II. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação. III.
Título.

À Andrea e ao Daniel.
Toda esta minha luta
foi por vocês.

“Confia no Senhor de todo teu coração, e não te estribes no teu próprio entendimento; reconhece-O em todos os teus caminhos, e Ele endireitará as tuas veredas. Não sejas sábio a teus próprios olhos; teme ao Senhor e aparta-te do mal.”
Provérbios 3:5-7

Agradecimentos

Devo a realização deste trabalho à contribuição, direta ou indireta, de muitos, entre professores, colegas e amigos. Em especial, gostaria de externar os meus sinceros agradecimentos a:

- Prof. Dr. Paulo Augusto Valente Ferreira, um grande exemplo de dedicação e competência, cuja orientação impecável e paciência foram primordiais para a viabilização deste trabalho;
- À minha Andrea, por ter estado sempre presente, dando o equilíbrio e apoio em todos os momentos;
- Aos meus pais, a quem eu muito amo e cujo exemplo será sempre essencial, sejam quais forem os meus caminhos;
- Aos meus irmãos, pelo torcida que faziam por mim;
- Aos amigos e colegas do Laboratório de Telemática, pelo ótimo ambiente de trabalho proporcionaram;
- Aos professores da FEEC que ministraram os cursos que auxiliaram no desenvolvimento deste trabalho;
- Aos que, através da FAPESP, financiaram este trabalho.
- E, acima de tudo, Àquele a quem sempre será devido todos os méritos, pois dEle provêm todas as coisas; Deus.

Sumário

O presente trabalho apresenta uma abordagem integrada para o problema de controle multicritério de sistemas dinâmicos. A partir de um procedimento genérico, aqui denominado de *Algoritmo Básico*, são tratados problemas de controle tanto no domínio do tempo quanto no domínio da frequência. Este procedimento decompõe o problema em dois níveis hierárquicos, sendo que o nível superior consiste sempre de um problema de programação matemática simples, enquanto que o nível inferior consiste de um problema min-max que trata as características particulares de cada classe de problemas. A metodologia proposta é baseada em programação convexa e explora o fato de que para um grande número de problemas de controle existentes na literatura existe um problema convexo equivalente. O trabalho inclui inúmeros exemplos numéricos que ilustram e atestam a viabilidade e a eficiência da abordagem proposta.

Abstract

In this work, an integrated approach for the multicriteria control problem of dynamic systems is presented. Using a generic procedure called *Basic Algorithm*, different control problems are treated, both in time and frequency domains. This procedure decomposes the problem into hierarchical levels, being the higher level represented by a simple mathematical programming problem. The lower level consists of a min-max problem that treats the characteristics of a given class of control problems. The methodology is based on convex programming and exploits the property that many different control problems admit an equivalent convex formulation. The work includes several examples that illustrate and attest the viability and efficiency of the approach proposed.

Conteúdo

Introdução Geral	1
PARTE I. Controle no Domínio do Tempo	5
1 Fundamentos Teóricos	6
1.1 Introdução	6
1.2 Convexidade de Conjuntos e Funções	6
1.3 Equivalência Primal-Dual	9
1.4 Solução Através de Aproximação Tangencial	12
1.5 Existência do Gradiente: Teorema de Daskin	13
1.6 Programação Multicritério: Aspectos Básicos	15
1.7 Controle Ótimo Escalar	17
1.7.1 Equação de Hamilton-Jacobi-Bellman	18
1.7.2 Problema Linear-Quadrático	19
1.7.3 Princípio do Mínimo de Pontryagin	20
1.8 Problema Linear-Quadrático Gaussiano	21
1.8.1 Solução do LQG via Problema Determinístico	22
1.8.2 Filtro de Kalman-Bucy	23
1.8.3 Princípio da Separação	24
1.9 Conclusão	25
2 Controle Multicritério Determinístico	26
2.1 Introdução	26
2.2 Formulação do Problema	26
2.3 Algoritmo Básico	28
2.4 Convergência e Implementação Realimentada	33
2.5 Problema Linear-Quadrático Multicritério	34
2.6 Critérios Não-Quadráticos: Um Exemplo	42
2.7 Conclusão	47
3 Controle Multicritério Estocástico	48
3.1 Introdução	48
3.2 Formulação do Problema	48

3.3	Solução do Nível de Análise	49
3.4	Um Exemplo Ilustrativo	50
3.5	Conclusão	52
4	Controle Multicritério Hierárquico	53
4.1	Introdução	53
4.2	Formulação do Problema	53
4.3	Estratégia de Decomposição	55
4.4	Exemplo	58
4.5	Conclusão	61
	PARTE II. Controle no Domínio da Frequência	64
5	Fundamentos Teóricos	65
5.1	Introdução	65
5.2	Arquitetura do Sistema	65
5.3	Exemplo Ilustrativo	67
5.4	Especificações de Projeto	69
5.5	Otimização Multicritério	70
5.6	Paradigma do Controlador Ótimo	71
5.7	Normas de Sinais	72
5.8	Normas de Sistemas	74
5.9	Especificações de Projeto	75
5.9.1	Realizabilidade	77
5.9.2	Estabilidade	78
5.9.3	Desempenho	78
5.9.4	Comentários	79
5.10	Formulação do Problema	80
5.11	Parametrização de Youla	81
5.11.1	Resultados Preliminares	82
5.12	Estabilidade	83
5.13	Parametrização dos Controladores Estabilizantes	84
5.14	Parametrização de Youla em Espaço de Estados	85
5.15	Conclusão	87
6	Projeto de Controladores Dinâmicos	88
6.1	Introdução	88
6.2	Formulação do Problema via Parametrização de Youla	88
6.3	Aproximação de Ritz	89
6.4	Método de Planos de Corte	89
6.4.1	Método de Planos de Corte: Problema Irrestrito	91
6.4.2	Método de Planos de Corte: Problema Restrito	91
6.5	Exemplo Ilustrativo	92

6.6	Subgradientes na Aproximação de Ritz	94
6.7	Problema de Projeto Multicritério	94
6.7.1	Formulação do Problema	94
6.7.2	Estrutura Multinível - Decisão e Análise	95
6.7.3	Algoritmo Básico	96
6.8	Solução do Problema Min-max	97
6.9	Solução do Problema Multicritério	99
6.9.1	Nível de Análise	99
6.9.2	Nível de Decisão	100
6.9.3	Exemplo Ilustrativo	101
6.10	Método das Desigualdades	102
6.11	Conclusão	105
7	Projeto de Controladores l_1	107
7.1	Introdução	107
7.2	Transformada λ	107
7.3	Formulação do Problema	108
7.4	Condições de Interpolação	109
7.5	Solução do Problema Mono-bloco	111
7.5.1	Aspectos Computacionais da Condição de Interpolação de Zero	111
7.6	Formulação Através de Programação Linear	114
7.6.1	Problema de Dimensão Infinita	114
7.6.2	Problema de Dimensão Finita	117
7.7	Solução do Problema Multi-Bloco	118
7.8	Comentário sobre a Utilização do Algoritmo Básico	120
7.9	Solução Via Programação Alvo	120
7.10	Exemplos Numéricos	121
7.10.1	Problema do Caça X29	121
7.10.2	Problema da Haste Flexível	125
7.11	Aspectos Computacionais	131
7.12	Conclusão	132
	Conclusão Geral	133
	Bibliografia	135

Lista de Figuras

1.1	Minimização simultânea de duas funções.	15
1.2	Problema no espaço dos critérios.	16
2.1	Pólo do sistema em função de α	39
2.2	(a): 1ª Iteração; (b): 2ª Iteração; (c): 3ª Iteração.	40
2.3	Convergência de Θ - $p = 1$, $p = 2$ e $p = \infty$	41
2.4	Solução no espaço dos critérios - $v_1 = v_2 = 1$	45
2.5	Trajetórias ótimas do estado e do controle - $v_1 = v_2 = 1$	45
2.6	Solução no espaço dos critérios - $v_1 = 5$, $v_2 = 1$	46
2.7	Trajetórias ótimas do estado e do controle - $v_1 = 5$, $v_2 = 1$	46
3.1	Evolução de $\Theta(y)$ a cada iteração.	52
4.1	Geometria de $\mathcal{V}(y)$	60
4.2	Quatro iterações no espaço dos critérios.	60
4.3	Evolução de $\mathcal{V}(y)$ e $\Theta(y)$	62
4.4	Evolução do controle e dos estados de $t = 0$ a $t = 1$	62
4.5	Diagrama do método.	63
5.1	Arquitetura do Sistema.	65
5.2	Arquitetura em malha fechada	67
5.3	(a) Diagrama de blocos, (b) Diagrama considerando perturbações.	67
5.4	Sistema de controle segundo a arquitetura adotada.	68
5.5	Arquitetura incluindo a matriz Q	81
6.1	Aproximação de $\phi(H)$ por planos de corte.	90
6.2	Representação clássica do sistema 1-DOF.	92
6.3	Resposta ao degrau de (a) H_{13} e (b) H_{23} para norma infinito.	102
6.4	Resposta ao degrau de (a) H_{13} e (b) H_{23} para norma Euclideana.	103
6.5	Curva de <i>trade-off</i>	104
6.6	Ilustração do método - caso bi-critério.	105
7.1	Problema do caça X29 dentro da arquitetura padrão.	122
7.2	Caça X29: Resposta ao degrau de H via PA sem considerar overshoot	124

7.3	Caça X29: Resposta ao degrau de H via PA considerando overshoot	126
7.4	Haste flexível.	126
7.5	Problema da haste flexível dentro da arquitetura padrão.	127
7.6	Haste flexível: Resposta ao degrau de H sem considerar over/undershoot.	127
7.7	Haste flexível: Resposta ao degrau de H sem considerar over/undershoot (planta perturbada).	128
7.8	Haste flexível: Resposta ao degrau de H via PA com ordenação léxica dos desvios. .	129
7.9	Haste flexível: Resposta ao degrau de H via PA com desvios compostos em um único critério.	130

Lista de Tabelas

2.1	Comportamento do Algoritmo Básico para $\alpha = 0.5$	38
2.2	Valores Ótimos para $p = 1$, $p = 2$, $p = \infty$	41
2.3	Evolução do Algoritmo Básico - $p = 2$	42
2.4	Solução para $v_1 = 1$, $v_2 = 1$	44
2.5	Solução para $v_1 = 5$, $v_2 = 1$	47
3.1	Helicóptero CH-47: Solução do exemplo a cada iteração.	51
4.1	Controle Hierárquico: Solução do exemplo a cada iteração.	61
5.1	Exemplos de especificações via conjuntos.	76
6.1	Solução do Algoritmo Básico para a norma infinito.	101
6.2	Solução do Algoritmo Básico para a norma Euclideana.	103
7.1	Caça X29: Resultados encontrados com o método DA.	122
7.2	Caça X29: Resultados encontrados via PA sem considerar overshoot.	124
7.3	Caça X29: Resultados encontrados via PA considerando overshoot.	125
7.4	Haste flexível: Resultados via PA com ordenação léxica dos desvios.	129
7.5	Haste flexível: Resultados via PA com desvios compostos em um único critério.	130

Introdução Geral

A busca de uma solução de compromisso entre diferentes critérios de desempenho associados a um mesmo sistema dinâmico é uma necessidade que surge naturalmente em diversos contextos, tais como, uso de recursos hídricos (Haimes et al., 1975) ou projeto de controladores (Kresselmeier e Steinhauser, 1983). O problema de controle multicritério, que consiste na otimização simultânea de vários critérios de desempenho foi primeiramente introduzido por Zadeh (1963). Desde então, vários autores vêm propondo soluções para diferentes classes de problemas de controle multicritério, como pode ser visto no decorrer desta introdução. O objetivo deste trabalho é o de apresentar uma abordagem integrada para o problema de controle multicritério de sistemas dinâmicos formulados tanto no domínio do tempo quanto no domínio da frequência. Observa-se na literatura (Francis, 1987; Boyd e Barrat, 1991; Dahleh e Diaz-Bobillo, 1995; Dorato et al., 1995) que importantes classes de problemas de controle possuem uma formulação convexa. O procedimento aqui proposto, denominado *Algoritmo Básico*, explora este fato e aplica estratégias de programação convexa na solução de problemas de controle multicritério.

A primeira parte do trabalho trata o problema de controle multicritério de sistemas dinâmicos no domínio do tempo. Entende-se por controle no domínio do tempo a busca da solução do problema de controle no espaço das funções contínuas por partes da variável independente *tempo*. Neste contexto, Salukvadze (1979) desenvolveu um método que utiliza o valor ótimo escalar resultante da otimização individual de cada critério para introduzir o conceito de *ponto ideal* no espaço dos critérios. Em seguida, a distância entre este ponto e o conjunto de valores alcançáveis através de controles factíveis é minimizada. Na prática, qualquer norma l_p pode ser usada para medir a distância e as escolhas mais comuns são $p = 1$ (combinação linear), $p = 2$ (norma Euclideana) e $p = \infty$ (norma infinito). Medanić e Andjelić (1971,1972) resolvem o problema de controle min-max de sistemas lineares-quadráticos usando Teoria de Jogos, mostrando que a convexidade garante a existência de ponto de sela para esta classe de problemas. Li (1990a, 1990b) propõe a solução do problema min-max usando uma abordagem baseada na equalização dos critérios de desempenho.

O problema linear-quadrático gaussiano multicritério (LQMG) tem recebido atenção especial na literatura de controle multicritério. Em Toivonen (1984), uma ordenação hierárquica dos critérios quadráticos é usada para relacionar a minimização das variâncias dos estados com o critério de desempenho padrão do regulador. O método do ponto ideal de Salukvadze foi estendido para o caso discreto do problema LQMG por Koussoulas e Leondes (1984) e um procedimento para calcular uma solução *eficiente* ou *não inferior* (Zadeh, 1963) para o problema LQMG foi desenvolvido por Toivonen e Mäkilä (1989). Mais recentemente, atenção tem sido direcionada para a solução do problema LQMG em espaços de Hilbert, através da introdução de vetores de critérios $\mathcal{H}_2$ e $\mathcal{H}_\infty$.

(Khargonekar e Rotea, 1991).

Uma das consequências do emprego das idéias de decomposição (Dantzig e Wolfe, 1960) e controle multinível (Mesarovic *et al*, 1970) tem sido o acentuado desenvolvimento de abordagens hierárquicas para sistemas de grande porte, confirmado pela grande variedade de modelos e procedimentos de otimização obtidos nas últimas décadas para esta classe de sistemas. Tradicionalmente restrita à consideração de apenas um critério por unidade de decisão individual, a teoria de sistemas hierárquicos tem sido progressivamente integrada com a área de otimização multicritério. Atualmente, um grande número de métodos que integram abordagens multicritério e hierárquicas encontram-se disponíveis. A otimização multicritério de sistemas dinâmicos descentralizados constitui um campo que merece particular atenção. Problemas que envolvam sistemas lineares descentralizados com critérios quadráticos são de especial interesse e suficientemente gerais para descreverem de forma adequada muitas situações práticas. O controle de sistemas descentralizados com estrutura linear quadrática múltipla, isto é, sistemas onde o critério global é uma função não-linear de índices de desempenho quadráticos, foi introduzido e resolvido por Li (1993) através de uma metodologia hierárquica multinível, onde a estratégia de separação empregada é a mesma apresentada em Li e Haimes (1990).

Apesar do grande número de contribuições fundamentais para o domínio do controle multicritério, a questão básica de resolver eficientemente o problema permanece aberta. Além disso, embora a convexidade desempenhe um papel essencial para a obtenção de muitas propriedades relacionadas com o projeto de controle de sistemas lineares (Boyd e Barrat, 1991), os métodos disponíveis não exploram tal propriedade em sua completa extensão.

A abordagem proposta neste trabalho baseia-se em estratégias de projeção e relaxação para compor a solução do problema de controle multicritério de acordo com uma estrutura multinível (Carvalho e Ferreira, 1992, 1993; Carvalho, 1993). No *nível inferior* ou de *análise*, resolve-se um subproblema *min-max* similar ao introduzido por Medanić e Andjelić (1971), porém parametrizado por informações vindas do *nível superior* ou de *decisão*, onde uma versão relaxada e estática do problema é considerada. A solução do nível de análise fornece informações para que seja feita uma melhor aproximação da região de soluções eficientes do problema. O *decisor*, dispondo destas informações, encontra uma nova solução que será avaliada pelo nível de análise. O procedimento prossegue até que o decisor disponha de informações suficientes para poder encontrar uma solução que seja considerada satisfatória pelo nível de análise. Em particular, demonstra-se que se os funcionais são critérios quadráticos, a solução do problema de análise consiste, do ponto de vista da teoria de controle, na solução iterativa de equações do tipo Riccati.

O fato do problema ser formulado no espaço dos critérios, apesar de apresentar certas vantagens metodológicas, não preserva a convexidade exibida pelo problema no espaço das variáveis de controle. O espaço dos critérios é então limitado a uma região de interesse e, através de propriedades fundamentais de conjuntos convexos, usa-se uma técnica de relaxação para fazer uma aproximação gradual da região factível do problema. Este tipo de abordagem evita uma limitação habitual da formulação no espaço dos critérios, que é a inviabilidade de se fazer um mapeamento completo das soluções do espaço das variáveis de controle.

Como contribuições à classe de problemas de controle multicritério no domínio do tempo, este trabalho aplica a metodologia ao problema de controle linear quadrático gaussiano e ao problema de controle de sistemas hierárquicos. Os Capítulos que compõem a primeira parte do trabalho estão

assim organizados.

Capítulo 1 - Neste capítulo será feita uma breve revisão sobre os fundamentos teóricos que darão suporte ao trabalho proposto, tais como conjuntos convexos, programação convexa escalar e controle multicritério.

Capítulo 2 - Apresentará a solução do problema de controle multicritério determinístico de sistemas dinâmicos, introduzindo a abordagem que servirá de base para todo o trabalho.

Capítulo 3 - Neste capítulo a abordagem será usada para resolver o problema de controle linear quadrático multicritério gaussiano, dentro da classe de sistemas de controle estocásticos.

Capítulo 4 - Este capítulo tratará da solução de problemas de controle hierárquicos. Neste caso, o nível inferior da estrutura proposta (nível de análise) será formada por uma outra estrutura bi-nível, tratada através de técnicas de decomposição-coordenação.

A segunda parte do trabalho aborda o problema de projeto multicritério de controladores no domínio da frequência. Publicações recentes (Boyd e Barrat, 1991; Diaz-Bobillo e Dahleh, 1993) utilizam uma arquitetura padrão para representar sistemas onde a matriz de transferência resultante descreve o comportamento da dinâmica do sistema entre *entradas exógenas* (ruídos e qualquer outro sinal que não faça parte das entradas de controle) e *saídas reguladas* da planta.

Esta arquitetura tem-se mostrado de grande funcionalidade em se tratando de problemas de projeto de controladores, uma vez que a caracterização do conjunto de controladores estabilizantes pode ser obtida através da parametrização de Youla (Vidyasagar, 1991). Com isso, a relação entre a matriz de transferência em malha fechada e a variável de busca muda de uma dependência não-linear, quando a busca é feita sobre a matriz de transferência do controlador, para uma relação afim, quando a busca é feita sobre a matriz de parametrização. Portanto, as relações convexas que geralmente existem entre a matriz de transferência em malha fechada do sistema e as especificações de projeto são preservadas e o problema resultante é um problema de otimização convexo.

Uma abordagem que tem recebido muita atenção consiste em minimizar a norma l_1 da matriz de transferência em malha fechada do sistema, o que implica em minimizar o máximo ganho pico-a-pico entre as entradas exógenas e as saídas reguladas. Este problema possui características que permitem transformá-lo em um problema de programação linear (Diaz-Bobillo e Dahleh, 1995).

Os Capítulos que constituem esta parte do trabalho estão estruturados como segue.

Capítulo 5 - Define o paradigma que será usado para a representação de sistemas de controle, além dos fundamentos teóricos para o entendimento do trabalho a ser realizado nesta segunda parte, tais como, fatoração coprima e parametrização de Youla.

Capítulo 6 - Neste Capítulo, o Algoritmo Básico será aplicado ao problema de projeto de controladores. Para tanto, será adotada a parametrização de Youla para a representação em malha fechada do sistema de controle, obtendo-se com isso um problema equivalente convexo, porém de dimensão infinita. O aspecto da dimensão será tratado através da aproximação de Ritz. O Algoritmo Básico é implementado através de planos de cortes.

Capítulo 7 - Diferentemente do Capítulo 6, que aborda sistemas contínuos, neste Capítulo são considerados sistemas discretos no tempo, enfocados através de metodologias baseadas na norma l_1 . O problema original de dimensão infinita é manipulado matematicamente até transformar-se em um problema de programação linear de dimensão finita, geralmente aproximado, sendo que para alguns casos obtém-se um problema de dimensão finita equivalente. A consideração de múltiplos critérios gera um problema de programação linear multicritério, que pode ser resolvido eficientemente por inúmeros métodos. Em particular, propõe-se uma técnica baseada em Programação Alvo, ilustrada através de exemplos.

Parte I

**Controle no Domínio
do Tempo**

Capítulo 1

Fundamentos Teóricos

1.1 Introdução

O objetivo deste Capítulo é fornecer uma rápida revisão de temas considerados relevantes para um bom entendimento da proposta apresentada neste trabalho. Inicialmente, serão revistos conceitos e propriedades importantes para a caracterização de um problema convexo. Dentro ainda do escopo da programação convexa, serão apresentados resultados referentes à equivalência entre os problemas primal e dual a partir da existência do ponto de sela e aspectos de diferenciabilidade e solução numérica do problema dual. Alguns resultados referentes a problemas multicritérios convexos, como caracterizações do conjunto de soluções *eficientes*, são discutidos. Serão ainda revistos tópicos referentes à teoria de controle ótimo. Neste sentido, serão apresentadas a formulação do problema e algumas abordagens conhecidas para a sua solução, como a equação de Hamilton-Jacobi-Bellman e as condições de otimalidade estabelecidas pelo Princípio do Mínimo de Pontryagin. Finalmente será apresentada a solução do problema linear-quadrático gaussiano.

1.2 Convexidade de Conjuntos e Funções

Em virtude do trabalho estar baseado em técnicas de otimização convexa, os conceitos e propriedades referentes a conjuntos e funções convexas serão de fundamental importância. Para uma discussão completa sobre o assunto, veja por exemplo Lasdon (1970) ou Luenberger (1984). Considera-se que os subconjuntos de interesse estão definidos em espaços vetoriais lineares sobre o campo dos reais.

Definição 1.1 - (Conjuntos Convexos) Diz-se que um conjunto Ω de vetores é convexo se,

$$\forall x^1, x^2 \in \Omega \rightarrow wx^1 + (1-w)x^2 \in \Omega, \quad \forall 0 \leq w \leq 1. \quad (1.1)$$

Dados quaisquer dois pontos pertencentes ao conjunto, este será convexo se o segmento de reta que os une também pertencer ao conjunto. Algumas propriedades básicas de conjuntos convexos, que podem ser demonstradas diretamente da Definição 1.1, são apresentadas a seguir.

Proposição 1.1 *Sejam C^i , $i = 1, 2, \dots, r$ conjuntos convexos. Então*

- a) wC^i , $\forall i = 1, 2, \dots, r$ também é um conjunto convexo, $\forall w \in \mathbb{R}$.
- b) $\sum_{i=1}^r C^i \triangleq \{x : x = \sum_{i=1}^r x^i, x^i \in C^i\}$ é um conjunto convexo.
- c) $\bigcap_{i=1}^r C^i$ é um conjunto convexo.
- d) $\forall x^1, x^2, \dots, x^n \in C^i \rightarrow \left(\sum_{j=1}^n w_j x^j \right) \in C^i$, $\forall i = 1, 2, \dots, r$, $\forall w \in W$, onde

$$W \triangleq \{w \in \mathbb{R}^n : w \geq 0, \sum_{i=1}^n w_i = 1\}. \quad (1.2)$$

A combinação convexa de vetores pertencentes a um conjunto convexo também pertence ao conjunto.

Definição 1.2 - (Funções Convexas) *Diz-se que uma função $f(\cdot) : \Omega \rightarrow \mathbb{R}$ é convexa sobre um conjunto convexo Ω se para quaisquer $x^1, x^2 \in \Omega$*

$$f(wx^1 + (1-w)x^2) \leq wf(x^1) + (1-w)f(x^2), \quad \forall 0 \leq w \leq 1. \quad (1.3)$$

Uma função é estritamente convexa se valer a desigualdade estrita em (1.3) para $0 < w < 1$

Se uma função é convexa, então a reta que liga quaisquer dois valores da função sempre fica acima, podendo tocar, a curva dos valores da função. A seguir serão apresentadas algumas propriedades de funções convexas que, como no caso de conjuntos convexos, poderão ser demonstradas usando-se a definição.

Proposição 1.2 *Seja $f(\cdot)$ um vetor cujas componentes $f_j(\cdot) : \Omega^j \rightarrow \mathbb{R}$ são funções convexas sobre conjuntos convexos Ω^j , $j = 1, 2, \dots, m$. Então*

- a) $wf_j(\cdot)$ é uma função convexa sobre Ω^j , $\forall w > 0$.
- b) $\sum_{i=1}^m w_i f_i(\cdot)$ é uma função convexa sobre $\bigcap_{j=1}^m \Omega^j$, $\forall w \in W$.
- c) $\forall x^1, x^2, \dots, x^n \in \Omega^j$ e $f_j(\cdot)$, $j = 1, 2, \dots, m$

$$f_j \left(\sum_{i=1}^n w_i x^i \right) \leq \sum_{i=1}^n w_i f_j(x^i), \quad \forall w \in W$$

Seja $\mathcal{C}^1$ o conjunto de funções que possuem até primeira derivada contínua e $\mathcal{C}^2$ o conjunto de funções que possuem até segunda derivada contínua.

Proposição 1.3 *Seja $f(\cdot) \in \mathcal{C}^1$ uma função convexa definida sobre Ω convexo. Então*

$$f(x) \geq f(y) + \nabla f(y)'(x - y), \quad \forall x, y \in \Omega \quad (1.4)$$

Se uma função $f(\cdot)$ é convexa, então o seu valor supera o da sua *aproximação linear* no ponto y , quaisquer que sejam os pontos $x, y \in \Omega$ considerados.

Proposição 1.4 *Seja $f(\cdot) \in \mathcal{C}^2$ uma função convexa definida sobre Ω convexo. Então a matriz*

$$F(x) \triangleq \nabla^2 f(x) \quad (1.5)$$

é semi-definida positiva sobre Ω . Se $f(\cdot)$ for estritamente convexa sobre Ω , então $F(x)$ é definida positiva sobre Ω .

A seguir, algumas relações importantes entre conjuntos e funções convexas.

Proposição 1.5 *Seja $f(\cdot)$ uma função convexa sobre Ω convexo. Então o conjunto*

$$\mathcal{C}_\alpha \triangleq \{x \in \Omega : f(x) \leq \alpha, \quad \alpha \in \mathbb{R}\} \quad (1.6)$$

é um conjunto convexo.

Proposição 1.6 - (Epígrafo) *O epígrafo de uma função $f(\cdot)$, definido como*

$$[\Omega, f(\cdot)] \triangleq \{(x, \alpha) : f(x) \leq \alpha, \quad x \in \Omega, \alpha \in \mathbb{R}\} \quad (1.7)$$

é um conjunto convexo se e somente se $f(\cdot)$ é uma função convexa.

Sejam $f_i(\cdot) : \Omega \rightarrow \mathbb{R}, i = 1, 2, \dots, m$ funções convexas sobre um conjunto convexo Ω . Considerando $f(\cdot)$ um vetor cujas componentes são as funções $f_1(\cdot), f_2(\cdot), \dots, f_m(\cdot)$, é fácil mostrar que o conjunto $\mathcal{G}$ definido como

$$\mathcal{G} \triangleq \{y \in \mathbb{R}^m : f(x) \leq y, \text{ para algum } x \in \Omega\} \quad (1.8)$$

é convexo, uma vez que $\mathcal{G}$ é a interseção dos epígrafos de $f_j(\cdot), j = 1, 2, \dots, m$.

O resultado a seguir é de grande importância ao longo de todo o trabalho. Trata-se de uma propriedade fundamental de conjuntos convexas.

Teorema 1.1 *Seja $\mathcal{G}$ definido como em (1.8). Então $y \in \mathcal{G}$ se e somente se y satisfaz o sistema linear com infinitas restrições*

$$\min_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0, \quad \forall w \in W \quad (1.9)$$

Prova - (Necessidade) Seja um vetor $x^0 \in \Omega$ tal que $f(x^0) \leq y$ para algum $y \in \mathbb{R}^m$ (ou seja, $y \in \mathcal{G}$). Como $w \in W$ é sempre não-negativo, então

$$\langle w, f(x^0) - y \rangle \leq 0, \quad \forall w \in W \quad (1.10)$$

Pode-se então substituir o lado esquerdo de (1.10) pelo menor valor da função $\langle w, f(x) - y \rangle$ para $x \in \Omega$, mantendo o sinal da desigualdade. Então

$$\min_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0, \quad \forall w \in W \quad (1.11)$$

(Suficiência) Note que a desigualdade (1.9) vale para todo $w \in W$, inclusive para o que produz o maior valor da função, ou seja,

$$\max_{w \in W} \min_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0 \quad (1.12)$$

Como a restrição $\sum_{i=1}^n w_i = 1$ presente em W é para atender a uma normalização, a desigualdade continuará valendo se a maximização for feita apenas sobre $w \geq 0$. Então

$$\max_{w \geq 0} \min_{x \in \Omega} \langle w, f(x) - y \rangle \leq 0 \quad (1.13)$$

e uma solução é $w = 0$. Agora, note que (1.13) é o *problema dual* de

$$\begin{aligned} \min_{x \in \Omega} \quad & 0'x \\ \text{s.a.} \quad & f(x) - y \leq 0 \end{aligned} \quad (1.14)$$

É possível mostrar que se o problema *dual* for limitado, como no presente caso, então o problema primal será factível (Lasdon, 1970). Consequentemente, a solução de (1.14) é factível, ou seja existe $x \in \Omega$ tal que $f(x) \leq y$ e, portanto, $y \in \mathcal{G}$. $\square$

1.3 Equivalência Primal-Dual

Agora que os conceitos de conjuntos e funções convexas necessários para definir um problema convexo são conhecidos, pode-se formalizar a solução desta classe de problemas.

Seja o problema de otimização não-linear

$$\begin{aligned} \min_{x \in \Omega} \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0 \end{aligned} \quad (1.15)$$

onde $f(\cdot)$ e $g(\cdot) = [g_1(\cdot) \dots g_q(\cdot)]'$ são funções convexas sobre Ω convexo. Problemas convexos possuem, como será visto no decorrer desta seção, uma série de propriedades que fazem com que sua abordagem seja relativamente simples. Estes resultados são amplamente conhecidos na literatura, podendo suas demonstrações ser encontradas, por exemplo, em Lasdon (1970).

A primeira característica apresentada diz respeito à relação entre soluções globais e soluções locais em problemas convexos.

Teorema 1.2 *Seja $\alpha = \min_{x \in \Omega} f(x)$. Então*

- $\Omega^* \triangleq \{x \in \Omega : f(x) = \alpha\}$ é um conjunto convexo.
- Se x^* é um mínimo local de $f(\cdot)$, então $f(x^*) = \alpha$ e portanto x^* é também um mínimo global de $f(\cdot)$.

Se $x^* \in \Omega$ resolve (1.15), então existem multiplicadores $w^* \in \mathbb{R}^q$ tais que (x^*, w^*) satisfazem às condições necessárias de otimalidade de Kuhn-Tucker:

- i) $w_i^* \geq 0$, $i = 1, 2, \dots, q$
- ii) $w_i^* g_i(x^*) = 0$, $i = 1, 2, \dots, q$
- iii) $g_i(x^*) \leq 0$, $i = 1, 2, \dots, q$
- iv) $\nabla l(x, w)|_{\substack{x=x^* \\ w=w^*}} = 0$

onde $l(\cdot)$ é a função Lagrangeano de (1.15), definida como

$$l(x, w) = f(x) + \langle w, g(x) \rangle \quad (1.16)$$

sendo w o vetor formado pelos multiplicadores de Lagrange w_i , $i = 1, 2, \dots, q$.

Apesar das condições de Kuhn-Tucker serem apenas condições necessárias, a convexidade inerente ao problema tratado garante a sua suficiência, o que pode ser verificado a partir da definição de ponto de sela.

Definição 1.3 Diz-se que um par (x^0, w^0) , $x^0 \in \Omega$, $w^0 \geq 0$ é um ponto de sela da função Lagrangeano (1.16), se satisfaz

$$\begin{aligned} l(x^0, w^0) &\leq l(x, w^0), & \forall x \in \Omega \\ l(x^0, w^0) &\geq l(x^0, w), & \forall w \geq 0 \end{aligned}$$

Um par (x^0, w^0) é um ponto de sela se x^0 minimiza o Lagrangeano sobre Ω e w^0 maximiza o Lagrangeano sobre $w \geq 0$.

Teorema 1.3 Seja o par (x^0, w^0) , $x^0 \in \Omega$, $w^0 \geq 0$. Então (x^0, w^0) é um ponto de sela de $l(x, w)$ se e somente se:

- a) $g_i(x^0) \leq 0$, $i = 1, 2, \dots, q$
- b) $w_i^0 g_i(x^0) = 0$, $i = 1, 2, \dots, q$
- c) x^0 minimiza $l(x, w^0)$ sobre $\mathbb{R}^n$

Perceba que as condições (a) e (b) são respectivamente as condições de factibilidade (iii) e de folga complementar (ii) de Kuhn-Tucker supondo-se que Ω é todo o $\mathbb{R}^n$. A diferença está na condição de estacionariedade (iv) que aqui é representada pela condição (c). Note porém que, se (c) for satisfeita então (iv) também será.

Este Teorema é de grande importância, pois uma condição suficiente para que um problema satisfaça as condições (a)-(c), possuindo um ponto de sela, é que seja convexo. O Teorema a seguir fundamenta a afirmação de que a convexidade de um problema garante a suficiência das condições de Kuhn-Tucker.

Teorema 1.4 *Suponha que $f(\cdot)$, $g_i(\cdot)$, $i = 1, 2, \dots, q$ são funções convexas sobre Ω convexo e que exista $x \in \Omega$ tal que $g(x) < 0$. Então x^0 resolve (1.15) se e somente se existem multiplicadores de Lagrange $w_i^0 \geq 0$, $i = 1, 2, \dots, q$, tais que o par (x^0, w^0) satisfaça as condições de Kuhn-Tucker.*

A condição $g(x) < 0$, chamada de *qualificação de restrição*, indica que o conjunto de restrições deve ter um ponto interior. Como os problemas convexas sempre possuem um ponto de sela, a prova deste teorema torna-se imediata, através da similaridade entre as condições (i)-(iv) e as condições (a)-(c).

Uma consequência imediata do Teorema é que se (x^0, w^0) é um ponto de sela de $l(x, w)$, então, pelas condições (a)-(c), x^0 é a solução de (1.15).

Considere agora a função $\phi(\cdot)$ definida como

$$\phi(w) \triangleq \min_{x \in \Omega} l(x, w) \quad (1.17)$$

Chama-se de *problema dual* de (1.15) o problema

$$\max_{w \in \mathcal{W}} \phi(w) \quad (1.18)$$

onde $\phi(\cdot)$, denominada *função dual*, é côncava sobre o domínio $\mathcal{W}$, dado por

$$\mathcal{W} \triangleq \{w \geq 0 : \exists \min_{x \in \Omega} l(x, w)\} \quad (1.19)$$

Os problemas dual e primal estão relacionados entre si por diversos aspectos, como será visto no decorrer desta seção. A primeira relação apresentada diz respeito à factibilidade.

Proposição 1.7 *Sejam os problema primal e dual definidos como em (1.15) e (1.18). Se a solução do primal for limitada (o valor da função objetivo na solução é finito), então o dual será factível; se o primal for factível, a solução do dual será limitada.*

Verificou-se anteriormente que problemas convexas possuem uma característica bastante desejável que é a existência de um ponto de sela (x^0, w^0) e que, em consequência, x^0 resolve o problema primal. O objetivo agora será o de verificar que no ponto de sela também está a solução do problema dual e que os valores das funções $f(\cdot)$ e $\phi(\cdot)$ nas respectivas soluções ótimas são idênticos.

Teorema 1.5 *Se existem x^0, w^0 soluções factíveis primal e dual, tais que $\phi(w^0) = f(x^0)$ então x^0 resolve o primal e w^0 resolve o dual.*

Uma condição suficiente para que tenham sido encontradas as soluções dos problemas primal e dual é que as funções que estão sendo otimizadas possuam o mesmo valor para as respectivas soluções. O próximo Teorema mostra onde é satisfeita a condição apresentada pelo Teorema 1.5.

Teorema 1.6 *O par (x^0, w^0) é um ponto de sela de $l(x, w)$ se e somente se x^0 e w^0 são factíveis e $\phi(w^0) = f(x^0)$.*

Observe a importância do papel desempenhado pelo ponto de sela (x^0, w^0) na solução de um problema convexo. Além de ter sua existência garantida nesta classe de problemas, os Teoremas 1.5 e 1.6 mostram que no ponto de sela encontra-se tanto a solução do problema primal quanto a do problema dual. A consequência deste fato é que, na busca da solução de um problema convexo, a abordagem pelo problema primal e a abordagem pelo problema dual são equivalentes.

1.4 Solução Através de Aproximação Tangencial

Na seção anterior foi formulado o problema dual de um problema convexo (primal) e verificada a equivalência da solução do dual em relação ao primal a partir da existência de um ponto de sela. Nada foi dito porém quanto a encontrar a solução do dual. É deste assunto que trata esta seção. Antes, porém, é necessário definir o conjunto de subgradientes da função dual $\phi(\cdot)$.

Definição 1.4 *Seja $\phi(\cdot)$ definida como em (1.17). Então o vetor $\mu \in \mathbb{R}^q$ é um subgradiente de $\phi(\cdot)$ no ponto $w^0 \in \mathcal{W}$ se*

$$\phi(w) \leq \phi(w^0) + \langle \mu, w - w^0 \rangle, \quad \forall w \in \mathcal{W} \quad (1.20)$$

O conjunto de todos os subgradientes de $\phi(\cdot)$ no ponto w^0 será denotado por $\partial\phi(w^0)$. Seja

$$X(w) \triangleq \left\{ x : x = \arg \min_{x \in \Omega} l(x, w) \right\}. \quad (1.21)$$

Pode-se mostrar que para qualquer $x \in X(w)$, $g(x)$ é um subgradiente de $\phi(\cdot)$. É possível, portanto, fazer uma aproximação tangencial de $\phi(\cdot)$ através de hiperplanos suportes nos pontos $w^i \in \mathcal{W}$, $i = 1, 2, \dots, q$, cujas inclinações serão dadas pelos subgradientes $g(x^i)$, $x^i \in \Omega$, $i = 1, 2, \dots, q$.

O problema gerado por uma aproximação de $\phi(\cdot)$ a partir de r subgradientes consistiria em

$$\max_{w \in \mathcal{W}} \phi^r(w) \quad (1.22)$$

onde

$$\phi^r(w) = \min_{1 \leq j \leq r} \{ f(x^j) + \langle w, g(x^j) \rangle \} \quad (1.23)$$

caracteriza a aproximação de $\phi(\cdot)$.

Note que (1.22) é equivalente ao seguinte problema de programação linear:

$$\begin{aligned} & \max_{w \in \mathcal{W}} \sigma \\ & \text{s.a.} \quad \sigma \leq f(x^i) + \langle w, g(x^i) \rangle, \quad i = 1, 2, \dots, r \end{aligned} \quad (1.24)$$

O método de Aproximação Tangencial consiste em:

Passo 1 - A partir de uma aproximação com r subgradientes, resolva (1.24), obtendo w^{r+1} ;

Passo 2 - Encontre o valor de $\phi(\cdot)$ para $w = w^{r+1}$:

$$\phi(w^{r+1}) = \min_{x \in \Omega} \{f(x) + \langle w^{r+1}, g(x) \rangle\} \quad (1.25)$$

obtendo uma solução x^{r+1} e um subgradiente $g(x^{r+1})$.

Passo 3 - Se $\sigma - \phi(w^{r+1}) < \epsilon$, para $\epsilon > 0$ suficientemente pequeno, pare. Caso contrário, com os valores de w^{r+1} , x^{r+1} e $g(x^{r+1})$, adicione mais uma restrição ao problema (1.24), faça $r = r + 1$ e volte ao Passo 1.

Uma vantagem de se trabalhar com esta metodologia é que, independentemente da forma da função dual original, o problema (1.24) é sempre um problema de programação linear. O que o método faz, quando a cada iteração acrescenta uma restrição ao problema (1.24), é melhorar a aproximação da função dual. Note que a aproximação feita por subgradientes sempre superestima o valor de $\phi(\cdot)$ e, assim sendo, o valor da diferença $\sigma - \phi(w^{r+1})$ será sempre um valor positivo.

1.5 Existência do Gradiente: Teorema de Daskin

Uma análise que sempre deve ser levada em conta quando da abordagem do problema dual refere-se à diferenciabilidade de $\phi(\cdot)$. Caso a diferenciabilidade seja garantida, encontrar o valor de $d\phi/dw$ é uma operação simples.

Note que a solução do problema dual não depende desta análise, uma vez que na seção anterior foi apresentado um método que encontra a solução, bastando apenas que $\phi(\cdot)$ seja contínua. O que interessa aqui é determinar sob que condições a propriedade de diferenciabilidade de $\phi(\cdot)$ para todo $w \in \mathcal{W}$ é válida.

Teorema 1.7 *Seja o problema primal (1.15) onde $f(\cdot)$ e $g_i(\cdot)$ são funções contínuas sobre Ω . A função dual $\phi(\cdot)$ é diferenciável no ponto w^0 se e somente se $g(\cdot)$ for constante sobre $X(w)$ e, neste caso, o valor de cada derivada parcial será dado por*

$$\left. \frac{\partial \phi}{\partial w_i} \right|_{w=w^0} = g_i(x), \quad \forall x \in X(w^0) \quad (1.26)$$

Prova: Este Teorema é demonstrado a partir de um outro resultado: dadas as condições impostas sobre $f(\cdot)$, $g_i(\cdot)$ e Ω , a derivada direcional unilateral de $\phi(\cdot)$ existe em qualquer direção h e em

qualquer ponto $w \in \mathcal{W}$; seu valor é dado por

$$\begin{aligned} D\phi(w, h) &= \lim_{w \rightarrow 0^+} \frac{\phi(w + \alpha h) - \phi(w)}{\alpha} \\ &= \min_{x \in X(w)} \sum_{i=1}^m h_i \frac{\partial l(x, w)}{\partial w_i} \end{aligned} \quad (1.27)$$

e como $\frac{\partial l(x, w)}{\partial w_i} = g_i(x)$, então

$$D\phi(w, h) = \min_{x \in X(w)} \sum_{i=1}^m h_i g_i(x) \quad (1.28)$$

Para calcular a derivada parcial de $\phi(\cdot)$ em relação a w_i , basta fazer h igual ao i -ésimo vetor unitário e^i em (1.28). Assim sendo, tem-se que

$$D\phi(w, e^i) = \left(\frac{\partial \phi}{\partial w_i} \right)^+ = \min_{x \in X(w)} g_i(x) \quad (1.29)$$

No caso da derivada parcial à esquerda tem-se, com $h = -e^i$,

$$\begin{aligned} \left(\frac{\partial \phi}{\partial w_i} \right)^- &= \lim_{w \rightarrow 0^+} \frac{\phi(w) - \phi(w - \alpha h)}{\alpha} = - \lim_{w \rightarrow 0^+} \frac{\phi(w + \alpha(-e^i)) - \phi(w)}{\alpha} \\ &= -D\phi(w, -e^i) = - \min_{x \in X(w)} -g_i(x) = \max_{x \in X(w)} g_i(x) \end{aligned} \quad (1.30)$$

Sabe-se que a condição para que uma função seja diferenciável é que a sua derivada à esquerda seja igual a sua derivada à direita. Note que, no caso de $\phi(\cdot)$, a partir de (1.29) e (1.30), isto só será possível se $g_i(\cdot)$ for constante para todo $x \in X(w)$.

Sendo $g_i(\cdot) = k(w)$ constante sobre $x \in X(w)$, então a expressão da derivada direcional de $\phi(\cdot)$ em qualquer direção h (1.28) torna-se

$$D\phi(w, h) = k(w) \sum_{i=1}^m h_i \quad (1.31)$$

que é linear com relação às componentes de h .

Lema 1.1 (Teorema de Fenchel) *Seja $\phi(\cdot)$ uma função côncava sobre um conjunto convexo Ω . Seja $y^0 \in \text{int } \Omega$ e suponha que $Df(y^0, h)$ é linear em h . Então $\phi(\cdot)$ é diferenciável em y^0 .*

Está provada a necessidade. A suficiência é imediata, uma vez que, como foi mencionado anteriormente, a igualdade das derivadas laterais é condição para a diferenciabilidade. $\square$

Um caso de particular interesse, e que na verdade motivou esta análise de diferenciabilidade, é quando $\min_{x \in \Omega} l(x, w^0)$ tem apenas uma única solução x^0 . Neste caso, $X(w^0)$ é o conjunto unitário

(*Singleton*) composto apenas por x^0 . Consequentemente, $g(x^0)$ será constante sobre $X(w^0)$, fazendo com que $\phi(\cdot)$ seja sempre diferenciável.

A vantagem de garantir a diferenciabilidade de $\phi(\cdot)$ é a possibilidade de usar métodos baseados no gradiente $\nabla\phi(\cdot)$ para maximizar a função dual.

Uma classe de problemas em que $X(w^0)$ é sempre um conjunto unitário é obviamente a dos problemas estritamente convexos. Um exemplo de problema convexo bastante conhecido e explorado na teoria de controle - o regulador linear-quadrático - será também tratado neste trabalho, porém no contexto de programação multicritério.

1.6 Programação Multicritério: Aspectos Básicos

Em Programação Multicritério procura-se otimizar não uma função, mas um vetor de funções agregadas por uma *função valor* ou *utilidade* (Yu, 1985). Considere então o problema multicritério

$$\min_{x \in \Omega} \mathcal{V}(f(x)) \quad (1.32)$$

onde Ω é um conjunto convexo qualquer e $f(\cdot) : \Omega \rightarrow \mathbb{R}^m$ é um vetor composto pelos critérios a serem minimizados $f_i(\cdot) : \Omega \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$, agregados pela função valor $\mathcal{V}(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}$, suposta crescente em relação a cada critério. Uma característica importante da programação multicritério é que $\mathcal{V}(\cdot)$ não precisa ser explicitamente conhecida. Em geral, apenas propriedades globais como diferenciabilidade, convexidade e crescimento em relação a cada critério são suficientes para garantir uma solução eficiente. As funções aqui consideradas satisfazem estas condições.

Como geralmente os critérios são conflitantes, ou seja, numa situação de equilíbrio a diminuição do valor de um critério implicará no aumento de pelo menos um outro, não existe para (1.32) uma solução ótima, no sentido empregado em problemas escalares. A Figura 1.1 ilustra esta situação. Por exemplo, ao variar-se x dentro do intervalo $[a, b]$ de forma a reduzir do valor de $f_1(\cdot)$ provoca-se um aumento no valor de $f_2(\cdot)$. Em Programação Multicritério, a solução é caracterizada através do conceito de eficiência.

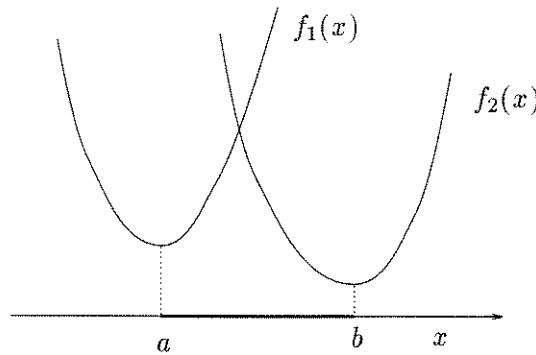


Figura 1.1: Minimização simultânea de duas funções.

Definição 1.5 (Solução Eficiente) Diz-se que $x^* \in \Omega$ é uma solução eficiente de (1.32) se não existe nenhum outro ponto $x \in \Omega$ tal que $f(x) \leq f(x^*)$ e $f(x) \neq f(x^*)$.

O conjunto de soluções eficientes será denotado por Ω^* (Na Figura 1.1, Ω^* seria o próprio intervalo $[a, b]$). O problema (1.32) formulado no *espaço dos critérios* assume a forma

$$\min_{y \in \mathcal{F}} \mathcal{V}(y) \quad (1.33)$$

onde $\mathcal{F}$, definido como

$$\mathcal{F} \triangleq \{y \in \mathbb{R}^m : y = f(x), \text{ para algum } x \in \Omega\} \quad (1.34)$$

é o mapeamento de Ω no espaço dos critérios, representado por simplicidade por $\mathcal{F} = f(\Omega)$; Γ denotará o conjunto de soluções eficientes no espaço dos critérios, ou seja $\Gamma = f(\Omega^*)$. É possível mostrar (Ferreira e Geromel, 1990) que pontos interiores de $\mathcal{F}$ não satisfazem a Definição 1.5 e consequentemente $\Gamma \subset \partial\mathcal{F}$, onde $\partial\mathcal{F}$ representa a fronteira de $\mathcal{F}$ (ver Figura 1.2).

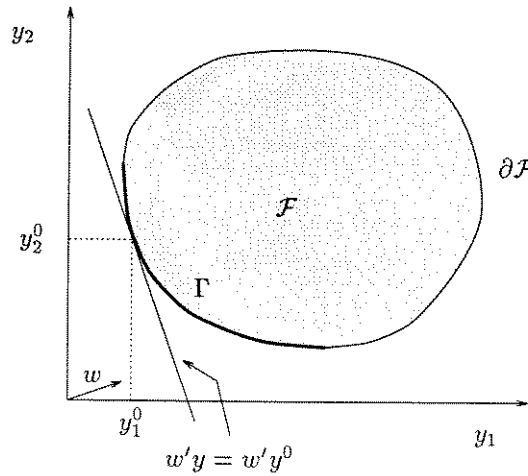


Figura 1.2: Problema no espaço dos critérios.

Condições necessárias para caracterizar a eficiência de uma solução x^* serão derivadas a seguir através de problemas escalares, os quais estão relacionados ao problema original multicritério.

Teorema 1.8 Se existe um $j \in [1, 2, \dots, m]$ e números reais ϵ_i , $i = 1, 2, \dots, m$ ($i \neq j$) tais que $f_j(x^*) \leq f_j(x)$ para todo $x \in \Omega$ tal que

$$f_i(x) \leq \epsilon_i, \quad i = 1, 2, \dots, m \quad (i \neq j) \quad (1.35)$$

então $x^* \in \Omega$ é uma solução eficiente de (1.32).

Prova: Por contradição, suponha que x^* é eficiente mas não satisfaz às hipóteses do Teorema. Então, para todo $j \in [1, \dots, m]$ e todos os números reais ϵ_i , $i = 1, 2, \dots, m (i \neq j)$, x^* não minimiza $f_j(\cdot)$ sujeito a (1.35). Em particular, se $j = 1$ e $\epsilon_i = f_i(x^*)$, $i = 2, \dots, m$, então existe um vetor $\hat{x} \in \Omega$ tal que $f_1(\hat{x}) < f_1(x^*)$ e $f_i(\hat{x}) \leq f_i(x^*)$, $i = 2, \dots, m$, o que contradiz, pela definição de solução eficiente, a hipótese de eficiência de x^* . $\square$

Através do Teorema acima é possível então determinar quando x^* é uma solução eficiente de (1.32).

Teorema 1.9 *Se para todo $j \in [1, 2, \dots, m]$, x^* é a solução do seguinte problema:*

$$\begin{aligned} \min_{x \in \Omega} \quad & f_j(x) \\ \text{s.a.} \quad & f_i(x) \leq f_i(x^*), \quad i = 1, 2, \dots, m \quad (i \neq j) \end{aligned} \tag{1.36}$$

então x^ é eficiente.*

A prova deste teorema segue diretamente como consequência da prova anterior. Uma outra técnica usada para a caracterização de soluções eficientes, e que será de grande importância para o desenvolvimento deste trabalho, é a que faz uso do problema paramétrico

$$\min_{x \in \Omega} \langle w, f(x) \rangle \tag{1.37}$$

Proposição 1.8 *Seja $x(w)$ uma solução ótima de (1.37) para $w \in W$. Então $x(w) \in \Omega$ é uma solução eficiente de (1.32) se*

$$\text{a) } w_i > 0, \quad i = 1, 2, \dots, m$$

ou

$$\text{b) } x(w) \text{ é solução única de (1.37)}$$

Sob a hipótese de convexidade dos critérios de desempenho, Ω^* pode ser completamente gerado variando-se w sobre W definido em (1.2). Em outras palavras, se $y^* \in \Gamma$, então existe $w \in W$ tal que $y^* = f(x(w))$, onde $x(w)$ resolve (1.37).

Este resultado pode ser interpretado através do conceito de dualidade visto anteriormente. Se o problema for convexo, então Γ admite um hiperplano suporte em qualquer solução eficiente de (1.32), uma vez que se $x(w^*)$ é solução eficiente de (1.37), então $f(x(w^*)) \in \Gamma \subset \partial \mathcal{F}$ (ver Figura 1.2). Assim resolver (1.32) é equivalente ao problema de encontrar o vetor $w = w^*$ que fornece $x^* = x(w^*)$.

1.7 Controle Ótimo Escalar

A busca de uma estratégia ótima de controle para um sistema dinâmico qualquer é naturalmente motivada pela existência de um ou mais critérios de desempenho a partir dos quais se deseja avaliar a ação de controle. Esta avaliação pode ter como objetivo, por exemplo, encontrar um controlador

que leve o sistema ao estado desejado no menor tempo possível ou com um mínimo de energia despendida. Tradicionalmente, resolver um problema de controle ótimo consiste em encontrar uma lei de controle para um sistema dinâmico que minimize um critério de desempenho pré-estabelecido.

Seja o sistema cuja dinâmica é regida pela equação

$$\dot{x} = \phi(x(t), u(t)), \quad x(0) = x_0 \quad (1.38)$$

onde $x(\cdot) : [0, T] \rightarrow \mathcal{X} \subset \mathbb{R}^n$ é o estado do sistema, e $\mathcal{X}$ é um conjunto aberto e $u(\cdot) : [0, T] \rightarrow U$ é uma função contínua por partes que representa o controle, sendo U uma esfera q -dimensional de raio fixo. O problema de controle ótimo para o sistema dinâmico acima é formulado como

$$\min_{u(\cdot) \in \mathcal{U}} J(u(\cdot)) = \left\{ \int_0^T f(x(t), u(t)) dt + g(x(T)) \right\} \quad (1.39)$$

$$\dot{x} = \phi(x(t), u(t)), \quad x(0) = x_0$$

onde $J(\cdot)$ representa o critério de desempenho que avaliará o controle sobre o conjunto de todos os controles factíveis $\mathcal{U} \subset U$, definidos no intervalo de tempo $[0, T]$. Supõe-se que as funções $f(\cdot)$ e $g(\cdot)$ que compõem $J(\cdot)$ são continuamente diferenciáveis com respeito a seus argumentos. Note que $f(\cdot)$ considera o estado e o controle durante todo o intervalo de tempo, enquanto $g(\cdot)$ leva em conta apenas o valor do estado final na solução do problema.

1.7.1 Equação de Hamilton-Jacobi-Bellman

Dentre as diversas formas de se resolver o problema de controle ótimo apresentado na seção anterior, encontra-se a abordagem baseada na *Equação de Hamilton-Jacobi-Bellman*. Através do *Princípio da Otimalidade de Bellman* sabe-se que, se $u(t) \in \mathcal{U}$, $t \in [0, T]$ resolve (1.39), então $u(t) \in \mathcal{U}$, $t \in [t^*, T]$ é solução de

$$V(x^*, t^*) = \min_{u(\cdot) \in \mathcal{U}} \left\{ \int_{t^*}^T f(x(t), u(t)) dt + g(x(T)) \right\} \quad (1.40)$$

$$\dot{x} = \phi(x(t), u(t)), \quad x(t^*) = x^*$$

onde $V(\cdot)$ é a *função valor ótima* associada ao problema. Se os valores de $V(x(T), T)$ forem conhecidos para todos os valores de $x(T)$ possíveis e supondo um instante de tempo ΔT suficientemente pequeno, então

$$\begin{aligned} V(x(T-\Delta), T-\Delta) &= \min_{u(\cdot) \in \mathcal{U}} \left\{ \int_{T-\Delta}^T f(x(t), u(t)) dt + V(x(T), T) \right\} \\ &= \min_{u(\cdot) \in \mathcal{U}} \{ f(x(T-\Delta), u(T-\Delta)) \Delta T + V(x(T), T) \} \end{aligned} \quad (1.41)$$

onde, por simplicidade, foi introduzida a notação $T - \Delta T = T_{-\Delta}$ e $T + \Delta T = T_{+\Delta}$.

Como por hipótese $V(x(T), T)$ é conhecido, a solução do problema consiste em encontrar o controle que transfere o sistema do estado $x(T_{-\Delta})$ para $x(T)$. Através de (1.38) tem-se que

$$\begin{aligned} x(T) &= x(T_{-\Delta}) + \int_{T_{-\Delta}}^T \phi(x(t), u(t)) dt \\ &\simeq x(T_{-\Delta}) + \phi(x(T_{-\Delta}), u(T_{-\Delta})) \Delta T \end{aligned} \quad (1.42)$$

Para um t genérico, pode-se então reescrever (1.41) como

$$V(x(t), t) = \min_{u(\cdot) \in \mathcal{U}} \{f(x(t), u(t)) \Delta T + V(x(T_{+\Delta}), T_{+\Delta})\} \quad (1.43)$$

onde $x(T_{+\Delta}) = x(t) + \phi(x(t), u(t)) \Delta T$. Expandindo em série de Taylor, obtém-se então

$$\begin{aligned} V(x(T_{+\Delta}), T_{+\Delta}) &= V(x(t), t) + \frac{\partial V}{\partial x} \phi(x(t), u(t)) \Delta T + \\ &+ \frac{\partial V}{\partial t} \Delta T + \mathcal{O}(\Delta T^2) \end{aligned} \quad (1.44)$$

Substituindo (1.44) em (1.41), dividindo-se por ΔT , reagrupando-se os termos e tomando o limite quando $\Delta T \rightarrow 0$, encontra-se a *Equação de Hamilton-Jacobi-Bellman* que fornece $V(x(t), t)$:

$$\frac{\partial V(x(t), t)}{\partial t} + \min_{u(\cdot) \in \mathcal{U}} \left\{ f(x(t), u(t)) + \frac{\partial V(x(t), t)}{\partial x} \phi(x(t), u(t)) \right\} = 0 \quad (1.45)$$

Após encontrada, a função $V(\cdot)$ caracteriza a solução do problema (1.39).

1.7.2 Problema Linear-Quadrático

Em geral, (1.45) não admite uma solução analítica. Uma exceção importante, cuja solução é bastante conhecida e explorada na teoria de controle ótimo, é a do problema linear-quadrático (PLQ), onde o critério de desempenho para (1.39) é do tipo

$$J(u(\cdot)) = \frac{1}{2} \int_0^T (x'(t) Q x(t) + u(t)' R u(t)) dt + \frac{1}{2} x(T)' F x(T), \quad (1.46)$$

onde $Q = Q' \geq 0$, $F = F' \geq 0$ e $R = R' > 0$. Neste caso, tem-se como função valor

$$V(x(t), t) = \frac{1}{2} x(t)' P(t) x(t), \quad P(t) = P(t)' > 0, \quad t \in [0, T] \quad (1.47)$$

e $P(\cdot)$ é solução da *equação matricial de Riccati*

$$\dot{P}(t) + P(t)A + A'P(t) - P(t)SP(t) + Q = 0, \quad P(T) = F, \quad (1.48)$$

$$S = BR^{-1}B' \quad (1.49)$$

O controle ótimo, solução de (1.45), tem como expressão

$$u^*(t) = -R^{-1}B'P(t)x(t), \quad t \in [0, T], \quad (1.50)$$

Note que a solução para o controle é do tipo $u(t) = -K(t)x(t)$, caracterizando uma realimentação de estados. Para a trajetória ótima tem-se

$$\dot{x} = (A - SP(t))x(t), \quad x(0) = x_0, \quad (1.51)$$

e o valor do critério de desempenho é dado por

$$J(u^*(\cdot)) = \frac{1}{2}x_0'P(0)x_0 \quad (1.52)$$

1.7.3 Princípio do Mínimo de Pontryagin

A solução de problemas de controle ótimo pode ser caracterizada através das condições de otimalidade estabelecidas pelo *Princípio do Mínimo de Pontryagin*. Nesta seção serão apresentadas as condições para sistemas de controle com tempo e estado final livres. Este é um resultado de grande importância para a teoria de controle e sua demonstração pode ser encontrada, por exemplo, em Athans e Falb (1966).

Considere o problema de controle ótimo

$$\min_{u(\cdot) \in \mathcal{U}} \int_0^T f(x(t), u(t)) dt + g(x(T)) \quad (1.53)$$

$$\dot{x} = \phi(x(t), u(t)), \quad x(0) = x_0$$

Suponha que não existe restrição quanto ao estado final do sistema (estado final livre) e que o sistema trabalha com horizonte de tempo livre.

As condições de otimalidade apresentadas a seguir foram desenvolvidas a partir de um *sistema canônico* associado ao problema original (1.53), formulado através da função *Hamiltoniano*

$$H(x(t), p(t), u(t)) \triangleq f(x(t), u(t)) + p(t)' \phi(x(t), u(t)) \quad (1.54)$$

onde o vetor $p(\cdot)$ é denominado de *vetor de co-estados* do sistema. Note que a derivada em relação a $p(\cdot)$ é

$$\frac{\partial H}{\partial p(t)} = \phi(x(t), u(t)), \quad (1.55)$$

e que, devido às hipóteses anteriores, $H(\cdot)$ possui derivada parcial contínua em relação a $x(\cdot)$.

Definição 1.6 *Seja a função Hamiltoniano definida em (1.54). Chama-se então de sistema canônico ou sistema hamiltoniano associado ao problema de controle ótimo (1.53), o sistema*

$$\begin{aligned}\dot{x} &= \frac{\partial H}{\partial p} = \phi(x(t), u(t)) \\ -\dot{p} &= \frac{\partial H}{\partial x} = \frac{\partial f(x(t), u(t))}{\partial x} + \left(\frac{\partial \phi(x(t), u(t))}{\partial x} \right)' p(t)\end{aligned}\tag{1.56}$$

A partir da definição do sistema canônico é possível estabelecer condições necessárias de otimalidade para um controle $u^*(\cdot)$ relacionado ao par $(x^*(\cdot), p^*(\cdot))$, solução do sistema canônico (1.56).

Teorema 1.10 (Princípio do Mínimo de Pontryagin) *Seja $u^*(\cdot) \in \mathcal{U}$ o controle que transfere o sistema considerado do estado $x(0)$ para o estado $x(T)$, cuja trajetória correspondente é dada por $x^*(t)$, $t \in [0, T]$. Se $u^*(\cdot)$ é o controle ótimo do sistema em termos do problema (1.53) então existe uma função $p^*(\cdot)$ tal que:*

i) $p^*(\cdot)$ e $x^*(\cdot)$ são soluções do sistema canônico

$$\begin{aligned}\dot{x}^* &= \frac{\partial H}{\partial p} \\ -\dot{p}^* &= \frac{\partial H}{\partial x}\end{aligned}\tag{1.57}$$

satisfazendo às condições de contorno

$$x^*(0) = x_0, \quad p^*(t) = \frac{dg(x^*(T))}{dx(T)}\tag{1.58}$$

ii) $u^*(t)$ minimiza $H(x^*(t), p^*(t), u(t))$ sobre $\mathcal{U}$ para $t \in [0, T]$, ou seja

$$\min_{u(\cdot) \in \mathcal{U}} H(x^*(t), p^*(t), u(t)) = H(x^*(t), p^*(t), u^*(t))\tag{1.59}$$

iii) $H(x^*(t), p^*(t), u^*(t))$ deve satisfazer à relação

$$H(x^*(t), p^*(t), u^*(t)) = 0, \quad \forall t \in [0, T]\tag{1.60}$$

Vale a pena ressaltar que se forem mudadas as hipóteses iniciais sobre o problema (1.53), as condições derivadas do Princípio do Mínimo sofrem alterações.

1.8 Problema Linear-Quadrático Gaussiano

O problema linear-quadrático gaussiano é um caso particularmente interessante de problemas de controle de sistemas sujeito a ruídos, por possuir uma solução analítica e de fácil implementação.

Como normalmente é feito nos estudos preliminares deste problema, os sistemas lineares considerados serão invariantes no tempo. Por simplicidade, os resultados aqui apresentados concentrar-se-ão nas propriedades do problema linear-quadrático gaussiano propriamente dito; uma exposição geral sobre controle estocástico pode ser encontrada, por exemplo, em Dorato et al. (1995). O problema linear-quadrático gaussiano consiste em, dado o sistema linear estocástico

$$\begin{aligned}\dot{x} &= Ax(t) + bu(t) + v(t) \\ y(t) &= Cx(t) + \psi(t)\end{aligned}\tag{1.61}$$

onde os vetores aleatórios $v(t)$ e $\psi(t)$ representam ruídos brancos gaussianos de média nula e não correlatos, com as seguintes matrizes de auto-correlação

$$\begin{aligned}E\{v(t)v'(\tau)\} &= \Upsilon\delta(t - \tau) \\ E\{\psi(t)\psi'(\tau)\} &= \Psi\delta(t - \tau),\end{aligned}\tag{1.62}$$

encontrar o controle que minimiza um funcional quadrático do tipo

$$J(u(\cdot)) = \lim_{t \rightarrow \infty} E\{x(t)'Qx(t) + u(t)'Ru(t)\}\tag{1.63}$$

onde $Q = Q' \geq 0$ e $R = R' > 0$. As informações sobre o controle $u(t)$ estão disponíveis de forma direta, enquanto que as informações sobre o estado estão disponíveis de forma indireta, através do vetor de medidas $y(t)$, que por sua vez está sujeito a ruídos.

1.8.1 Solução do LQG via Problema Determinístico

A análise do problema linear-quadrático mostra que existe uma grande semelhança entre o caso estocástico e o determinístico, de forma que são facilmente identificáveis elementos que podem ser chamados de duais. Esta propriedade será utilizada para encontrar a solução do estimador ótimo para o caso estocástico, como será visto ainda nesta seção.

Sabe-se que o problema determinístico

$$\min_x \int_0^\infty x'Qxdt\tag{1.64}$$

$$\dot{x} = Ax(t), \quad x(0) = x_0$$

tem como solução ótima (Seção 1.7.2)

$$\begin{cases} V_d^* = \text{Tr}\{Px_0x_0'\} \\ A'P + PA + Q = 0 \end{cases}\tag{1.65}$$

Por outro lado, o problema estocástico

$$\min_x \lim_{t \rightarrow \infty} E\{x'cc'xdt\}\tag{1.66}$$

$$dx = Axdx + dw$$

onde o vetor c é dado e dw é um processo de Wiener, com matriz de covariância Wdt , tem como solução ótima (Dorato et al., 1995)

$$\begin{aligned} V_s^* &= \text{Tr}\{Scc'\} \\ AS + SA' + W &= 0 \end{aligned} \quad (1.67)$$

Note que estes problemas, apesar de fazerem parte de domínios diferentes, possuem soluções semelhantes. Comparando-se (1.65) e (1.67) tem-se as seguintes dualidades:

1. x_0 é dual em relação a c
2. A matriz ponderação de estados Q é dual em relação a matriz de intensidade de ruído W
3. A matriz de sistema A é dual em relação a A'

Neste sentido a solução do problema determinístico pode ser usada para resolver o problema estocástico, como será visto a seguir.

1.8.2 Filtro de Kalman-Bucy

Como os estados do sistema estocástico (1.61) não estão diretamente disponíveis e as medidas y estão sujeitas a ruídos, nasce naturalmente a necessidade de se derivar um sistema dinâmico que seja capaz de fornecer um valor estimado $\hat{x}(t)$ o mais próximo possível do estado original $x(t)$. Para tanto, define-se uma variável *erro de estimativa* $\tilde{x} = x - \hat{x}$ e a seguinte função a ser minimizada:

$$\tilde{V} = \lim_{t \rightarrow \infty} E\{\tilde{x}'(t)cc'\tilde{x}(t)\} \quad (1.68)$$

onde o vetor linha c é dado.

O estimador ótimo terá a seguinte estrutura

$$\dot{\hat{x}} = A\hat{x} + Bu + K_f(y - C\hat{x}) \quad (1.69)$$

onde K_f representa a *matriz de ganho do filtro*. A partir de (1.61) e (1.69) obtém-se a seguinte expressão para o erro de estimação.

$$\dot{\tilde{x}} = (A - K_fC)\tilde{x} + \xi \quad (1.70)$$

onde o processo $\xi = v - K_f\psi$ representa um ruído branco gaussiano de média nula, com a seguinte matriz de correlação:

$$E\{\xi(t)\xi'(\tau)\} = (K_f\Psi K_f' + \Upsilon)\delta(t - \tau) \quad (1.71)$$

Neste caso, a solução para (1.68) pode ser encontrada a partir de

$$\begin{aligned} \tilde{V} &= \text{Tr}\{Scc'\} \\ (A - K_fC)S + S(A - K_fC)' + K_f\Psi K_f' + \Upsilon &= 0 \end{aligned} \quad (1.72)$$

que apresenta características similares às do problema do regulador linear-quadrático determinístico, visto na Seção 1.7.2. Observando a equação de Riccati dada por (1.48) para o caso de regulação em horizonte de tempo infinito ($\dot{P} = 0$), tem-se as seguintes equivalências:

- Υ é dual em relação a Q e Ψ é dual em relação a R
- c é dual em relação a x_o
- A é dual em relação a A' e C' é dual em relação a B
- K'_f é dual em relação a K
- S é dual em relação a P

Portanto, através do dualismo determinístico-estocástico, o valor do ganho do estimador K_f que minimiza (1.68) é dado por

$$K_f = SC'\Psi^{-1} \quad (1.73)$$

onde S satisfaz à equação algébrica de Riccati do filtro

$$AS + SA' - SC'\Psi^{-1}CS + \Upsilon = 0 \quad (1.74)$$

Ainda por dualismo, pode-se inferir que existe solução estável caso o par (A', C') seja controlável, (A', E) seja observável, com $\Upsilon = E'E$ e $\Psi > 0$.

1.8.3 Princípio da Separação

Considere o problema de controle linear-quadrático gaussiano

$$\min_x J(u(\cdot)) = \lim_{t \rightarrow \infty} E \{x(t)'Qx(t) + u(t)'Ru(t)\} \quad (1.75)$$

sujeito ao sistema estocástico (1.61). Assuma que estimador ótimo anteriormente encontrado possui as seguintes propriedades

- O vetor de estados estimados $\hat{x}$ e o vetor erro de estimativa $\tilde{x}$ são não-correlatos;
- O processo aleatório $y - C\hat{x}$ é um ruído branco com média nula e matriz de correlação $\Psi\delta(t - \tau)$.

Então é possível encontrar a solução do problema aplicando-se o resultado a seguir, conhecido como o Princípio da Separação.

Teorema 1.11 *A solução do problema LQG consiste na solução em separado do problema do estimador ótimo (Filtro de Kalman-Bucy)*

$$K_f = SC'\Psi^{-1}$$

onde S satisfaz à equação de Riccati

$$AS + SA' - SC'\Psi^{-1}CS + \Upsilon = 0$$

e do problema de controle ótimo determinístico equivalente,

$$K_c = R^{-1}B'P$$

onde P satisfaz à equação de Riccati

$$A'P + PA - PBR^{-1}B'P + Q = 0$$

Prova: Dorato et al. (1995).

Como consequência deste resultado, o problema LQG é resolvido através de duas equações de Riccati desacopladas. O valor ótimo do critério de desempenho é dado por

$$J^* = \text{Tr}\{PK_f\Psi K_f'\} + \text{Tr}\{SQ\} \quad (1.76)$$

1.9 Conclusão

Neste capítulo foram revistos os conceitos necessários para a caracterização de problemas convexos. Foi enfatizada a importância da existência do ponto de sela e a equivalência dos problemas primal e dual. Discutiu-se ainda a questão da diferenciabilidade da função dual e da solução numérica do problema dual. Finalmente, foram apresentados os aspectos básicos de um problema multicritério convexo, com a caracterização do conjunto de soluções eficientes. Foi também abordada a solução do problema de controle ótimo determinístico.

Dentro da classe de sistemas estocásticos, foi discutida a solução do problema linear-quadrático gaussiano, que possui uma solução analítica e de fácil implementação, em função da aplicabilidade do Princípio da Separação.

Capítulo 2

Controle Multicritério Determinístico

2.1 Introdução

A proposta deste capítulo é tratar o problema de controle multicritério através da programação matemática. Neste caso, como será visto adiante, propriedades relacionadas a conjuntos convexos terão um papel fundamental no desenvolvimento, pois garantem a decomposição do problema numa estrutura de dois níveis (decisão e análise). No nível de decisão, os funcionais são agregados em uma *função valor*, que reúne as preferências do decisor sobre os critérios. Por outro lado, cada decisão tomada é avaliada pelo nível de análise, em termos de desempenho e factibilidade, através de um subproblema de controle ótimo escalar, que preserva as propriedades do problema tratado. Em particular, se os funcionais forem quadráticos, a solução do problema de controle ótimo multicritério consiste em soluções iterativas de equações de Riccati.

2.2 Formulação do Problema

Em sistemas reais de controle é comum que exista mais de um aspecto que mereça avaliação quanto aos resultados da lei de controle aplicada, implicando na existência de vários critérios de desempenho. Neste caso, a teoria de controle ótimo apresentada nas seções anteriores é insuficiente, por tratar o problema com apenas um critério. A busca da solução de problemas de controle ótimo com mais de um critério de desempenho será abordada nesta seção. Será verificado que na formulação do problema de controle multicritério, os conceitos vistos na Seção 1.6, que trata do problema multicritério estático, são válidos, bastando que sejam guardadas as peculiaridades quanto à formulação de um problema dinâmico.

Seja $J(u(\cdot)) = (J_1(u(\cdot)), \dots, J_m(u(\cdot)))$ um vetor de m critérios de desempenho, onde cada critério $J_i(\cdot) : \mathcal{U} \rightarrow \mathbb{R}$ é função do controle $u(\cdot) \in \mathcal{U}$, sendo $\mathcal{U}$ o conjunto de controles admissíveis para o sistema, definidos no intervalo de tempo $[0, T]$. O problema de controle multicritério consiste em procurar minimizar todos os critérios simultaneamente, ou seja, dado $J(\cdot) : \mathcal{U} \rightarrow \mathbb{R}^m$

$$\min_{u(\cdot) \in \mathcal{U}} J(u(\cdot)) \quad (2.1)$$

$$\dot{x} = Ax(t) + Bu(t), \quad x(0) = x_0,$$

onde $x(t) \in \mathbb{R}^n$ é o estado do sistema, $u(t) \in \mathbb{R}^q$ é a variável de controle, A e B são matrizes de dimensões apropriadas.

Como foi visto anteriormente (Seção 1.6), geralmente os objetivos (critérios de desempenho) são conflitantes, não existindo uma solução *ótima* para (2.1) no sentido empregado em otimização escalar. Para o problema de controle multicritério, a solução será caracterizada sob o conceito de *controle eficiente*.

Definição 2.1 Diz-se que um controle $u^*(\cdot) \in \mathcal{U}$ é eficiente se não existe nenhum outro controle $u(\cdot) \in \mathcal{U}$ tal que $J(u(\cdot)) \leq J(u^*(\cdot))$ e $J(u(\cdot)) \neq J(u^*(\cdot))$.

O conjunto de controles eficientes será denotado por $\mathcal{U}^*$. De forma similar ao que foi visto para problemas estáticos (1.37), o problema paramétrico

$$\min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) \rangle \quad (2.2)$$

$$\dot{x} = Ax(t) + Bu(t), \quad x(0) = x_0,$$

assume grande importância para o problema de controle multicritério, pois também neste caso, se os critérios de desempenho forem convexos, o conjunto de controles eficientes $\mathcal{U}^*$ pode ser completamente gerado variando-se w sobre W definido em (1.2).

Como (2.1) possui um número infinito de soluções eficientes é necessário então que haja interação com um decisor para que este decida qual dentre os controles eficientes será a solução do problema.

Supondo que as preferências do decisor sobre os critérios possam ser representadas através de uma *função valor* (ou *função utilidade*) $\mathcal{V}(\cdot) : \mathcal{U} \rightarrow \mathbb{R}$, o problema (2.1) pode ser reformulado como segue:

$$\min_{u(\cdot) \in \mathcal{U}} \mathcal{V}(J(u(\cdot))) \quad (2.3)$$

$$\dot{x} = Ax(t) + Bu(t), \quad x(0) = x_0,$$

Assim como no problema multicritério estático, apresentado na Seção 1.6, a função $\mathcal{V}(\cdot)$ não precisa ser explicitamente conhecida. Em geral, apenas propriedades globais como diferenciabilidade, convexidade e crescimento com relação a cada critério são suficientes para viabilizar o emprego deste tipo de formulação. Uma escolha usual para $\mathcal{V}(\cdot)$ é

$$\mathcal{V}(J(u(\cdot))) = \|J(u(\cdot)) - \underline{y}\|_p, \quad p \geq 1 \quad (2.4)$$

onde $\underline{y}$ é um vetor de dimensão m composto pelos valores resultantes da minimização individual de cada critério de desempenho, ou seja,

$$\underline{y}_i = \min_{u(\cdot) \in \mathcal{U}} J_i(u(\cdot)), \quad i = 1, 2, \dots, m \quad (2.5)$$

O vetor $\underline{y}$ é conhecido como *solução utópica* ou *solução ideal* porque geralmente $\underline{y}$ não é atingível. Para $p = 2$ e $p \rightarrow \infty$ tem-se os problemas de Salukvadze e min-max, respectivamente. Note que $\mathcal{V}(\cdot)$ para $p \rightarrow \infty$ é não-decrescente, contrariando a afirmação feita na Seção 1.6 de que $\mathcal{V}(\cdot)$ deve ser crescente com respeito a cada critério. Porém, este caso não requer hipóteses adicionais, sendo naturalmente tratada pela metodologia proposta. A representação explícita de $\mathcal{V}(\cdot)$ não é realmente necessária. Contudo, assume-se que o decisor seja capaz de fornecer avaliações locais de $\mathcal{V}(\cdot)$ em termos de *trade-offs* implícitos e explícitos (Hwang e Masud, 1979). A hipótese de uma função crescente garante que o controle $u^*(\cdot) \in \mathcal{U}$ que faz com que $\mathcal{V}(\cdot)$ alcance o seu mínimo seja *eficiente*.

Seguindo uma abordagem similar, proposta por Ferreira e Geromel (1990), o método proposto explora as propriedades anteriormente apresentadas, embora trabalhe diretamente no *espaço dos critérios*

$$\mathcal{F} \triangleq \{y : y = J(u(\cdot)), \text{ para algum } u(\cdot) \in \mathcal{U}\}. \quad (2.6)$$

2.3 Algoritmo Básico

Note que o problema (2.1) está formulado sobre o espaço das variáveis de decisão. Considerando-se agora a formulação do problema no espaço dos critérios, tem-se

$$\min_{y \in \mathcal{F}} \mathcal{V}(y) \quad (2.7)$$

onde $\mathcal{F}$ caracteriza o mapeamento de $\mathcal{U}$ neste novo espaço, ou seja $\mathcal{F} = J(\mathcal{U})$. Assim como no caso estático, se $\mathcal{U}^*$ representa o conjunto de controles eficientes, então $\Gamma = J(\mathcal{U}^*)$ representa a sua imagem no espaço dos critérios.

Observe que apesar do problema original ser dinâmico, a formulação no espaço dos critérios permanece estática. Apesar disso, e além de outras vantagens que esta reformulação do problema apresenta, $\mathcal{F}$ é apenas uma abstração, tornando a sua implementação computacional inviável. Além disso, $\mathcal{F}$ em geral não preserva a convexidade do problema, isto é, o problema (2.7) pode ser não-convexo mesmo que o problema original (2.1) seja convexo. Portanto, tornou-se necessário buscar uma representação convexa para $\mathcal{F}$. Usando-se técnicas básicas de relaxação e projeção serão feitas aproximações sucessivas de $\mathcal{F}$ a partir da seguinte definição.

Definição 2.2 Diz-se que um vetor $y \in \mathbb{R}^m$ de trade-off's implícitos é satisfatório se existe $u(\cdot) \in \mathcal{U}$, tal que $J(u(\cdot)) \leq y$.

A projeção da restrição $J(u(\cdot)) \leq y$, $u(\cdot) \in \mathcal{U}$ sobre o espaço dos trade-off's implícitos é representado pelo conjunto convexo

$$\mathcal{C} \triangleq \{y \in \mathbb{R}^m : J(u(\cdot)) \leq y, \text{ para algum } u(\cdot) \in \mathcal{U}\} \quad (2.8)$$

o qual também define todo o conjunto de soluções satisfatórias.

Ao invés de se trabalhar com o conjunto $\mathbb{R}^m$ como domínio para os trade-off's implícitos, podem se considerar alguns subconjuntos $\mathcal{Y} \subset \mathbb{R}^m$, desde que não se exclua nenhum elemento de $\mathcal{C}$. Por exemplo, um caso limite é $\mathcal{Y} \triangleq \{y : y \geq \underline{y}\}$, onde $\underline{y}$ é a solução ideal dada por (2.5). Na prática os vetores $\underline{y}$ mais usados são a solução ideal e a origem do espaço $\mathbb{R}^m$, ou seja, $\underline{y} \triangleq \{y \in \mathbb{R}^m : y = [0 \ 0 \ \dots \ 0]'\}$

Considere agora o problema de se determinar a melhor solução satisfatória de (2.1), ou seja,

$$\min_{y \in \mathcal{C} \cap \mathcal{Y}} \mathcal{V}(y) \quad (2.9)$$

Teorema 2.1 *Seja y^* a solução ótima de (2.9). Então, y^* resolve o problema de controle multi-critério (2.7).*

Prova: Se $y^* \in \mathcal{C} \cap \mathcal{Y}$ resolve (2.9) então existe $u^*(\cdot) \in \mathcal{U}$ tal que $y^* = J(u^*(\cdot))$. Caso contrário, se $y^* \geq J(u^*(\cdot))$ e $y^* \neq J(u^*(\cdot))$ então $y^0 = J(u^*(\cdot))$ contradiz a otimalidade de y^* , pois $y^0 \in \mathcal{C} \cap \mathcal{Y}$ e $\mathcal{V}(y^0) < \mathcal{V}(y^*)$, devido às hipóteses prévias sobre $\mathcal{V}(\cdot)$. Note finalmente que $\mathcal{F} \subset \mathcal{C} \cap \mathcal{Y}$, permitindo concluir que $J(u^*(\cdot)) \in \mathcal{F}$ também resolve (2.7). $\square$

Consequentemente, os problemas (2.9) e (2.7) são equivalentes. Porém, ao contrário de (2.7), as restrições de (2.9) podem ser caracterizadas através da seguinte propriedade de conjuntos convexos, que re-edita o resultado enunciado no Teorema 1.1 para problemas de controle.

Teorema 2.2 *Seja $\mathcal{C}$ definido por (2.8). Então, $y \in \mathcal{C}$ se e somente se y satisfaz o sistema linear com infinitas restrições*

$$\min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y \rangle \leq 0, \quad \forall w \in W \quad (2.10)$$

$\square$

Observe que a minimização indicada em (2.10) é independente de y . O Teorema 2.2 permite representar (2.9) através do seguinte problema com infinitas restrições

$$\min_{y \in \mathcal{Y}} \mathcal{V}(y) \quad (2.11)$$

$$\langle w, y \rangle \geq \langle w, J(u(w, \cdot)) \rangle, \quad \forall w \in W.$$

onde

$$u(w, \cdot) = \arg \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) \rangle \quad (2.12)$$

É óbvio que alguma estratégia de relaxação deverá ser empregada para resolver (2.11). Técnicas de relaxação são baseadas em duas etapas principais:

- i) determinar se a solução corrente do problema satisfaz a todas as restrições ignoradas e, caso contrário,
- ii) gerar uma nova restrição ou, de preferência, a restrição mais violada, para ser adicionada à aproximação da região factível anterior.

Consequentemente, uma questão básica é determinar se um dado $y \in \mathcal{Y}$ é factível para (2.11).

Proposição 2.1 Dado $y \in \mathcal{Y}$, então $y \in \mathcal{C}$ se e somente se $\Theta(y) \leq 0$, onde

$$\begin{aligned}\Theta(y) &= \max_{w \in W} \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y \rangle \\ &= \max_{w \in W} \phi(w)\end{aligned}\tag{2.13}$$

□

sendo

$$\phi(w) \triangleq \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y \rangle\tag{2.14}$$

A verificação deste resultado é imediata. Claramente, se $\Theta(y) > 0$ então a restrição mais violada do problema (2.11) é automaticamente identificada. Antes de partir para a implementação prática da Proposição 2.1 são necessárias algumas definições e propriedades adicionais sobre o problema min-max.

Teorema 2.3 Seja $\mathcal{U}$ o conjunto de controlos admissíveis e defina

$$\begin{aligned}\mathcal{U}(w) &\triangleq \{u(\cdot) : u(\cdot) = \arg \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y \rangle\} \\ \xi &\triangleq J(u^0(\cdot)) - y, \quad u^0(\cdot) \in \mathcal{U}(w)\end{aligned}\tag{2.15}$$

Então, para qualquer $u^0(\cdot) \in \mathcal{U}(w)$, tem-se que $\xi \in \partial\phi(w)$, onde $\partial\phi(w)$ representa o conjunto de todos os subgradiantes de $\phi(\cdot)$ em w .

Prova: Novamente a prova é imediata, uma vez que

$$\begin{aligned}\phi(v) &\leq \langle v, J(u(\cdot)) - y \rangle, \quad \forall u(\cdot) \in \mathcal{U}, v \in W \\ \phi(w) &= \langle w, J(u^0(\cdot)) - y \rangle, \quad \forall u^0(\cdot) \in \mathcal{U}(w), w \in W\end{aligned}$$

Fazendo-se $u(\cdot) = u^0(\cdot)$ e subtraindo-se as duas equações encontra-se

$$\phi(v) \leq \phi(w) + \xi'(v - w), \quad \forall v \in W$$

que é precisamente a definição do *subgradiente* de $\phi(\cdot)$ em w (Lasdon, 1970).

□

A existência de um subgradiente em qualquer ponto $w \in W$ é suficiente para garantir a implementação de um método baseado em Aproximação Tangencial, com o objetivo de resolver (2.13). A solução ótima de (2.11) é obtida através do seguinte procedimento, chamado de Algoritmo Básico (Carvalho e Ferreira (1993;1995a)).

Algoritmo Básico

Faça $k = 0$ e $\mathcal{Y}^0 = \mathcal{Y}$.

Passo 1. Resolva o problema relaxado

$$\min_{y \in \mathcal{Y}^k} \mathcal{V}(y) \quad (2.16)$$

Seja y^k a solução relaxada

Passo 2. Resolva o problema min-max

$$\Theta(y^k) = \max_{w \in W} \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y^k \rangle$$

Escolha $w^0 \in W$ e faça $\ell = 0$.

Passo 2.1 Resolva o problema de controle ótimo (*min*)

$$\min_{u(\cdot) \in \mathcal{U}} \langle w^\ell, J(u(\cdot)) - y^k \rangle$$

obtendo $u^\ell(\cdot) \in \mathcal{U}(w^\ell)$, $\xi^\ell = J(u^\ell(\cdot)) - y^k$ e $\phi(w^\ell)$.

Passo 2.2 Resolva o problema linear (*max*)

$$\max_{\substack{\sigma \\ w \in W}} \sigma$$

$$\sigma \leq \langle w, \xi^i \rangle, \quad i = 1, 2, \dots, \ell.$$

Seja $\sigma^{\ell+1}, w^{\ell+1}$ a solução ótima encontrada.

Se $\sigma^{\ell+1} - \phi(w^{\ell+1}) < \epsilon$, para $\epsilon > 0$ arbitrariamente pequeno, faça $w^k = w^{\ell+1}$, $u^k(\cdot) = u^\ell(\cdot)$ e $\Theta(y^k) = \sigma^{\ell+1}$ e vá ao Passo 3. Caso contrário, faça $\ell = \ell + 1$ e volte ao Passo 2.1.

Passo 3. Se $\Theta(y^k) \leq 0$, então $u^*(\cdot) = u(w^k, \cdot)$ resolve (2.1) e $y^* = J(u^*(\cdot))$ resolve (2.11) ((2.9)). Caso contrário, adicione a restrição mais violada a $\mathcal{Y}^k$

$$\mathcal{Y}^{k+1} = \{y \in \mathcal{Y}^k : \langle w^k, y \rangle \geq \langle w^k, J(u(\cdot)) \rangle\}$$

faça $k = k + 1$ e volte ao Passo 1.

Com relação ao Algoritmo Básico, várias observações e interpretações úteis são possíveis.

- i) Observe que o Algoritmo Básico decompõe o problema de controle multicritério dinâmico original em uma estrutura hierárquica. O nível superior desta estrutura é representado pelo Passo 1 e envolve apenas os aspectos de decisão do problema. Este nível diz respeito à solução de um problema multicritério relaxado e estático sob condições extremamente favoráveis, pois os critérios e as restrições são lineares e o número de variáveis (critérios) é igual a m , o qual

frequentemente é muito pequeno. Um método de programação matemática adequado pode ser aplicado, dependendo da natureza de $\mathcal{V}(\cdot)$. Observe também que através da projeção do problema no espaço dos critérios, a possível *não-separabilidade* do problema original no sentido da Programação Dinâmica (Larsen e Casti, 1978) é superada. Observe também que, mesmo quando todos os critérios do problema são separáveis no tempo, o efeito de agregação causado pela função $\mathcal{V}(\cdot)$ pode em princípio levar à perda desta importante propriedade;

- ii) O nível inferior da estrutura é representado pelo Passo 2 e compreende apenas os aspectos de controle ótimo do problema. Para cada $w \in W$ fixo, o problema interno é definido pela minimização de uma combinação convexa dos funcionais sujeita à dinâmica e restrições originais do sistema. Consequentemente, as propriedades de uma dada classe de funcionais são preservadas e evita-se a necessidade de uma abordagem baseada no aumento de estados, como proposta por Salukvadze (1982). Tal característica é particularmente importante para a classe de problemas lineares-quadráticos, que será discutida mais adiante na Seção 2.5.
- iii) Embora tenha sido originalmente desenvolvido para problemas convexos, o método aqui proposto pode também ser aplicado em problemas não convexos. Com efeito, a hipótese de convexidade é usada para assegurar que todo o ponto $y \in \Gamma$ (e, consequentemente, a solução y^* do problema) admita um hiperplano suporte que possa ser gerado através do problema paramétrico. Em geral, o método minimiza $\mathcal{V}(\cdot)$ sobre a *casca convexa* de $\mathcal{C}$ e uma hipótese mais fraca para a obtenção do controle $u^*(\cdot)$ que resolve (2.1) poderia ser a existência de um hiperplano suporte na solução ótima y^* .

Uma propriedade adicional e interessante deste método é estabelecida no Teorema 2.4.

Teorema 2.4 *O controle ótimo $u^k(\cdot) = u(w^k, \cdot)$ obtido no Passo 2 do Algoritmo Básico é a solução do Problema min-max*

$$\min_{u(\cdot) \in \mathcal{U}} \max_{1 \leq i \leq m} \{J(u(\cdot)) - y_i^k\} \quad (2.17)$$

Prova: Este Teorema é verificado notando-se que (2.17) pode ser reescrito como

$$\begin{aligned} & \min_{\substack{u(\cdot) \in \mathcal{U} \\ \sigma}} \sigma \\ & \sigma \leq J_i(u(\cdot)) - y_i^k, \quad i = 1, 2, \dots, m. \end{aligned}$$

cujo *dual* é precisamente o problema min-max do Passo 2 (ver Geoffrion (1972) para a verificação de um resultado similar) e, obviamente, os problemas são equivalentes, devido às hipóteses de convexidade.

□

Vale a pena notar que, se $y = 0$, então o problema (2.13) é exatamente o problema min-max de Medanić e Andjelić (1971, 1972) estendido para um problema multicritério convexo. Portanto, o Algoritmo Básico pode ser visto como uma generalização deste procedimento para uma função valor $\mathcal{V}(\cdot)$ arbitrária.

2.4 Convergência e Implementação Realimentada

Observe que, como $\mathcal{V}(\cdot)$ é uma função crescente em relação aos seus argumentos, o problema (2.16) sempre tem um mínimo. Sob tais condições, fica simples demonstrar que o Algoritmo Básico converge.

Teorema 2.5 *Qualquer ponto limite da sequência $\{y^k\}$ gerada pelo Algoritmo Básico resolve (2.9).*

Prova: Considere a sequência $\{y^k\}$ ou, se necessário, uma subsequência desta, com y^* sendo o ponto limite. Em uma iteração arbitrária k , a última restrição linear adicionada a $\mathcal{Y}^k$ pode ser escrita como

$$\begin{aligned} 0 &\geq \langle w^k, J(u^k(\cdot)) - y^k \rangle + \langle w^k, y^k - y \rangle \\ 0 &\geq \Theta(y^k) + \langle w^k, y^k - y \rangle \end{aligned} \quad (2.18)$$

e em qualquer iteração posterior $r > k$, a solução y^r deve satisfazer (2.18), ou seja,

$$\begin{aligned} \Theta(y^k) &\leq \langle w^k, y^r - y^k \rangle \\ &\leq \|w^k\|_2 \cdot \|y^r - y^k\|_2 \end{aligned}$$

implicando que $\Theta(y^*) \leq 0$ uma vez que $\|w^k\|_2 \leq 1$, $\forall w^k \in W$ e $\|y^r - y^k\|_2 \rightarrow 0$ quando $k \rightarrow \infty$. Em outras palavras, y^* é viável para (2.9). Agora, supondo que $\mathcal{V}^*$ é o valor ótimo de (2.9), então $\mathcal{V}^k \leq \mathcal{V}^*$ uma vez que $(\mathcal{C} \cap \mathcal{Y}) \subset \mathcal{Y}^k$ para todo $k = 0, 1, 2, \dots$. Consequentemente, $\mathcal{V}(y^*) \leq \mathcal{V}^*$, mostrando que y^* resolve (2.9). $\square$

Torna-se claro que a convergência do Algoritmo Básico vai depender de quão precisa é a determinação da solução do problema relaxado, quando $\mathcal{V}(\cdot)$ não é perfeitamente conhecida. É claro que em tais casos, a determinação de y^k depende da convergência do método usado para resolver o problema relaxado. Alguns aspectos sobre a convergência de métodos multicritérios iterativos são discutidos em Geoffrion (1972).

Um tópico de interesse potencial é a implementação do procedimento de forma a assegurar a correção da solução na presença de perturbações sobre a trajetória do sistema (*open-loop feedback*). Uma versão do Algoritmo Básico em malha fechada será função tanto do estado corrente como dos valores acumulados dos m critérios sobre o tempo decorrido. Definindo os valores acumulados no intervalo $0 \leq t \leq \tau$ como $J(\tau)$, o problema de determinar a política de controle ótimo para o horizonte de tempo remanescente é semelhante ao que foi feito para desenvolver o Algoritmo Básico. Para tal é necessário apenas definir o conjunto convexo

$$\mathcal{C}(\tau) \triangleq \{y \in \mathbb{R}^m : J(u(\cdot)) \leq y - J(\tau), \text{ para algum } u(\cdot) \in \mathcal{U}\} \quad (2.19)$$

onde agora $u(\cdot)$ é o controle no intervalo $[\tau, T]$. O problema min-max torna-se então

$$\begin{aligned} \max_{w \in W} \min_{u(\cdot) \in \mathcal{U}} &\langle w, J(\tau, T)(u(\cdot)) - (y - J(\tau)) \rangle \\ \dot{x} &= Ax(t) + Bu(t), \quad x(\tau) = x_\tau \end{aligned} \quad (2.20)$$

e uma restrição a ser incorporada ao problema de decisão assume a forma

$$\langle w^k, y \rangle \geq \langle w^k, J^{(\tau, T)}(u^k(\cdot)) \rangle + \langle w^k, J^{(\tau)} \rangle \quad (2.21)$$

onde a dependência da solução no nível de decisão com os valores acumulados é considerada explicitamente. Observe que na ausência de perturbações, $J^{(\tau)} = J^{(\tau)}(u^k(\cdot))$, onde $u^k(\cdot)$ é definido no intervalo $[0, \tau]$, e o problema se reduz ao caso já tratado (malha aberta). O Algoritmo Básico pode ser implementado em malha fechada introduzindo-se um terceiro nível na estrutura que incrementará o tempo decorrido τ de uma quantidade fixa ΔT , por exemplo. Como para pequenos valores de ΔT espera-se que os subconjuntos $\mathcal{C}^{(\tau)}$ e $\mathcal{C}^{(\tau+\Delta T)}$ não sofram grandes alterações, uma boa heurística é inicializar a próxima iteração com a solução ótima da iteração anterior.

Em geral, o valor de w^* dependerá do estado corrente e dos valores acumulados dos critérios. Entretanto, na ausência de perturbações e com perfeita descrição da planta, o valor de w^* permanece constante e só depende do estado inicial $x(0)$:

$$w^*(x(\tau), J^{(\tau)}) = w^*(x(0)) = \text{constante}. \quad (2.22)$$

De fato, observe que se $u(w^*, t)$, $t \in [\tau, T]$ resolve

$$\min_{u(\cdot) \in \mathcal{U}} \mathcal{V}(J(\cdot)) \quad (2.23)$$

$$\dot{x} = Ax(t) + Bu(t), \quad x(\tau) = x_\tau$$

para quaisquer $\tau \in [0, T]$ e então

$$y^{\tau, T} = J^{(\tau, T)}(u(w^*, \cdot)) = y^* - J^{(\tau)}(u(w^*, \cdot)) \quad (2.24)$$

resolve o mesmo problema no espaço dos critérios, indicando que w^* e y^* resolvem o problema multicritério para qualquer valor de τ .

2.5 Problema Linear-Quadrático Multicritério

Nesta seção, demonstra-se que o Algoritmo Básico é especialmente adequado para tratar o problema de controle linear-quadrático multicritério. Considere o problema de controle multicritério (2.1) com um conjunto finito de m critérios quadráticos

$$J_i(u(\cdot)) = \frac{1}{2} \int_0^T \{x(t)' Q_i x(t) + u(t)' R_i u(t)\} dt + \frac{1}{2} x(T)' F_i x(T) \quad (2.25)$$

$$i = 1, 2, \dots, m$$

onde $Q_i = Q_i' \geq 0$, $F_i = F_i' \geq 0$ e $R_i = R_i' > 0$. Obviamente, a menos de y , o problema interno de minimização no Passo 2 assume a forma

$$\min_{u(\cdot)} \frac{1}{2} \int_0^T \{x(t)' Q x(t) + u(t)' R u(t)\} dt + \frac{1}{2} x(T)' F x(T) \quad (2.26)$$

$$\dot{x} = Ax + Bu, \quad x(0) = x_0$$

onde $Q = \sum_{i=1}^m w_i Q_i$, $R = \sum_{i=1}^m w_i R_i$ e $F = \sum_{i=1}^m w_i F_i$, $w \in W$.

Em vista dos resultados já existentes para esta classe de problemas (Seção 1.7.2; ver também Athans e Falb, 1966), a propriedade apresentada a seguir torna-se evidente.

Proposição 2.2 *Os elementos do conjunto de controles eficientes $\mathcal{U}^*$ de (2.26) têm a forma*

$$u(w, t) = -K(w, t)x(t) \quad (2.27)$$

onde $K(w, t) = R^{-1}B'P(w, t)$ e $P(w, t)$ é a solução da equação matricial de Riccati

$$\dot{P}(w, t) + A'P(w, t) + P(w, t)A - P(w, t)BR^{-1}B'P(w, t) + Q = 0 \quad (2.28)$$

tendo como condição de contorno $P(w, T) = F$.

Menos evidente, e aparentemente ainda não explorada na literatura de controle multicritério, é a seguinte consequência da Proposição 2.2.

Proposição 2.3 *O conjunto de todos os ganhos eficientes é independente da condição inicial x_0 .*

Em outras palavras, todo ganho K calculado por (2.27)-(2.28) permanece eficiente para o mesmo valor de $w \in W$, independente do valor que x_0 assume. Entretanto, como Γ **depende** de x_0 , uma solução ótima encontrada a partir de uma determinada condição inicial não permanece ótima se a condição inicial muda (embora o ganho permaneça eficiente). A utilidade prática da Proposição 2.3 é que se x_0 muda, os semi-espacos suportes que definem a região viável $\mathcal{Y}^{k+1}$ no Passo 3 do Algoritmo Básico podem ser facilmente atualizados e o procedimento trabalhará a partir de uma aproximação razoável de Γ . De grande importância prática para problemas multicritérios lineares quadráticos é também o seguinte resultado.

Teorema 2.6 *A função $\phi(\cdot)$ é diferenciável sobre W e*

$$\nabla \phi(w) = \frac{\partial \phi}{\partial w} = J(u(w, \cdot)), \quad \forall w \in W \quad (2.29)$$

□

A unicidade da solução do problema (2.26) leva-nos a $\mathcal{U}(w) = \{u(w, \cdot)\}$, um singleton e um caso especial de aplicação do Teorema de Daskin (Teorema 1.5). Consequentemente, como foi visto na Seção 1.5, o algoritmo de aproximação tangencial do Passo 2 do Algoritmo Básico (etapa de maximização do nível de análise) pode ser eficientemente substituído por um método de gradiente ótimo baseado no algoritmo iterativo

$$w^{\ell+1} = w^\ell + \alpha_\ell s^\ell \quad (2.30)$$

onde s^ℓ é a projeção ortogonal de $\nabla\phi(w^\ell)$ sobre o espaço nulo gerado pelas restrições ativas de W e α_ℓ é um tamanho de passo tal que $w^{\ell+1} \in W$ e $\phi(w^{\ell+1}) > \phi(w^\ell)$, ambos facilmente calculados neste caso. A seguir apresenta-se um exemplo da implementação do Algoritmo Básico. Para tal foi usado o pacote para operações matriciais *MATLAB*.

O nível de análise (Passo 2) consiste, como visto anteriormente, em um problema que possui como função objetivo

$$\phi(w) = \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) - y \rangle \quad (2.31)$$

Uma vez garantida a diferenciabilidade de $\phi(\cdot)$ com respeito ao vetor w , o cálculo do gradiente da função dual é imediato, bastando solucionar a etapa de minimização, que consiste em um problema de controle ótimo padrão.

Uma formulação semelhante ao problema acima, mas que considera $y = 0$, foi utilizada por Medanić e Andjelić (1971). A implementação do nível de análise baseou-se na mesma estratégia usada por aqueles autores, onde a direção de busca em cada iteração da etapa de maximização era calculada projetando-se o gradiente sobre as restrições ativas do problema.

$$w^{k+1} = w^k + \alpha Proj_W \nabla \phi \quad (2.32)$$

Com relação ao nível de decisão (Passo 1), o método a ser usado dependerá da função valor $\mathcal{V}(\cdot)$ escolhida. Neste trabalho, considerou-se como função valor a norma l_p para $p = 1$ (combinação linear), $p = 2$ (norma Euclideana) e $p = \infty$ (norma infinito). Vale a pena salientar que, como foi visto anteriormente, o nível de decisão possui características que lhe garantem grande facilidade de implementação qualquer que seja a função valor adotada.

Para $p = 1$, tem-se o seguinte problema no nível de decisão

$$\min_{y \in \mathcal{Y}^k} \sum_{i=1}^m (y_i - \underline{y}_i) \quad (2.33)$$

e para $p = \infty$ tem-se

$$\min_{\substack{\sigma \\ y^k}} \sigma \quad (2.34)$$

$$\sigma \geq (y_i - \underline{y}_i), \quad i = 1, 2, \dots, m$$

Observe que ambos são problemas de otimização lineares, para os quais dispõem-se de diversos procedimentos eficientes de solução (Luenberger, 1969). Para encontrar a solução do nível de decisão quando $p = 2$, resolve-se

$$\min_{y \in \mathcal{Y}^k} \frac{1}{2} \sum_{i=1}^m (y_i - \underline{y}_i)^2 \quad (2.35)$$

que é um problema quadrático, também bastante simples (Luenberger, 1969).

As funções valores aqui apresentadas são todas explícitas e como foi dito antes, devido à simplicidade intrínseca do problema no nível de decisão, não houve nenhuma dificuldade para a sua implementação. A seguir apresenta-se a solução de um problema envolvendo um sistema de 1ª ordem, que apesar de extremamente simples, é de grande valia no sentido de mostrar como o algoritmo se comporta durante a busca da solução.

Seja o problema de controle ótimo linear quadrático

$$\min_{u(\cdot)} J(u) = \frac{1}{2} \int_0^\infty (q x^2(t) + r u^2(t)) dt \quad (2.36)$$

$$\dot{x} = -ax(t) + u(t), \quad x(0) = x_0$$

onde a , q e r são quantidades escalares positivas. A partir da equação de Riccati para o caso de horizonte de tempo infinito

$$ap + pa - pr^{-1}p + q = 0 \quad (2.37)$$

encontra-se o seguinte valor factível para p :

$$p = ar + \sqrt{a^2 r^2 + qr} \quad (2.38)$$

Com isso, o ganho de realimentação e o valor do critério na solução são dados por

$$\begin{aligned} k &= r^{-1}p \\ &= a + \sqrt{a^2 + \frac{q}{r}} \end{aligned} \quad (2.39)$$

$$J(u^*) = \frac{1}{2} p x_0^2$$

Fez-se então uma divisão do critério em duas partes, uma relativa ao estado e a outra ao controle, resultando no seguinte problema bi-critério:

$$\begin{aligned} \min_{u(\cdot)} J_1(u) &= \frac{1}{2} \int_0^\infty x^2(\cdot) dt \\ J_2(u) &= \frac{1}{2} \int_0^\infty u^2(t) dt \end{aligned} \quad (2.40)$$

$$\dot{x} = -ax(t) + u(t), \quad x(0) = x_0$$

Todas as soluções eficientes do problema acima podem ser geradas a partir do problema ponderado

$$\min_{u(\cdot)} \frac{1}{2} \int_0^\infty (w x^2(t) + (1-w) u^2(t)) dt, \quad 0 \leq w \leq 1 \quad (2.41)$$

$$\dot{x} = -ax(t) + u(t)$$

cuja solução para cada valor de w é imediatamente determinada pela solução de (2.36) bastando-se que naquelas expressões q seja substituído por w e r por $(1-w)$. Para ilustrar o comportamento do Algoritmo Básico usou-se, como função valor,

$$\mathcal{V}(y) = y_1 + \frac{1-\alpha}{\alpha} y_2, \quad 0 < \alpha \leq 1 \quad (2.42)$$

onde α é um valor definido pelo decisor e que procura refletir suas preferências quanto aos aspectos de regulação e esforço de controle que compõem o problema bi-critério.

Na Tabela 2.1 são apresentados os resultados encontrados pelo Algoritmo Básico para $a = 2$ e $x_0 = 1$, considerando-se que ambos os critérios possuem a mesma importância do ponto de vista do decisor, o que se procura representar através de $\alpha = 0.5$. O critério de parada escolhido foi $\Theta(w) < 10^{-5}$.

Tabela 2.1: Comportamento do Algoritmo Básico para $\alpha = 0.5$.

k	y_1	y_2	w_1	w_2	$J_1(u)$	$J_2(u)$	$\Theta(y)$
0	0	0	0.833	0.167	0.0833	0.0833	0.0833
1	0.1000	0	0.554	0.446	0.1092	0.0092	0.0092
2	0.1166	0	0.323	0.677	0.1181	0.0016	0.0016
3	0.1135	0.0038	0.449	0.551	0.1140	0.0043	$4.68 \cdot 10^{-4}$
4	0.1115	0.0063	0.504	0.496	0.1116	0.0064	$1.27 \cdot 10^{-4}$
5	0.1128	0.0052	0.477	0.523	0.1128	0.0053	$3.04 \cdot 10^{-5}$
6	0.1122	0.0058	0.490	0.510	0.1122	0.0058	$0.78 \cdot 10^{-5}$

Note que, a cada iteração o nível de decisão fornece soluções (vetor y) cada vez mais próximas da solução dada pelo nível de análise (vetor $J(u)$). Como é garantida a eficiência das soluções do nível de análise, o vetor y se aproxima da curva de *trade-off* do problema, o que pode ser constatado pelo decréscimo do valor de $\Theta(w)$. Esta “aproximação” é medida pelo nível de análise. Quando y estiver suficientemente próximo, segundo o critério de parada adotado, o controle u^* determinado pelo nível de análise é a solução do problema. Como era de se esperar, o valor de w^* tende para o valor de α dado, significando que a solução $u^*(\cdot)$ fornece os valores de critérios de desempenho que satisfazem à preferência pré-estabelecida pelo decisor. Qualquer que seja o valor de α fornecido, espera-se que a solução final encontrada seja compatível, fornecendo valores que espelham suas preferências.

De fato, note que tem-se como expressão para o sistema em malha fechada

$$\dot{x} = (a - k)x \tag{2.43}$$

que, por sua vez, tem como autovalor

$$\lambda = a - k = -\sqrt{a^2 + \frac{q}{r}} = -\sqrt{a^2 + \frac{\alpha}{1 - \alpha}} \tag{2.44}$$

indicando que, ao preferir um determinado valor de α , o decisor influencia diretamente na dinâmica do sistema. Esta influência se dá, como pode ser visto na Figura 2.1, na localização do pólo do sistema. Quanto menor o valor de α maior o valor do pólo, que estando diretamente ligado à constante de tempo do sistema, indicará que o decisor dá maior importância à economia de energia, optando por uma regulação mais lenta. Caso contrário, quanto maior o valor de α menor será a constante de tempo, mais rápida será a regulação, e em contrapartida, maior o esforço de controle.

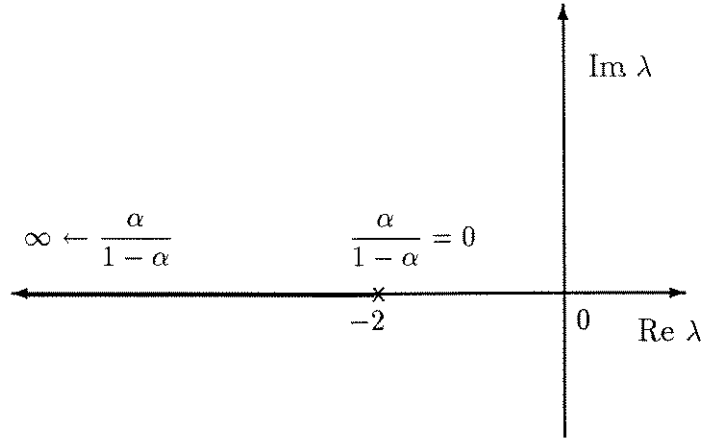


Figura 2.1: Pólo do sistema em função de α .

A Figura 2.2 mostra as três primeiras iterações do Algoritmo Básico durante a resolução do exemplo. Pode-se ver como a cada passo o decisor vai melhorando a sua aproximação de Γ através dos planos suportes, tendo assim mais condições para fornecer uma solução y^k melhor. Pode-se mostrar que $\Theta(y^k)$ representa a distância da solução (infactível) encontrada pelo decisor e a solução da curva de *trade-off* (logo factível) mais próxima, determinada pelo analista. Com isso, decisor e analista estão sempre trocando informações até que o decisor tenha condições (hiperplanos suportes) suficientes para fornecer uma solução que seja *satisfatória* para o analista ($\Theta(y) \leq \epsilon$, para $\epsilon > 0$).

A aplicação da abordagem no caso linear-quadrático demonstra a grande flexibilidade do algoritmo proposto, onde o prévio conhecimento de determinadas propriedades do sistema a ser tratado simplifica a solução do problema.

A seguir apresentam-se alguns resultados obtidos a partir do Algoritmo Básico. Os problemas de controle min-max e de múltiplos alvos considerados por Medanić e Andjelić (1971, 1972) podem ser formulados de acordo com a abordagem proposta como

$$\min_{u(\cdot)} \mathcal{V}(J(u(\cdot))) \quad (2.45)$$

$$\dot{x} = Ax(t) + Bu(t), \quad x(0) = x_0,$$

com $\mathcal{V}(J(u(\cdot))) = \|J(u(\cdot)) - \underline{y}\|_\infty$. Na realidade, os autores consideraram o caso em que $\underline{y} = 0$. Em algumas formas especiais do problema de Salukvadze (Salukvadze, 1982), a função valor é dada por

$$\mathcal{V}(J(u(\cdot))) = \|J(u(\cdot)) - \underline{y}\|_p, \quad p \geq 1 \quad (2.46)$$

sendo que no espaço dos critérios, o problema assume a forma

$$\min_{y \in C \cap Y} \|y - \underline{y}\|_p, \quad p \geq 1 \quad (2.47)$$

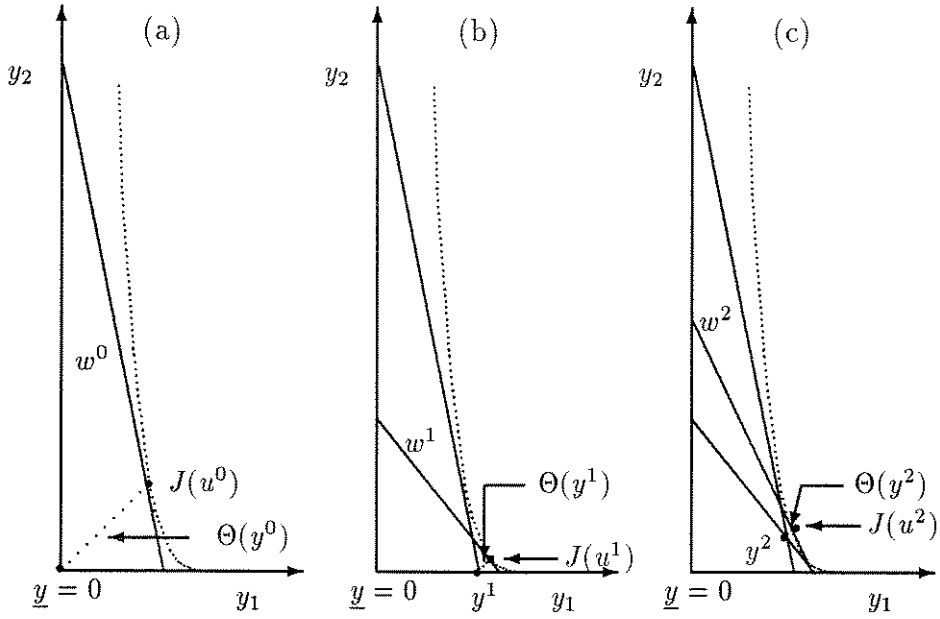


Figura 2.2: (a): 1ª Iteração; (b): 2ª Iteração; (c): 3ª Iteração.

Os valores $p = 1$, $p = 2$ e $p = \infty$, quando (2.11) representa um problema linear, quadrático e min-max, respectivamente, são de particular interesse. Como o número de variáveis é m , o qual é frequentemente bem pequeno, o problema (2.47) pode ser facilmente resolvido. Com o objetivo de ilustrar alguns aspectos numéricos do procedimento proposto, o seguinte exemplo foi considerado.

Exemplo Problema de Salukvadze (Li, 1990a, 1990b)

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \quad T = 1,$$

$$Q_1 = \text{diag}(3, 0, 0) \quad Q_2 = \text{diag}(0, 2, 0) \quad Q_3 = \text{diag}(0, 0, 1)$$

$$F_1 = \text{diag}(1, 0, 0) \quad F_2 = \text{diag}(0, 1, 0) \quad F_3 = \text{diag}(0, 0, 2)$$

$$R_1 = R_2 = R_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \underline{y} = [22.39 \quad 12.39 \quad 8.32]'$$

O exemplo foi resolvido para $p = 1$, $p = 2$ e $p = \infty$. A Tabela 2.2 apresenta os valores correspondentes dos critérios e a Tabela 2.3 ilustra a performance do Algoritmo Básico para o caso $p = 2$. A Figura 2.3 descreve o comportamento do algoritmo para os três valores de p considerados. Perceba que o número de iterações varia com o tipo de norma utilizada. O critério de parada

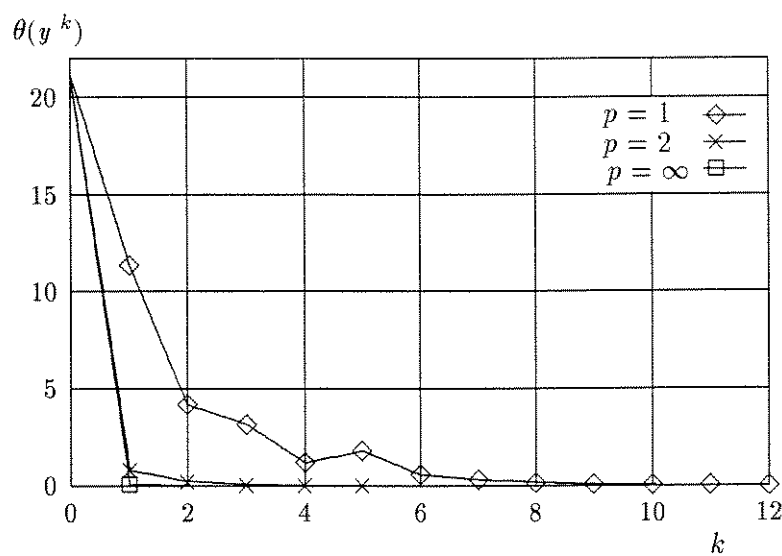


Figura 2.3: Convergência de Θ - $p = 1$, $p = 2$ e $p = \infty$.

Tabela 2.2: Valores Ótimos para $p = 1$, $p = 2$, $p = \infty$.

p	w_1	w_2	w_3	$J_1(u)$	$J_2(u)$	$J_3(u)$
1	0.337	0.334	0.328	29.49	41.65	29.67
2	0.231	0.418	0.351	36.14	37.29	29.09
∞	0.181	0.482	0.337	43.48	33.63	29.53

Tabela 2.3: Evolução do Algoritmo Básico - $p = 2$.

k	y_1	y_2	y_3	w_1	w_2	w_3	$J_1(u)$	$J_2(u)$	$J_3(u)$	$\Theta(y)$
0	22.39	12.39	8.32	0.181	0.483	0.336	43.48	33.55	29.61	21.1910
1	32.50	39.40	27.11	0.263	0.366	0.371	33.24	40.20	27.93	0.7916
2	36.44	36.17	29.64	0.227	0.435	0.348	36.58	36.38	29.92	0.2206
3	35.79	38.07	28.27	0.233	0.403	0.364	35.84	38.10	28.36	0.0565
4	36.05	37.22	29.19	0.232	0.419	0.359	36.03	37.21	29.26	0.0163
5	36.17	37.32	29.04	0.231	0.418	0.351	36.14	37.29	29.09	-0.0019

adotado foi $\Theta(y^k) < 10^{-2}$. Observe que, quando $p = \infty$, o método sempre encontrará a solução ótima na primeira iteração.

Com o objetivo de ilustrar a generalidade da abordagem proposta com relação a escolha dos critérios, funções valores e métodos utilizados para resolver o problema no espaço dos critérios, considera-se na próxima seção um problema com critérios não-quadráticos.

Em particular, demonstra-se que de acordo com a abordagem proposta, o problema de controle multicritério é visto pelo decisor como um problema multicritério estático, o qual pode ser eficientemente tratado por um grande número de técnicas existentes.

2.6 Critérios Não-Quadráticos: Um Exemplo

Nesta seção, o problema de minimizar tanto o tempo necessário para transferir um sistema linear de um estado inicial $x(0)$ para a origem e a energia despendida para tal é usado para demonstrar a generalidade da abordagem proposta. Por conveniência, é considerado o mesmo exemplo numérico sugerido em Kirk (1970):

$$\begin{aligned} \min_{-1 \leq u(\cdot) \leq 1} J_1(u(\cdot)) &= \int_0^T 1 \cdot dt \\ \min_{-1 \leq u(\cdot) \leq 1} J_2(u(\cdot)) &= \int_0^T u(t)^2 dt \end{aligned} \tag{2.48}$$

$$\dot{x} = -x(t) + u(t), \quad x(0) = x_0, \quad x(T) = 0$$

onde $J_1(u(\cdot))$ representa tempo decorrido e $J_2(u(\cdot))$ o gasto de energia. Vale a pena ressaltar o caráter conflitante destes dois critérios, pois quanto mais rápida deva ser a ação do controle sobre o sistema, no sentido de deslocá-lo do seu estado inicial ao estado desejado, mais energia terá que ser gasta.

A combinação de $J_1(u(\cdot))$ e $J_2(u(\cdot))$, que define o Passo 2 do Algoritmo Básico, fornece o problema de controle ótimo

$$\min_{-1 \leq u(\cdot) \leq 1} \int_0^T \{w_1 + w_2 u(t)^2\} dt, \quad w \in W \quad (2.49)$$

$$\dot{x} = -x(t) + u(t), \quad x(0) = x_0, \quad x(T) = 0$$

Observe que a unicidade da solução de (2.49) garantirá a diferenciabilidade de $\phi(\cdot)$ sobre W . Através do Princípio do Mínimo de Pontryagin obtém-se

$$u^*(\beta, t) = \begin{cases} x(t) - \sqrt{(x(t))^2 + \beta}, & \text{se } 0 < \frac{\beta - 1}{2} < x(t) \\ -1.0, & \text{se } 0 < x(t) \leq \frac{\beta - 1}{2} \\ +1.0, & \text{se } -\frac{\beta - 1}{2} \leq x(t) < 0 \\ x(t) + \sqrt{(x(t))^2 + \beta}, & \text{se } x(t) < -\frac{\beta - 1}{2} < 0 \\ 0, & \text{se } x(t) = 0 \end{cases} \quad (2.50)$$

onde $\beta = \frac{w_1}{w_2}$.

Observe que, como $w_2 = 1 - w_1$, o problema de maximização poderia ser resolvido através de qualquer procedimento de busca unidimensional. Contudo, para manter a generalidade do procedimento, foi empregado o método do gradiente projetado.

Uma atenção considerável tem sido dedicada para a solução de problemas de otimização bi-critério. Geoffrion (1966) determinou propriedades básicas para esta classe de problemas e discutiu algumas formas especiais para funções valores ou utilidade. Uma forma adequada para o caso aqui considerado parece ser (Geoffrion, 1966)

$$\mathcal{V}(J(u(\cdot))) = \alpha_1 e^{J_1(u(\cdot))} + \alpha_2 e^{J_2(u(\cdot))} \quad (2.51)$$

onde α_1 e α_2 são constantes positivas dadas. O problema geral a ser resolvido no espaço dos critérios é simplesmente

$$\min_{y \in \mathcal{Y}^k} \alpha_1 e^{y_1} + \alpha_2 e^{y_2} \quad (2.52)$$

Vários métodos iterativos podem ser utilizados para resolver (2.52). Devido a sua associação com o conhecido procedimento de Geoffrion, Dyer and Feinberg (GDF) para problemas de otimização multicritério (Geoffrion et al., 1972), o método de Frank-Wolfe foi escolhido. O método pode ser resumido como segue: escolha uma solução inicial $y^0 \in \mathcal{Y}^k$ e faça $l = 0$.

Passo 1. Resolva o problema linear

$$\min_{\xi \in \mathcal{Y}^k} \langle \nabla \mathcal{V}(y^k), \xi \rangle$$

obtendo ξ^l . Faça $d^l = y^l - \xi^l$.

Passo 2. Selecione um passo α^l através da solução do problema unidimensional

$$\min_{0 \leq \alpha \leq 1} \mathcal{V}(y^l + \alpha d^l)$$

e faça $y^{l+1} = y^l + \alpha^l d^l$.

Passo 3. Se $\|y^{l+1} - y^l\| < \epsilon$, para $\epsilon > 0$ arbitrariamente pequeno, pare. Caso contrário, faça $l = l + 1$ e volte ao Passo 1.

O exemplo foi resolvido para $x(0) = 7$, $\alpha_1 = \alpha_2 = 1$ e para $\alpha_1 = 5$, $\alpha_2 = 1$. Considerou-se como critério de parada para o método do gradiente projetado (nível de análise)

$$(G(w^k) - G(w^{k-1})) < 10^{-2}(1 + G(w^{k-1})) \text{ ou } (\|\nabla G_p^k\| < 10^{-2}(\|\nabla G^k\|)^{\frac{1}{3}}) \quad (2.53)$$

e para o método de Frank-Wolfe, no nível de decisão, o critério de parada foi $(y^k - y^{k-1}) < 10^{-6}$, enquanto que o Algoritmo Básico termina quando $\Theta < 10^{-3}$. Nas Tabelas 2.4 e 2.5 encontram-se os resultados obtidos para $\alpha = [1 \ 1]'$ e $\alpha = [5 \ 1]'$ respectivamente. Observe que, para o segundo caso, a parcela relativa ao tempo possui um peso maior do que a relativa à energia. Consequentemente, o decisor estará disposto a um maior gasto de energia para obter uma economia de tempo. As Figuras 2.4 e 2.6 apresentam as curvas de nível de $\mathcal{V}(\cdot)$ e uma aproximação das curvas de trade-off Γ do problema, para cada valor de α , bem como a solução gerada pelo Algoritmo básico, ilustrando mais claramente a influência da forma de $\mathcal{V}(\cdot)$ sobre a solução do problema. Observe que $y^0 = J(u^k(\cdot))$ é uma condição inicial adequada para (2.52) e que um número relativamente pequeno de iterações do Algoritmo Básico foi necessária para resolver os exemplos. A solução de compromisso $J(u^*(\cdot))$ para cada caso também está apresentada na Figura 2.4. As Figuras 2.5 e 2.7 mostram as trajetórias ótimas do controle e das variáveis de estado nos dois casos, podendo-se ver os instantes em que ocorrem saturações do controle e a sua dependência com o estado. Para o segundo caso, a saturação deu-se antes que no primeiro caso, resultando em um maior gasto de energia. Em compensação, o sistema chegou mais rápido ao estado final, sendo este resultado esperado uma vez que um peso maior foi dado ao critério que representa o tempo.

Tabela 2.4: Solução para $v_1 = 1$, $v_2 = 1$.

k	y_1	y_2	w_1	w_2	$J_1(u)$	$J_2(u)$	$\Theta(y)$
0	2.0786	0	0.5343	0.4657	2.5783	0.5688	0.5319
1	2.0786	1.1422	0.7845	0.2155	2.2065	1.3085	0.1362
2	2.2915	0.9992	0.6817	0.3183	2.3379	0.9378	0.0121
3	2.2660	1.0918	0.7435	0.2565	2.2567	1.1476	0.0073
4	2.2451	1.1810	0.7657	0.2343	2.2273	1.2292	-0.0024

Muitos métodos multiobjetivos não exigem a definição de uma função valor explícita e, embora o exemplo possa ser resolvido com o decisor interagindo diretamente no espaço dos critérios (o conjunto $\mathcal{Y}^k$ pode ser representado graficamente no caso bidimensional), o método GDF foi adotado

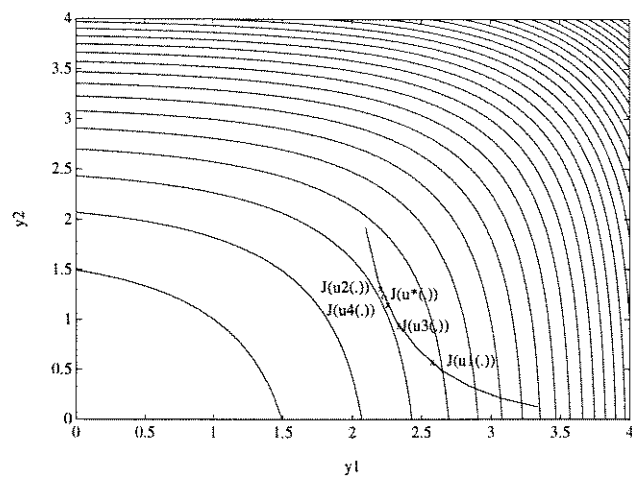


Figura 2.4: Solução no espaço dos critérios - $v_1 = v_2 = 1$

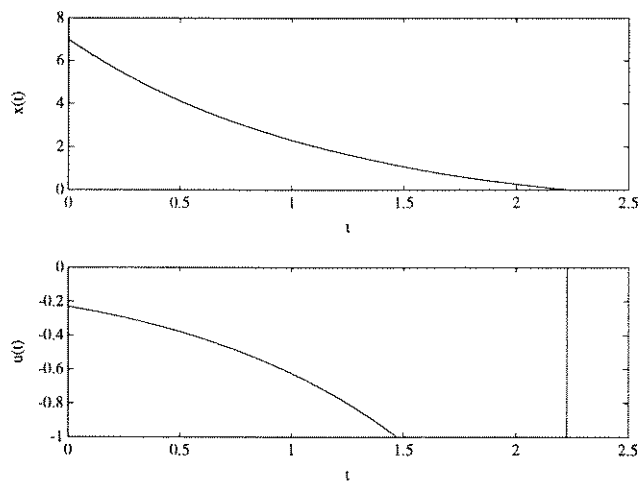


Figura 2.5: Trajetórias ótimas do estado e do controle - $v_1 = v_2 = 1$

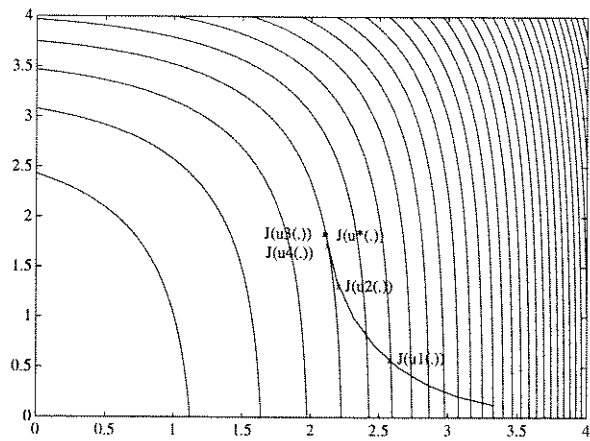


Figura 2.6: Solução no espaço dos critérios - $v_1 = 5$, $v_2 = 1$

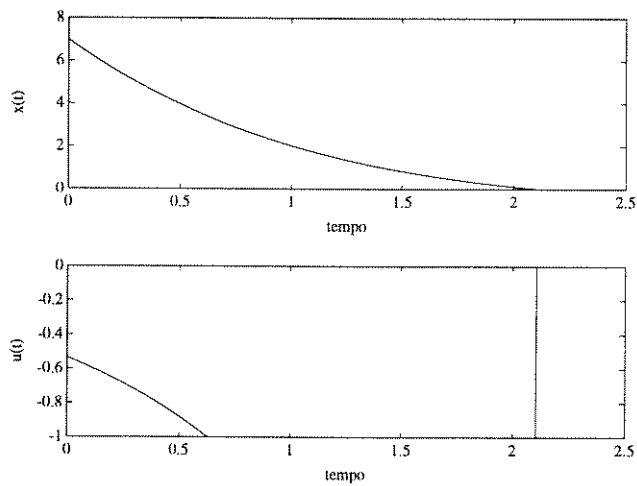


Figura 2.7: Trajetórias ótimas do estado e do controle - $v_1 = 5$, $v_2 = 1$

Tabela 2.5: Solução para $v_1 = 5$, $v_2 = 1$.

k	y_1	y_2	w_1	w_2	$J_1(u)$	$J_2(u)$	$\Theta(y)$
0	2.0786	0	0.5343	0.4657	2.5789	0.5660	0.5308
1	2.0786	1.1400	0.7845	0.2155	2.2063	1.3065	0.1361
2	2.0786	1.7715	0.8857	0.1143	2.1055	1.8342	0.0310
3	2.1246	1.6860	0.8652	0.1348	2.1226	1.7123	0.0018
4	2.1146	1.7637	0.8857	0.1143	2.1055	1.8342	$-2.3 \cdot 10^{-13}$

para ilustrar o caso geral. De acordo com este método, quando $\mathcal{V}(\cdot)$ não é conhecida explicitamente, suas derivadas parciais são substituídas pelas chamadas *taxas marginais de substituição* e o procedimento evolui com a assistência do decisor. Veja Geoffrion (1972) para maiores detalhes sobre a implementação, neste caso.

2.7 Conclusão

A abordagem proposta possui grande aplicabilidade e, como foi observado anteriormente, permite o aproveitamento de características do problema para aprimorar o seu desempenho. Outra grande vantagem é a capacidade do Algoritmo Básico de se adequar a diversos tipos de problemas, sejam estes diferenciáveis, não-diferenciáveis, com funções valores explícitas ou implícitas.

Através do uso adequado dos resultados provenientes da programação convexa e manipulações básicas tais como projeção e relaxação, foi possível decompor o problema multicritério em uma estrutura bi-nível, em que tanto os aspectos de decisão quanto os subproblemas de controle podem ser convenientemente manipulados. Esta decomposição em níveis permite que a tomada de decisão seja sempre feita através de um problema de otimização de características bastante favoráveis.

Apresentou-se um tratamento geral de problemas de controle multicritério, estendendo-se os trabalhos de Medanić e Andjelić (1971, 1972), associados a sistemas de controle min-max lineares-quadráticos, para problemas com funcionais convexos agregados em uma função valor $\mathcal{V}(\cdot)$ arbitrária. No caso de problemas convexos, evita-se o uso das estratégias baseadas no aumento do número de estados, como proposto por Salukvadze (1982).

O próximo passo será o estudo de problemas de controle ótimo que envolvam perturbações sobre a dinâmica do sistema.

Capítulo 3

Controle Multicritério Estocástico

3.1 Introdução

As próximas seções tratarão do problema de controle linear-quadrático gaussiano multicritério (LQGM), onde uma solução de compromisso para o sistema dinâmico linear estocástico é obtida minimizando-se uma *função valor* (implícita ou explícita), que agrega um dado conjunto de critérios de desempenho (Carvalho e Ferreira, 1995b).

O problema LQGM tem recebido atenção especial na literatura de controle multicritério. Em Looze et al. (1983) é proposta uma solução minimax para o problema e em Toivonen (1984), uma ordenação hierárquica dos critérios quadráticos é usada para gerar uma relação de compromisso entre a minimização das variâncias dos estados e os critérios de desempenho padrão do regulador. O método do ponto ideal de Salukvadze foi estendido para o problema LQGM discreto por Koussoulas e Leondes (1984) e um procedimento para calcular uma *solução eficiente* ou *não inferior* (Zadeh, 1963) para o problema LQGM foi desenvolvido por Toivonen e Mäkilä (1989). Mais recentemente, as pesquisas tem sido direcionadas para a solução dos problemas LQGM em espaços de Hilbert através da introdução de vetores de critérios com normas $\mathcal{H}_2$ e $\mathcal{H}_\infty$ (Khargonekar e Rotea, 1991).

A solução do problema LQGM apresentada neste capítulo segue a mesma metodologia vista no Capítulo anterior.

3.2 Formulação do Problema

Considere o seguinte problema de controle multicritério estocástico

$$\begin{aligned} \min_{u(\cdot) \in \mathcal{U}} \mathcal{V}(J(u(\cdot))) \\ \dot{x} = Ax(t) + Bu(t) + v(t) \\ y = Cx(t) + \psi(t) \end{aligned} \tag{3.1}$$

onde v e ψ são vetores aleatórios definidos na Seção 1.8 e $J(\cdot) : \mathcal{U} \rightarrow \mathcal{F} \subset \mathbb{R}^m$ é um vetor formado por m funcionais quadráticos do tipo

$$J_i(u(\cdot)) = \lim_{t \rightarrow \infty} \mathcal{E} \{x(t)' Q_i x(t) + u(t)' R_i u(t)\}, \quad i = 1, 2, \dots, m, \quad (3.2)$$

onde $Q_i = Q'_i \geq 0$ e $R_i = R'_i > 0$. Finalmente, $\mathcal{V}(\cdot) : \mathcal{F} \rightarrow \mathbb{R}$ é uma *função valor* contínua, por hipótese crescente em relação a cada componente de $J(\cdot)$. As características de (3.1) tornam o problema perfeitamente adequado para ser resolvido através do Algoritmo Básico apresentado na Seção 2.3. Neste caso, vale notar que apesar do sistema estar sujeito a ruídos, o nível de decisão não muda, sendo ainda representado por um problema relaxado e estático, em condições muito favoráveis para ser resolvido. Além disso, todas as demais vantagens metodológicas, já discutidas quando da aplicação do algoritmo a sistemas determinísticos, continuam válidas para o LQGM. Portanto, a aplicação do Algoritmo Básico ao problema LQGM vai requerer considerações apenas no nível de análise, que serão discutidas na próxima seção.

3.3 Solução do Nível de Análise

Considere o Algoritmo Básico aplicado ao problema (3.1), com um conjunto finito de m critérios quadráticos. A parte de minimização do Passo 2 do algoritmo (Seção 2.3) assume a forma

$$\min_{u(\cdot)} \lim_{t \rightarrow \infty} E \{x(t)' Q x(t) + u(t)' R u(t)\} \quad (3.3)$$

$$\begin{aligned} \dot{x} &= Ax(t) + Bu(t) + v(t) \\ y &= Cx(t) + \psi(t) \end{aligned}$$

$$Q = \sum_{i=1}^m w_i Q_i; \quad R = \sum_{i=1}^m w_i R_i; \quad w \in W$$

Note que para cada vetor w , (3.3) consiste em um problema LQG escalar cuja solução foi apresentada na Seção 1.8. Para o caso do problema do estimador ótimo, tem-se o filtro de *Kalman-Bucy*, cuja matriz de ganho K_f é dada por

$$K_f = SC' \Psi^{-1} \quad (3.4)$$

onde S é a matriz de covariância para o erro de estimativa, $S = E\{\tilde{x}' \tilde{x}\}$, que satisfaz a equação matricial algébrica de Riccati dada por (1.74).

O regulador ótimo possui como matriz de ganho em malha fechada

$$K_c(w) = R^{-1} B' P(w) \quad (3.5)$$

onde $P(w)$ satisfaz a equação matricial algébrica de Riccati

$$A' P(w) + P(w) A - P(w) B R^{-1} B' P(w) + Q = 0 \quad (3.6)$$

O valor ótimo de (3.3) é

$$\text{Tr}\{P(w) K_f \Psi K_f\} + \text{Tr}\{SQ\} \quad (3.7)$$

O Princípio da Separação desempenha um importante papel na solução do problema LQG multicritério. Note que a solução do problema do estimador ótimo não depende do vetor w . Portanto, o filtro de *Kalman-Bucy* pode ser calculado fora do laço principal do Algoritmo Básico. Em vista destes resultados, a propriedade a seguir torna-se evidente.

Proposição 3.1 *Os elementos do conjunto de controles eficientes $\mathcal{U}^*$ de (3.3) são caracterizados por*

$$u(w, t) = -K_c(w)\hat{x}(t) \quad (3.8)$$

onde $K_c(w)$ é dado por (3.5) e $\hat{x}(t)$ é gerado pelo estimador ótimo caracterizado pelo filtro com matriz de ganho (3.4). Além disso, o valor ótimo de (3.3) é dado por (3.7).

Para resolver o problema de maximização do Passo 2, é necessário calcular os valores individuais de cada funcional, para que se possa obter o gradiente generalizado de $\phi(w)$ (ver Algoritmo Básico). Desde que (3.7) não fornece nenhuma informação sobre a contribuição de cada critério $J_i(\cdot)$, $i = 1, 2, \dots, m$, é necessário resolver as m equações de Riccati acopladas (Li, 1990b)

$$\begin{aligned} A'P_i(w) + P_i(w)A - P(w)BR^{-1}B'P_i(w) - P_i(w)BR^{-1}B'P(w) + \\ + P(w)BR^{-1}R_iR^{-1}B'P(w) + Q_i = 0; \end{aligned} \quad (3.9)$$

onde $P(w) = \sum_i^m w_i P_i(w)$. Então, os valores correspondentes aos funcionais são dados por

$$J_i(u(w, \cdot))^* = \text{Tr}\{P_i(w)K_f\Psi K_f\} + \text{Tr}\{SQ_i\}, \quad i = 1, 2, \dots, m \quad (3.10)$$

3.4 Um Exemplo Ilustrativo

Considere o modelo linearizado de um helicóptero CH-47 em movimento horizontal, submetido a um deslocamento de ar nominal de aproximadamente 40 nós (Dorato et al. (1995), p. 114):

$$\begin{aligned} \dot{x} &= Ax(t) + Bu(t) + v \\ z &= Cx(t) + \psi \end{aligned}$$

onde

$$\begin{aligned} A &= \begin{bmatrix} -0.02 & 0.005 & 2.4 & -32 \\ -0.14 & 0.44 & -1.3 & -30 \\ 0 & 0.018 & -1.6 & 1.2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ B &= \begin{bmatrix} 0.14 & -0.12 \\ 0.36 & -8.6 \\ 0.35 & 0.009 \\ 0 & 0 \end{bmatrix} \\ C &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 57.3 \end{bmatrix} \end{aligned}$$

Além disso, assuma que $\Upsilon = BB'$ e $\Psi = I_{2 \times 2}$. Por simplicidade, o funcional original

$$J(u) = \lim_{t \rightarrow \infty} E \{x(t)'C'Cx(t) + u(t)'u(t)\}; \quad (Q = C'C; \quad R = I_{2 \times 2})$$

foi decomposto em três funcionais

$$J_1(u) = \lim_{t \rightarrow \infty} E \{x(t)'C'_1C_1x(t)\};$$

$$J_2(u) = \lim_{t \rightarrow \infty} E \{x(t)'C'_2C_2x(t)\};$$

$$J_3(u) = \lim_{t \rightarrow \infty} E \{u(t)'u(t)\}$$

onde $C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$. De acordo com esta decomposição, os funcionais J_1 , J_2 e J_3 estão associados com a regulação da velocidade vertical (z_1), atitude de pitch (z_2) e o esforço de controle, respectivamente.

O algoritmo foi implementado usando-se rotinas *MATLAB*. A função valor $\mathcal{V}(\cdot)$ usada para agregar os funcionais foi a norma l_2 (norma Euclideana), i.e.,

$$\mathcal{V}(J(u(\cdot))) = \|J(u(\cdot))\|_2$$

No espaço dos critérios, o problema do Passo 1 do Algoritmo Básico assume a forma

$$\min_{y \in \mathcal{Y}^k} \|y\|_2,$$

muito mais simples que a do problema original.

O critério de convergência adotado foi $\Theta(y^k) < 10^{-2}$. A Tabela 1 ilustra a performance do Algoritmo Básico e a Figura 1 descreve o seu comportamento.

Tabela 3.1: Helicóptero CH-47: Solução do exemplo a cada iteração.

k	y_1	y_2	y_3	w_1	w_2	w_3	$J_1(u)$	$J_2(u)$	$J_3(u)$
0	0	0	0	0.26	0.31	0.43	14.29	14.30	14.21
1	10.52	12.54	17.69	0.65	0.24	0.11	11.99	13.85	19.14
2	12.96	12.49	16.28	0.42	0.33	0.25	13.30	12.88	16.58
3	13.70	13.42	15.18	0.34	0.32	0.34	13.75	13.52	15.26
4	14.08	13.74	14.73	0.29	0.32	0.38	14.06	13.79	14.76
5	13.84	13.76	14.95	0.32	0.32	0.36	13.86	13.70	15.00

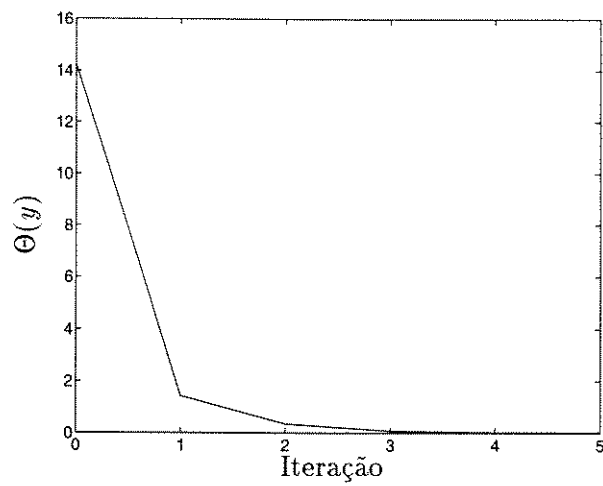


Figura 3.1: Evolução de $\Theta(y)$ a cada iteração.

3.5 Conclusão

Neste Capítulo, a abordagem multinível usada para resolver problemas de controle multicritério determinísticos foi eficientemente aplicada ao problema LQGM. Em particular, demonstrou-se que a solução do subproblema de filtragem independe do vetor de ponderação $w \in W$, o que permite tratar o problema LQGM como um problema multicritério determinístico equivalente. Com isso, fica mais uma vez constatado que a abordagem fornece um tratamento geral para problemas de controle multicritérios.

Capítulo 4

Controle Multicritério Hierárquico

4.1 Introdução

O desenvolvimento das idéias de decomposição (Dantzig e Wolfe, 1960) e de controle multinível (Mesarovic et al., 1970) tem permitido tratar problemas multicritérios envolvendo sistemas de grande porte. A importância deste tratamento é confirmada pela grande variedade de modelos e procedimentos de otimização que têm sido desenvolvidos para este tipo de sistema, através de metodologias de análise hierárquica multicritério (Li e Haimes, 1988).

Contudo, o tratamento de sistemas dinâmicos multicritérios de grande porte constitui um campo que ainda merece atenção, particularmente os chamados problemas de controle hierárquicos de grande porte não-separáveis com critérios de desempenho quadráticos.

O problema de controle de sistemas de grande porte com critérios quadráticos agregados em uma função não-linear foi resolvido por Li (1993) através de uma metodologia hierárquica de três níveis onde, da mesma forma que em Li e Haimes (1990), a otimização multicritério baseia-se numa estratégia de separação. O problema de otimização não-separável é tratado através de decomposição Lagrangeana, separável com relação a subsistemas associados aos dois níveis inferiores. No terceiro nível, um vetor ponderação é atualizado através de um algoritmo apropriado. O valor ótimo do vetor ponderação leva ao controle ótimo.

Este Capítulo aborda o mesmo problema tratado por Li (1993), porém usando uma estratégia de separação baseada no Algoritmo Básico (Oliveira et al., 1995 e 1997; Carvalho et al., 1996). O nível de decisão não se altera, enquanto que o nível de análise obtém as soluções dos subsistemas via técnicas de decomposição e coordenação.

4.2 Formulação do Problema

Considere o seguinte problema de controle ótimo de grande porte, com uma estrutura geral linear-quadrática, consistindo de n subsistemas interconectados:

$$\min_{u(\cdot) \in \mathcal{U}} \mathcal{V}(J(u(\cdot))) \quad (4.1)$$

sujeito a

$$\dot{x}^i(t) = A^i x^i(t) + B^i u^i(t) + C^i z^i(t), \quad x^i(0) = x_0^i, \quad \forall i \in \{1, 2, \dots, n\} \quad (4.2)$$

$$z^i(t) = \sum_{j=1}^n D_{ij} x^j(t), \quad \forall i \in \{1, 2, \dots, n\} \quad (4.3)$$

onde $x^i : [0, T] \rightarrow \mathbb{R}^{n_i}$, $u^i : [0, T] \rightarrow \mathbb{R}^{p_i}$, $z^i : [0, T] \rightarrow \mathbb{R}^{q_i}$ representam o estado, o controle e a entrada de interação do subsistema i , respectivamente, para cada $i \in \{1, 2, \dots, n\}$; T é o horizonte de tempo considerado, $A^i \in \mathbb{R}^{n_i \times n_i}$, $B^i \in \mathbb{R}^{n_i \times p_i}$, $C^i \in \mathbb{R}^{n_i \times q_i}$ e $D_{ij} \in \mathbb{R}^{q_i \times n_j}$, $\forall i, j \in \{1, 2, \dots, n\}$. Finalmente, $J_k : \mathcal{U} \rightarrow \mathbb{R}$, $k = 1, 2, \dots, m$, são os critérios de desempenho do sistema, os quais possuem a forma quadrática:

$$J_k(u(t)) = \sum_{i=1}^n \frac{1}{2} \|x^i(T)\|_{S_k^i}^2 + \int_0^T \left(\frac{1}{2} \|x^i(t)\|_{Q_k^i}^2 + \frac{1}{2} \|u^i(t)\|_{R_k^i}^2 \right) dt \quad (4.4)$$

e $\mathcal{U}$ é o conjunto convexo de controles admissíveis. Observe que $x(t) = (x^1(t), \dots, x^n(t))$ e $z(t) = (z^1(t), \dots, z^n(t))$ são unicamente determinados por $u(t) = (u^1(t), \dots, u^n(t))$; $S_k^i, Q_k^i \in \mathbb{R}^{n_i \times n_i}$ são matrizes semidefinidas positivas e $R_k^i \in \mathbb{R}^{p_i \times p_i}$ é uma matriz definida positiva, para todo $i \in \{1, 2, \dots, n\}$ e todo $k \in \{1, 2, \dots, m\}$.

Finalmente, $\mathcal{V}$ é a função valor que agrega os critérios quadráticos, por hipótese contínua, não-linear e crescente em relação a cada componente de $y \in \mathcal{F} \triangleq J(\mathcal{U}) \subseteq \mathbb{R}^m$.

O uso de técnicas de otimização multicritério para abordar o problema de controle de grande porte não-separável (4.1) é possível devido ao seguinte resultado.

Proposição 4.1 *Uma solução ótima $u^*(\cdot)$ de (4.1) é também uma solução eficiente para o problema*

$$\begin{aligned} & \min_{u(\cdot) \in \mathcal{U}} J(u(\cdot)) \\ & \text{s.a.} \quad \begin{cases} (4.2) \\ (4.3) \end{cases} \end{aligned} \quad (4.5)$$

Prova: A existência de $u^*(\cdot)$ é assegurada pela continuidade de $J(\cdot)$ e $\mathcal{V}(\cdot)$. O fato de u^* ser uma solução eficiente de (4.5) é uma consequência imediata da hipótese de monotonicidade de $\mathcal{V}(\cdot)$. $\square$

O resultado acima mostra que $u^*(\cdot)$ será encontrado dentre as soluções eficientes de (4.5). O conjunto das soluções eficientes de (4.5) é denotado por $\mathcal{U}^*$. Assim como nos casos anteriores, devido à hipótese de convexidade de (4.5), sabe-se que $\mathcal{U}^*$ pode ser gerado pelo problema ponderado

$$\begin{aligned} & \min_{u(\cdot) \in \mathcal{U}} \langle w, J(u(\cdot)) \rangle \\ & \text{s.a.} \quad \begin{cases} (4.2) \\ (4.3) \end{cases} \end{aligned} \quad (4.6)$$

Uma formulação equivalente ao problema original é obtida no espaço dos critérios através de técnicas de projeção e relaxação, o que naturalmente leva à aplicação do Algoritmo Básico a esta classe de problemas.

4.3 Estratégia de Decomposição

Do ponto de vista da classe de problemas aqui tratados, a abordagem através do Algoritmo Básico apresenta duas grandes facilidades metodológicas. Primeiro, o problema no nível de análise (Passo 2) possui estrutura aditiva, e portanto, separável com relação a cada subsistema. Segundo, a característica linear-quadrática fica preservada, o que permite a aplicação do método interação-predição (Singh e Titli, 1978) para resolver o problema hierárquico, por exemplo.

O problema de controle ótimo no nível de análise consiste em

$$\begin{aligned} \min_{u(\cdot) \in \mathcal{U}} & \quad \langle w^{l+1}, J(u(\cdot)) \rangle \\ \text{s.a.} & \quad \begin{cases} (4.2) \\ (4.3) \end{cases} \end{aligned} \quad (4.7)$$

onde

$$\begin{aligned} \langle w^{l+1}, J(u) \rangle &= \sum_{k=1}^m w_k^{l+1} J_k(u) \\ &= \sum_{k=1}^m w_k^{l+1} \left[\sum_{i=1}^n \frac{1}{2} \|x^i(T)\|_{S_k^i}^2 + \int_0^T \left(\frac{1}{2} \|x^i(t)\|_{Q_k^i}^2 + \frac{1}{2} \|u^i(t)\|_{R_k^i}^2 \right) dt \right] \end{aligned}$$

Logo, a função Lagrangeana de (4.7) é

$$\begin{aligned} L(x, u, z, \lambda, \pi) &= \sum_{k=1}^m w_k^{l+1} \left[\sum_{i=1}^n \frac{1}{2} \|x^i(T)\|_{S_k^i}^2 + \int_0^T \left(\frac{1}{2} \|x^i(t)\|_{Q_k^i}^2 + \frac{1}{2} \|u^i(t)\|_{R_k^i}^2 \right) dt \right] + \\ &\quad \sum_{i=1}^n \int_0^T \left[\lambda^i(t)' \left(z^i(t) - \sum_{j=1}^n D_{ij} x^j(t) \right) + \pi^i(t)' \left(A^i x^i(t) + B^i u^i(t) + c^i z^i(t) - \dot{x}(t) \right) \right] dt \end{aligned} \quad (4.8)$$

onde $(\lambda^i(t), \pi^i(t)) \in \mathbb{R}^{q_i} \times \mathbb{R}^{n_i}$ são os vetores de co-estados, sendo $\lambda^i(t)$ um vetor de multiplicadores de Lagrange de dimensão q_i e $\pi^i(t)$ um vetor adjunto de dimensão n_i , para todo $i \in \{1, 2, \dots, n\}$.

Para que se possam construir os subproblemas, deve-se usar o fato de que, devido a estrutura do sistema, a função Lagrangeano (4.8) pode ser rearranjada em uma soma de n Lagrangeanos, cada um correspondente a um subsistema:

$$L(x, u, z, \lambda, \pi) = \sum_{i=1}^n L_i(x^i, u^i, z^i, \lambda, \pi^i) \quad (4.9)$$

onde $\lambda(t) = (\lambda^1(t), \dots, \lambda^n(t))$, $\pi(t) = (\pi^1(t), \dots, \pi^n(t))$, e

$$L(x^i, u^i, z^i, \lambda, \pi^i) = \sum_{k=1}^m w_k^{l+1} \left[\frac{1}{2} \|x^i(T)\|_{S_k^i}^2 + \int_0^T \left(\frac{1}{2} \|x^i(t)\|_{Q_k^i}^2 + \frac{1}{2} \|u^i(t)\|_{R_k^i}^2 \right) dt \right] + \int_0^T \left[\lambda^i(t)' z^i(t) - \sum_{j=1}^n \lambda^j(t)' D_{ji} x^i(t) + \pi^i(t)' \left(A^i x^i(t) + B^i u^i(t) + c^i z^i(t) - \dot{x}(t) \right) \right] dt \quad (4.10)$$

Logo, aproveitando-se a estrutura, obtêm-se n subproblemas interconectados associados aos subsistemas, cada um tendo a forma linear-quadrática. Observe que o *vetor de coordenação* é $(\lambda(t), z^i(t))$, e com isso, de acordo com o primeiro nível do método interação-predição (encontrar um controle u^i que satisfaz (4.2) e (4.3), $\forall i \in \{1, 2, \dots, n\}$), assume-se $(\lambda(t), z^i(t))$ como conhecido em (4.8). Neste caso, cada um dos subproblemas pode ser resolvido separadamente.

O Hamiltoniano do subsistema i é

$$H_i(x^i, u^i, z^i, \lambda, \pi^i) = \sum_{k=1}^m w_k^{l+1} \left(\frac{1}{2} \|x^i(t)\|_{Q_k^i}^2 + \frac{1}{2} \|u^i(t)\|_{R_k^i}^2 \right) + \lambda^i(t)' z^i(t) - \sum_{j=1}^n \lambda^j(t)' D_{ji} x^i(t) + \pi^i(t)' \left(A^i x^i(t) + B^i u^i(t) + C^i z^i(t) \right)$$

e as condições necessárias derivadas do Princípio do Mínimo para que u^i seja a solução ótima do subproblema i são:

a) Condições sobre o estado:

$$\dot{\pi}^i(t) = -\frac{\partial H_i}{\partial x^i} = -\sum_{k=1}^m w_k^{l+1} Q_k^i x^i(t) - \pi^i(t)' A^i + \sum_{j=1}^n \lambda^j(t)' D_{ji} \quad (4.11)$$

$$\pi^i(T) = \sum_{k=1}^m w_k^{l+1} S_k^i x^i(T) \quad (4.12)$$

b) Condições sobre o co-estado:

$$\frac{\partial H_i}{\partial \pi^i} = \dot{x}^i(t), \quad x^i(0) = x_0^i \quad (4.13)$$

c) Condições sobre o controle:

$$\frac{\partial H_i}{\partial u^i} = \sum_{k=1}^m w_k^{l+1} R_k^i u^i(t) + (B^i)' \pi^i(t) = 0 \quad (4.14)$$

De (4.14) obtém-se $u^i(t) = -\left(\sum_{k=1}^m w_k^{l+1} R_k^i\right)^{-1} (B^i)' \pi^i(t)$, que substituído em (4.11)-(4.13) gera um problema de valor de contorno linear acoplado por $x^i(t)$ e $\pi^i(t)$ desconhecidos, o qual pode ser desacoplado introduzindo-se

$$\pi^i(t) = K^i(t)x^i(t) + g^i(t), \quad (4.15)$$

onde $g^i(t)$ é um vetor adjunto de dimensão n_i .

Diferenciando-se ambos os lados de (4.15) e substituindo-se (4.11) e (4.13), após algumas manipulações, obtém-se a equação de Riccati

$$\dot{K}^i(t) = -K^i(t)A^i - (A^i)'K^i(t) + K^i(t)B^i \left(\sum_{k=1}^m w_k^{l+1} R_k^i\right)^{-1} (B^i)' K^i(t) - \sum_{k=1}^m w_k^{l+1} Q_k^i \quad (4.16)$$

e a equação adjunta

$$\dot{g}^i(t) = \left[K^i(t)B^i \left(\sum_{k=1}^m w_k^{l+1} R_k^i\right)^{-1} (B^i)' - (A^i)'\right] g^i(t) - K^i(t)C^i z^i(t) + \sum_{j=1}^n D'_{ji} \lambda^j(t) \quad (4.17)$$

com as condições de contorno, que surgem a partir de (4.12) e (4.15),

$$K^i(T) = \sum_{k=1}^m \lambda^k(T)' S_k^i \quad (4.18)$$

$$g^i(T) = 0 \quad (4.19)$$

respectivamente. O controle ótimo torna-se portanto,

$$u^i(w^{l+1}) = -\left(\sum_{k=1}^m w_k^{l+1} R_k^i\right)^{-1} (B^i)' [K^i(t)x^i(t) + g^i(t)] \quad (4.20)$$

O segundo nível do método consiste em atualizar o vetor de coordenação. Os valores de $\lambda(t)$ e $z^i(t)$ podem ser obtidos a partir de

$$\frac{\partial L}{\partial z^i} = \lambda^i(t) + (C^i)' \pi^i(t) = 0 \quad (4.21)$$

$$\frac{\partial L}{\partial \lambda^i} = z^i(t) - \sum_{j=1}^n D_{ij} x^j(t) = 0 \quad (4.22)$$

$\forall i \in \{1, 2, \dots, n\}$. O problema de controle ótimo do nível de análise do Algoritmo Básico é, portanto, resolvido através do procedimento de dois níveis do método interação-predição, da seguinte forma:

Passo 1:

1.1 (Inicialização) Escolha o vetor coordenador inicial: $\lambda^0(t)$, $z_0^i(t)$, para todo $i \in \{1, 2, \dots, n\}$, faça $s = 0$.

1.2: Resolva para cada um dos n subsistemas:

- a) a equação (4.16) com a condição final (4.18), obtendo $K^{i,s}$
- b) a equação (4.17) com a condição final (4.19), obtendo $g^{i,s}$
- c) a equação

$$\begin{aligned} \dot{x}^i(t) = & \left[A^i - B^i \left(\sum_{k=1}^m w_k^{l+1} R_k^i \right)^{-1} (B^i)' K^{i,s}(t) \right] x^i(t) - \\ & B^i \left(\sum_{k=1}^m w_k^{l+1} R_k^i \right)^{-1} (B^i)' g^{i,s}(t) + C^i z^i(t) \end{aligned} \quad (4.23)$$

com $x^i(0) = x_0^i$, obtendo $x^{i,s}(t)$.

Passo 2:

2.1: Atualize o vetor coordenador através de (4.21) e de (4.22)

$$\lambda^{i,s+1}(t) = -(C^i)' \pi^{i,s}(t) = -(C^i)' \left[K^{i,s}(t) x^{i,s}(t) + g^{i,s}(t) \right] \quad (4.24)$$

$$z^{i,s+1} = \sum_{j=1}^n D_{ij} x^{j,s}(t), \quad \forall i \in \{1, \dots, n\} \quad (4.25)$$

2.2: Avaliação do erro global da iteração, onde Δt é o passo de integração. Se

$$\sum_{i=1}^n \int_0^T \left(z^{i,s+1}(t) - z^{i,s}(t) \right)' \left(z^{i,s+1}(t) - z^{i,s}(t) \right) dt / \Delta t < \varepsilon \quad (4.26)$$

para $\varepsilon > 0$ arbitrariamente pequeno, faça

$$u^i(w^{l+1}) = - \left(\sum_{k=1}^m w_k^{l+1} R_k^i \right)^{-1} (B^i)' \left[K^{i,s}(t) x^{i,s}(t) + g^{i,s}(t) \right], \quad (4.27)$$

para todo $i \in \{1, 2, \dots, n\}$. Caso contrário, faça $s = s + 1$ e volte ao Passo 1.2

A convergência deste algoritmo foi estudada em Singh e Hassan (1976). Observe que o procedimento apresentado, na verdade, é uma extensão para o caso de uma soma ponderada de m critérios de desempenho (i.e., (4.7)) da abordagem iterativa preditiva tratada na literatura.

4.4 Exemplo

Para ilustrar a metodologia proposta, considere o mesmo exemplo de Li (1993). O problema de controle ótimo com estrutura linear-quadrática considerado tem dois subsistemas interconectados:

$$\min_{u \in \mathcal{U}} \mathcal{V}(J(u)) = (J_1(u))^2 J_2(u) + (J_2(u))^{1.5}$$

sujeito a

$$\begin{bmatrix} \dot{x}_1^1(t) \\ \dot{x}_2^1(t) \end{bmatrix} = \begin{bmatrix} 2 & 0.1 \\ 0.2 & -1 \end{bmatrix} \begin{bmatrix} x_1^1(t) \\ x_2^1(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0.1 \end{bmatrix} u^1(t) + \begin{bmatrix} 0.01 & 0 \\ 0.1 & -0.5 \end{bmatrix} \begin{bmatrix} z_1^1(t) \\ z_2^1(t) \end{bmatrix}$$

com $x^1(0) = (x_1^1(0), x_2^1(0)) = (-1, 0.1)$,

$$\begin{bmatrix} \dot{x}_1^2(t) \\ \dot{x}_2^2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0.05 \\ -0.25 & -1.2 \end{bmatrix} \begin{bmatrix} x_1^2(t) \\ x_2^2(t) \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix} u^2(t) + \begin{bmatrix} 0.05 & 0.15 \\ 0 & -0.2 \end{bmatrix} \begin{bmatrix} z_1^2(t) \\ z_2^2(t) \end{bmatrix}$$

com $x^2(0) = (x_1^2(0), x_2^2(0)) = (1, -0.5)$,

$$\begin{bmatrix} z_1^1(t) \\ z_2^1(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1^2(t) \\ x_2^2(t) \end{bmatrix}$$

$$\begin{bmatrix} z_1^2(t) \\ z_2^2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1^1(t) \\ x_2^1(t) \end{bmatrix}$$

Os dois critérios de desempenho são

$$\begin{aligned} J_1(u(t)) &= \frac{1}{2} \int_0^1 \left(\begin{bmatrix} x_1^1(t) & x_2^1(t) \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1^1(t) \\ x_2^1(t) \end{bmatrix} + u^1(t)[1]u^1(t) \right) dt + \\ &\quad \frac{1}{2} \int_0^1 \left(\begin{bmatrix} x_1^2(t) & x_2^2(t) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1^2(t) \\ x_2^2(t) \end{bmatrix} + u^2(t)[2]u^2(t) \right) dt \end{aligned} \quad (4.28)$$

$$\begin{aligned} J_2(u(t)) &= \frac{1}{2} \begin{bmatrix} x_1^1(t) & x_2^1(t) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1^1(t) \\ x_2^1(t) \end{bmatrix} + \frac{1}{2} \int_0^1 u^1(t)[2]u^1(t) dt + \\ &\quad \frac{1}{2} \begin{bmatrix} x_1^2(t) & x_2^2(t) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1^2(t) \\ x_2^2(t) \end{bmatrix} + \frac{1}{2} \int_0^1 u^2(t)[1]u^2(t) dt \end{aligned} \quad (4.29)$$

Observe que, de acordo com (4.1), $n = 4$, $n_1 = n_2 = 2$, $p_1 = p_2 = 1$, $q_1 = q_2 = 2$, $m = 2$, D_{11} , D_{22} , S_1^1 , S_2^1 , Q_1^2 e Q_2^2 são matrizes nulas de segunda ordem e $y^0 = \underline{y} = (0, 0)$.

O problema foi resolvido para os mesmos vetores iniciais π e z usados em Li (1993). O procedimento converge quando $\Theta(y) \leq 10^{-3}$, o que ocorreu na 7ª iteração. A Figura 4.1 apresenta tanto a forma 3-D da função valor $\mathcal{V}(y) = (y_1)^2 y_2 + (y_2)^{1.5}$, quanto o espaço dos critérios (o plano $y_1 \times y_2$) com as curvas de contorno de $\mathcal{V}(y)$ e de trade-off. A Figura 4.2 mostra as primeiras quatro iterações do procedimento no espaço dos critérios. É possível observar utilização dos hiperplanos suportes para aproximar a curva de trade-off e encontrar a solução.

Note que a complexidade da função valor $\mathcal{V}(\cdot)$ é grandemente reduzida quando o procedimento de otimização é transferido do espaço das variáveis de controle

$$\mathcal{V}(J(u)) = (J_1(u))^2 J_2(u) + (J_2(u))^{1.5}$$

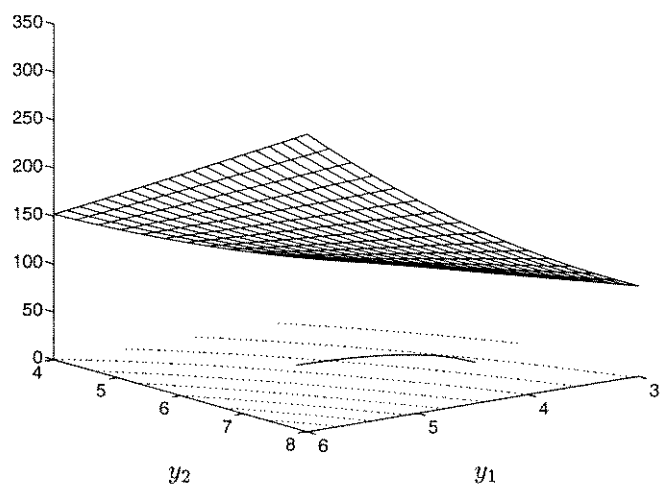


Figura 4.1: Geometria de $\mathcal{V}(y)$.

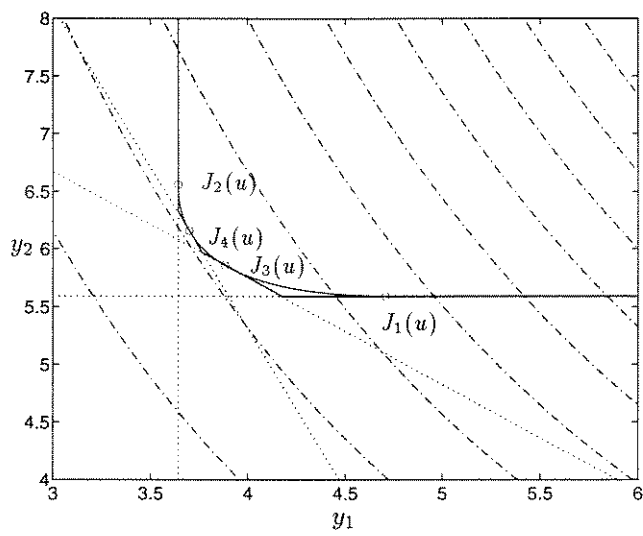


Figura 4.2: Quatro iterações no espaço dos critérios.

Tabela 4.1: Controle Hierárquico: Solução do exemplo a cada iteração.

k	w_1	w_2	$J_1(u)$	$J_2(u)$	y_1	y_2	Θ
1	0	1.0000	4.7088	5.5893	0	0	5.58929
2	1.0000	0	3.6413	6.5576	0.0000	5.5893	3.64134
3	0.4799	0.5201	3.8817	5.8568	3.6413	5.5893	0.25447
4	0.7361	0.2639	3.6972	6.1559	3.6413	6.0786	0.06154
5	0.6069	0.3931	3.7724	5.9936	3.7663	5.9633	0.01566
6	0.6712	0.3288	3.7307	6.0716	3.7343	6.0524	0.00388
7	0.7036	0.2964	3.7129	6.1130	3.7185	6.0965	0.00097

($J_1(u)$ e $J_2(u)$ dados por (4.28) e (4.29)), para o espaço dos critérios

$$\mathcal{V}(y) = (y_1)^2 y_2 + (y_2)^{1.5}$$

A Tabela 4.1 mostra os valores de w , J_1 , J_2 , y_1 , y_2 e Θ a cada iteração. A Figura 4.3 ilustra como $\mathcal{V}(y)$ e $\Theta(y)$ decrescem com as sucessivas iterações.

A Figura 4.4 mostra a evolução das variáveis de estado e de controle para os dois subsistemas. O comportamento das variáveis de estado reflete o horizonte de tempo finito ($T = 1$) e a ausência de restrições sobre o estado final.

4.5 Conclusão

Neste Capítulo abordou-se uma importante classe de problemas de grande porte não-separáveis. A metodologia hierárquica proposta, baseada no uso apropriado de conceitos de programação convexa, permitiu reduzir consideravelmente a complexidade do problema. O aspecto fundamental explorado pela abordagem foi o de que, embora não-separável, o problema tem sua solução no conjunto de soluções eficientes Γ . Neste caso, o problema pode ser resolvido através da geração do valor adequado do vetor de ponderação $w \in W$. Esta propriedade viabiliza o emprego de métodos de coordenação especialmente desenvolvidos para tratar problemas com estrutura aditiva, em problemas que em princípio não apresentam esta estrutura. A partir da solução do problema com estrutura linear-quadrática, atestou-se a facilidade de implementação e a eficiência do procedimento.

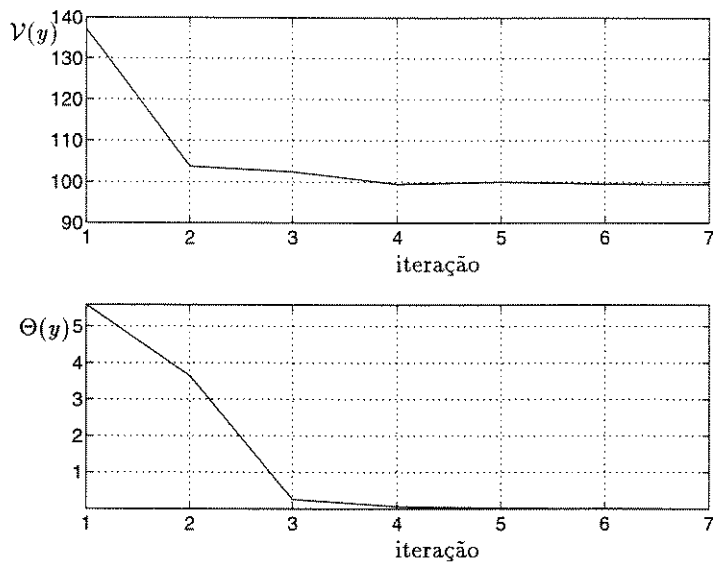


Figura 4.3: Evolução de $V(y)$ e $\Theta(y)$.

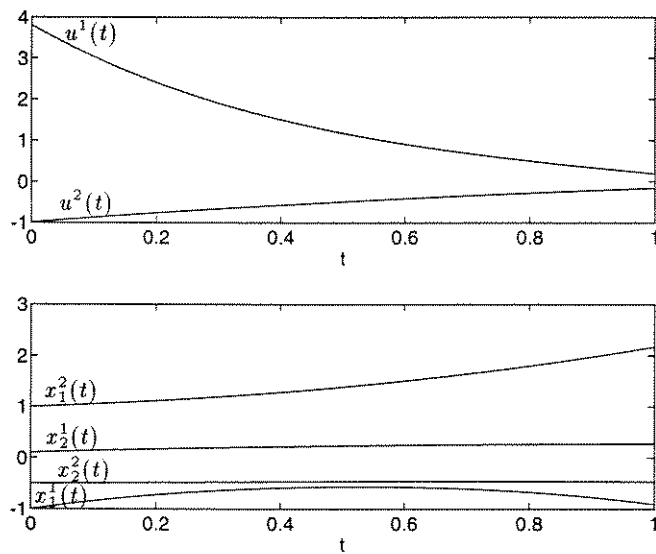


Figura 4.4: Evolução do controle e dos estados de $t = 0$ a $t = 1$.

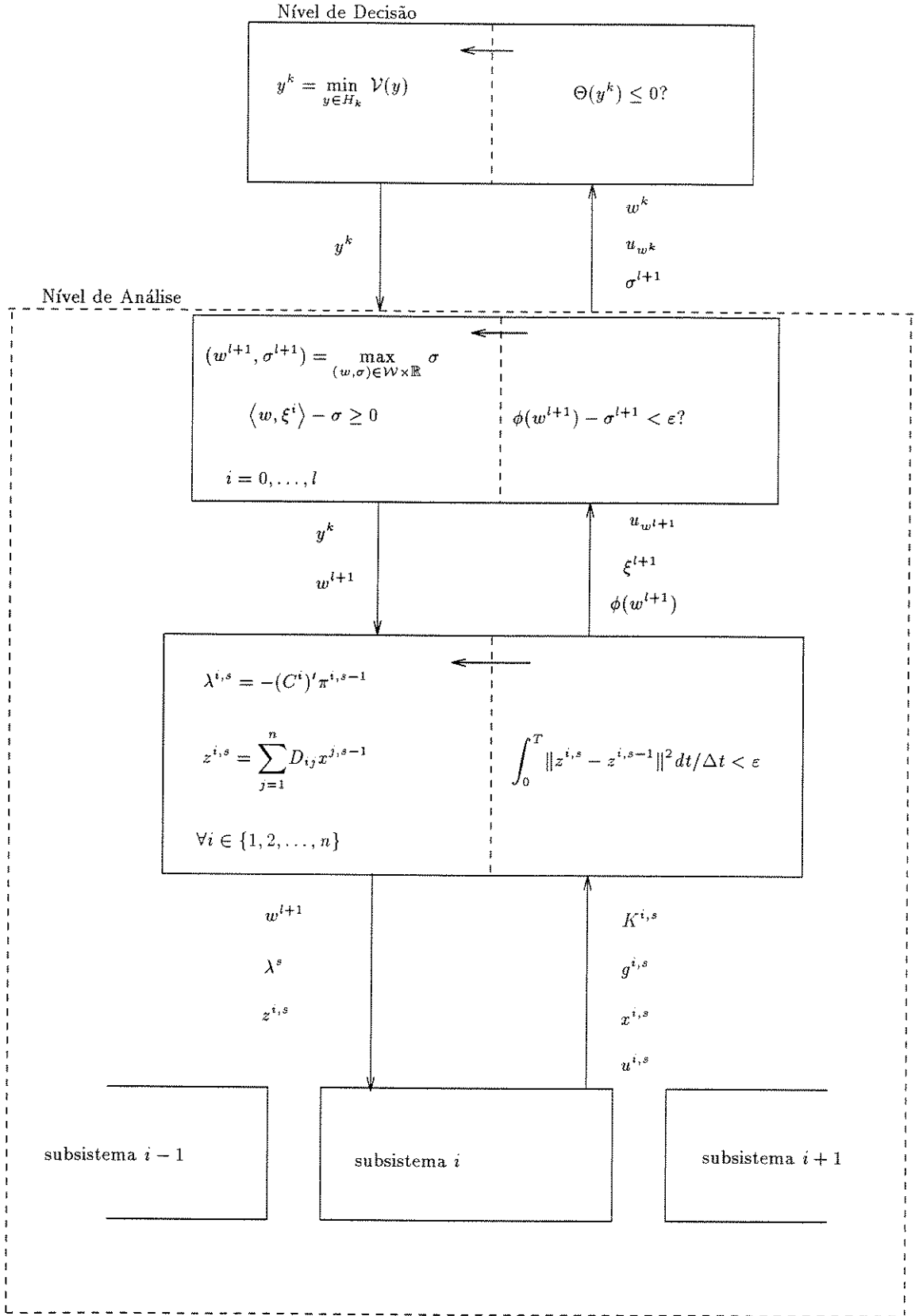


Figura 4.5: Diagrama do método.

Parte II

**Controle no Domínio
da Frequência**

Capítulo 5

Fundamentos Teóricos

5.1 Introdução

Neste Capítulo inicia-se um estudo sobre a aplicabilidade da abordagem apresentada no Capítulo anterior para o caso de projeto de controladores sujeitos a um conjunto de especificações de desempenho. Para uma discussão mais extensa ver Boyd e Barrat (1991).

5.2 Arquitetura do Sistema

Por conveniência, todos os sistemas estudados serão representados a partir da arquitetura padrão da Figura 5.1, onde P reúne os parâmetros da planta; as entradas e saídas v , u , y e z são definidas como segue.

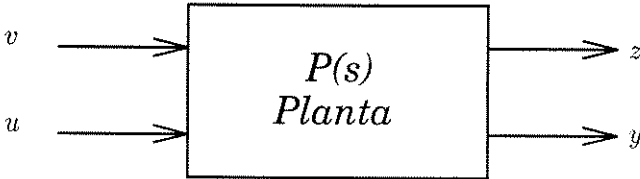


Figura 5.1: Arquitetura do Sistema.

Entradas

- $u \in \mathbb{R}^{n_u}$ representa todas as entradas do sistema que são manipuladas pelo controlador, ou seja, u é o sinal gerado pelo controlador do sistema;
- $v \in \mathbb{R}^{n_v}$ representa as entradas exógenas do sistema. A partir de v são representadas todas as perturbações e ruídos, além de qualquer entrada de comando que atue diretamente sobre o sistema.

Saídas

- $y \in \mathbb{R}^{n_y}$ representa os sinais que são acessíveis ao controlador, incluindo sinais de sensores e entradas de comando sobre o sistema;
- $z \in \mathbb{R}^{n_z}$ incorpora todos os sinais de saída que possuam algum interesse para o projeto, como por exemplo, as variáveis do sistema que estão sendo controladas.

Como este Capítulo trata de controle no domínio da frequência, as seguintes suposições são necessárias para garantir a representação do sistema através de matrizes de transferência.

Hipóteses

1. Os sinais v , u , z e y são representados por vetores e cada componente destes vetores é uma função real e contínua do tempo, ou seja:

$$v : \mathbb{R}^+ \rightarrow \mathbb{R}^{n_v}, \quad z : \mathbb{R}^+ \rightarrow \mathbb{R}^{n_z}, \quad u : \mathbb{R}^+ \rightarrow \mathbb{R}^{n_u}, \quad y : \mathbb{R}^+ \rightarrow \mathbb{R}^{n_y};$$

2. Os sistemas tratados são lineares, invariantes no tempo e estacionários;
3. O controlador a ser projetado é linear, invariante no tempo e estacionário.

Considerando a Figura 5.1, tem-se a seguinte matriz de transferência entre as saídas z , y e as entradas v , u :

$$\begin{bmatrix} z \\ y \end{bmatrix} = P \begin{bmatrix} v \\ u \end{bmatrix} \quad (5.1)$$

sendo

$$P = \left[\begin{array}{c|c} P_{zv} & P_{zu} \\ \hline P_{yv} & P_{yu} \end{array} \right] \quad (5.2)$$

onde cada partição representa a matriz de transferência entre as combinações de entradas e saídas da planta.

Considerando-se agora o bloco de controle incorporado ao sistema, tem-se a arquitetura da Figura 5.2. Note que a presença do bloco de controle, processando o vetor y e gerando o vetor u , indica que o sistema agora está em malha fechada. Portanto, tem-se que

$$u = Ky \quad (5.3)$$

Colocando (5.1) apenas como função de z e v , obtêm-se

$$z = H_{zv}v \quad (5.4)$$

onde

$$H_{zv} = P_{zv} + P_{zu}K(I - P_{yu}K)^{-1}P_{yv} \quad (5.5)$$

é a matriz de transferência em malha fechada entre z e v . Assume-se que $\det(I - P_{yu}K) \neq 0$ para garantir a existência de $(I - P_{yu}K)^{-1}$, o que se conhece por "*well-posedness condition*".

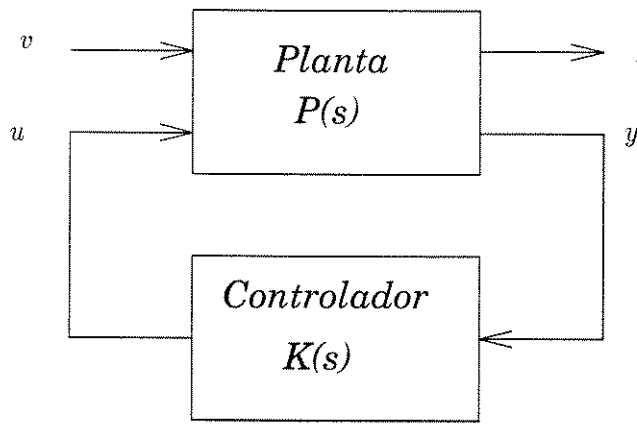


Figura 5.2: Arquitetura em malha fechada

5.3 Exemplo Ilustrativo

Para exemplificar como se processa a determinação dos vetores v , u , z e y , considere o caso do regulador clássico, que consiste em manter a saída próxima de zero através de um sinal de controle finito. A Figura 5.3 representa o regulador através do modelo em diagrama de blocos, largamente utilizado na teoria de controle, sendo que 5.3-a apresenta o modelo sem as perturbações, enquanto que 5.3-b considera perturbações no atuador do sistema η_p e no sinal η_s medido pelo sensor. A função de transferência P_0 representa a planta do sistema e K , o controlador.

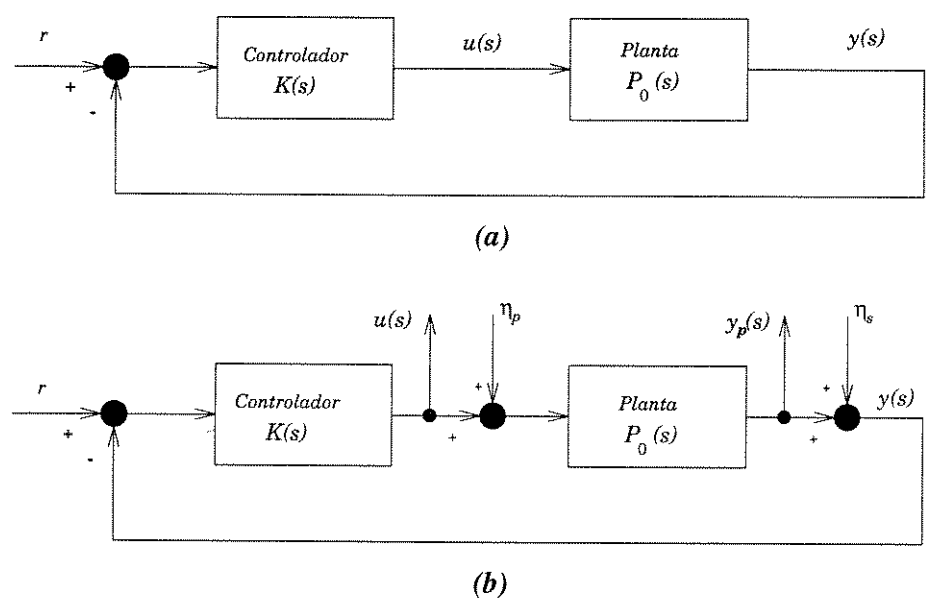


Figura 5.3: (a) Diagrama de blocos, (b) Diagrama considerando perturbações.

1. As realimentações de estados e de saída são tratadas pela arquitetura padrão como implementações em malha fechada diferentes, uma vez que os vetores y serão distintos;
2. As perturbações fazem parte da especificação do sistema e são consideradas em v ;
3. A arquitetura padrão facilita a determinação do sistema em malha fechada mais adequado, uma vez que qualquer especificação poderá ser expressa através de v e z , apenas.

5.4 Especificações de Projeto

Nesta seção serão discutidos alguns aspectos referentes ao projeto de controladores. Para tanto, será assumido que existe uma planta dada, com os vetores v , u , z e y definidos e sua matriz de transferência P_0 determinada.

Durante a realização de um projeto é necessário que sejam observadas especificações que determinarão se o resultado final é satisfatório ou não. Para o caso do projeto de sistemas de controle, tem-se como aspectos a serem considerados

- Realizabilidade;
- Estabilidade;
- Desempenho;
- Robustez.

Para que um controlador seja considerado satisfatório para uma determinada planta, a matriz de transferência em malha fechada correspondente deve satisfazer à especificação ou ao conjunto de especificações dado. Como apresentado em Boyd e Barrat (1991), especificações de projeto são definidas da seguinte maneira

Definição 5.1 *Uma especificação de projeto $\mathcal{D}$ é uma função booleana, ou um teste, sobre a matriz de transferência em malha fechada H . Portanto $\mathcal{D}$ é uma função que possui uma matriz $H_{n_z \times n_v}$ como argumento e retorna um valor booleano como resultado, ou seja:*

$$\mathcal{D} : \mathcal{H} \rightarrow (\text{passa}, \text{falha})$$

onde $\mathcal{H}$ representa o conjunto de todas as matrizes de transferências $n_z \times n_v$.

As especificações $\mathcal{D}$ podem também estar relacionadas entre si, ou seja, dadas várias especificações de projeto $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_{n_D}$ pode-se, por exemplo,

- Ordená-las parcialmente. A ordenação pode sugerir que existem especificações que são mais rígidas do que outras, isto é: se $\mathcal{D}_1$ é mais rígida do que $\mathcal{D}_2$, então todas as matrizes H que satisfizerem $\mathcal{D}_1$ satisfarão $\mathcal{D}_2$ e se existir pelo menos uma matriz que satisfaz $\mathcal{D}_2$ e não $\mathcal{D}_1$, então diz-se que $\mathcal{D}_1$ é *estritamente mais rígida* que $\mathcal{D}_2$. A ordenação é parcial porque existem especificações que não podem ser comparadas;

- Agrupá-las através de *operadores booleanos*, de forma a construir uma expressão de especificações. Exemplos:

- O operador $\wedge$ denominando satisfação conjunta ($\mathcal{D}_1 \wedge \mathcal{D}_2$);
- O operador $\vee$ denominando satisfação disjunta ($\mathcal{D}_1 \vee \mathcal{D}_2$);

O uso de operadores booleanos cria uma nova especificação. No caso de $\mathcal{D}_1 \wedge \mathcal{D}_2$ tem-se uma especificação mais rígida do que $\mathcal{D}_1$ ou $\mathcal{D}_2$.

Uma especificação de projeto é *consistente* ou *factível* se for satisfeita para pelo menos uma matriz de transferência. Caso contrário, ou seja, se a especificação não for satisfeita por nenhuma matriz de transferência, então a especificação é inconsistente ou infactível.

Um aspecto que aparece naturalmente é o *problema de factibilidade*, que consiste em verificar se um conjunto de especificações de projeto $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_{n_D}$ pode ser satisfeita para alguma matriz de transferência do sistema. Na prática, o problema de factibilidade consiste em verificar a factibilidade de apenas uma especificação, $\mathcal{D}_1 \wedge \dots \wedge \mathcal{D}_{n_D}$.

Geralmente, o uso de uma parametrização para as especificações de projeto, e não apenas a utilização de valores fixos, é uma forma mais adequada de refletir as necessidades do projeto de um controlador para uma determinada planta. Esta parametrização é feita da seguinte forma: sejam uma especificação $\mathcal{D}_1$ e um valor $a_1 \geq 0$. Então

$$\mathcal{D}_1^{(a_1)} : \text{especificação} \leq a_1$$

onde *especificação* pode ser, por exemplo, sobre-sinal de y_p devido a uma determinada entrada r .

Caso seja atribuído valores numéricos à uma família de especificações de projeto $\mathcal{D}^{(a)}$, ao invés de apenas um valor booleano (*passa* ou *falha*), tem-se então uma *especificação por inequação funcional*, que consiste em

$$\mathcal{D}_\phi^{(a)} : \phi(H) \leq a \quad (5.7)$$

onde

$$\phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\} \quad (5.8)$$

é um funcional em $H \in \mathcal{H}$, que atribui para cada H um valor real, possivelmente infinito.

Com isso, as famílias de especificações definidas desta forma podem ser linearmente ordenadas, uma vez que se $a^1 \leq a^2$ então $\mathcal{D}_\phi^{(a^1)}$ é mais rígida do que $\mathcal{D}_\phi^{(a^2)}$, e quantificadas pois, se $a^2 = 2a^1$ então $\mathcal{D}_\phi^{(a^1)}$ é duas vezes mais rígida do que $\mathcal{D}_\phi^{(a^2)}$, no sentido da medida fornecida pelo funcional ϕ .

5.5 Otimização Multicritério

Nesta seção, os conceitos de *especificação de projeto*, de *famílias de especificações* e de *funcionais* serão utilizados no escopo da otimização multicritério de sistemas.

O problema de controle multicritério com especificações de projeto consiste em satisfazer (no sentido booleano) uma especificação rígida $\mathcal{D}_R$ (que pode ser a conjunção de várias especificações)

e, ao mesmo tempo, minimizar vários critérios (funcionais ϕ) simultaneamente. Desta forma, considera-se a família de especificações parametrizadas pelas constantes $a_1, a_2, \dots, a_D$, dada por

$$\mathcal{D}^{(a_1, \dots, a_D)} : \mathcal{D}_R \wedge \mathcal{D}_{\phi_1}^{(a_1)} \wedge \dots \wedge \mathcal{D}_{\phi_D}^{(a_D)} \quad (5.9)$$

Portanto, qualquer que seja o valor de $a_1, a_2, \dots, a_D$, a especificação $\mathcal{D}^{(a_1, \dots, a_D)}$ será mais rígida do que $\mathcal{D}_R$. O objetivo é determinar valores para $a_1, a_2, \dots, a_D$ de forma que a especificação resultante esteja no limite do conjunto de especificações consistentes, de acordo com o conceito de especificação consistente visto anteriormente. Chega-se então a um conjunto de especificações que forma o limite do conjunto de especificações consistentes, que são as *especificações eficientes*.

Definição 5.2 Diz-se que uma especificação $\mathcal{D}^{(a_1, \dots, a_D)}$ é eficiente se não existir nenhuma especificação consistente $\mathcal{D}^{(\tilde{a}_1, \dots, \tilde{a}_D)}$ formada pelas constantes $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_D$ tal que $\tilde{a}_n \leq a_n, \forall n \in [1, 2, \dots, D]$ com pelo menos uma desigualdade estrita, ou seja, $\tilde{a}_n < a_n$ para algum $n \in [1, 2, \dots, D]$. A matriz de transferência em malha fechada que satisfaz a uma especificação eficiente é chamada de Matriz de Transferência Eficiente.

As especificações eficientes não podem ser comparadas entre si no sentido de uma ser mais rígida do que outra. O conjunto de especificações eficientes forma a *curva de trade-off* entre os critérios considerados.

5.6 Paradigma do Controlador Ótimo

Neste caso, a família de especificações de projeto é parametrizada por apenas uma constante a , ou seja

$$\mathcal{D}^{(a)} : \mathcal{D}_R \wedge \mathcal{D}_{\phi}^{(a)} \quad (5.10)$$

onde $\mathcal{D}_{\phi}^{(a)}$ é a especificação por inequação funcional.

No paradigma clássico de controle ótimo, as famílias de especificações podem ser linearmente ordenadas e portanto o problema consiste em encontrar o menor valor de a que torna $\mathcal{D}^{(a)}$ o mais rígida possível, mantendo é claro a factibilidade da especificação, ou seja, tem-se o problema

$$a_{\star} = \min_{H \in \mathcal{H}} \phi(H) \quad (5.11)$$

onde $\mathcal{H}$ denota o conjunto das matrizes H que satisfazem $\mathcal{D}_R$.

O funcional ϕ pode ser uma combinação de vários outros funcionais. Em particular, ϕ pode ser a soma de vários funcionais, multiplicados por constantes que atribuem relações de importância entre si, ou seja

$$\phi(H) = w_1 \phi_1(H) + \dots + w_D \phi_D(H) \quad (5.12)$$

onde $w_i \geq 0, i = 1, 2, \dots, D$.

Nas próximas seções serão revistos alguns tipos de normas usadas para medir sinais e sistemas que serão utilizados no decorrer do trabalho.

Uma norma deve satisfazer as propriedades da definição a seguir (Luenberger, 1969).

Definição 5.3 *Seja a função $\|\cdot\| : V \rightarrow \mathbb{R}$, que transforma um vetor $v \in V$ em um escalar não-negativo. Então $\|\cdot\|$ é uma norma se satisfaz*

- $\|v\| \geq 0, \forall v \in V$ (não-negatividade)
- Para $\|v\| < \infty$, $\|\alpha v\| = |\alpha|\|v\|, \forall \alpha \in \mathbb{R}$ (homogeneidade)
- $\|v + w\| \leq \|v\| + \|w\|, \forall v, w \in V$ (desigualdade triangular)

5.7 Normas de Sinais

A seguir serão relacionadas algumas normas comumente utilizadas para sinais (Boyd e Barrat, 1991; Dahleh e Diaz-Bobillo, 1995).

Pico

Esta norma nada mais é que a norma l_∞ vista na Seção 2.2 para vetores, sendo para sinais definida como

$$\|u\|_\infty \triangleq \sup_{t \geq 0} |u(t)| \quad (5.13)$$

Esta norma fornece o maior valor alcançado por $u(t)$ no intervalo de tempo considerado.

Raiz Quadrada do Valor Quadrático Médio

Esta medida é usualmente conhecida como valor *RMS* (do inglês, *root-mean-square*) e fornece o valor médio alcançado pelo sinal no intervalo de tempo considerado, sendo definida como

$$\|u\|_{rms} \triangleq \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t)^2 dt \right)^{\frac{1}{2}} \quad (5.14)$$

Se u for um vetor, a expressão fica na forma

$$\|u\|_{rms} \triangleq \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t)' u(t) dt \right)^{\frac{1}{2}} \quad (5.15)$$

Valor Médio Absoluto

Esta norma atribui menor ênfase a valores transitórios. É definida como

$$\|u\|_{ma} \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |u(t)| dt \quad (5.16)$$

Enquanto que $\|u\|_{rms}$ é mais usada como norma de sinais de energia, a norma $\|u\|_{ma}$ é mais utilizada como medida de sinais que indicam gastos de recursos ou combustíveis.

No caso em que u é um vetor com n componentes, tem-se

$$\|u\|_{ma} \triangleq \lim_{T \rightarrow \infty} \sup \frac{1}{T} \int_0^T \sum_{i=1}^n |u_i(t)| dt \quad (5.17)$$

Norma de Sinais Estocásticos

A norma de sinais estocásticos é calculada de acordo com as características do processo que gera sinal.

- (a) Se o processo estocástico for estacionário, então a norma mais utilizada é

$$\|u\|_{rms} \triangleq (\mathcal{E}[u(t)^2])^{\frac{1}{2}} \quad (5.18)$$

onde $\mathcal{E}$ é o operador *esperança matemática*, que não depende de t devido a estacionariedade. A expressão torna-se

$$\|u\|_{rms} \triangleq (\mathcal{E}[u(t)'u(t)])^{\frac{1}{2}} \quad (5.19)$$

caso u seja um vetor.

- (b) Se o processo tender a ser estacionário com o decorrer do tempo, então tem-se

$$\|u\|_{rms} \triangleq \lim_{t \rightarrow \infty} (\mathcal{E}[u(t)^2])^{\frac{1}{2}} \quad (5.20)$$

- (c) Se for considerada a autocorrelação de u

$$R_u(\tau) \triangleq \mathcal{E}[u(t)u(t+\tau)] \quad (5.21)$$

ou sua densidade espectral

$$S_u(\omega) \triangleq \int_{-\infty}^{\infty} R_u(\tau) e^{-j\omega\tau} d\tau, \quad (5.22)$$

a norma $\|u\|_{rms}$ pode ser computada como segue

$$\|u\|_{rms}^2 = R_u(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_u(\omega) d\omega, \quad (5.23)$$

No caso em que u é um sinal estocástico estacionário ergódico tem-se

$$\|u\|_{rms}^2 = \mathcal{E}\|u(t)\|_2^2 = \text{Tr} R_u(0) = \text{Tr} \frac{1}{2\pi} \int_{-\infty}^{\infty} S_u(\omega) d\omega, \quad (5.24)$$

onde, para vetores, a autocorrelação é dada por

$$R_u(\tau) \triangleq \mathcal{E} u(t)u(t+\tau)' \quad (5.25)$$

Para estes sinais, tem-se que

$$\|u\|_{rms} = \left(\sum_{i=1}^n \|u_i\|_{rms}^2 \right)^{\frac{1}{2}} \quad (5.26)$$

(d) A norma absoluta média para sinais estocásticos estacionários é definida como

$$\|u\|_{ma} \triangleq \mathcal{E}|u(t)| \quad (5.27)$$

a qual, para sinais ergódigos, é equivalente em valor à definição determinística.

Quando um sinal decai para zero com o tempo, as normas $\|u\|_{rms}$ e $\|u\|_{ma}$ possuem valores estocásticos nulos. Para medir a norma destes sinais são usadas as normas L_2 e L_1 definidas como segue.

Norma L_2

$$\|u\|_2 \triangleq \left(\int_0^\infty u(t)^2 dt \right)^{\frac{1}{2}} \quad (5.28)$$

que representa a energia total do sinal.

Se u for um vetor, tem-se

$$\|u\|_2 \triangleq \left(\int_0^\infty \sum_{i=1}^n u_i(t)^2 dt \right)^{\frac{1}{2}} = \left(\sum_{i=1}^n \|u_i(t)\|_2^2 \right)^{\frac{1}{2}} \quad (5.29)$$

Norma L_1

$$\|u\|_1 \triangleq \int_0^\infty |u(t)| dt \quad (5.30)$$

que mede o consumo total de recursos ou combustível do sinal, e que para vetores assume a forma

$$\|u\|_1 \triangleq \int_0^\infty \sum_{i=1}^n |u_i(t)| dt = \sum_{i=1}^n \|u_i(t)\|_1 \quad (5.31)$$

5.8 Normas de Sistemas

A seguir serão consideradas medidas de normas para a matriz de transferência entre z e v , H_{zv} .

Norma L_1

A norma L_1 , também conhecida como o ganho de pico-a-pico máximo, é definida como

$$\begin{aligned} \|H\|_1 &\triangleq \sup_{\|v\|_\infty \neq 0} \frac{\|Hv\|_\infty}{\|v\|_\infty} \\ &= \max_{1 \leq i \leq n_z} \int_0^\infty \sum_{j=1}^{n_v} |h_{ij}(t)| dt \end{aligned} \quad (5.32)$$

onde h_{ij} é a resposta ao impulso do elemento H_{ij} de H .

Resposta RMS para um Ruído Particular

Sejam H uma matriz de transferência estável e S_v a matriz densidade espectral de potência de v . A resposta RMS devido ao ruído v é dada por

$$\|H\|_{rms,v} = \left(\text{Tr} \frac{1}{2} \int_{-\infty}^{\infty} H(j\omega) S_v(j\omega) H(j\omega)^* d\omega \right)^{\frac{1}{2}} \quad (5.33)$$

Norma H_2

Mede a resposta RMS, considerando uma entrada tipo ruído branco, isto é $S_v = I$:

$$\|H\|_2 = \left(\text{Tr} \frac{1}{2} \int_{-\infty}^{\infty} H(j\omega) H(j\omega)^* d\omega \right)^{\frac{1}{2}} \quad (5.34)$$

Esta norma é expressa no tempo como

$$\begin{aligned} \|H\|_2 &= \left(\text{Tr} \int_0^{\infty} h(t) h(t)' dt \right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^{n_z} \sum_{k=1}^{n_v} \|H_{ik}\|_2^2 \right)^{\frac{1}{2}} \end{aligned} \quad (5.35)$$

onde h é a resposta ao impulso de H .

Norma H_{∞}

A norma H_{∞} é dada por

$$\|H\|_{\infty} \triangleq \sup_{\text{Re}\{s\} \geq 0} \sigma_{max}(H(s)) \quad (5.36)$$

onde σ_{max} é o valor singular máximo de H . Neste sentido, para que $\|H\|_{\infty} < \infty$, a matriz de transferência H deve ser estável, e neste caso

$$\|H\|_{\infty} = \sup_{\omega} \sigma_{max}(H(j\omega)) \quad (5.37)$$

5.9 Especificações de Projeto

Em problemas de otimização, algumas características dos funcionais e restrições envolvidas são desejáveis para que a busca da solução seja eficiente. Dentre estas características, uma que torna o problema mais tratável é a *convexidade*, uma vez que esta garante que a solução encontrada é realmente a melhor solução para o problema.

Para que o problema de controle envolvendo especificações de projeto seja adequadamente formulado, e para que se possa avaliar as consequências da convexidade e de outras características sobre o problema, as especificações de projeto devem ser representadas na forma de conjuntos.

Dada uma especificação de projeto $\mathcal{D}_i$, o conjunto associado será o conjunto $\mathcal{H}_i \subset \mathcal{H}$ de matrizes de transferência em malha fechada que a satisfazem, isto é

$$\mathcal{H}_i = \{H \in \mathcal{H} : H \text{ satisfaz } \mathcal{D}_i\} \quad (5.38)$$

Com isso, a qualquer relação entre especificações de projeto corresponde uma relação entre conjuntos. Alguns exemplos são apresentados na Tabela 5.1.

Tabela 5.1: Exemplos de especificações via conjuntos.

Especificação de Projeto	Conjunto de Matrizes de Transferência
H satisfaz $\mathcal{D}_1$	$H \in \mathcal{H}_1$
$\mathcal{D}_1$ é mais rígida do que $\mathcal{D}_2$	$\mathcal{H}_1 \subseteq \mathcal{H}_2$
$\mathcal{D}_1$ é mais fraca do que $\mathcal{D}_2$	$\mathcal{H}_1 \supseteq \mathcal{H}_2$
$\mathcal{D}_1 \wedge \mathcal{D}_2$	$\mathcal{H}_1 \cap \mathcal{H}_2$
$\mathcal{D}_1$ é infactível (ou inconsistente)	$\mathcal{H}_1 = \emptyset$
$\mathcal{D}_1$ é factível (ou consistente)	$\mathcal{H}_1 \neq \emptyset$
$\mathcal{D}_1$ é estritamente mais rígida do que $\mathcal{D}_2$	$\mathcal{H}_1 \subseteq \mathcal{H}_2, \mathcal{H}_1 \neq \mathcal{H}_2$

A seguir demonstra-se que um grande número de especificações de projeto, quando convenientemente formuladas em termos de matrizes de transferência, são convexas ou afins. Algumas destas especificações são aplicadas a partições específicas de H_{zv} de acordo com a seguinte divisão dos elementos dos vetores v e z :

$$v = \begin{bmatrix} v_c \\ v_d \\ v_{etc} \end{bmatrix} \quad \begin{array}{l} \text{comandos} \\ \text{perturbações} \\ \text{demais componentes de } v \end{array}$$

$$z = \begin{bmatrix} z_c \\ z_a \\ z_o \\ z_{etc} \end{bmatrix} \quad \begin{array}{l} \text{variáveis comandadas} \\ \text{sinais dos atuadores} \\ \text{outros sinais críticos} \\ \text{demais componentes de } z \end{array}$$

Neste sentido, tem-se as seguintes partições da matriz H_{zv} relacionadas às divisões apresentadas para z e v .

$$z = \begin{bmatrix} H_{cc} & H_{cd} & * \\ H_{ac} & H_{ad} & * \\ H_{oc} & H_{od} & * \\ * & * & * \end{bmatrix} v$$

5.9.1 Realizabilidade

Uma matriz de transferência H realizável deve pertencer ao conjunto

$$\mathcal{H}_{real} \triangleq \{H : H = P_{zv} + P_{zu}K(I - P_{yu}K)^{-1}P_{yv}, \text{ para algum } K\}. \quad (5.39)$$

A partir da seguinte troca de variável:

$$R = K(I - P_{yu}K)^{-1}, \text{ onde } R \in \mathbb{R}^{n_u \times n_y} \quad (5.40)$$

obtem-se uma representação afim de (5.39) em R .

$$\mathcal{H}_{real} = \{H : H = P_{zv} + P_{zu}RP_{yv}, \text{ para algum } R \in \mathbb{R}^{n_u \times n_y}\} \quad (5.41)$$

Vale a pena notar que R e K possuem uma relação bi-unívoca. Logo, a representação em (5.41) não degenera a correspondência entre H (matriz de transferência) e K (controlador), uma vez que, encontrado H e R basta fazer

$$K = (I + RP_{yu})^{-1}R, \quad (5.42)$$

para obter K .

Agora, supondo-se duas matrizes $H, \tilde{H} \in \mathcal{H}_{real}$, existem duas matrizes $R, \tilde{R} \in \mathbb{R}^{n_u \times n_y}$ tais que

$$H = P_{zv} + P_{zu}RP_{yv}$$

$$\tilde{H} = P_{zv} + P_{zu}\tilde{R}P_{yv}$$

Então, qualquer que seja $w \in \mathbb{R}$, a matriz de transferência H_w dada por

$$H_w = wH + (1 - w)\tilde{H},$$

será tal que

$$H_w = P_{zv} + P_{zu}R_wP_{yv}$$

onde

$$R_w = wR + (1 - w)\tilde{R}$$

Portanto, $H_w \in \mathcal{H}_{real}$ e, conseqüentemente, $\mathcal{H}_{real}$ é um conjunto afim. Para encontrar K_w que realiza a matriz de transferência em malha fechada, basta fazer

$$K_w = (A + wB)^{-1}(C + wD) \quad (5.43)$$

onde

$$\begin{aligned} A &= I + \tilde{K}(I - P_{yu}\tilde{K})^{-1}P_{yu}, \\ B &= K(I - P_{yu}K)^{-1}P_{yu} - \tilde{K}(I - P_{yu}\tilde{K})^{-1}P_{yu}, \\ C &= \tilde{K}(I - P_{yu}\tilde{K})^{-1}, \\ D &= K(I - P_{yu}K)^{-1} - \tilde{K}(I - P_{yu}\tilde{K})^{-1} \end{aligned}$$

5.9.2 Estabilidade

Além da realizabilidade, uma outra especificação primordial para o projeto de um sistema de controle é a estabilidade, ou seja, a matriz de transferência em malha fechada a ser determinada deve pertencer ao conjunto

$$\mathcal{H}_{estvl} \triangleq \{H : H = P_{zv} + P_{zu}K(I - P_{yu}K)^{-1}P_{yv}, \text{ para algum } K \text{ que estabilize } P\}. \quad (5.44)$$

Pode-se mostrar, de uma maneira semelhante à usada para $\mathcal{H}_{reavl}$, que $\mathcal{H}_{estvl}$ também é um conjunto afim, ou seja, se K e $\tilde{K}$ estabilizam P , então K_w , dado por (5.43), também estabiliza P .

5.9.3 Desempenho

Dado um conjunto $\mathcal{H}_{estvl}$ de matrizes de transferência estáveis em malha fechada, ainda é necessário estabelecer especificações que dirão qual o comportamento desejado do sistema com relação aos comandos e perturbações atuantes. Estas especificações são chamadas especificações de desempenho e a seguir serão apresentados alguns exemplos.

Especificações de Entrada/Saída

As especificações a seguir referem-se ao elemento H_{cc} de H .

Overshoot e Undershoot: Definidos pelos funcionais

$$\begin{aligned} \phi_{ov}(H_{cc}) &\triangleq \sup_{t \geq 0} s(t) - 1 \\ \phi_{ud}(H_{cc}) &\triangleq \sup_{t \geq 0} -s(t) \end{aligned}$$

onde $s(t)$ é a resposta ao degrau de H_{cc} . Estes funcionais são convexos, como também o são as especificações

$$\begin{aligned} \mathcal{H}_{ov} &\triangleq \{H : \phi_{ov} \leq \alpha\} \\ \mathcal{H}_{ud} &\triangleq \{H : \phi_{ud} \leq \alpha\} \end{aligned}$$

Se dadas duas respostas ao degrau, nenhuma excede 10% ($\alpha = 0.1$) de overshoot, então a média das duas respostas, por exemplo, também não excederá.

Tempo de Subida e de Estabilização: Definidos, por exemplo¹, pelos funcionais

$$\begin{aligned} \phi_{tsub}(H_{cc}) &\triangleq \inf\{T : s(t) > 0.8 \text{ para } t \geq T\} \\ \phi_{test}(H_{cc}) &\triangleq \inf\{T : |s(t) - 1| < 0.05 \text{ para } t \geq T\} \end{aligned}$$

¹Existem várias definições de tempo de subida e de estabilização em uso.

que são *funcionais quasi-convexos*. Porém as especificações por inequações funcionais

$$\mathcal{H}_{tsub} \triangleq \{H : \phi_{tsub} \leq T_{max}\}$$

$$\mathcal{H}_{test} \triangleq \{H : \phi_{test} \leq T_{max}\}$$

são especificações convexas.

Especificações de Regulação

As especificações de regulação referem-se à partição H_{cd} de H_{zv} . Ou seja, a submatriz de transferência das entradas de perturbações para as saídas referentes às variáveis comandadas. Logo, procura-se manter o valor de H_{cd} pequeno. As especificações de regulação são do tipo

$$\|H_{cd}\| \leq \alpha \quad (5.45)$$

onde $\|\cdot\|$ é qualquer norma apropriada. Como exemplo de especificação de regulação tem-se a regulação RMS.

Regulação RMS: Definida como

$$\mathcal{H}_{rgrms} \triangleq \{H : \|H_{cd}\Omega\|_2 \leq \alpha\} \quad (5.46)$$

onde Ω é um fator espectral de densidade espectral de potência conhecida. Portanto esta especificação limita o desvio *RMS* da variável comandada a um valor menor do que α .

Esforço do Atuador

Devido a várias razões (aquecimento excessivo, saturação, gasto de energia) qualquer sistema de controle deve manter a magnitude do sinal do atuador limitada, ou seja,

$$\|u\|_{atd} \leq L_{max} \quad (5.47)$$

onde $\|\cdot\|_{atd}$ é alguma norma apropriada e L_{max} é o limite máximo permitido para o sinal.

5.9.4 Comentários

Nesta seção foram discutidos apenas alguns aspectos a serem observados no projeto de um controlador. Discutiu-se a formulação de algumas especificações relacionadas à realizabilidade, estabilidade e desempenho. Não foram exemplificadas especificações relacionadas com sensibilidade e robustez do sistema, pois seria necessário uma discussão mais aprofundada sobre normas de sistemas. O objetivo foi fazer uma introdução para a próxima seção, que vai tratar da formulação do problema de controle.

5.10 Formulação do Problema

Como foi visto na Seção 5.5, a formulação de um problema de controle multicritério a partir de especificações de projeto consiste em, satisfeita uma determinada especificação rígida $\mathcal{D}_R$, encontrar um controlador K que minimize um conjunto de funcionais ϕ relacionados a diversas outras especificações.

Como parte de $\mathcal{D}_R$ pode-se imaginar, por exemplo, que antes de qualquer coisa, o controlador a ser determinado deve fornecer uma matriz de transferência em malha fechada realizável. Porém, de nada adianta um controlador que não seja estabilizante. Portanto, uma matriz de transferência em malha fechada estável é uma condição básica para qualquer projeto. É fácil perceber que a especificação H_{estvl} é mais rígida do que H_{reavl} (ver Seções 5.9.1 e 5.9.2), ou seja, $\mathcal{H}_{estvl} \subseteq \mathcal{H}_{reavl}$. Logo, na composição de $\mathcal{D}_R$, a especificação H_{estvl} deve obrigatoriamente ser levada em conta, pois nela está embutida a realizabilidade e estabilidade do sistema em malha fechada. Supondo que todas as outras especificações que compõem $\mathcal{D}_R$ estejam relacionadas no conjunto $\mathcal{H}_D$ tem-se o seguinte problema de otimização

$$\min_{H_{zv} \in \mathcal{H}} \Phi(H_{zv}) \quad (5.48)$$

onde $\mathcal{H} = \mathcal{H}_{estvl} \cap \mathcal{H}_D$ e $\Phi(H_{zv})$ é composto pelos funcionais (relacionados às especificações) que se deseja minimizar.

Um fator complicador na solução de (5.48) é a relação entre H_{zv} e o controlador K (ver equações (5.5) e (5.44)). Note que trata-se de uma relação do tipo linear-fracional, e conseqüentemente, qualquer restrição sobre H_{zv} , por mais simples que seja, se refletirá em uma restrição bem mais complexa em K .

Por esta razão, o problema é modificado a partir do conceito de *parametrização* Q (Vidyasagar, 1985). Em linhas gerais a parametrização Q consiste em fazer a busca sobre uma determinada matriz Q (parâmetro livre). O conjunto $\mathcal{H}_{estvl}$ é caracterizado em função de Q da seguinte forma:

$$\mathcal{H}_{estvl} = \{H_{zv} = T_1 + T_2 Q T_3 : Q \in \mathbb{R}^{n_u \times n_y} \text{ é uma matriz estável}\} \quad (5.49)$$

onde T_1 , T_2 e T_3 são matrizes de transferência estáveis, derivadas da planta a partir de *fatoração coprima* (Vidyasagar, 1985) de um controlador nominal que estabilize a planta, K_{nom} .

É possível mostrar (Vidyasagar, 1985) que T_1 , T_2 e T_3 dependem da planta e que T_1 nada mais é do que H_{zv} encontrado para o controlador K_{nom} . A Figura 5.5 mostra como fica a arquitetura básica com a presença de Q no sistema. A essência da parametrização está no fato de que, como Q é estável e K_{nom} estabiliza H , a matriz resultante da introdução de Q também será estável. De fato, pode-se fazer com que a matriz Q não *enxergue* nenhuma realimentação e portanto não desestabilize a planta. Neste caso, variando-se Q sobre o conjunto $\mathcal{RH}_\infty$ das matrizes racionais próprias estáveis, através da expressão

$$H_{zv} = T_1 + T_2 Q T_3 \quad (5.50)$$

é feito um mapeamento a partir de Q de todo o conjunto $\mathcal{H}_{estvl}$.

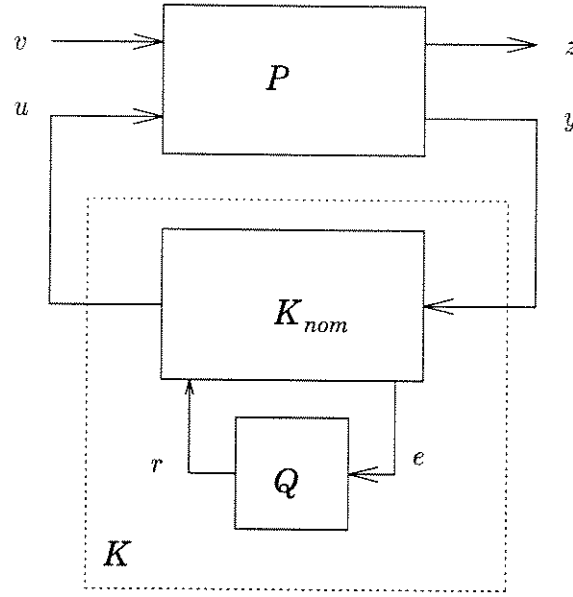


Figura 5.5: Arquitetura incluindo a matriz Q

Consequentemente, o problema formulado em termos da matriz Q

$$\min_{Q \in \mathcal{Q}} \Phi(Q) \quad (5.51)$$

onde

$$\mathcal{Q} \triangleq \{Q : T_1 + T_2 Q T_3 \in \mathcal{H}_D, Q \in \mathcal{RH}_\infty\},$$

é agora um problema em que a relação entre a variável de otimização Q e o conjunto de busca $\mathcal{Q}$ é afim. Como foi visto na Seção 5.9, muitos critérios de desempenho são convexos, e com isso o problema pode ser tratado através da abordagem via programação convexa introduzida no Capítulo 2. Tendo-se encontrado a matriz Q solução do problema, o controlador K é obtido facilmente.

5.11 Parametrização de Youla

Na seção anterior, a parametrização Q para o projeto de controladores sujeito a especificações de projeto apenas foi introduzida. Nesta seção será mostrado em que consiste Q -parametrizar todos os controladores estabilizantes de uma determinada planta. Em seguida, obtêm-se a parametrização Q a partir da sua representação em espaço de estados, que é necessária para viabilizar a implementação computacional do procedimento, uma vez que a solução através da representação através de matrizes de transferência mostra-se de difícil tratamento numérico.

Vale a pena lembrar que, do ponto de vista da abordagem multinível proposta, está sendo tratada apenas a solução do nível de análise, uma vez que é neste nível que se concentram as características dinâmicas do problema considerado.

5.11.1 Resultados Preliminares

Torna-se necessário apresentar alguns conceitos algébricos relevantes (Francis, 1987).

Matrizes Coprimas

Sejam G_1, G_2 elementos do espaço formado por todas as matrizes de transferências próprias e estáveis, $\mathcal{RH}_\infty$, e que possuem o mesmo número de colunas. Então G_1, G_2 são *matrizes coprimas à direita* se existem matrizes X, Y também pertencentes a $\mathcal{RH}_\infty$, tais que

$$\begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = XG_1 + YG_2 = I \quad (5.52)$$

o que é equivalente a dizer que a matriz $\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$ é *invertível à esquerda* em $\mathcal{RH}_\infty$.

Do mesmo modo, sejam $G_1, G_2 \in \mathcal{RH}_\infty$ com mesmo número de linhas. Então G_1, G_2 são *matrizes coprimas à esquerda* se existem matrizes $X, Y \in \mathcal{RH}_\infty$, tais que

$$\begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = G_1X + G_2Y = I \quad (5.53)$$

ou seja, $\begin{bmatrix} G_1 & G_2 \end{bmatrix}$ é *invertível à direita* em $\mathcal{RH}_\infty$.

Fatoração Coprima

Considere agora uma matriz P racional própria com coeficientes reais. Então P pode ser fatorada na forma

$$P = NM^{-1} \quad (5.54)$$

onde N e M são matrizes coprimas à direita em $\mathcal{RH}_\infty$. Esta fatoração é chamada de *fatoração coprima à direita* de P . Existem também matrizes $\tilde{M}$ e $\tilde{N}$, coprimas à esquerda, tais que,

$$P = \tilde{M}^{-1}\tilde{N} \quad (5.55)$$

que definem a *fatoração coprima à esquerda* de P .

Observe que como M e N são coprimas à direita, então existem matrizes $X, Y \in \mathcal{RH}_\infty$ que satisfazem (5.52) e, como $\tilde{M}$ e $\tilde{N}$ são coprimas à esquerda, então existe outro par de matrizes, $\tilde{X}, \tilde{Y} \in \mathcal{RH}_\infty$ que satisfazem (5.53). Para cada matriz racional P existem oito matrizes em $\mathcal{RH}_\infty$ que satisfazem as identidades

$$P = NM^{-1} = \tilde{M}^{-1}\tilde{N} \quad (5.56)$$

$$\begin{bmatrix} \tilde{X} & -\tilde{Y} \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & Y \\ N & X \end{bmatrix} = I \quad (5.57)$$

e que constituem a *fatoração duplamente coprima* de P .

5.12 Estabilidade

O objetivo desta seção é mostrar a utilidade dos conceitos de fatoração coprima para a caracterização da estabilidade da matriz de transferência de uma determinada planta P a partir de um controlador K , ambas por hipótese matrizes racionais próprias com coeficientes reais.

Teorema 5.1 *Sejam*

$$P = NM^{-1} = \tilde{M}^{-1}\tilde{N} \quad (5.58)$$

$$K = UV^{-1} = \tilde{V}^{-1}\tilde{U} \quad (5.59)$$

fatorações coprimas à direita e à esquerda de P e K , respectivamente. Então, afirmar que K estabiliza P é equivalente a afirmar que

$$\begin{bmatrix} M & \begin{bmatrix} 0 \\ I \end{bmatrix} U \\ \begin{bmatrix} 0 & I \end{bmatrix} N & V \end{bmatrix} \in \mathcal{RH}_\infty \quad (5.60)$$

ou que

$$\begin{bmatrix} \tilde{M} & \tilde{N} \begin{bmatrix} 0 \\ I \end{bmatrix} \\ \tilde{U} \begin{bmatrix} 0 & I \end{bmatrix} & \tilde{V} \end{bmatrix} \in \mathcal{RH}_\infty \quad (5.61)$$

Prova: Francis (1987). □

Uma consequência imediata deste resultado é a obtenção de um teste que indica se K é um controlador que estabiliza P .

Teorema 5.2 *Seja P uma matriz de transferência racional própria com coeficientes reais. Então P é estabilizável se e somente se*

$M, \begin{bmatrix} 0 & I \end{bmatrix} N$ são coprimas a direita e

$M, \begin{bmatrix} 0 \\ I \end{bmatrix}$ são coprimas a esquerda.

ou se

$\tilde{M}, \tilde{N} \begin{bmatrix} 0 \\ I \end{bmatrix}$ são coprimas a esquerda e

$\tilde{M}, \begin{bmatrix} 0 & I \end{bmatrix}$ são coprimas a direita.

Prova: Francis (1987). □

Note que para que P seja estabilizável, seus pólos instáveis devem ser alcançados pelo controlador. Como os pólos instáveis de P devem estar contidos na partição P_{yu} , pode-se mostrar que se K estabiliza P_{yu} , então K estabiliza P .

5.13 Parametrização dos Controladores Estabilizantes

Uma vez estabelecidas as condições para que uma determinada matriz de transferência P seja estabilizável e as condições para que um controlador K seja estabilizante, o próximo passo será parametrizar o conjunto de todos os controladores estabilizantes a partir do conceito de fatoração coprima.

Sabe-se que é necessário trabalhar apenas sobre a partição P_{yu} . Seja a fatoração duplamente coprima de P_{yu}

$$P_{yu} = N_{yu} M_{yu}^{-1} = \tilde{M}_{yu}^{-1} \tilde{N}_{yu} \quad (5.62)$$

$$\begin{bmatrix} \tilde{X}_{yu} & -\tilde{Y}_{yu} \\ -\tilde{N}_{yu} & \tilde{M}_{yu} \end{bmatrix} \begin{bmatrix} M_{yu} & Y_{yu} \\ N_{yu} & X_{yu} \end{bmatrix} = I \quad (5.63)$$

e as fatorações coprimas de K dadas por (5.59). Pode-se então estabelecer o seguinte resultado, consequência do Teorema 5.1.

Lema 5.1 K estabiliza P_{yu} se e somente se

$$\begin{bmatrix} M_{yu} & U \\ N_{yu} & V \end{bmatrix} \in \mathcal{RH}_\infty \quad (5.64)$$

ou

$$\begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N}_{yu} & \tilde{M}_{yu} \end{bmatrix} \in \mathcal{RH}_\infty \quad (5.65)$$

□

A partir deste resultado e das matrizes que fatoram P e K , obtêm-se a parametrização de todos os controladores que estabilizam P_{yu} e, consequentemente, P .

Teorema 5.3 O conjunto de matrizes racionais próprias K que estabilizam P_{yu} está parametrizado pela matriz de matriz de transferência Q através das expressões

$$\begin{aligned} K &= (Y_{yu} - M_{yu}Q)(X_{yu} - N_{yu}Q)^{-1} \\ &= (\tilde{X}_{yu} - Q\tilde{N}_{yu})^{-1}(\tilde{Y}_{yu} - Q\tilde{M}_{yu}) \end{aligned} \quad (5.66)$$

para $Q \in \mathcal{RH}_\infty$

□

Considere as fatorações de P_{yu} dadas por (5.63) e a Q -parametrização de K dada por (5.66) e defina

$$\begin{aligned} T_1 &\triangleq P_{yu} + P_{yu} M_{yu} \tilde{Y}_{yu} P_{yu} \\ T_2 &\triangleq P_{yu} M_{yu} \\ T_3 &\triangleq \tilde{M}_{yu} P_{yu} \end{aligned} \quad (5.67)$$

Teorema 5.4 Com K dado por (5.66) e com T_1, T_2 e T_3 dados por (5.67), a matriz de transferência em malha fechada de v para z é dada por

$$H_{zv} = T_1 + T_2 Q T_3 \quad (5.68)$$

Prova: Francis (1987). □

Este resultado é de fundamental importância para a solução do problema via programação convexa, pois, ao contrário do que ocorria com a caracterização do conjunto de todas as matrizes de transferência estáveis $\mathcal{H}_{estvl}$ através de K , a caracterização a partir de Q mantém a convexidade do problema. Uma vez encontrado o valor ótimo de Q , basta aplicar (5.66) para encontrar o controlador K que implementa a solução do problema.

5.14 Parametrização de Youla em Espaço de Estados

Após ter-se definido a arquitetura a ser utilizada, caracterizado o conceito de especificação e o seu papel no projeto do controlador, definiu-se qual o conjunto no qual será realizada a busca da solução do problema de otimização ($\mathcal{H}_{estvl}$). Neste ponto do trabalho, todos os requisitos para a formulação do problema no contexto de matrizes de transferência foram satisfeitos. Porém, ainda é necessário implementar a parametrização Q através de espaço de estados, por duas razões: é uma formulação melhor comportada em termos computacionais (Francis, 1987; Boyd e Barrat, 1981), e os procedimentos que serão utilizados para a solução do problema estão baseados na formulação através de espaço de estados.

Uma matriz de transferência, descrita através da realização em espaço de estados

$$G(s) = C(sI - A)^{-1}B + D \quad (5.69)$$

será representada através da seguinte notação:

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \quad (5.70)$$

Sendo assim, considere que P possui a seguinte realização em espaço de estados:

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \quad (5.71)$$

Como a matriz P está definida de acordo com as suas partições de entrada e saída

$$\begin{bmatrix} v \\ u \end{bmatrix}, \begin{bmatrix} z \\ y \end{bmatrix}$$

a realização de P levando-se em conta as suas partições fica na forma

$$P = \left[\begin{array}{c|cc} A & B_v & B_u \\ \hline C_z & D_{zv} & D_{zu} \\ C_y & D_{yv} & D_{yu} \end{array} \right] \quad (5.72)$$

Assume-se que a partição P_{yu} é estritamente própria ($D_{yu} = 0$) por uma questão de simplicidade no desenvolvimento das equações. Outro resultado que será levado em consideração é que, como foi visto anteriormente, a estabilizabilidade de P é equivalente à estabilizabilidade de P_{yu} e dizer que P_{yu} é estabilizável é o mesmo que afirmar que o par (A, B_u) é estabilizável e que o par (C_y, A) é observável. Portanto, existem matrizes K_{sfb} e L_{est} que fazem com que as matrizes

$$A - B_u K_{sfb} \text{ e } A - L_{est} C_y \quad (5.73)$$

sejam estáveis.

Com isso, pode-se obter (Francis, 1987; Boyd e Barrat, 1991) as realizações das oito matrizes da fatoração duplamente coprima de P_{yu} :

$$M_{yu} = \left[\begin{array}{c|c} A - B_u K_{sfb} & B_u \\ \hline -K_{sfb} & I \end{array} \right] , \quad N_{yu} = \left[\begin{array}{c|c} A - B_u K_{sfb} & B_u \\ \hline C_y & 0 \end{array} \right] \quad (5.74)$$

$$\tilde{M}_{yu} = \left[\begin{array}{c|c} A - L_{est} C_y & -L_{est} \\ \hline C_y & I \end{array} \right] , \quad \tilde{N}_{yu} = \left[\begin{array}{c|c} A - L_{est} C_y & B_u \\ \hline C_y & 0 \end{array} \right] \quad (5.75)$$

$$X_{yu} = \left[\begin{array}{c|c} A - B_u K_{sfb} & L_{est} \\ \hline C_y & I \end{array} \right] , \quad Y_{yu} = \left[\begin{array}{c|c} A - B_u K_{sfb} & L_{est} \\ \hline -K_{sfb} & 0 \end{array} \right] \quad (5.76)$$

$$\tilde{X}_{yu} = \left[\begin{array}{c|c} A - L_{est} C_y & -B_u \\ \hline -K_{sfb} & I \end{array} \right] , \quad \tilde{Y}_{yu} = \left[\begin{array}{c|c} A - L_{est} C_y & L_{est} \\ \hline -K_{sfb} & 0 \end{array} \right] \quad (5.77)$$

Considerando que a realização da matriz de parâmetros livres Q tem como realização

$$Q = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array} \right] \quad (5.78)$$

e substituindo estas realizações em (5.66) e em (5.67), e em seguida em (5.68), obtêm-se as representações em espaço de estados para o controlador

$$K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array} \right] \quad (5.79)$$

e para a matriz de transferência em malha fechada

$$H = \left[\begin{array}{cc|c} A + B_u D_K C_y & B_u C_K & B_v + B_u D_K D_{yv} \\ B_K C_y & A_K & B_K D_{yv} \\ \hline C_z + D_{zu} D_K C_y & D_{zu} C_K & D_{zv} + D_{zu} D_K D_{yv} \end{array} \right] \quad (5.80)$$

onde

$$\begin{aligned}
 A_K &= \begin{bmatrix} A - B_u K_{sfb} - L_{est} C_y - B_u D_Q C_y & B_u C_Q \\ -B_Q C_y & A_Q \end{bmatrix} \\
 B_K &= \begin{bmatrix} L_{est} + B_u D_Q \\ B_Q \end{bmatrix} \\
 C_K &= \begin{bmatrix} -(K_{sfb} + D_Q C_y) & C_Q \end{bmatrix} \\
 D_K &= D_Q.
 \end{aligned} \tag{5.81}$$

Com isso, obtêm-se o mesmo problema Q -parametrizado da seção anterior agora representado através de espaço de estados. O próximo Capítulo discutirá um método de solução para o problema de minimização proposto.

5.15 Conclusão

Neste Capítulo foram discutidos aspectos relacionados ao problema básico de projeto de sistemas de controle no domínio da frequência. Foi apresentado o conceito de fatoração coprima de matrizes e discutida sua importância e aplicabilidade na parametrização de todos os controladores estabilizantes da planta. Mostrou-se que esta parametrização gera um problema de otimização equivalente ao problema original, com a diferença de que a relação entre a variável de decisão e a matriz de transferência em malha fechada passa a ser afim.

Tendo sido formulado o problema de otimização para o sistema descrito por matrizes de transferência, optou-se por representá-lo através da descrição por espaço de estados. Para tanto as realizações das matrizes de transferências envolvidas foram apresentadas.

Capítulo 6

Projeto de Controladores Dinâmicos

6.1 Introdução

Neste Capítulo será abordado o problema de projeto de controladores no domínio da frequência de acordo com a parametrização introduzida no Capítulo anterior. Demonstra-se que é possível resolver o problema de projeto através do Algoritmo Básico estendendo-se alguns conceitos, como o de subgradiente, para espaços de dimensão infinita. Através de aproximações finitas, obtém-se um problema equivalente, tratável pelas técnicas introduzidas neste trabalho. Relações entre a abordagem convexa proposta e uma conhecida técnica para projetos de controladores - o Método das Desigualdades - são discutidas no final do Capítulo.

6.2 Formulação do Problema via Parametrização de Youla

Para simplificar a notação o índice zv da matriz de transferência em malha fechada das entradas exógenas para as saídas reguladas (H_{zv}) será suprimido sempre que possível. O problema de projeto, formulado originalmente em termos do controlador K , pode ser definido em termos de Q como

$$\min_{Q \in \mathcal{Q}} \Phi(H(Q)) \quad (6.1)$$

onde $\Phi(\cdot)$ é em geral um vetor de critérios (funcionais) de desempenho convexos sobre o conjunto das matrizes de transferência H e

$$\mathcal{Q} = \{Q \in \mathcal{RH}_\infty : H(Q) \leq 0\} \quad (6.2)$$

representa o conjunto definido por inequações funcionais que descrevem especificações de projeto a serem satisfeitas. A escolha das especificações de projeto que definirão os critérios de desempenho ou que comporão as restrições fica a critério do projetista, podendo inclusive ser intercambiadas para uma melhor avaliação do desempenho do sistema. Observe também que através da parametrização de Youla, as especificações de realizabilidade e estabilidade da matriz de transferência da solução do problema estão satisfeitas, por construção.

O problema Q -parametrizado é de dimensão infinita, uma vez que Q está contido em $\mathcal{RH}_\infty$, que por sua vez é um espaço de dimensão infinita. Torna-se portanto necessário o uso de alguma técnica de aproximação finita adequada ao $\mathcal{RH}_\infty$, o que será visto em seguida.

6.3 Aproximação de Ritz

No contexto deste trabalho, a aproximação de Ritz (Boyd e Barrat, 1991) consiste em, dado um conjunto $\mathcal{H}$ de matrizes de transferência, gerar uma sequência de n elementos deste conjunto que irão compor uma aproximação $\mathcal{H}^n$ de dimensão n , e assim proceder a busca da solução do problema em questão sobre $\mathcal{H}^n$. Como a aproximação será sempre de dimensão finita, o problema de otimização passa a ser de dimensão finita.

Caso o problema na aproximação n seja infactível, gera-se um elemento a mais na sequência, de forma que a nova sequência gere uma aproximação de dimensão $n + 1$. O aumento de dimensão é progressivo até que a solução do problema seja encontrado no conjunto corrente.

Dada uma sequência $R_0, R_1, \dots, R_n$ de matrizes de transferências estáveis, a expressão básica da aproximação é

$$\mathcal{H}^n \triangleq \left\{ H : H = R_0 + \sum_{i=1}^n x_i R_i, \quad x \in \mathbb{R}^n \right\} \quad (6.3)$$

que fornece, para cada $x \in \mathbb{R}^n$, um elemento de $\mathcal{H}^n$. A aproximação não compromete a natureza convexa dos funcionais e especificações de projeto, pois a dependência de H com $x \in \mathbb{R}^n$ é afim.

No caso do problema Q -parametrizado, tem-se o seguinte espaço de matrizes de transferência para aproximar:

$$\mathcal{H}_{estvl} = \{ H : H = T_1 + T_2 Q T_3 : Q \in \mathcal{RH}_\infty \} \quad (6.4)$$

Para obter uma aproximação de Ritz, seleciona-se uma sequência de n matrizes de transferências estáveis Q_i , $i = 1, \dots, n$ e faz-se $R_0 = T_1$ e $R_i = T_2 Q_i T_3$, $i = 1, \dots, n$ e com isso a aproximação terá a forma

$$\mathcal{H}_{estvl}^n \triangleq \left\{ H : H = T_1 + T_2 \left(\sum_{i=1}^n x_i Q_i \right) T_3, \quad x \in \mathbb{R}^n \right\} \quad (6.5)$$

Agora a busca passa a ser em $x \in \mathbb{R}^n$. A próxima seção abordará o método a ser utilizado para resolver o problema de otimização resultante.

6.4 Método de Planos de Corte

Na seção anterior foi discutida uma aproximação finita para o problema de projeto de sistemas de controle Q -parametrizado. Esta seção discute o método de Planos de Corte, escolhido por ser aplicável a problemas de otimização convexos não-diferenciáveis. Antes de passar a descrever o método, é necessário aprofundar o conceito de subgradiente de um funcional convexo.

Definição 6.1 Seja $\phi(\cdot)$ um funcional convexo sobre $\mathcal{H}$. Então o funcional $\phi^{sg}(\cdot)$ é um subgradiente de $\phi(\cdot)$ no ponto $H_0 \in \mathcal{H}$ se ϕ^{sg} for linear sobre $\mathcal{H}$ e

$$\phi(H) \geq \phi(H_0) + \phi^{sg}(H - H_0), \quad \forall H \in \mathcal{H} \quad (6.6)$$

O conjunto de todos os subgradientes de $\phi(\cdot)$ no ponto H_0 será denotado por $\partial\phi(H_0)$. É possível verificar que, se uma função for convexa então possui pelo menos um subgradiente em todo o seu domínio, e se além de convexa for diferenciável, então o conjunto $\partial\phi(H_0)$ terá sempre um único elemento, chamado de *gradiente* de $\phi(\cdot)$ no ponto H_0 .

Se $\mathcal{H} \subseteq \mathbb{R}^n$, então o subgradiente em qualquer ponto de $\mathcal{H}$ é uma função linear na forma $g'H$, onde $g \in \mathbb{R}^n$ depende de H_0 , e a expressão (6.6) fica sendo

$$\phi(H) \geq \phi(H_0) + \langle g, H - H_0 \rangle, \quad \forall H \in \mathcal{H} \quad (6.7)$$

onde o lado direito da desigualdade descreve a equação de um hiperplano suporte a $\phi(\cdot)$ no ponto H_0 .

É fácil verificar que, dados os subgradientes $g^i \in \partial\phi(H^i)$ nos pontos H^i , $i = 1, 2, \dots, k$ então

$$\phi(H) \geq \max_{i \in [1, \dots, k]} \left\{ \phi(H^i) + \langle g^i, H - H^i \rangle \right\}, \quad \forall H \in \mathcal{H} \quad (6.8)$$

De fato, cada subgradiente constitui numa aproximação de primeira ordem da função, sendo esta aproximação exata no ponto considerado. Portanto, é possível fazer a minimização não mais sobre a função $\phi(\cdot)$, mas sobre a aproximação baseada nos subgradientes encontrados. O valor do funcional será sempre menor ou igual ao valor encontrado minimizando-se $\phi(\cdot)$, uma vez que os planos de corte constituem limitantes inferiores para $\phi(\cdot)$ (ver Figura 6.1). Um aspecto importante é o de que a aproximação via planos de corte é linear por partes, podendo-se obter o limitante inferior de $\phi(\cdot)$ através de programação linear.

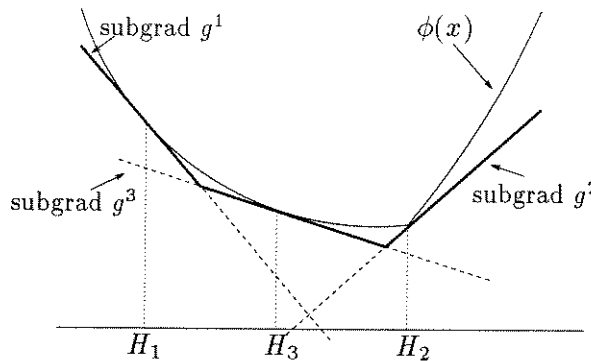


Figura 6.1: Aproximação de $\phi(H)$ por planos de corte.

6.4.1 Método de Planos de Corte: Problema Irrestrito

O problema consiste em

$$\min_{H \in \mathcal{H}} \phi^k(H) \quad (6.9)$$

onde

$$\phi^k(H) = \max_{i \in [1, \dots, k]} \{ \phi(H^i) + \langle g^i, H - H^i \rangle \} \quad (6.10)$$

é a aproximação de $\phi(\cdot)$ a partir da existência de k subgradientes.

Note que (6.9) é equivalente ao seguinte problema linear:

$$\begin{aligned} & \min_{H \in \mathcal{H}} \sigma \\ & \text{s.a.} \quad \phi(H^i) + \langle g^i, H - H^i \rangle \leq \sigma, \quad i = 1, \dots, k \end{aligned} \quad (6.11)$$

O método de planos de corte consiste em:

Passo 1 - A partir de uma aproximação com k subgradientes, resolva (6.11), obtendo H^{k+1} .

Passo 2 - Encontre o valor de $\phi(\cdot)$ para $H = H^{k+1}$.

Passo 3 - Se $\phi(H^{k+1}) - \sigma < \epsilon$, para $\epsilon > 0$ suficientemente pequeno, pare. Caso contrário, encontre o subgradiente de $\phi(\cdot)$ no ponto H^{k+1} e acrescente o semi-espço correspondente ao problema (6.11). Faça $k = k + 1$ e volte ao Passo 1.

Como observado, uma vantagem de se trabalhar com este método é a de que o problema (6.11) é sempre um problema de programação linear. O que o método faz quando a cada iteração acrescenta uma restrição ao problema (6.11) é melhorar a aproximação da função. Note que a aproximação por subgradientes sempre subestima o valor de $\phi(\cdot)$ e sendo assim, o valor da diferença $\phi(H^{k+1}) - \sigma$ será sempre um valor positivo.

6.4.2 Método de Planos de Corte: Problema Restrito

Para o caso de otimização restrita deve-se acrescentar às restrições do problema (6.11) a aproximação via planos de corte do conjunto de restrições do problema. Para tanto, é necessário calcular os subgradientes também dos funcionais $\psi(\cdot)$ que compõem o conjunto de especificações por inequações funcionais (6.2).

Considere, portanto, $h^i \in \partial\psi(H^i)$ subgradientes do funcional $\psi(\cdot)$ nos pontos H^i , $i = [1, \dots, k]$. O problema (6.11) passa a ser

$$\begin{aligned} & \min_{H \in \mathcal{H}} \sigma \\ & \text{s.a.} \quad \phi(H^i) + \langle g^i, H - H^i \rangle \leq \sigma, \quad i = 1, 2, \dots, k \\ & \quad \psi(H^i) + \langle h^i, H - H^i \rangle \leq 0 \end{aligned} \quad (6.12)$$

6.5 Exemplo Ilustrativo

Para ilustrar o funcionamento do método de planos de corte será utilizado o seguinte exemplo (Boyd e Barrat, 1991; Cruz, 1994). Trata-se de um sistema conhecido como *sistema de controle com um grau de liberdade* (1-DOF) (Figura 6.2). Neste sistema o sinal de referência r é uma entrada externa que varia com o tempo. O principal objetivo neste sistema é manter a saída y_p da planta o mais próximo possível do sinal de referência r , apesar das perturbações n_{proc} e n_{sen} , ao mesmo tempo em que se tenta evitar que o esforço do atuador u cresça muito.

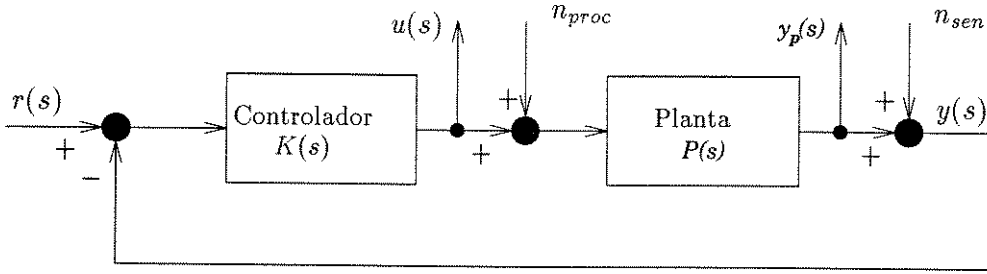


Figura 6.2: Representação clássica do sistema 1-DOF.

Este mesmo sistema, considerando-se a arquitetura da Figura 5.1, possui quatro entradas e três saídas, que são

$$v = \begin{bmatrix} n_{proc} \\ n_{sen} \\ r \end{bmatrix}, \quad z = \begin{bmatrix} y_p \\ u \end{bmatrix}, \quad y = r - y_p - n_{sen} \quad (6.13)$$

A matriz de transferência da planta é

$$P = \begin{bmatrix} P_{zv} & P_{zu} \end{bmatrix} = \begin{bmatrix} P_0 & 0 & 0 & P_0 \\ 0 & 0 & 0 & 1 \\ -P_0 & -1 & 1 & -P_0 \end{bmatrix} \quad (6.14)$$

A matriz de transferência em malha fechada H_{zv} tem a forma

$$H_{zv} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{bmatrix} = \begin{bmatrix} \frac{P_0}{(1 + P_0 K)} & -\frac{P_0 K}{(1 + P_0 K)} & \frac{P_0 K}{(1 + P_0 K)} \\ -\frac{P_0 K}{(1 + P_0 K)} & -\frac{K}{(1 + P_0 K)} & \frac{K}{(1 + P_0 K)} \end{bmatrix} \quad (6.15)$$

Os funcionais podem representar os valores RMS das variáveis de saída medida (y_p) e de controle (u), overshoot e ganho de pico, por exemplo. No capítulo anterior foram discutidas algumas especificações de projeto, e especificações mais rígidas, tais como realizabilidade e estabilidade. A seguir serão apresentadas algumas especificações de projeto de interesse e as expressões dos seus

subgradientes.

• Overshoot

O funcional overshoot possui a seguinte definição matemática:

$$\phi_{ov}(H) = \sup_{t \geq 0} s(t) - 1,$$

onde $s(t)$ é a resposta ao degrau unitário da função de transferência H .

Este funcional possui como subgradiente no ponto H_0

$$\phi^{sg}(H) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\omega t_0}}{j\omega} H(j\omega) d\omega$$

onde t_0 é qualquer instante de tempo em que o overshoot ocorre. Note que este subgradiente é um funcional linear de H .

• Norma H_2

O funcional norma H_2 é definido como

$$\phi(H) = \|H\|_2 = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega \right)^{1/2}$$

ou, no domínio do tempo

$$\phi(H) = \|H\|_2 = \left(\int_0^{\infty} h^2(t) dt \right)^{1/2}$$

A expressão para o subgradiente de ϕ no ponto H_0 tem a forma

$$\phi^{sg}(H) = \frac{1}{\phi(H_0)} \int_{-\infty}^{\infty} h_0(t) h(t) dt$$

• Ganho de Pico

O funcional ganho de pico, por sua vez, é definido como

$$\phi(H) = \|H\|_{gp} = \int_0^{\infty} |h(t)| dt.$$

onde $h(t)$ é a resposta ao impulso da função de transferência H . Para a expressão do subgradiente tem-se

$$\phi^{sg}(H) = \int_0^{\infty} \text{sgn}(h_0(t)) h(t) dt.$$

onde $h_0(t)$ é a resposta ao impulso da função de transferência H no ponto $H = H_0$.

6.6 Subgradientes na Aproximação de Ritz

As expressões dos subgradientes apresentadas anteriormente são funcionais lineares sobre um espaço de dimensão infinita. Para adequá-las ao problema de dimensão finita gerado pela aproximação de Ritz é necessário obter os subgradientes em termos das variáveis x . Considere o funcional

$$\varphi(x) = \phi \left(R_0 + \sum_{i=1}^n x_i R_i \right) \quad (6.16)$$

Um subgradiente $g \in \partial\varphi(\tilde{x})$ será um vetor cuja i -ésima componente é o subgradiente do funcional $\phi(\cdot)$ no ponto $x = \tilde{x}$. Então

$$g = \begin{bmatrix} \phi^{sg}(R_1) \\ \vdots \\ \phi^{sg}(R_n) \end{bmatrix} \in \partial\varphi(\tilde{x}) \quad (6.17)$$

No caso do problema parametrizado, tem-se $R_0 = T_1$ e $R_i = T_2 Q_i T_3$, sendo Q_i , $i = 1, \dots, n$ matrizes estáveis pré-estabelecidas. Assim sendo, o subgradiente de dimensão finita do funcional overshoot fica por exemplo na forma

$$g = \begin{bmatrix} \phi^{sg}(R_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\omega t_0}}{j\omega} R_1(j\omega) d\omega \\ \vdots \\ \phi^{sg}(R_n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\omega t_0}}{j\omega} R_n(j\omega) d\omega \end{bmatrix} \quad (6.18)$$

onde t_0 é qualquer instante de tempo em que ocorre o overshoot para

$$\tilde{H} = T_1 + T_2 \left(\sum_{i=1}^n \tilde{x} Q_i \right) T_3$$

6.7 Problema de Projeto Multicritério

Esta seção tratará da solução do problema de projeto de controladores com múltiplos critérios de desempenho. O que se buscará é a re-edição da estrutura multinível derivada para o problema multicritério no domínio do tempo, às condições do problema no domínio da frequência.

6.7.1 Formulação do Problema

Como foi visto anteriormente, uma especificação de projeto pode ter um tratamento matemático adequado a partir do conceito de *especificação por inequação funcional*

$$\mathcal{D}_\phi^{(a)} : \phi(H) \leq a$$

A especificação de projeto definida deste modo pode ser *otimizada*, bastando-se que para isto seja minimizado o valor de a , mantendo-se a factibilidade do problema. Portanto, do ponto de vista de controle ótimo, um problema de projeto de controladores poderia ser formulado como

$$\min_{H \in \mathcal{H}} \phi(H)$$

onde $\mathcal{H}$ representa genericamente quaisquer outras especificações ou restrições rígidas do problema.

Supondo-se que se deseja otimizar simultaneamente vários critérios, o problema acima deixa de ser escalar para ser vetorial, ou seja,

$$\min_{H \in \mathcal{H}} \Phi(H)$$

sendo $\Phi(H)$ um vetor cujas componentes são os vários critérios a serem otimizados

$$\Phi(H)' = [\phi_1(H) \ \phi_2(H) \ \dots \ \phi_m(H)]$$

A definição a seguir interpreta o conceito de solução eficiente em termos de matrizes de transferência de malha fechada.

Definição 6.2 *Diz-se que uma matriz de transferência de malha fechada H^* é eficiente se não existir nenhuma outra matriz de transferência em malha fechada $\tilde{H}$ tal que $\Phi(\tilde{H}) \leq \Phi(H^*)$, com pelo menos uma desigualdade estrita, ou seja, $\phi_i(\tilde{H}) < \phi_i(H^*)$ para algum $i \in [1, 2, \dots, n]$.*

O problema de projeto de controladores dentro do escopo da programação multicritério consiste em escolher dentro da curva de trade-off uma matriz de transferência em malha fechada que minimize a função valor $\mathcal{V}(\cdot)$. Logo, tem-se o seguinte problema

$$\min_{H \in \mathcal{H}} \mathcal{V}(\Phi(H)) \quad (6.19)$$

O problema acima está formulado no espaço das matrizes de transferência. Considere agora sua formulação no espaço dos critérios:

$$\min_{y \in \mathcal{F}} \mathcal{V}(y) \quad (6.20)$$

onde

$$\mathcal{F} \triangleq \{y \in \mathbb{R}^m : y = \Phi(H), \text{ para algum } H \in \mathcal{H}\} \quad (6.21)$$

Esta formulação é equivalente à apresentada na Seção 1.6 e exhibe as mesmas características e vantagens já observadas.

6.7.2 Estrutura Multinível - Decisão e Análise

A metodologia apresentada a seguir foi discutida no Capítulo 2 quando se tratou o problema no domínio do tempo, sendo aqui re-editada apenas para adequar a notação ao caso de matrizes de transferência.

Definição 6.3 Diz-se que um vetor $y \in \mathbb{R}^m$ de trade off's implícitos é satisfatório se existe $H \in \mathcal{H}$, tal que $\Phi(H) \leq y$.

Seja o conjunto $\mathcal{C}$ definido como

$$\mathcal{C} \triangleq \{y \in \mathbb{R}^m : \Phi(H) \leq y, \text{ para algum } H \in \mathcal{H}\} \quad (6.22)$$

Considere agora o problema formulado no espaço dos critérios,

$$\min_{y \in \mathcal{C} \cap \mathcal{Y}} \mathcal{V}(y) \quad (6.23)$$

onde

$$\mathcal{Y} \triangleq \{y : y \geq \underline{y}\} \quad (6.24)$$

sendo $\underline{y}$ a solução utópica.

Teorema 6.1 Seja y^* a solução ótima de (6.23). Então y^* resolve o problema de controle multicritério definido por (6.20).

Prova: Segue a mesma demonstração do Teorema 2.1. □

O problema a ser resolvido computacionalmente fica na forma

$$\begin{aligned} & \min_{y \in \mathcal{Y}} \mathcal{V}(y) \\ & \text{s.a.} \quad w'y \geq w'\Phi(H(w)), \quad \forall w \in W. \end{aligned} \quad (6.25)$$

onde

$$H(w) = \arg \min_{H \in \mathcal{H}} w'\Phi(H) \quad (6.26)$$

6.7.3 Algoritmo Básico

Para resolver o problema (6.26) de projeto no domínio da frequência usou-se o Algoritmo Básico apresentado na Seção 2.3.

Algoritmo Básico

Faça $k = 0$ e $\mathcal{Y}^0 = \mathcal{Y}$.

Passo 1. Resolva o problema relaxado

$$\min_{y \in \mathcal{Y}^k} \mathcal{V}(y) \quad (6.27)$$

Seja y^k a solução relaxada

Passo 2. Resolva o problema min-max

$$\Theta(y) = \max_{w \in W} \min_{H \in \mathcal{H}} w'(\Phi(H) - y^k)$$

Escolha $w^0 \in W$ e faça $\ell = 0$.

Passo 2.1 Resolva o problema de controle ótimo (*min*)

$$\lambda(w) = \min_{H \in \mathcal{H}} (w^\ell)'(\Phi(H) - y^k)$$

obtendo $H^\ell \in \mathcal{H}(w^\ell)$, $\xi^\ell = \Phi(H^\ell) - y^k$ e $\lambda(w^\ell)$.

Passo 2.2 Resolva o problema linear (*max*)

$$\max_{\sigma} \sigma$$

$$\sigma \leq w' \xi^i, \quad i = 1, 2, \dots, \ell.$$

se $\sigma^\ell - \lambda(w^\ell) < \epsilon$, para $\epsilon > 0$ arbitrariamente pequeno, faça $w^k = w^\ell$, $H^k = H^\ell$ e $\Theta(y^k) = \sigma^\ell$ e vá ao Passo 3. Caso contrário, faça $l = l + 1$ e volte ao Passo 2.1

Passo 3. Se $\Theta(y^k) \leq 0$, então $H^* = H(w^k)$ resolve (6.19) e $y^* = \Phi(H^*)$ resolve (6.25) ((6.23)). Caso contrário, adicione a restrição mais violada a $\mathcal{Y}^k$

$$\mathcal{Y}^{k+1} = \{y \in \mathcal{Y}^k : (w^k)'y \geq (w^k)'\Phi(H^k)\}$$

faça $k = k + 1$ e volte ao Passo 1.

A utilização do Algoritmo Básico para resolver o problema de projeto de controladores apresenta as mesmas vantagens já enunciadas quando do seu desenvolvimento para problemas de controle no domínio do tempo. Em particular, a função valor pode ser livremente arbitrada, sem que a solução do nível de análise seja modificada. O problema a ser tratado no nível de decisão será sempre estático, de pequena dimensão, com restrições lineares e geralmente convexo.

6.8 Solução do Problema Min-max

O problema min-max tem uma importância particular no desenvolvimento deste trabalho por concentrar a maior parte do esforço computacional do procedimento. A sua análise é fundamental para se estabelecer estratégias que visem o aperfeiçoamento do Algoritmo Básico.

O problema min-max consiste em

$$\max_{w \in W} \min_{H \in \mathcal{H}} w' \Phi(H) \tag{6.28}$$

onde $\mathcal{H}$ representa o conjunto de matrizes de transferências factíveis e

$$\Phi(H)' = [\phi_1(H) \ \phi_2(H) \ \dots \ \phi_m(H)]$$

é um vetor cujas componentes são os m critérios de projeto.

Através da parametrização de Youla, obtém-se o problema equivalente

$$\max_{w \in W} \min_{Q \in \mathcal{Q}} w' \Phi(T_1 + T_2 Q T_3) \quad (6.29)$$

e uma vez que a dependência com w é linear, a convexidade do problema é preservada.

Por conveniência, o problema é representado como

$$\max_{w \in W} \lambda(w) \quad (6.30)$$

onde

$$\lambda(w) = \min_{H \in \mathcal{H}} w' \Phi(H) \quad (6.31)$$

Note que, para um determinado valor de w , o produto $w' \Phi(H)$ dá origem a um critério escalar e o problema (6.31) se reduz ao caso tratado no Capítulo anterior. Neste caso, cada subgradiente da combinação linear de critérios obtida tem a forma

$$\Phi^{sg}(H) = w_1 \phi(H)_1^{sg} + w_2 \phi(H)_2^{sg} + \dots + w_m \phi(H)_m^{sg} \quad (6.32)$$

onde $\phi(H)_i^{sg}$, $i = 1, 2, \dots, m$ representa o subgradiente de cada critério. Portanto, para resolver (6.31) basta aplicar o procedimento apresentado, considerando que o subgradiente da função é uma soma ponderada dos subgradientes dos critérios do problema.

Encontrada a solução de (6.31), o problema de maximização (6.30) fica caracterizado. Para resolver (6.30) usou-se também o método de planos de corte. Note que, considerando H^l a solução para o problema de mínimo na iteração l , cada subgradiente que aproximará $\lambda(w)$ terá a forma:

$$\Psi^{sg}(w) = \begin{bmatrix} \phi_1(H^l) \\ \vdots \\ \phi_m(H^l) \end{bmatrix} = \Phi(H^l)$$

A seguir apresenta-se o algoritmo que encontra a solução do problema min-max.

Passo 1 - A partir do vetor w^l resolva (6.31) pelo método de planos de cortes, obtendo H^l .

Passo 2 - Resolva o problema

$$\begin{aligned} & \max_{\sigma, w \in W} \sigma \\ & \text{s.a. } \sigma \leq w' \Phi(H^i), \quad i = 1, \dots, l \end{aligned}$$

obtendo σ^{l+1} e w^{l+1} . Se $(\sigma^{l+1} - (w^l)' \Phi(H^l)) < \epsilon$, para ϵ suficientemente pequeno, pare. Caso contrário, faça $l = l + 1$ e volte ao passo 1.

Com o objetivo de ilustrar o procedimento, a seguir será apresentado uma variação do exemplo abordado em Boyd e Barrat (1991). O exemplo original mono-critério consiste em obter a solução do problema 1-DOF discutido na Seção 6.5 através da seguinte formulação:

$$\begin{aligned} \min \quad & \phi_{rms-u}(H) \\ \text{s. a} \quad & \phi_{rms-y_p}(H) \leq 0.1 \end{aligned} \tag{6.33}$$

onde $\phi_{rms-y_p}(H)$ e $\phi_{rms-u}(H)$ são dados por

$$\phi_{rms-y_p}(H) = (\|H_{11}W_{proc}\|_2^2 + \|H_{12}W_{sen}\|_2^2)^{1/2} \tag{6.34}$$

$$\phi_{rms-u}(H) = (\|H_{21}W_{proc}\|_2^2 + \|H_{22}W_{sen}\|_2^2)^{1/2} \tag{6.35}$$

O problema min-max derivado do exemplo acima consiste em fazer com que $\phi_{rms-y_p}(H)$ não caracterize uma restrição a ser satisfeita no problema, mas sim um critério a ser minimizado. Tem-se então um problema bi-critério, onde os critérios envolvidos são conflitantes (melhorar a regulação implica em aumentar o esforço de controle). Portanto o problema min-max resultante consiste em

$$\max_{w \in W} \min_{H \in \mathcal{H}} [w_1 \phi_{rms-u}(H) + w_2 \phi_{rms-y_p}(H)] \tag{6.36}$$

6.9 Solução do Problema Multicritério

Esta seção tratará dos aspectos referentes aos métodos matemáticos e computacionais usados na implementação do Algoritmo Básico. O método de otimização escolhido foi o método de planos de corte, devido a sua facilidade de implementação. Na solução do nível de análise foi utilizada a parametrização de Youla em conjunto com o método de planos de corte para obter a solução do problema de minimização; para resolver o problema de maximização foi também utilizado o método de planos de corte. Nas próximas seções serão discutidas as implementações de cada nível e a solução de um exemplo ilustrativo.

6.9.1 Nível de Análise

Como foi visto anteriormente, o nível de análise consiste de dois problemas de otimização: um problema de mínimo e um problema de máximo. A solução de cada problema será tratada separadamente a seguir.

O problema de minimização consiste em

$$\min_{H \in \mathcal{H}} w'(\Phi(H) - y^k)$$

onde y^k é fornecido pelo nível de decisão e w é o vetor solução do problema de maximização. Nesta etapa é resolvido um problema de projeto de controladores clássico e, portanto o problema é reformulado em termos de matrizes de transferência estáveis Q . Apenas lembrando, a parametrização de Youla permite a mudança da variável de otimização, da matriz de transferência do controlador K , para uma matriz de transferência estável Q . Esta reformulação permite preservar as características de convexidade que a maioria das especificações de projeto apresenta em malha fechada, pois a dependência de H com Q é afim e, desta forma, funcionais convexos em H permanecem convexos em Q .

O problema em termos da variável Q fica, portanto, formulado como

$$\min_{Q \in \mathcal{Q}} w'(\Phi(T_1 + T_2 Q T_3) - y^k) \quad (6.37)$$

O subgradiente a cada iteração, necessário para a aplicação do método de planos de cortes, é idêntico ao do problema min-max, pois a parte escalar $(-w'y^k)$ do funcional é eliminada no subgradiente.

No problema de maximização, que encontra o vetor w solução do problema, também é utilizado o método de planos de corte.

6.9.2 Nível de Decisão

O nível de decisão é facilmente implementado, sendo sempre um problema estático, de pequena dimensão, definido no $\mathbb{R}^m$, com restrições lineares. Para ilustrar o método, foram utilizadas duas funções valores $\mathcal{V}(\cdot)$.

Norma Infinito: Neste caso, a função valor $\mathcal{V}(\cdot)$ assume a forma

$$\mathcal{V}(y) = \max_{1 \leq i \leq m} (y_i - \underline{y}_i) \quad (6.38)$$

onde $\underline{y}$ representa a solução utópica. Portanto, através de $\mathcal{V}(y)$ se procura encontrar uma solução que minimize a distância entre o conjunto de soluções factíveis e o ponto $\underline{y}$, medida pela norma infinito. No caso particular do Algoritmo Básico, a norma infinito possui uma característica importante para a validação do procedimento pois, é possível mostrar (Carvalho, 1993) que neste caso, o problema a ser resolvido no nível de decisão é equivalente ao do nível de análise e portanto o procedimento termina em apenas uma iteração.

Norma Euclideana: Através da norma

$$\mathcal{V}(y) = \left(\sum_{i=1}^n (y_i - \underline{y}_i)^2 \right)^{1/2} \quad (6.39)$$

se procura também minimizar a distância entre a solução utópica e o conjunto de soluções factíveis, neste caso medida pela norma Euclideana. Esta norma gera um problema não-linear, de fácil solução e interpretação física.

A seguir um exemplo numérico foi resolvido com a finalidade de ilustrar o desempenho do Algoritmo Básico.

6.9.3 Exemplo Ilustrativo

O exemplo selecionado foi o mesmo sistema 1-DOF da Seção 6.5, sendo que

$$P_0(s) = \frac{1}{s^2} \frac{10 - s}{10 + s}$$

e o vetor de critérios

$$\Phi(H) = \begin{bmatrix} \phi_{rms-y_p}(H) \\ \phi_{rms-u}(H) \end{bmatrix}$$

onde $\phi_{rms-y_p}(H)$ e $\phi_{rms-u}(H)$ são dados por (6.34) e (6.35), respectivamente.

Como vetor $\underline{y}$, foi escolhido a origem, ou seja, $\underline{y} = [0 \ 0]'$. O critério de parada do Algoritmo Básico foi $\Theta(y) \leq 10^{-3}$. O controlador nominal K_{nom} (o mesmo utilizado por Boyd e Barrat (1991)) é caracterizado pelos seguintes ganhos:

$$L_{est} = \begin{bmatrix} 0.0000 \\ -3.1623 \\ -1.1115 \end{bmatrix}, \quad K_{sfb} = [1.4276 \ 10.2948 \ 2.4495]$$

Cada termo da aproximação de Ritz para a matriz Q , dado por (6.5), foi construído a partir do termo geral

$$Q_i = \left(\frac{1}{s+1} \right)^i$$

e o problema resolvido para aproximações usando três termos ($n = 3$).

Todo o procedimento foi implementado usando-se rotinas MATLAB. A Tabela 6.1 apresenta os resultados obtidos pelo Algoritmo Básico considerando-se a *norma infinito* como função valor e a Figura 6.3 mostra as respostas ao degrau de $H_{13}(s)$ e $H_{23}(s)$. Como era de se esperar, o algoritmo convergiu em apenas uma iteração, sendo que o controlador encontrado foi de ordem 9.

Tabela 6.1: Solução do Algoritmo Básico para a norma infinito.

k	y'	w'	$\phi_{rms-u}(H)$	$\phi_{rms-y_p}(H)$	$\Theta(y)$
0	[0 0]	[0.4801 0.5199]	0.0630	0.0643	0.0636
1	[0.0636 0.0636]	[0.4753 0.5247]	0.0634	0.0639	$3.27 \cdot 10^{-6}$

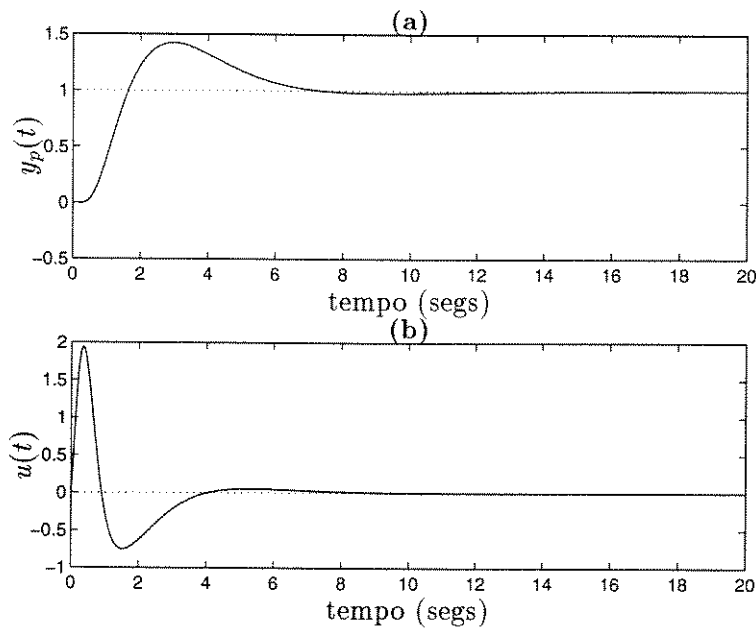


Figura 6.3: Resposta ao degrau de (a) H_{13} e (b) H_{23} para norma infinito.

Na Tabela 6.2 apresenta-se os resultados para a *norma Euclideana*, cuja solução foi encontrada com o mesmo número de iterações da norma infinito. A Figura 6.4 mostra as respostas ao degrau de $H_{13}(s)$ e $H_{23}(s)$ para este caso, que apresentou como controlador

$$A_K = \begin{bmatrix} -11.428 & -10.295 & -2.449 & 0 & 0 & 0 & 0 & 0 & 9.837 \\ 1 & 3.162 & -31.623 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.112 & -11.115 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 10 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 29.696 & -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 67.467 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad B_K = \begin{bmatrix} 0 \\ -3.162 \\ -1.112 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$
$$C_K = \begin{bmatrix} -1.428 & -10.2948 & -2.449 & 0 & 0 & 0 & 0 & 0 & 9.837 \end{bmatrix} \quad D_K = 0$$

A Figura 6.5 apresenta os valores dos funcionais para as normas Euclideana e infinito dentro da curva de *trade-off* do problema.

6.10 Método das Desigualdades

Um dos mais conhecidos métodos para projeto de controladores através de técnicas de otimização é o chamado Método das Desigualdades (Zakian e Al-Naib, 1973; Maciejowski, 1989). O método

Tabela 6.2: Solução do Algoritmo Básico para a norma Euclideana.

k	y'	w'	$\phi_{rms-u}(H)$	$\phi_{rms-y_p}(H)$	$\Theta(y)$
0	[0 0]	[0.4512 0.5488]	0.0668	0.0609	0.0636
1	[0.0568 0.0691]	[0.5182 0.4818]	0.0581	0.0692	$6.58 \cdot 10^{-4}$

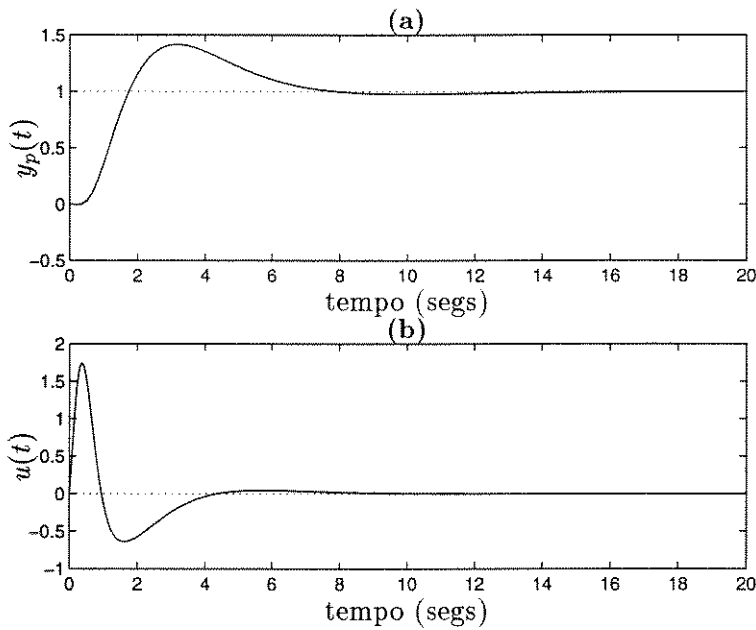


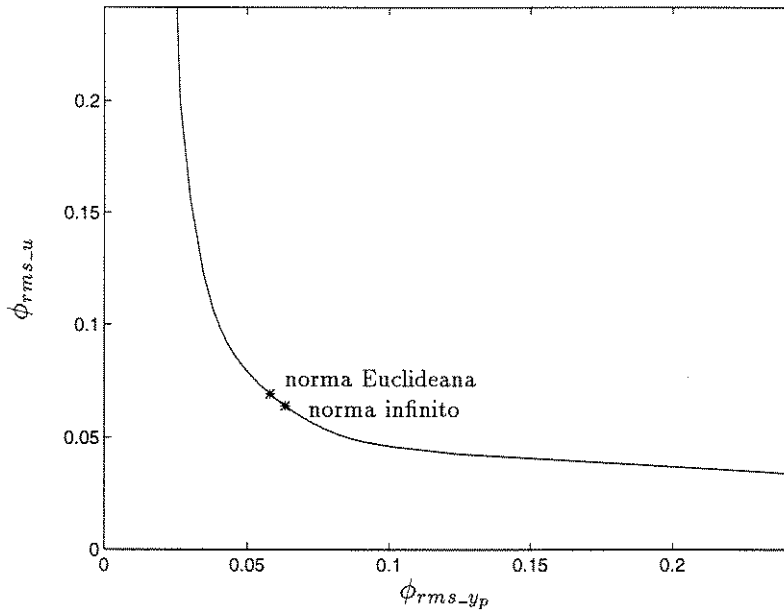
Figura 6.4: Resposta ao degrau de (a) H_{13} e (b) H_{23} para norma Euclideana.

requer o estabelecimento de limites para as especificações de desempenho. De acordo com o método, a solução para o problema de projeto é qualquer controlador K que satisfaça as desigualdades

$$\begin{aligned} \phi_1(K) &\leq a_1 \\ \phi_2(K) &\leq a_2 \\ &\vdots \\ \phi_m(K) &\leq a_m \end{aligned} \tag{6.40}$$

onde $\phi_1(K)$, $\phi_2(K)$, ..., $\phi_m(K)$ são os critérios de desempenho e a_1 , a_2 , ..., a_m são os limites ou níveis de satisfação do projetista. Em Zakian e Al-Naib (1973), uma variante do algoritmo de Resenbrock para problemas de otimização não-diferenciáveis é utilizado no sentido de obter uma solução para (6.40).

O Método das Desigualdades apresenta algumas desvantagens importantes. O problema de otimização derivado do método é não-convexo, não-diferenciável e extremamente sensível aos limites a_1 , a_2 , ..., a_m . Para um dado conjunto de valores de a_1 , a_2 , ..., a_m pode não haver solução

Figura 6.5: Curva de *trade-off*.

factível e, mesmo havendo solução factível, não existe nenhum procedimento que permita eventualmente melhorar os limitantes. Por outro lado, por trabalhar diretamente sobre os parâmetros do controlador K , o método permite definir a estrutura (PID, por exemplo) e a ordem do controlador.

Se interpretado no contexto deste Capítulo, uma formulação equivalente do método das desigualdades seria encontrar $Q \in \mathcal{RH}_\infty$ tal que

$$\begin{aligned} \phi_1(Q) &\leq a_1 \\ \phi_2(Q) &\leq a_2 \\ &\vdots \\ \phi_m(Q) &\leq a_m \end{aligned} \tag{6.41}$$

O problema associado agora é convexo e existe uma solução factível para as desigualdades *se e somente se*

$$\max_{w \in W} \min_{Q \in \mathcal{Q}} w'(\Phi(Q) - a) \tag{6.42}$$

for menor ou igual a zero, onde $a' = [a_1 \ a_2 \ \dots \ a_m]$. Em outras palavras, através da abordagem convexa, é possível testar a factibilidade de (6.41). Se a é tal que (6.41), a abordagem convexa informa *como* os novos valores de a devem ser escolhidos de forma a se tender à factibilidade, isto é, a deve satisfazer

$$(w^*)'a \geq (w^*)'\Phi(Q^*) \tag{6.43}$$

onde (w^*, Q^*) resolve (6.42). Assim, a abordagem convexa permite melhorar os limitantes de acordo com alguma função valor $V(\cdot)$ que agregue as preferências do projetista (decisor). A função $V(\cdot)$ não precisa ser explícita e, em projetos que envolvam apenas dois critérios, o processo de escolha e adaptação dos valores de a pode ser representado no plano — espaço dos critérios — conforme ilustração da Figura 6.6.

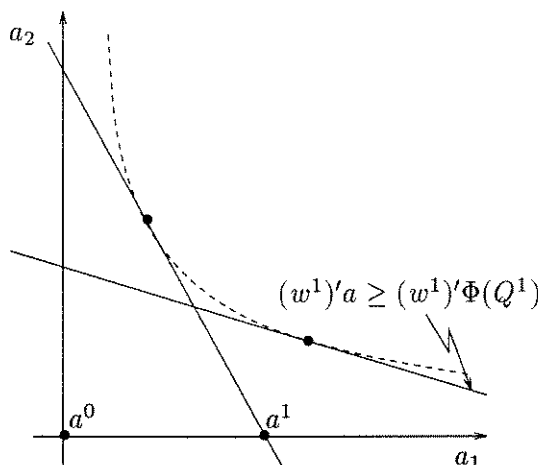


Figura 6.6: Ilustração do método - caso bi-critério.

Se, apesar de não explicitar sua função valor $V(\cdot)$, o projetista seleciona valores de a respeitando as hipóteses globais sobre $V(\cdot)$, como crescimento em relação a cada critério, então seguramente a solução final será eficiente, isto é, estará sobre a curva de trade-off do problema de projeto. Neste sentido, a abordagem convexa proposta conduz aos limites de desempenho do sistema considerado.

6.11 Conclusão

Como o problema resultante da parametrização de Youla (ou parametrização Q) possui dimensão infinita, foi necessário desenvolver uma aproximação finita do conjunto de matrizes de transferência. Esta aproximação, feita a partir do método de Ritz, gera um problema finito, convexo e computacionalmente tratável.

Foi discutida a aplicação do método de planos de corte para resolver o problema de otimização, que faz uso do conceito de subgradiente. Foi discutida ainda a implementação do Algoritmo Básico para resolver o problema de projeto de controladores com múltiplos critérios de projeto.

A formulação do problema no espaço dos critérios, e a subsequente divisão do seu tratamento em dois níveis tornam o procedimento bastante versátil. O problema no nível de decisão oferece grande flexibilidade no que diz respeito à escolha da função valor mais adequada ao problema. O problema no nível de análise fica completamente caracterizado pela solução do nível de decisão e, com isso, a dinâmica do sistema pode ser tratada isoladamente. O problema min-max foi apresentado e resolvido através das técnicas de parametrização de Youla.

Um exemplo numérico bi-critério com estrutura 1-DOF foi resolvido como forma de ilustrar o funcionamento do Algoritmo Básico.

Finalmente, estabeleceu-se uma comparação entre a abordagem convexa e o chamado Método das Desigualdades, através do qual foi possível evidenciar algumas vantagens da técnica proposta.

Capítulo 7

Projeto de Controladores l_1

7.1 Introdução

Este Capítulo tratará da solução de problemas formulados a partir da utilização da norma l_1 . Em função das características do método que será adotado, o sistema dinâmico deverá ser discreto. No decorrer do Capítulo, alguns conceitos referentes a sistemas discretos serão revistos. Vale a pena ressaltar que esta classe de problemas tem recebido muita atenção em trabalhos recentes, o que pode ser atestado pelo crescente número de publicações a respeito (veja, por exemplo, Diaz-Bobillo e Dahleh, 1993; Dahleh e Diaz-Bobillo, 1995).

As seções a seguir tratarão da formulação do problema de norma l_1 , bem como da caracterização de seu conjunto de soluções factíveis. Em seguida será discutida a formulação do problema resultante de dimensão infinita como um problema de programação linear de dimensão finita.

7.2 Transformada λ

Antes de se partir para a formulação do problema, serão apresentadas a seguir notações e conceitos relativos a sistemas discretos (Dahleh e Diaz-Bobillo, 1995).

l_1 Espaço das sequências absolutamente somáveis sobre os inteiros não-negativos. Se $x \in l_1$, então $\|x\|_1 = \sum_{k=0}^{\infty} |x(k)| < \infty$.

$l_1^{p \times q}$ Espaço das matrizes $p \times q$ com elementos no l_1 . Se $M \in l_1^{p \times q}$, então $\|M\|_1 \triangleq \max_{1 \leq i \leq p} \sum_{j=1}^q \|m_{ij}\|_1$.

- l_∞ Espaço de todas as seqüências limitadas de números reais sobre os inteiros não-negativos. Se $x \in l_\infty$ então $\|x\|_1 = \sup_k |x(k)| < \infty$.
- $l_\infty^{p \times q}$ Espaço das matrizes $p \times q$ com elementos no l_∞ . Se $M \in l_\infty^{p \times q}$, então $\|M\|_1 \triangleq \sum_{j=1}^p \max_{1 \leq i \leq q} \|m_{ij}\|_\infty$.
- $c_0^{p \times q}$ Subespaço de $l_\infty^{p \times q}$ formado por matrizes cujos elementos (seqüências) decaem para zero, i.e., $\lim_{k \rightarrow \infty} m_{ij}(k) = 0$ para todo i, j .

Trunc Operador Truncamento: dada uma seqüência infinita N , $\hat{N} = \text{Trunc}_T\{N\}$

é tal que $\hat{N} = N$, para $0 \leq t \leq T$, e $\hat{N} = 0$, para $t > T$.

Dada uma matriz M , M_i e M^j denotarão a i -ésima linha e a j -ésima coluna de M , respectivamente.

Definição 7.1 (Transformada λ): Para $N \in l_1^{m \times n}$, a transformada λ de N (denota-se $\tilde{N}$) é definida como

$$\tilde{N}(\lambda) = \sum_{i=0}^{\infty} N(i) \lambda^i$$

onde λ é a variável complexa que representa o atraso unitário. Portanto, $\tilde{N}$ é analítica no disco unitário aberto $\mathcal{D}$ e contínua na fronteira.

Por comodidade, não será usado o símbolo $\sim$ para diferenciar sistemas discretos de sistemas contínuos; todos os sistemas deste Capítulo devem ser considerados discretos, a não ser que se especifique o contrário.

O uso de transformada λ deve-se ao fato dela apresentar uma estrutura polinomial, o que viabiliza o desenvolvimento do problema tratado como um problema de programação linear e facilita a caracterização da solução a partir de seqüências do tipo FIR (resposta ao impulso finita).

7.3 Formulação do Problema

O problema foi formulado usando-se a modelagem discutida no Capítulo 5 (ver Figura 5.2). O modelo é representado por um operador linear invariante no tempo P , de dimensão infinita, que mapeia as entradas v e u para as saídas z e y .

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{zv} & P_{zu} \\ P_{yv} & P_{yu} \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} \quad (7.1)$$

Portanto, a solução é encontrada levando-se em conta a matriz de transferência entre as saídas reguladas z e as entradas exógenas v .

$$H = P_{zv} + P_{zu} K (I - P_{yu} K)^{-1} P_{yv} \quad (7.2)$$

O problema de controle ótimo l_1 consiste em encontrar o controlador estabilizante que minimiza o ganho de pico-a-pico máximo de H sobre o espaço de perturbações limitadas com norma unitária (Dahleh e Diaz-Bobillo, 1995). Em outras palavras, consiste em

$$\min_{K_{estvl}} \max_{\max_{1 \leq i \leq n_v} \|v_i\|_\infty = 1} \left(\max_{1 \leq k \leq n_z} \|(Hv)_k\|_\infty \right) \quad (7.3)$$

$$\min_{K_{estvl}} \|H\|_1$$

Como foi discutido no Capítulo 6, o conjunto das matrizes de transferência em malha fechada H estáveis pode ser mapeado variando-se a matriz estável $Q \in l_1^{n_u \times n_y}$ sobre a expressão afim

$$H = T_1 + T_2 Q T_3 \quad (7.4)$$

onde T_1 , T_2 e T_3 são mapeamentos (matrizes racionais) estáveis que dependem unicamente de P . A formulação do problema em termos de Q fica sendo

$$\min_{Q \in l_1^{n_u \times n_y}} \|T_1 + T_2 Q T_3\|_1 \quad (7.5)$$

Convencionando que $T_1 - T_2(-Q)T_3$, e incorporando o sinal negativo em Q , o problema anterior pode ser redefinido como um problema de distância mínima entre sistemas estáveis (*model matching problem*):

$$\min_{R \in \mathcal{L}} \|T_1 - R\|_1 = \min_{T_1 - H \in \mathcal{L}} \|H\|_1 \quad (7.6)$$

onde

$$\mathcal{L} \triangleq \{R \in l_1^{n_z \times n_w} : R = T_2 Q T_3 \text{ para algum } Q \in l_1^{n_u \times n_y}\} \quad (7.7)$$

representa o conjunto de R 's factíveis.

As próximas seções discutirão formas adequadas do problema de distância mínima (7.6) que possam ser manipuladas computacionalmente. Neste sentido é necessário caracterizar $\mathcal{L}$ de forma explícita, o que será feito a seguir através das *condições de interpolação*.

7.4 Condições de Interpolação

O problema original de minimização da norma l_1 de H sobre o conjunto dos controladores estabilizantes foi reformulado como um problema clássico de *model matching* em $\mathcal{L}$ através da Q -parametrização de todas as matrizes de transferências estáveis ($H = T_1 - T_2 Q T_3$). Portanto,

torna-se necessário estabelecer condições explícitas para que o valor ótimo computado R^* pertença a $\mathcal{L}$, sendo então possível encontrar Q^* estável, que por sua vez levará a K^* . Estas condições são conhecidas por *condições de interpolação*.

De uma forma simples, fica claro a partir do caso escalar (SISO), que como $R = (T_2 T_3)Q = TQ$ (ou seja $Q = R/T$), os pólos de Q possuem uma estreita relação com os zeros de T . Portanto, para garantir que Q seja estável, basta fazer com que os termos relativos aos zeros instáveis de T façam parte dos zeros de R , para que possam ser cancelados. Como o conceito de zeros para o caso de matrizes de transferência envolve informações frequenciais e direcionais - para uma mesma frequência, uma matriz de transferência pode ter um zero numa direção mas não em outra - as condições de interpolação são divididas em duas: condições de zero e condições de rank (Dahleh e Diaz-Bobillo, 1995).

Toda a estrutura dos zeros de uma matriz racional e suas direções são diretamente visualizadas através de sua *decomposição de Smith-McMillan* (Dahleh e Diaz-Bobillo, 1995) onde, supondo-se uma matriz racional genérica N de rank r , tem-se

$$N = L_N M_N R_N$$

sendo que L_N e R_N são matrizes unimodulares e M tem a forma

$$M_N = \begin{pmatrix} \frac{\epsilon_1}{\psi_1} & 0 & \dots & 0 & \dots & 0 \\ 0 & \ddots & & \vdots & \ddots & \vdots \\ \vdots & & \frac{\epsilon_r}{\psi_r} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

As informações sobre os pólos, zeros e suas respectivas multiplicidades σ estão em ϵ_i e ψ_i , $i = 1, \dots, r$ e as informações sobre as direções, tanto à direita quanto à esquerda estão em L_N e R_N .

Para o caso em questão (MIMO), é necessário garantir a manutenção da estrutura de zeros à esquerda de T_2 , bem como a estrutura de zeros à direita de T_3 . A seguir apresentam-se as condições de interpolação do problema.

Teorema 7.1 *Seja uma matriz de transferência em malha fechada $H = T_1 - T_2 Q T_3$, e considere a decomposição de Smith-McMillan de T_2 e T_3 :*

$$T_2 = L_{T_2} M_{T_2} R_{T_2}$$

$$T_3 = L_{T_3} M_{T_3} R_{T_3}$$

Defina os vetores-linha

$$\alpha_i(\lambda) = (L_{T_2}^{-1})_i(\lambda) \quad i = 1, 2, \dots, n_z$$

os vetores-coluna

$$\beta_j(\lambda) = (R_{T_3}^{-1})^j(\lambda) \quad j = 1, 2, \dots, n_v$$

e o conjunto $\Lambda_{T_2 T_3}$ formado pelos zeros instáveis de T_2 e T_3 . Então, para uma dada matriz $R \in l_1^{n_z \times n_v}$, existe $Q \in l_1^{n_u \times n_y}$ tal que $R = T_2 Q T_3$ se e somente se, para todo $\lambda_0 \in \Lambda_{T_2 T_3} \subset \mathcal{D}$, as seguintes condições são satisfeitas:

1. *Condição de Interpolação de Zeros:*

$$(\alpha_i R \beta_j)^{(k)}(\lambda_0) = 0, \quad \text{para } \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ k = 0, \dots, \sigma_i(\lambda_0) + \sigma_j(\lambda_0) - 1 \end{cases}$$

2. *Condição de Interpolação de Rank:*

$$\begin{cases} (\alpha_i R)(\lambda) \equiv 0, & \text{para } i = n_u + 1, \dots, n_z \\ (R \beta_j)(\lambda) \equiv 0, & \text{para } j = n_y + 1, \dots, n_v \end{cases}$$

onde o índice k representa derivada de ordem k e σ a multiplicidade de λ_0 na cadeia i ou j .

Prova: Ver Dahleh e Diaz-Bobillo (1995)

□

Este resultado é de grande importância para o cálculo de R , dando espaço para uma extensa análise de sua aplicação. Porém, este Capítulo vai se ater apenas ao caso de sistemas mono-bloco, onde $n_z = n_u$ e $n_v = n_y$, pois caso o sistema seja multi-bloco, o método selecionado pode ainda ser aplicado através de uma aproximação mono-bloco do problema. Para um estudo detalhado das condições de interpolação e suas aplicações, ver Dahleh e Diaz-Bobillo (1995) e referências internas.

7.5 Solução do Problema Mono-bloco

A primeira implicação no tratamento de sistemas mono-blocos é a utilização apenas da *condição de interpolação de zeros*, uma vez que a *condição de interpolação de rank* é naturalmente satisfeita pela estrutura do problema.

7.5.1 Aspectos Computacionais da Condição de Interpolação de Zero

Na seção anterior foi feita a caracterização algébrica de $\mathcal{L}$ a partir da decomposição de Smith-McMillan, que é a forma mais natural de se determinar a estrutura de zeros de uma matriz racional. Porém, como o cálculo desta decomposição é muito mal comportado do ponto de vista computacional, especialmente quando ocorrem zeros múltiplos, optou-se por seguir a metodologia

proposta por Dahleh e Diaz-Bobillo (1995), que consiste em descrever as condições de interpolação de zeros a partir de uma cadeia de vetores nulos. Esta representação é equivalente à anterior com a vantagem de não ter problemas numéricos, assumindo apenas que a posição dos zeros de T_2 e T_3 são conhecidas. Para tanto, algumas definições serão feitas.

Definição 7.2 *Seja uma matriz racional genérica $N(\lambda)$, analítica em λ_0 , um número inteiro positivo σ e a expansão de Taylor de N*

$$N(\lambda) = N_0 + (\lambda - \lambda_0)N_1 + (\lambda - \lambda_0)^2 N_2 + \dots + (\lambda - \lambda_0)^{\sigma-1} N_{\sigma-1}$$

onde $N_i = \frac{1}{i!}(N)^{(i)}(\lambda_0)$. *Define-se a seguinte matriz Toeplitz bloco-diagonal inferior*

$$T_{\lambda_0, \sigma}(N) = \begin{bmatrix} N_0 & 0 & 0 & \dots & 0 \\ N_1 & N_0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ N_{\sigma-1} & N_{\sigma-2} & N_{\sigma-3} & \dots & N_0 \end{bmatrix}$$

O que será feito é explorar a relação que existe entre as condições de interpolação e a estrutura do espaço nulo de $T_{\lambda_0, \sigma}$.

Definição 7.3 *Seja uma matriz racional N de dimensão $m \times n$, analítica em λ_0 . Defina uma cadeia nula à direita de ordem σ em λ_0 , como o conjunto ordenado de vetores do $\mathbb{R}^n$, $X = \{x^1, x^2, \dots, x^\sigma\}$, tal que $x^1 \neq 0$ e*

$$T_{\lambda_0, \sigma}(N) \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^\sigma \end{bmatrix} = 0.$$

Do mesmo modo, defina uma cadeia nula à esquerda de ordem σ em λ_0 , como o conjunto ordenado de vetores do $\mathbb{R}^m$, $Y = \{y^1, y^2, \dots, y^\sigma\}$, tal que $y^1 \neq 0$ e

$$T_{\lambda_0, \sigma}(N') \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^\sigma \end{bmatrix} = 0.$$

Um resultado importante no estabelecimento das condições de interpolação de zeros a partir das cadeias nulas associadas aos zeros da matriz, diz que caso N seja quadrada, então a existência de uma cadeia nula à direita/esquerda de ordem σ em λ_0 é equivalente a existência de um zero em λ_0 de multiplicidade algébrica σ . Caso N possua zeros com multiplicidade geométrica $l > 1$, então existem l cadeias nulas à direita/esquerda diferentes, não necessariamente de mesma ordem, de forma que juntas representam uma base para o espaço nulo de $N(\lambda_0)$ (cadeias nulas à direita) e uma base para $N'(\lambda_0)$ (cadeias nulas à esquerda). Estes conceitos são usados na definição a seguir.

Definição 7.4 *Sejam X^i e Y^i as cadeias nulas de índice i à direita e à esquerda respectivamente, cada uma com sua ordem σ_i . Defina um conjunto canônico de cadeias nulas à direita de N em λ_0 , como um conjunto ordenado de cadeias nulas à direita (i.e., $X^i = (x^{i,1}, \dots, x^{i,\sigma_i})$) para $i = 1, \dots, l$, tal que*

1. $\{x^{1,1}, x^{2,1}, \dots, x^{l,1}\}$ são linearmente independentes;
2. $\text{span}\{x^{1,1}, x^{2,1}, \dots, x^{l,1}\} = \mathcal{N}[N(\lambda_0)]$;
3. $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_l$

e um conjunto canônico de cadeias nulas à esquerda de N em λ_0 , como um conjunto ordenado de cadeias nulas à esquerda (i.e., $Y^i = (y^{i,1}, \dots, y^{i,\sigma_i})$) para $i = 1, \dots, l$, tal que

1. $\{y^{1,1}, y^{2,1}, \dots, y^{l,1}\}$ são linearmente independentes;
2. $\text{span}\{y^{1,1}, y^{2,1}, \dots, y^{l,1}\} = \mathcal{N}[N'(\lambda_0)]$;
3. $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_l$

Para garantir sempre a existência de uma base com n cadeias nulas, mesmo quando o número de cadeias não é suficiente ($l < n$), será usado o conceito de cadeia nula estendida, que completa o número necessário para a formação da base com $n - l$ cadeias de multiplicidade $\sigma = 0$.

Definição 7.5 *Um conjunto estendido de cadeias nulas à direita de uma matriz racional N de rank completo $n \times n$ no ponto λ_0 é um conjunto canônico de cadeias nulas à direita aumentadas com $n - l$ vetores do $\mathbb{R}^n$, tal que $\text{span}\{x^{1,1}, x^{1,2}, \dots, x^{1,l}, x^{1,l+1}, \dots, x^n\} = \mathbb{R}^n$.*

O conjunto de cadeias nulas estendidas à esquerda é definido de forma semelhante. Pode-se agora enunciar a *condição de interpolação de zeros* a partir dos conceitos de cadeias nulas à direita e à esquerda. Vale a pena lembrar que, apesar de não representar a estrutura de zeros de uma matriz de uma forma natural, como a decomposição de Smith-McMillan, as cadeias nulas trazem toda a informação sobre os zeros e suas direções, sendo facilmente implementáveis e bem comportadas do ponto de vista computacional.

Teorema 7.2 *Para um problema mono-bloco dado, as condições de interpolação de zeros do Teorema 7.1 são as seguintes: para todo $\lambda_0 \in \Lambda_{T_2 T_3}$,*

$$(y_{\lambda_0}^i R x_{\lambda_0}^j)^{(k)}(\lambda_0) = 0, \quad \text{para } \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ k = 0, \dots, \sigma_i(\lambda_0) + \sigma_j(\lambda_0) - 1 \end{cases}$$

onde

$$x_{\lambda_0}^j \triangleq x_1^j + (\lambda - \lambda_0)x_2^j + \dots + (\lambda - \lambda_0)^{\sigma_j-1}x_{\sigma_j}^j, \quad \text{caso } \sigma_j > 0,$$

$$x_{\lambda_0}^j \triangleq x_1^j, \quad \text{caso } \sigma_j = 0$$

e

$$y_{\lambda_0}^i \triangleq y_1^i + (\lambda - \lambda_0)y_2^i + \dots + (\lambda - \lambda_0)^{\sigma_i-1}y_{\sigma_i}^i, \quad \text{caso } \sigma_i > 0,$$

$$y_{\lambda_0}^i \triangleq y_1^i, \quad \text{caso } \sigma_i = 0. \quad \square$$

Toda esta discussão preliminar tem como propósito caracterizar o conjunto $\mathcal{L}$. Ainda dentro do processo de reformulação de (7.6), R será substituída por $T_1 - H$. Neste caso a condição de interpolação do Teorema 7.2 fica sendo

$$(y_{\lambda_0}^i H x_{\lambda_0}^j)^{(k)}(\lambda_0) = (y_{\lambda_0}^i T_1 x_{\lambda_0}^j)^{(k)}(\lambda_0), \quad \text{para } \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ k = 0, \dots, \sigma_i(\lambda_0) + \sigma_j(\lambda_0) - 1 \end{cases} \quad (7.8)$$

Concluída a caracterização de $\mathcal{L}$, as próximas seções representarão (7.6) como um problema de programação linear. Note que o conjunto de restrições gerado por (7.8) possui infinitas variáveis, uma vez que $H, T_1 \in l_1$.

7.6 Formulação Através de Programação Linear

O objetivo desta seção é definir matrizes A , c e b associadas ao problema (7.6), de forma que este possa ser escrito como um problema de programação linear do tipo

$$\begin{aligned} \min_x & c'x \\ \text{s.a.} & Ax \leq b \end{aligned}$$

7.6.1 Problema de Dimensão Infinita

Como o problema trata de sequências infinitas, todas as matrizes e vetores envolvidos na formulação do problema serão de dimensão infinita.

Variáveis: Em primeiro lugar, a resposta ao impulso do sistema

$$H(t) = \begin{bmatrix} h_{11}(t) & \dots & h_{1n_v}(t) \\ \vdots & & \vdots \\ h_{n_z1}(t) & \dots & h_{n_zn_v}(t) \end{bmatrix}$$

será transformada em um vetor coluna h , onde

$$h(t) \triangleq \begin{bmatrix} h_{11}(t) \\ \vdots \\ h_{1n_v}(t) \\ \vdots \\ h_{n_z1}(t) \\ \vdots \\ h_{n_zn_v}(t) \end{bmatrix}$$

e, como cada elemento da matriz $H(t)$ representa uma sequência infinita, no vetor h os elementos da sequência estarão empilhados, i.e.,

$$h = \begin{bmatrix} h(0) \\ h(1) \\ h(2) \\ h(3) \\ \vdots \end{bmatrix}$$

Note que o procedimento acima apenas mudou a forma como os elementos de cada sequência são armazenados, não alterando a natureza da matriz H e suas características, especialmente quanto às medidas de norma, ou seja

$$\|H\|_1 = \|h\|_1 = \max_{1 \leq i \leq n_z} \sum_{j=1}^{n_v} \sum_{t=0}^{\infty} |h_{ij}(t)|$$

Funcional: O próximo passo é eliminar a não-linearidade do funcional $\|h\|_1$ fazendo a seguinte mudança de variável: $h = h^+ - h^-$, com $h^+, h^- \geq 0$. Com isso, tem-se que

$$\|H\|_1 = \max_{1 \leq i \leq n_z} \sum_{j=1}^{n_v} \sum_{t=0}^{\infty} h_{ij}^+(t) + h_{ij}^-(t),$$

Restrições - Matriz A: O operador linear que vai atuar sobre h associado ao conjunto de restrições (7.8) é definido com A_{zero} , onde cada linha da matriz A_{zero} está associada a uma das restrições definidas em (7.8). Os primeiros $n_z \cdot n_v$ elementos de cada linha vão atuar sobre $h(0)$, os próximos $n_z \cdot n_v$ vão atuar sobre $h(1)$, os próximos sobre $h(2)$, e assim por diante. Portanto, a lei de formação dos $n_z \cdot n_v$ elementos de cada linha de A_{zero} é a seguinte: sejam as matrizes $n_z \times n_v$

$$a_{ij\lambda_0} \triangleq y_{\lambda_0}^i x_{\lambda_0}^j, \quad \text{para cada } \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ \lambda_0 \in \Lambda_{T_2 T_3} \end{cases}$$

$$A_{zero_{i+j+k}}(t) \triangleq [(a_{11\lambda_0}, \dots, a_{1v\lambda_0}, \dots, a_{z1\lambda_0}, \dots, a_{zv\lambda_0}) \lambda^t]^{(k)} \Big|_{\lambda=\lambda_0},$$

$$\text{para } \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ k = 0, \dots, \sigma_i(\lambda_0) + \sigma_j(\lambda_0) - 1 \\ \lambda_0 \in \Lambda_{T_2 T_3} \end{cases}$$

Com isso, cada linha da matriz A_{zero} fica sendo

$$A_{zero_{i+j+k}} \triangleq [A_{zero_{i+j+k}}(0) \ A_{zero_{i+j+k}}(1) \ A_{zero_{i+j+k}}(2) \dots]$$

Note que, percorrendo as linhas de A_{zero} , a cada $n_z \cdot n_v$ posições os elementos são relativos a uma mesma subsequência da matriz $h(t)$. Para facilitar a preparação do problema de programação linear, cada subsequência será individualmente armazenada numa matriz M_{ij} , onde

$$M_{ij} = (A_{zero})_{n_z \cdot n_v t + n_v(i-1) + j} \quad \text{para} \quad \begin{cases} i = 1, \dots, n_u \\ j = 1, \dots, n_y \\ t = 0, 1, \dots \end{cases}$$

Restrições - Vetor b : O produto $A_{zero}T_1(t)$, $t = 0, 1, \dots$ gera o vetor b_{zero} do problema de programação linear, sendo que $T_1(t)$ é um vetor coluna de dimensão infinita formado a partir da matriz de transferência T_1 da mesma forma que $h(t)$ foi formado a partir de H .

Problema Primal Tendo-se gerado todas as matrizes e vetores necessários, o problema de programação linear equivalente ao problema (7.6) original é anunciado a seguir.

$$\min_{h^+, h^-} \left\{ \max_{i \in (1, \dots, n_z)} \sum_{j=1}^{n_v} \sum_{t=0}^{\infty} h_{ij}^+(t) + h_{ij}^-(t) \right\}$$

s.a.

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_v} M_{ij}(h_{ij}^+ - h_{ij}^-) = b_{zero}$$

$$h_{ij}^+, h_{ij}^- \geq 0$$

Para efeito de implementação computacional, o problema é ainda reescrito como um problema de minimizar um limitante superior v , usando-se variáveis de folga ξ para transformar todas as restrições em restrições de igualdade:

$$\min_{v, \xi, h^+, h^-} v$$

s.a.

$$\sum_{j=1}^{n_v} \sum_{t=0}^{\infty} h_{ij}^+(t) + h_{ij}^-(t) + \xi(i) = v, \quad \text{para } i = 1, \dots, n_z$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_v} M_{ij}(h_{ij}^+ - h_{ij}^-) = b_{zero}$$

$$\xi, h_{ij}^+, h_{ij}^- \geq 0$$

Problema Dual Ao problema primal associa-se o seguinte problema dual com infinitas restrições, que será de grande utilidade na determinação de um problema equivalente de dimensão finita:

$$\max_{\gamma_0, \gamma_1} \langle b_{zero}, \gamma_1 \rangle$$

s.a.

$$\sum_{i=1}^{n_z} \gamma_0(i) \leq 1$$

$$-\gamma_0(i) \leq (M'_{ij}\gamma_1)(t) \leq \gamma_0(i) \quad \text{para} \quad \begin{cases} i = 1, \dots, n_z \\ j = 1, \dots, n_v \\ t = 1, 2, \dots \end{cases}$$

$$\gamma_0 \geq 0$$

7.6.2 Problema de Dimensão Finita

Apesar do problema de minimizar a norma l_1 ser um problema de dimensão infinita, o problema linear resultante é estruturalmente de dimensão finita. Em primeiro lugar, o fato de se considerar problemas mono-blocos faz com que o problema primal tenha um número de restrições finito, embora ainda com infinitas variáveis. Agora, analisando-se o problema dual, é possível mostrar (Dahleh e Diaz-Bobillo, 1995) que, como os valores das sequências tendem a zero a medida que t tende para infinito, a partir de um determinado valor $t = T$, todas as restrições do problema dual relativas a $t > T$ nunca estarão ativas. Este resultado é consequência do seguinte Lema (Dahleh e Diaz-Bobillo, 1995).

Lema 7.1 *Seja M uma matriz com infinitas linhas e rank igual ao número de colunas, que mapeia $\mathbb{R}^n$ para c_0 . Então existe um inteiro positivo τ tal que*

$$\|(I - \text{Trunc}_\tau)M'x\|_\infty < \|\text{Trunc}_\tau\{M'\}x\|_\infty$$

para todo $x \in \mathbb{R}^n$ não nulo.

□

Consequentemente, a solução do problema dual considerando-se $t = 0, \dots, \tau_{max}$, $\tau_{max} = \max_{i,j} \tau_{ij}$ será equivalente à solução para $t \geq 0$. Para o caso mono-bloco, tem-se o seguinte par primal-dual de problemas de programação linear de dimensão finita.

Problema Primal

$$v^* = \min_{v, \xi, h^+, h^-} v$$

s.a.

$$\sum_{j=1}^{n_v} \sum_{t=0}^{\tau_{ij}} h_{ij}^+(t) + h_{ij}^-(t) + \xi(i) = v, \quad \text{para } i = 1, \dots, n_z$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_v} M_{ij}(h_{ij}^+ - h_{ij}^-) = b_{zero}$$

$$\xi, h_{ij}^+, h_{ij}^- \geq 0$$

Problema Dual

$$v^* = \max_{\gamma_0, \gamma_1} < b_{zero}, \gamma_1 >$$

s.a.

$$\sum_{i=1}^{n_z} \gamma_0(i) \leq 1$$

$$-\gamma_0(i) \leq (M'_{ij} \gamma_1)(t) \leq \gamma_0(i) \quad \text{para } \begin{cases} i = 1, \dots, n_z \\ j = 1, \dots, n_v \\ t = 1, \dots, \tau_{ij} < \infty \end{cases}$$

$$\gamma_0 \geq 0$$

7.7 Solução do Problema Multi-Bloco

Devido à necessidade de se considerar as restrições de interpolação de rank, o problema primal para o caso multi-bloco sempre terá infinitas restrições, enquanto que o problema dual terá infinitas variáveis. Portanto, de uma forma geral, a solução do problema multi-bloco não é exata.

O método discutido nesta seção para resolver o problema multi-bloco consiste basicamente em fazer uma aproximação mono-bloco, resolver o problema modificado e recuperar a solução do problema original. A aproximação para sistemas mono-blocos é feita aumentando-se as matrizes T_2 e T_3 com atrasos puros (deslocamentos à direita), sendo por isso chamado de método DA (*Delay Augmentation*). O desenvolvimento a seguir é baseado em Dahleh e Diaz-Bobillo (1995).

Em primeiro lugar, particiona-se as matrizes T_2 e T_3 em função das dimensões de u e y ,

$$T_2 = \begin{pmatrix} T_{2_1} \\ T_{2_2} \end{pmatrix}, \quad T_3 = \begin{pmatrix} T_{3_1} & T_{3_2} \end{pmatrix}$$

onde $T_{2_1} \in l_1^{n_u \times n_u}$ e $T_{3_1} \in l_1^{n_y \times n_y}$. Em seguida, aumenta-se estas matrizes com atrasos unitários de ordem τ_d de forma a obter-se o seguinte sistema aumentado:

$$H_{\tau_d} = \begin{pmatrix} H_{11,\tau_d} & H_{12,\tau_d} \\ H_{21,\tau_d} & H_{22,\tau_d} \end{pmatrix} = \begin{pmatrix} T_{1_{11}} & T_{1_{12}} \\ T_{1_{21}} & T_{1_{22}} \end{pmatrix} - \begin{pmatrix} T_{2_1} & 0 \\ T_{2_2} & S_{\tau_d} \end{pmatrix} Q_{\tau_d} \begin{pmatrix} T_{3_1} & T_{3_2} \\ 0 & S_{\tau_d} \end{pmatrix} \quad (7.9)$$

sendo

$$Q_{\tau_d} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \quad (7.10)$$

O sistema aumentado gera o seguinte problema mono-bloco:

$$\min_{Q_{\tau_d} \in l_1^{n_u \times n_z}} \|H_{\tau_d}\|_1 \quad (7.11)$$

Note que H_{τ_d} pode ser reescrito como

$$H_{\tau_d} = T_1 - T_2 Q_{11} T_3 - S_{\tau_d} \bar{R}_{\tau_d} = H - S_{\tau_d} \bar{R}_{\tau_d} \quad (7.12)$$

onde

$$\bar{R}_{\tau_d} \triangleq \begin{pmatrix} 0 & T_{2_1} Q_{12} \\ Q_{21} T_{3_1} & Q_{21} T_{3_2} + T_{2_2} Q_{12} + S_{\tau_d} Q_{22} \end{pmatrix}$$

O problema multi-bloco original pode ser formulado como

$$\|H\|_1^* = \min_{Q_{11} \in l_1^{n_u \times n_z}} \|H_{\tau_d}\|_1 \quad (7.13)$$

$$s.a \quad Q_{12} = Q_{21} = Q_{22} = 0$$

Com isso, (7.13) possui todas as restrições de (7.11) e restrições adicionais (referentes às partições de Q), implicando em $\|H_{\tau_d}\|_1^* \leq \|H\|_1^*$, que constitui um limitante inferior do problema original, denotado por $\underline{\eta}_{\tau_d}$.

O sistema aumentado foi construído de tal forma que os parâmetros livres acrescentados à matriz Q (i.e., Q_{12} , Q_{21} , Q_{22}) não afetam a solução para $t < \tau_d$ devido aos τ_d deslocamentos. Além disso, o termo H_{11} não é afetado por estes parâmetros extras (bloco de zeros em $\bar{R}_{\tau_d}$).

Finalmente, (7.13) fornece a menor norma dentre todas as matrizes Q tais que $Q_{12} = Q_{21} = Q_{22} = 0$, inclusive para a matriz

$$Q = \begin{bmatrix} Q_{\tau_d,11}^* & 0 \\ 0 & 0 \end{bmatrix},$$

que substituída em (7.12), fornece o valor

$$\bar{\eta}_{\tau_d} = \|T_1 - T_2 Q_{\tau_d,11}^* T_3\|_1 \geq \|H\|_1^* \quad (7.14)$$

Portanto, tem-se que $\bar{\eta}_{\tau_d}$ constitui-se em limitante superior de $\|H\|_1^*$ e Q_{11} em solução subótima do problema.

Através do método DA, aumenta-se iterativamente o valor da ordem τ_d dos atrasos, até que $\underline{\eta}_{\tau_d}$ esteja suficientemente próximo de $\bar{\eta}_{\tau_d}$.

7.8 Comentário sobre a Utilização do Algoritmo Básico

A partir do que foi exposto até o momento, fica claro que em se tratando de projeto de sistemas discretos através de critérios do tipo norma l_1 , o nível de análise do Algoritmo Básico (problema min-max) envolverá apenas problemas de programação linear. Entretanto, problemas multicritério lineares, como os induzidos pelo formalismo baseados na norma l_1 , envolvem conceitos e técnicas próprias, mais adaptadas à solução de problemas lineares. Assim, ao invés de um algoritmo geral como o Algoritmo Básico, nesta parte do trabalho utilizou-se uma técnica específica, particularmente adequada às condições do problema.

7.9 Solução Via Programação Alvo

Considerando que se deseja minimizar vários critérios de desempenho, tem-se então um problema de projeto de controladores multicritério. Uma técnica bastante eficiente para solucionar problemas multicritérios é a Programação Alvo (PA) (Ignizio, 1976), pois além de permitir a indicação de compromissos e prioridades entre os critérios, a técnica de solução via PA explora eficientemente os métodos baseados em programação linear.

De forma geral, seja o problema

$$\min_{K_{estvl}} \Phi$$

onde $\Phi = \{\phi_1, \phi_2, \dots, \phi_m\}$ é um vetor de critérios de desempenho. O problema de programação alvo consiste em

$$\min_{K_{estvl}} f(d_1^+, d_1^-, d_2^+, d_2^-, \dots, d_m^+, d_m^-)$$

$$\phi_1 - d_1^+ + d_1^- = \bar{\phi}_1$$

$$\phi_2 - d_2^+ + d_2^- = \bar{\phi}_2$$

$$\vdots$$

$$\phi_m - d_m^+ + d_m^- = \bar{\phi}_m$$

onde

- $\bar{\phi}_i$ representa a meta estabelecida para cada critério de desempenho;
- d_i^+, d_i^- representam os desvios com respeito à meta, para mais e para menos, respectivamente;

- $f(\cdot)$ representa uma função que indicará como será feita a ordenação e a composição dos desvios, ou seja,

$$f = Lex \left\{ \begin{array}{c} f_1(d_1^+, d_1^-, d_2^+, d_2^-, \dots, d_m^+, d_m^-) \\ f_2(d_1^+, d_1^-, d_2^+, d_2^-, \dots, d_m^+, d_m^-) \\ \vdots \\ f_m(d_1^+, d_1^-, d_2^+, d_2^-, \dots, d_m^+, d_m^-) \end{array} \right\}$$

onde *Lex* indica *ordenação lexicográfica*, isto é, a primeira função da lista é minimizada, a segunda função é então minimizada sujeito ao conjunto solução da primeira, e assim sucessivamente, até que a lista se esgote, ou que em alguma etapa seja encontrada uma solução única.

7.10 Exemplos Numéricos

A seguir são apresentados alguns resultados conseguidos a partir da aplicação do algoritmo de programação alvo. Os modelos tratados foram primeiramente resolvidos usando apenas a metodologia baseada no método DA, após o que foi aplicado a formulação via programação alvo. Portanto, para cada problema serão apresentadas as soluções encontradas para as duas formulações.

7.10.1 Problema do Caça X29

O primeiro problema resolvido foi baseado no modelo simplificado do caça X29, representado pela planta SISO (Dahleh e Diaz-Bobillo, 1995)

$$P(s) = \underbrace{\frac{s+3}{(s+10)(s-6)}}_{\text{airframe}} \underbrace{\frac{20}{(s+20)}}_{\text{atuador equiv overhead}} \underbrace{\frac{(s-26)}{(s+26)}}_{\text{overhead}}$$

O fator de *airframe* corresponde a um modelo simplificado da dinâmica de *pitch* do avião (considerado um corpo rígido) voando a baixa altitude e com velocidade do vento de 0.9 Mach.

Crítérios de Desempenho: O problema é originalmente formulado minimizando-se a norma l_1 da matriz de transferência discreta entre a entrada de ruído w , a sequência de controle ponderada z_1 e a saída ponderada z_2 (Figura 7.1):

$$\min_{K_{estvl}} \left\| \begin{array}{c} W_1 K S \\ W_2 S \end{array} \right\|_1 \quad (7.15)$$

As matrizes de transferência discretas foram encontradas a partir de um período de amostragem de 1/30 segundos.

Outro critério considerado foi a minimização do overshoot/undershoot do sinal de controle, uma vez que, tratando-se de um sistema de navegação é interessante evitar a saturação dos atuadores.

Primeiramente o método DA foi aplicado apenas ao problema de minimizar a norma l_1 (problema escalar). Os resultados encontrados são apresentados na Tabela 7.1. O critério de parada foi $\bar{\eta}/\underline{\eta} < 1.01$ (i.e., razão entre os limitantes menor que 1%).

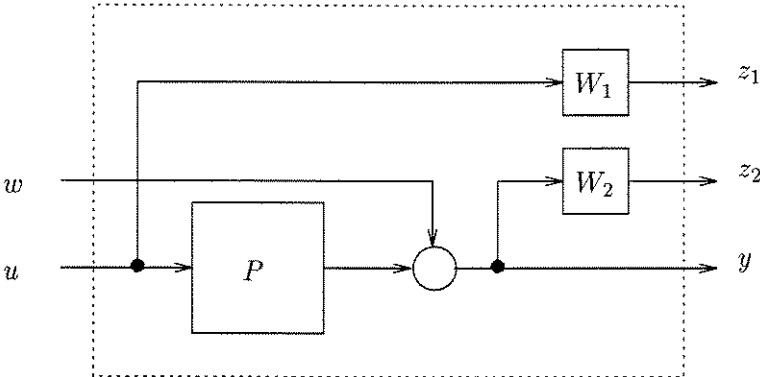


Figura 7.1: Problema do caça X29 dentro da arquitetura padrão.

Tabela 7.1: Caça X29: Resultados encontrados com o método DA.

τ_d	η	$\bar{\eta}$
4	3.2539	1256.4
5	4.0446	5.0591
6	4.0481	4.6523
7	4.0483	4.6300
8	4.0513	4.3163
9	4.0522	4.2240
10	4.0522	4.2189
20	4.0524	4.1982
30	4.0527	4.1747
40	4.0529	4.1488
50	4.0532	4.1250
60	4.0533	4.1060
70	4.0535	4.0938
80	4.0535	4.0861

O problema foi então resolvido dentro do contexto da programação alvo, usando-se os dois critérios de desempenho. Primeiramente foi considerada apenas a minimização do desvio superior da norma l_1 da matriz de transferência entre z e w . Como metas para os dois objetivos, considerou-se para a norma l_1 o valor de 1.0 e para o overshoot o valor de 10.0. O problema linear fica então sendo

$$\begin{aligned}
 & \min d_{l_1}^+ \\
 & \text{s.a} \\
 & v_{l_1} + d_{l_1}^- - d_{l_1}^+ = 1.0 \\
 & v_{ovsh} + d_{ovsh}^- - d_{ovsh}^+ = 10.0 \\
 & -v_{l_1} + \sum_{k=0}^{\tau_{d1}} (h_1^+(k) + h_1^-(k)) + d1^- = 0 \\
 & -v_{l_1} + \sum_{k=0}^{\tau_{d2}} (h_2^+(k) + h_2^-(k)) + d2^- = 0 \\
 & -v_{ovsh} + h_2(1) + d3^- = 0 \\
 & -v_{ovsh} + h_2(1) + h_2(2) + d4^- = 0 \\
 & -v_{ovsh} + h_2(1) + h_2(2) + h_2(3) + d5^- = 0 \\
 & \vdots \\
 & -v_{ovsh} + h_2(1) + \dots + h_2\left(\frac{\tau_{d2}}{div}\right) + d\tau_d^- = 0 \\
 & \quad \quad \quad A_{zero} \quad h = b
 \end{aligned} \tag{7.16}$$

onde τ_{d2}/div define até que instante a restrição sobre o overshoot será considerada, sendo que no problema em questão $div = 4$.

O resultados encontrados foram exatamente os mesmos do problema escalar anterior, com a diferença de que a Tabela 7.2 agora apresenta os valores do overshoot a cada valor de τ_d . O controlador encontrado foi o seguinte:

$$K(\lambda) = \frac{41.763\lambda^6 - 218.412\lambda^5 + 292.774\lambda^4 + 122.055\lambda^3 - 343.287\lambda^2 - 81.155\lambda + 187.220}{\lambda^6 + 1.975\lambda^5 - 1.769\lambda^4 - 3.033\lambda^3 - 0.846\lambda^2 + 0.145\lambda + 2.5289}$$

A Figura 7.2 apresenta a resposta ao degrau da matriz de transferência em malha fechada H encontrada, onde é possível visualizar o overshoot.

O problema foi novamente resolvido considerando-se os desvios relativos ao overshoot. Outra consideração a ser feita é que a solução do problema (7.16) é única e com isso não adianta fazer uma ordenação léxica. Usou-se então, como critério de desempenho $d_{l_1}^+ + d_{ovsh}^-$. A Tabela 7.3 apresenta os resultados para este novo caso, onde a função de transferência discreta do controlador encontrado é

$$\begin{aligned}
 K(\lambda) &= \\
 &= \frac{41.76\lambda^7 - 218.915\lambda^6 + 308.631\lambda^5 + 48.488\lambda^4 - 254.097\lambda^3 + 3.406\lambda^2 - 47.384\lambda + 119.148}{\lambda^7 + 1.961\lambda^6 - 1.471\lambda^5 - 2.482\lambda^4 - 0.830\lambda^3 - 0.759\lambda^2 + 0.068\lambda + 2.513}
 \end{aligned}$$

Tabela 7.2: Caça X29: Resultados encontrados via PA sem considerar overshoot.

τ_d	η	$\bar{\eta}$	overshoot
4	3.2539	1256.4	10.0
5	4.0446	5.0591	74.97
6	4.0481	4.6523	75.69
7	4.0483	4.6300	75.63
8	4.0513	4.3163	74.71
9	4.0522	4.2240	74.44
10	4.0522	4.2189	74.43
20	4.0524	4.1982	74.36
40	4.0529	4.1488	74.22
50	4.0532	4.1250	74.15
60	4.0533	4.1060	74.09
70	4.0535	4.0938	74.06
80	4.0535	4.0861	74.03

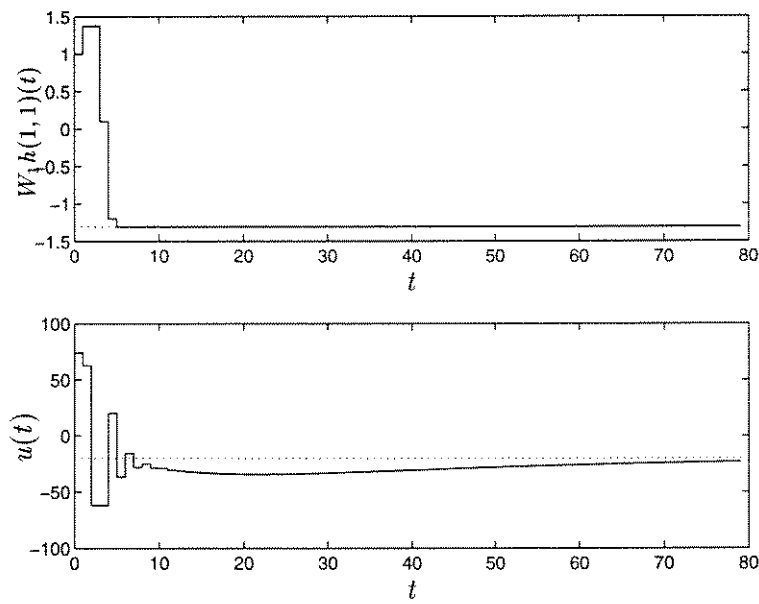


Figura 7.2: Caça X29: Resposta ao degrau de H via PA sem considerar overshoot

Note que a ordem do controlador aumenta de 6 para 7. É interessante também notar que o valor do overshoot foi consideravelmente reduzido com uma pequena degeneração da norma l_1 . Isto também mostra uma vantagem do uso de PA, pois neste caso o controlador que leva a este comportamento foi automaticamente encontrado. Caso o overshoot fosse considerado como uma restrição do problema (o que é mais usual), encontrar este controlador seria um processo empírico e interativo. A Figura 7.3 apresenta a resposta ao degrau de H , onde pode ser constatada a redução do overshoot.

Tabela 7.3: Caça X29: Resultados encontrados via PA considerando overshoot.

τ_d	$\underline{\eta}$	$\bar{\eta}$	overshoot
4	3.2539	1256.4	10.0
5	4.0925	5.9297	37.85
6	4.1424	4.7958	35.38
7	4.1427	4.7855	35.36
8	4.1402	4.3869	36.14
9	4.1407	4.3776	36.11
10	4.1018	4.3127	47.16
20	4.1014	4.2847	47.20
40	4.1006	4.2205	47.30
50	4.1002	4.1898	47.34
60	4.0998	4.1655	47.38
70	4.0996	4.1499	47.40
80	4.0995	4.1401	47.42

7.10.2 Problema da Haste Flexível

O segundo problema resolvido envolve uma haste flexível presa numa de suas extremidades a um motor DC, ficando a outra extremidade livre (Dahleh e Diaz-Bobillo, 1995). O objetivo do controlador é regular a posição da haste (Figura 7.4). Foi considerada como planta uma aproximação de quarta ordem do sistema:

$$P(s) = \frac{-6.475s^2 + 4.0302s + 175.77}{s(5s^3 + 3.5682s^2 + 139.5091s + 0.0929)}$$

Do mesmo modo que no problema do caça X29, em primeiro lugar o método DA foi aplicado ao problema

$$\min_{K_{estvl}} \left\| \begin{array}{c} W_1 S \\ W_2 K S \end{array} \right\|_1 \quad (7.17)$$

correspondendo à planta apresentada na Figura 7.5, sendo neste caso $W_1 = W_2 = 1$. Com $\tau_d = 30$, obtém-se $\underline{\eta} = \bar{\eta} = 2.0000$.

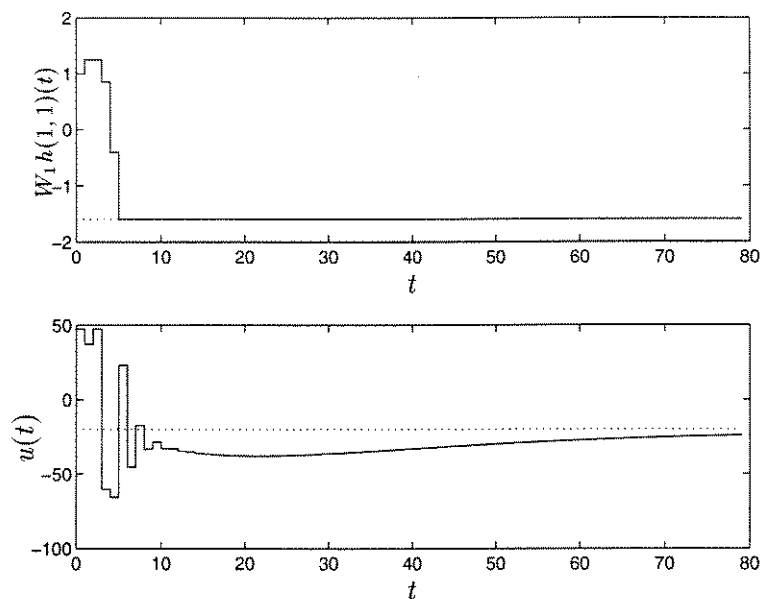


Figura 7.3: Caça X29: Resposta ao degrau de H via PA considerando overshoot

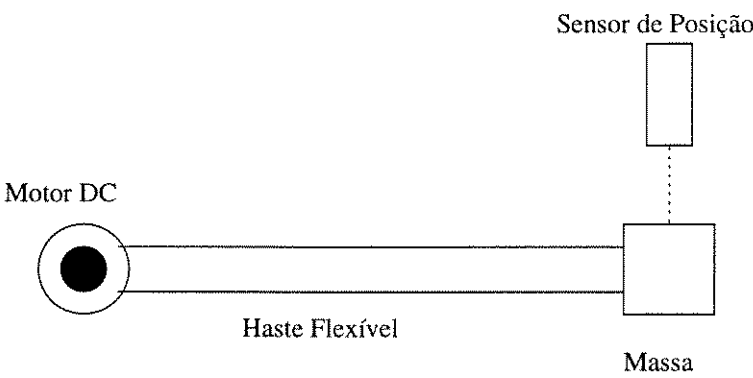


Figura 7.4: Haste flexível.

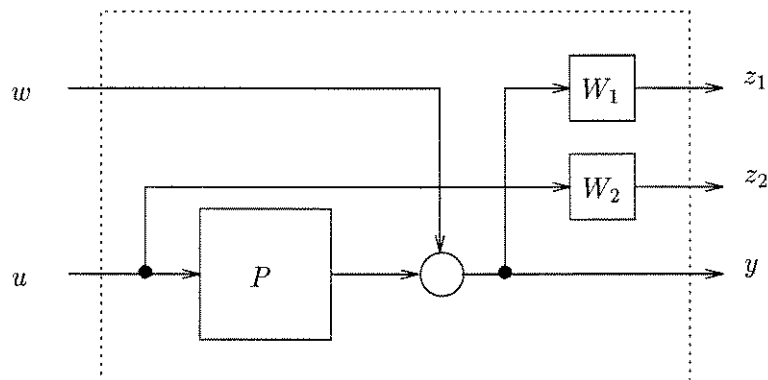


Figura 7.5: Problema da haste flexível dentro da arquitetura padrão.

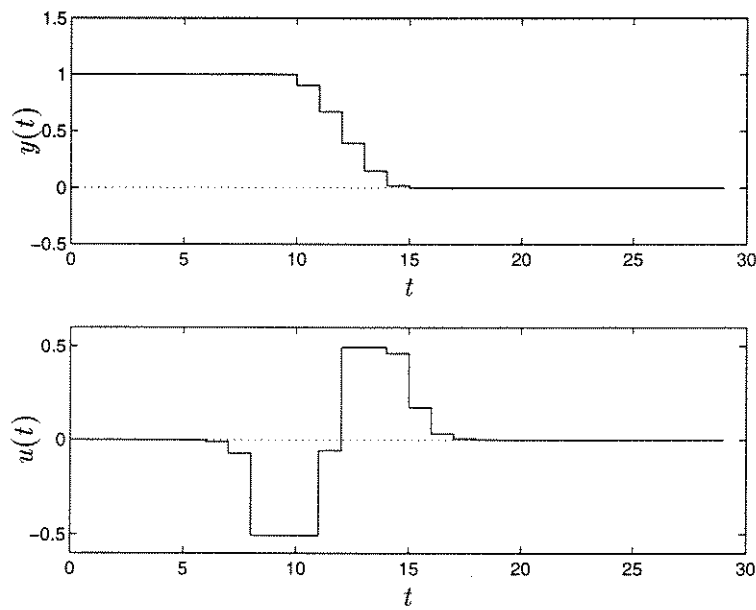


Figura 7.6: Haste flexível: Resposta ao degrau de H sem considerar over/undershoot.

A Figura 7.6 apresenta sequências ($u(t)$ e $y(t)$) da matriz de transferência em malha fechada encontrada. Note que nos primeiros instantes existe um atraso na ação do controle e, em consequência, a saída leva aproximadamente 10 instantes de tempo para começar a ser afetada. Neste caso, a solução adotada foi a mesma proposta por Dahleh e Dias-Bobillo (1995), perturbando o zero no círculo unitário para dentro da região de instabilidade, usando $\epsilon = 10^{-3}$. A solução encontrada Com $\tau_d = 30$, obtém-se $\underline{\eta} = \bar{\eta} = 2.006938$.

O controlador encontrado foi:

$$K(\lambda) = \frac{-21.295\lambda^9 - 88.666\lambda^8 - 95.38\lambda^7 - 119.873\lambda^6 - 6.666\lambda^5 + 98.18\lambda^4 + 98.082\lambda^3 + 117.242\lambda^2 + 15.816\lambda + 2.868}{\lambda^9 + 1.337\lambda^8 - 18.644\lambda^7 - 65.251\lambda^6 - 120.299\lambda^5 - 170.804\lambda^4 - 193.56\lambda^3 - 193.366\lambda^2 - 193.19\lambda - 236.792}$$

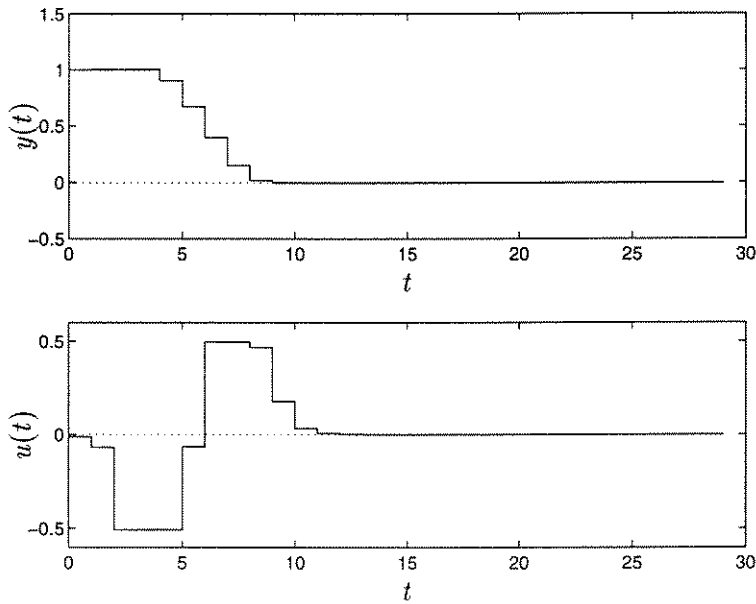


Figura 7.7: Haste flexível: Resposta ao degrau de H sem considerar over/undershoot (planta perturbada).

A seguir, o problema da haste flexível será tratado usando programação alvo. Neste caso, procura-se manter o valor da resposta ao degrau dentro de valores máximos e mínimos pré-definidos (overshoot/undershoot). A formulação do problema é a mesma de (7.16), acrescentando-se as restrições referentes ao undershoot.

O problema da haste flexível é originalmente um problema com múltiplas soluções. Quando perturbado, os custos marginais que antes eram nulos passam a ter valores da ordem de 10^{-3} . Para resolver esta questão, o tableaux simplex foi relaxado para considerar como nulos os custos marginais nesta ordem de grandeza. A Tabela 7.4 apresenta a solução do problema da haste flexível,

tendo com função de custo $C = d_{l_1}^+$ e $C = Lex\{d_{l_1}^+, d_{ov/ud}^+\}$, para $\tau_d = 30$. As metas usadas foram 1.0 (norma l_1), 0.2 (overshoot) e -0.2 (undershoot).

Tabela 7.4: Haste flexível: Resultados via PA com ordenação léxica dos desvios.

Função de Custo	τ_d	η	$\bar{\eta}$	overshoot	undershoot
$C = d_{l_1}^+$	30	2.006938	2.006938	0.4956	-0.4956
$C = Lex\{d_{l_1}^+, d_{ov/ud}^+\}$	40	2.007071	2.007071	0.4696	-0.4696

Note que a degeneração do valor da norma l_1 é devido à relaxação da solução, sendo da mesma ordem de grandeza. A Figura 7.8 mostra a resposta ao degrau para o caso em que se considera a minimização do overshoot/undershoot.

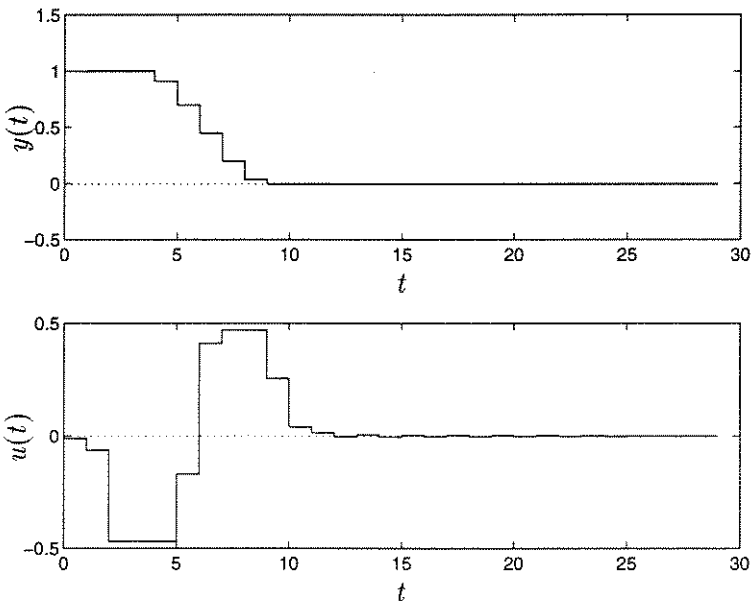


Figura 7.8: Haste flexível: Resposta ao degrau de H via PA com ordenação léxica dos desvios.

Em função da pequena melhora dos valores do overshoot/undershoot, estabeleceu-se uma estratégia de novamente compor uma função de custo a partir dos desvios dos dois objetivos, ou seja $C = d_{l_1}^+ + d_{ov/ud}^+$. Mantendo-se as mesmas metas, encontrou-se novos valores para $\tau_d = 40$. A Figura 7.9 mostra a resposta ao degrau para este novo caso, sendo que o controlador encontrado foi

Tabela 7.5: Haste flexível: Resultados via PA com desvios compostos em um único critério.

Função de Custo	η	$\bar{\eta}$	overshoot	undershoot
$C = d_{l_1}^+ + d_{ov/ud}^+$	2.008753	2.103612	0.2000	-0.2000

$$K(\lambda) = \frac{-21.295\lambda^{14} - 18.381\lambda^{13} - 23.703\lambda^{12} - 33.377\lambda^{11} - 33.097\lambda^{10} - 39.456\lambda^9 - 18.948\lambda^8 + 22.427\lambda^7 + 33.597\lambda^6 + 33.569\lambda^5 + 33.393\lambda^4 + 37.062\lambda^3 + 24.511\lambda^2 + 3.427\lambda + 0.538}{\lambda^{14} - 1.963\lambda^{13} - 12.681\lambda^{12} - 22.322\lambda^{11} - 33.242\lambda^{10} - 57.558\lambda^9 - 89.045\lambda^8 - 114.253\lambda^7 - 133.121\lambda^6 - 150.663\lambda^5 - 163.678\lambda^4 - 167.848\lambda^3 - 166.853\lambda^2 - 188.207\lambda - 112.8173}$$

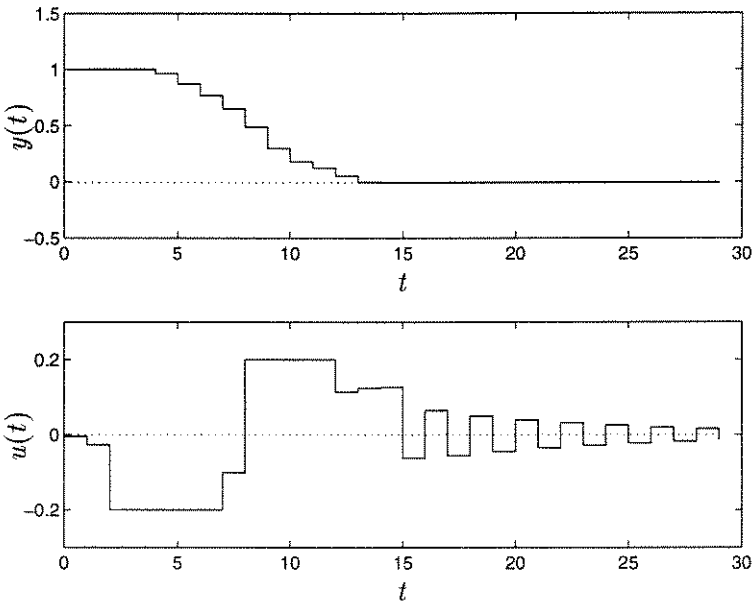


Figura 7.9: Haste flexível: Resposta ao degrau de H via PA com desvios compostos em um único critério.

Note que o overshoot/undershoot foi efetivamente reduzido, mas em compensação o tempo de regulação aumentou e a ação do controle é claramente oscilatória (Figura 7.9). Porém, estes aspectos não foram considerados no problema, sendo que do ponto de vista do projeto, o controlador encontrado é satisfatório.

Uma observação interessante sobre o método PA aplicado ao problema de projeto de controladores é que por construção o problema é sempre factível. Além disso, caso a solução não seja satisfatória, os desvios fornecem informações sobre como alterar as metas do projeto.

7.11 Aspectos Computacionais

A seguir serão feitas algumas observações com respeito à implementação dos métodos discutidos neste Capítulo.

Cálculo das cadeias nulas: Geralmente os métodos que calculam os zeros de transferência de uma matriz também fornecem a multiplicidade algébrica de cada um (Dahleh e Diaz-Bobillo, 1995). Com isso é necessário calcular o espaço nulo apenas para a matriz de Toeplitz completa. A melhor forma é usar representação em espaço de estados.

Cálculo de A_{zero} e b_{zero} : O método DA acrescenta zeros na origem, de multiplicidades geralmente elevadas, pelo menos nos casos estudados. Isto normalmente traria grandes transtornos numéricos não fosse a particularidade de que, como $\lambda_0 = 0$, restam em cada linha de A_{zero} apenas os termos independentes do polinômio, multiplicados por um termo fatorial. Porém, este termo fatorial é exatamente o mesmo que multiplica o elemento de b_{zero} correspondente, podendo ser eliminado. Com isso os i 's primeiros elementos da linha i de A_{zero} nada mais são do que os i 's primeiros elementos da cadeia nula para $\lambda_0 = 0$; os restantes são nulos.

Condicionamento do PL: Nesta parte deve-se tomar muito cuidado, pois apesar de, por construção, as restrições $Ax = b$ que formam o PL serem linearmente independentes, numericamente não o são. A melhor forma que se encontrou durante os trabalhos para resolver este problema foi calcular o rank de A e selecionar dentre as linhas de A e b aquelas que formam um conjunto de restrições com melhor número-condição, desconsiderando-se as restantes. O método é lento, porém encontra, dentre as restrições existentes, as melhores condicionadas. Mesmo assim, em alguns casos, o melhor condicionamento ainda é deficiente.

SIMPLEX a ser utilizado: O programa SIMPLEX deve estar preparado para lidar com problemas com matrizes esparsas, de grandes dimensões e rank deficiente. Dentre os problemas observados, um particularmente difícil de tratar foi a interrupção da execução do SIMPLEX por ter determinado um conjunto de variáveis básicas degeneradas, o que por definição não ocorreria assumindo matrizes de rank completo. Vale ressaltar que este problema ocorreu mesmo após a eliminação das linhas que do ponto de vista numérico são linearmente dependentes.

Cálculo das matrizes de transferência: Apesar de toda a teoria ter sido apresentada na representação via funções/matrizes de transferência, esta é definitivamente a forma menos adequada de representar numericamente os sistemas para uma implementação computacional. Para quase todos os cálculos, a melhor forma foi a representação via espaço de estados e para os cálculos onde as eliminações entre pólos e zeros foram importantes, a representação do tipo zero/pólo se mostrou muito eficaz.

7.12 Conclusão

Este Capítulo abordou o uso de técnicas de otimização para projeto de controladores através de critérios baseados na norma l_1 . Por definição, o projeto na norma l_1 consiste na minimização do máximo ganho de pico-a-pico da matriz de transferência entre as entradas exógenas v e as saídas reguladas z .

O Capítulo abordou a solução de problemas em tempo discreto usando a transformada- λ , em função desta fornecer facilidades metodológicas no desenvolvimento das condições de interpolação, que são de grande valia na determinação do problema de programação linear associado ao problema de projeto, além de facilitar a caracterização da solução.

O problema multicritério foi formulado agregando-se ao critério do tipo norma l_1 outras especificações de projeto. O método proposto para resolver o problema de norma l_1 multicritério foi uma combinação do método DA com técnicas de programação alvo. Esta combinação se mostrou eficiente nos casos estudados. Para ilustrar seu funcionamento, o método foi aplicado ao problema de projeto do controle de *pitch* do caça X29 e ao problema de controle de posição de uma haste flexível.

Conclusão Geral

Neste trabalho procurou-se abordar diferentes problemas de projeto de controladores para sistemas dinâmicos de uma forma integrada.

O objetivo da primeira parte do trabalho foi desenvolver uma abordagem geral para o problema de controle multicritério no domínio do tempo. Basicamente, a abordagem proposta trata as peculiaridades de cada problema no espaço dos critérios, isolando a dinâmica do sistema em um problema min-max. A estrutura multinível obtida tem demonstrado grande flexibilidade e facilidade de implementação.

Apesar do problema de controle no espaço dos critérios ser em geral não-convexo, foi possível derivar um problema convexo equivalente. Aliando às ferramentas revistas no Capítulo 1, manipulações básicas tais como projeção e relaxação, decompôs-se o problema multicritério em uma estrutura multinível, na qual tanto os aspectos de decisão quanto os de análise puderam ser trabalhados apropriadamente, transformando o problema no espaço dos critérios em um problema estático e convexo, com restrições lineares. Esta decomposição em níveis permite que a tomada de decisão fique isolada da dinâmica do sistema em questão.

No Capítulo 2, o procedimento chamado de *Algoritmo Básico*, foi estabelecido em termos da classe de problemas de controle multicritério determinísticos. Neste mesmo Capítulo foi possível verificar como se processa a troca de informações entre o nível de análise e o nível de decisão. Mostrou-se também que a metodologia aproveita as características intrínsecas de cada problema, como no caso linear-quadrático, onde as equações de Riccati e a existência do gradiente constituem parte integrante na busca da solução do problema. Esta característica do Algoritmo Básico ficou evidenciada também na solução do problema linear-quadrático gaussiano, apresentada no Capítulo 3. Naquele Capítulo, o Princípio da Separação exerceu um importante papel na solução do problema, mostrando que a introdução de múltiplos critérios modifica apenas o problema de controle, não alterando o problema de filtragem. Finalmente no Capítulo 4 a abordagem proposta foi usada na solução de problemas de controle de sistemas hierárquicos. Além de tratar de forma adequada uma classe importante de problemas de otimização de grande porte não separáveis, o emprego de ferramentas de otimização convexa reduziu a complexidade do problema considerado.

A segunda parte do trabalho tratou do problema de projeto multicritério de controladores no domínio da frequência. O problema foi tratado através do Algoritmo Básico, anteriormente aplicado a problemas no domínio do tempo.

Para tanto, o Capítulo 5 apresentou uma revisão dos conceitos necessários para a descrição do modelo e dos critérios de projeto, além de rever conceitos relativos a fatoração coprima, que são relevantes para o desenvolvimento da parametrização de Youla. O tratamento de problemas via

parametrização de Youla, apesar de trazer vantagens metodológicas consideráveis, é um problema de dimensão infinita e, conseqüentemente, fez-se necessário usar técnicas de aproximação progressiva da matriz de transferência.

Assim, no Capítulo 6, foi apresentada a técnica de aproximação da matriz de transferência utilizada, a aproximação de Ritz. Apesar de ser de fácil implementação, observou-se que esta técnica não permite limitar previamente a ordem do controlador. Uma vez formulado o problema parametrizado e com a aproximação de Ritz, o nível de análise do Algoritmo Básico foi implementado através do método de planos de cortes. Com isso evitaram-se problemas quanto à não-diferenciabilidade dos funcionais envolvidos. Finalmente, foi feita uma comparação com o Método das Desigualdades, que evidenciou as vantagens do procedimento proposto.

O Capítulo 7 teve como objetivo apresentar uma metodologia multicritério para o tratamento do problema de projeto baseado na norma l_1 . Como este Capítulo trata de sistemas discretos foi necessário rever alguns conceitos relativos, como o de Transformada- λ . Foi visto que através de manipulações adequadas do problema de projeto, este se transforma em um problema linear de dimensão finita equivalente. O algoritmo desenvolvido para resolver o problema de projeto l_1 multicritério combina o método DA (*Delay Augmentation*), apresentado em Dahleh e Diaz-Bobillo (1995), com técnicas de Programação Alvo. Para ilustrar o funcionamento do procedimento foram resolvidos dois exemplo de projeto, que evidenciaram as facilidades decorrentes do tratamento do problema via programação alvo.

Finalmente, vários pontos deixados em aberto podem ser explorados através de trabalhos futuros. Uma análise mais detalhada para problemas que não exibem estrutura linear-quadrática ampliaria a aplicabilidade da abordagem. Muitos problemas de controle ótimo não exibem esta estrutura, mas podem ser formulados como problemas convexos. Sistemas dinâmicos variantes no tempo podem ser considerados sem maiores complicações para o Algoritmo Básico.

No domínio da frequência, o aspecto principal a ser tratado é a redução de ordem dos controladores obtidos através de otimização convexa. Embora este aspecto ultrapasse em muito o contexto deste trabalho, a viabilidade dos controladores gerados pelo Algoritmo Básico depende em grande parte da ordem destes controladores. Neste sentido, estratégias baseadas em *satisfação*, como as sugeridas pelo Método das Desigualdades parecem mais promissoras.

Por último, os projetos baseados em critérios do tipo l_1 oferecem inúmeras possibilidades no que diz respeito ao emprego de técnicas eficientes de programação linear. Este tema é particularmente promissor em termos de trabalhos futuros.

Bibliografia

Athans, M. & P. Falb, (1966). **Optimal Control**, MacGraw-Hill.

Boyd, S. P. & C. H. Barrat (1991), **Linear Controller Design-Limits of Performance**, Prentice-Hall.

Carvalho, J. R. H. (1993), **Controle Ótimo Multicritério: Abordagem por Programação Convexa**, Tese de Mestrado FEE/UNICAMP.

Carvalho, J. R. H. & P. A. V. Ferreira (1992), Uma Abordagem para o Problema de Controle Multicritério de Sistemas Dinâmicos, IX. Congresso Brasileiro de Automática, Vitória-ES, Brasil, pp. 568-573.

Carvalho, J. R. H. & P. A. V. Ferreira (1993), Multiple-Criteria Control: A Convex Programming Approach, IFAC World Congress, Sidney, Australia, Vol. 5, pp. 29-32.

Carvalho, J.R.H. & P.A.V. Ferreira (1995a), Multiple-Criteria Control: A Convex Programming Approach, (extended version), Automatica, Vol. 31, No. 7, pp. 1025-1029.

Carvalho, J. R. H. & P. A. V. Ferreira (1995b), A Convex Approach for the Multiple-Criteria LQG Problem - Second Latin American Seminar on Advanced Control, Santiago, Chile, pp. 11-16.

Carvalho, J. R. H., S. L. C. de Oliveira & P. A. V. Ferreira (1996), A Convex Approach for Hierarchical Multicriteria Linear Quadratic Problems. XI Congresso Brasileiro de Automática, São Paulo-SP, Brasil, pp. 371-376.

Cruz Jr, O. A. (1994), **Síntese de Sistemas de Controle via Otimização Convexa**, Tese de Mestrado FEE/UNICAMP.

Dahleh, Munther A. & Ignacio J. Diaz-Bobillo (1995), **Control of Uncertain Systems: A Linear Programming Approach**, New Jersey, USA: Prentice Hall, Inc.

Dantzig, G. & P. Wolfe (1960). Decomposition principle for linear programs. *Ops. Res.*, Vol. 8,

101-111.

Diaz-Bobillo, Ignacio J. & Munther A. Dahleh (1993), Minimization of the Maximum Peak to Peak Gain: The General Multiblock Problem. *IEEE Trans. on Automatic Control*, Vol. 38, No. 10, pp. 1459-1482.

Dorato, P., Abdallah C. & V. Cerone (1995), *Linear Quadratic Control: An Introduction*, New Jersey, USA: Prentice Hall, Inc.

Francis, B. A. (1987), *A Course in H_∞ Control Theory*, Springer-Verlag.

Geoffrion, A.M. (1968), Proper Efficiency and the Theory of Vector Maximization. *Journal of Mathematical Analysis and Applications*, Vol. 22, 618-630.

Haimes, Y. Y., W. A. Hall & H. T. Freedman (1975), *Multiobjective Optimization in Water Resources Systems*, Elsevier.

Ignizio, J. P. (1976), *Goal Programming and Extensions*, D. C. Heath, Lexington, Massachusetts.

Khargonekar, P. P. & M. A. Rotea (1991), Multiple Objective Optimal Control of Linear Systems: The Quadratic Norm Case, *IEEE Trans. on Automatic Control*, Vol. 36, No. 1, pp. 14-24.

Kirk, D. E., (1970), *Optimal Control Theory - An Introduction*, Prentice Hall.

Koussoulas, N. T. & C. T. Leondes (1986), The Multiobjective Linear Quadratic Gaussian Problem, *International Journal of Control*, Vol. 43, pp. 337-349.

Kresselmeier, G. & R. Steinhauser (1983), Application of Vector Performance Optimization to a Robust Control Loop Design of a Fighter Aircraft, *International Journal of Control*, Vol. 37, No. 2, pp. 251-284.

Lasdon, L.S., (1970). **Optimization Theory for Large Systems**, London: MacMillan Co.

Li, D. & Y. Y. Haimes (1988), Hierarchical Multiobjective Analysis for Large-scale Systems: Review and Current Status. *Automatica*, Vol. 24, pp. 53-69.

Li, D. & Y. Y. Haimes (1990), Multilevel Methodology for a Class of Nonseparable Optimization Problems. *International Journal of Systems Science*, Vol. 21, pp. 2351-2360.

Li, D. (1990a), On the Minimax Solution of Multiple Linear-Quadratic Problems, *IEEE Trans. on Automatic Control*, Vol. 35, No. 10, pp. 1153-1156.

- Li, D. (1990b), A New Solution Approach to the Salukvadze Problem, *Proceedings of the American Control Conference*, San Diego, CA, pp. 409-414.
- Li, D. (1993), Hierarchical Control for Large Scale Systems with General Multiple Linear Quadratic Structure. *Automatica*, Vol. 29, No. 6.
- Looze, D. P., H. V. Poor, K. S. Vastola & J. C. Darragh (1983), Minimax Control of Linear Stochastic Systems with Noise Uncertainty, *IEEE Trans. on Automatic Control*, Vol. 28, No. 9.
- Luenberger, D. (1969), *Optimization by Vector Space Methods*, Wiley
- Maciejowski (1989), *Multivariable Feedback Design*. Addison-Wesley.
- Medanić, J. & M. Andjelić (1971), On a Class of Differential Games without Saddle-Point Solutions, *Journal of Optimization Theory and Applications*, Vol. 8, No. 6, pp. 413-430.
- Medanić, J. & M. Andjelić (1972), Minimax Solution of the Multiple-Target Problem, *IEEE Trans. on Automatic Control*, Vol. 17, No. 5, pp. 597-604.
- Mesarovic, M.D., D. Macko & Y. Takahara (1970), *Theory of Hierarchical, Multilevel Systems*. Academic Press, New York.
- Oliveira, S. L. C., J. R. H. Carvalho & P. A. V. Ferreira (1995), Uma Abordagem Convexa para uma Classe de Problemas de Otimização Não-separáveis. *Congresso Nacional de Matemática Aplicada e Computacional (CNMAC95)*, Curitiba-Pr, Brasil, pp. 219-223.
- Oliveira, S. L. C., J. R. H. Carvalho & P. A. V. Ferreira (1997), A Convex Approach for Multi-criteria Decision Making in Hierarchical Systems. *Journal of Multi-criteria Decision Analysis*, submetido.
- Salukvadze, M. (1982), An Approach to the Solution of the Vector Optimization Problem of Dynamic Systems, *Journal of Optimization Theory and Applications*, Vol. 38, pp. 409-422.
- Singh, M. G. & M. Hassan (1976), Comparison of two hierarchical optimization methods. *International Journal of Systems Science*, Vol. 7, pp. 603-611.
- Singh, M. G. & A. Titli (1978), *Systems: Decomposition, Optimization and Control*. Pergamon Press, Oxford.
- Toivonen, H. T. (1984), A Multiobjective Linear Quadratic Gaussian Control Problem, *IEEE Trans. on Automatic Control*, Vol. 29, No. 3, pp. 279-280.

Toivonen, H. T. & P. M. Mäkilä (1989), Computer-Aided Design Procedure for Multiobjective LQG Control Problems, *International Journal of Control*, Vol. 49, No. 2, pp. 655-666.

Vidyasagar, M. (1991), *Control System Synthesis: A Factorization Approach*, Cambridge, MA: M.I.T. Press.

Yu, P. L. (1985), *Multiple Criteria Decision Making*, Plenum Press.

Zadeh, L. A. (1963), Optimality and Non-Scalar Valued Performance Criteria, *IEEE Trans. on Automatic Control*, Vol. 8, No. 1, pp. 59-60.

Zakian, V. & U. Al-Naib (1973), Design of Dynamic and Control Systems by the Method of Inequalities, *Proceedings of the IEE*, Vol. 120, pp. 1421-1427.