

Universidade Federal do Maranhão

Centro Tecnológico

Departamento de Engenharia de Eletricidade

**PROJETO DE CONTROLADOR ROBUSTO
VIA OTIMIZAÇÃO CONVEXA**

Valeska Martins de Souza

São Luís, 2002

PROJETO DE CONTROLADOR ROBUSTO VIA OTIMIZAÇÃO CONVEXA

Dissertação de mestrado submetida à Coordenação do Curso de Pós-Graduação em Engenharia de Eletricidade da UFMA como parte dos requisitos para obtenção do título de Mestre em Automação e Controle.

Por
Valeska Martins de Souza
Dezembro, 2002

PROJETO DE CONTROLADOR ROBUSTO VIA OTIMIZAÇÃO CONVEXA

VALESKA MARTINS DE SOUZA

Dissertação aprovada em 13 de dezembro de 2002.

Prof. Dr. José Tarcisio Costa Filho
(Orientador)

Prof. Dr. Geraldo Leite Torres
(Membro da Banca Examinadora)

Prof. Dr. Marcos Antônio Ferreira de Araújo
(Membro da Banca Examinadora)

DEDICATÓRIA

Aos meus pais, Raimundo Nonato Martins de Souza e Maria Zuila Lopes Martins.

A Paolo Sanna.

A Leonardo Matias Campos de Souza.

Às minhas irmãs Valéria Martins e Verônica Martins.

AGRADECIMENTOS

A DEUS por nossa existência.

Aos meus pais pelo amor, educação, caráter, amizade e incentivo.

Ao Prof. Dr. José Tarcisio Costa Filho, pela amizade, paciência, compreensão e ensinamentos transmitidos durante o período de orientação deste trabalho.

À Profa. Dra. Maria da Guia, pelo apoio e incentivo.

A todos os funcionários e professores do DEE, pela colaboração durante o curso.

À Universidade Federal do Maranhão.

Aos amigos do Departamento de Matemática da UFMA.

Ao professor Eliúde Trovão Moraes, pela calorosa recepção no laboratório de eletrônica do CEFET-MA e pela contribuição dada neste trabalho.

Ao colega Manfred Fritz Aronés, pela ajuda no MATLAB.

Ao professor Francisco Chagas Sousa, pela ajuda no material didático.

A Ismael Godinho e Wagner Élvio, pelo apoio técnico.

À Profa. Veraluce Lima dos Santos pela revisão gramatical.

A todos que colaboraram para a realização deste trabalho.

“ ... E se não houver frutos valeu a beleza das flores;
se não houver flores valeu a sombra das folhas;
e se não houver folhas valeu a intenção da semente.”

Henfil

RESUMO

Nesta dissertação, é proposta uma nova metodologia de Otimização Convexa baseada em Desigualdades Matriciais Lineares (LMIs) como instrumento básico para a síntese de controladores robustos de sistemas dinâmicos discretos e lineares que atendam às especificações de atenuação de perturbações de pior caso. Nessa nova metodologia, obtém-se, inicialmente, a solução do problema de controle em malha aberta via Otimização Multinível, através de um esquema de decomposição paramétrica e de transformações duais de um conjunto de equações dinâmicas complexas em um conjunto de equações estáticas mais simples de resolver. A partir dessas equações estáticas, o problema de otimização multinível é modificado, redefinindo-o sobre uma região de confiança e resolvido via Programação Semidefinida (SDP), com a finalidade de obter maior estabilidade numérica e gerar uma família de controladores, com melhores propriedades de estabilidade e desempenho robusto. Além da simplicidade e consistência algébrica, uma das principais vantagens dessa metodologia é a disponibilidade de algoritmos computacionais eficientes, derivados dos métodos de pontos interiores para programação convexa. Desta forma, é proposto um novo algoritmo para a determinação de uma matriz de ganho de realimentação de estado, utilizado para o projeto de controle LQR e H_∞ discreto no domínio do tempo sem a necessidade do emprego das equações algébricas de Riccati (ARE). Todos os procedimentos de síntese de controladores são ilustrados através de alguns exemplos e são realizadas, também, análises comparativas.

ABSTRACT

In this work, a new convex optimization methodology based upon linear matrix inequalities (LMIs) that meet the attenuation requirements for the worst case, as basic tool for the synthesis of linear and dynamic discrete system's robust controllers, is proposed. Within this context, initially, the solution for the open-loop control problem is obtained by a multi-level optimization scheme using parametric decomposition and dual transformations of a set of complex dynamic equations into a set of easier-to-solve static equations. Thus, starting from those equations, the multi-level optimization problem is optimized redefining it over a thrust region, being solved by semidefinite programming (SDP), in order to obtain greater numerical stability and yield a family of controllers with enhanced stability properties and robust performance. Besides its simplicity and algebraic consistency, the main advantage of that LMIs based method is the availability of efficient computable algorithms derived from interior point methods for convex programming. Therefore, a new algorithm for determining a gain matrix of state feedback for the discrete LQR and H_∞ control Project in the time domain, is proposed, without handling the Riccati equations (ARE). All controller's synthesis processes are illustrated through samples and comparative analysis is also performed.

Sumário

1	Introdução e Fundamentos Teóricos	1
1.1	Introdução	1
1.1.1	Motivação e Objetivos do Trabalho	2
1.1.2	Organização da Dissertação	3
1.2	Fundamentos Teóricos	5
1.2.1	Otimização Convexa	5
1.2.2	Programação Quadrática	10
1.2.3	Regiões de Confiança	12
1.2.4	Métodos de Pontos Interiores	19
1.2.5	Sistemas Dinâmicos Lineares Discretos Invariantes no Tempo	22
1.2.6	Regulador Linear Quadrático	23
1.3	Comentário Final	26
2	Programação Semidefinida	27
2.1	Conceito de SDP	27
2.2	Cone das Matrizes Semidefinidas Positivas	28
2.3	Problemas SDPs	29
2.4	Dualidade em SDP	30
2.5	Solução dos Problemas SDPs	32
2.5.1	Métodos da Barreira Logarítmica	34

2.5.2	Métodos Seguidor de Trajetória Prima-Dual	35
2.6	Um breve histórico das LMIs na Teoria de Controle	36
2.7	O que são LMIs?	40
2.7.1	Complemento de Schur	43
2.7.2	Problemas da Programação Linear Expressos em LMIs	47
2.7.3	Problemas da Programação Quadrática Expressos em LMIs	47
2.8	Comentário Final	49
3	Controle Ótimo Via Programação Semidefinida	50
3.1	Formulação do Problema de Controle Ótimo de Sistemas Dinâmicos Discretos	50
3.2	Metodologia de Otimização Multinível	51
3.3	Estratégia de Coordenação	53
3.4	Reformulação do Algoritmo do Nível Superior	54
3.5	Modificação da Estratégia de Coordenação	57
3.6	Uma Proposta para Controle em Malha Fechada	58
3.7	Análise e Resultados	63
3.8	Comentário Final	69
4	Controle Robusto Via Programação Semidefinida	70
4.1	Introdução	70
4.2	Formulação do Problema	71
4.3	Projeto de Controladores H_∞ via ARE	73
4.4	Projeto de Controladores H_∞ via Otimização Convexa	76
4.5	Modificação da Estratégia de Coordenação	79
4.6	Proposta de Controle H_∞	83
4.7	Aplicações e Resultados	85
4.8	Comentário Final	91

5	Conclusão	92
5.1	Trabalhos Futuros	93
A	Normas e Valores Singulares	102
A.1	Normas de Vetores e Matrizes	102
A.2	Valores Singulares	103
A.3	Norma de Matrizes de Transferência	104
B	Topologia do R^n	106
B.1	Conjuntos Abertos e Fechados	106
B.2	Seqüências	108
B.3	Conjuntos Compactos	109
C	Álgebra Linear	110
C.1	Menores Principais	110
C.2	Matrizes Definidas	110
C.3	Desigualdades Matriciais	111
C.4	Inversas de Matrizes de Posto Completo	112
C.4.1	Inversa de Uma Matriz Quadrada	112
C.4.2	Inversa de Uma Matriz Retangular	112
C.5	Inversa Generalizada de Moore-Pensore	113
C.6	Número de Condição	113

Lista de Tabelas

1.1	Os três casos diferentes para o subproblema da região de confiança. . .	18
1.2	Algoritmo de Otimização Multinível	25
1.3	Regulador Linear Quadrático	26
2.1	Algoritmo de Pontos Interiores Primal-Dual para SDP	36
3.1	Algoritmo para o Regulador Linear Quadrático	63
3.2	Valores de p e t	65
3.3	Valores de p e t	66
3.4	Comportamento dos ganhos e autovalores do controlador	69
4.1	Algoritmo para Solução do Controle H_∞	75
4.2	Algoritmo para o controle H_∞ com realimentação de estado	85
4.3	Valores de p e t	87
4.4	Valores de p e t	88
4.5	Resultados numéricos do projeto via nova proposta com $\gamma = 0.51$. . .	89

Lista de Figuras

1.1	Estrutura do Controle Ótimo Multinível	25
2.1	LMI no plano: $y - x > 0$	46
2.2	LMI no plano: $y > x^2$	46
2.3	LMI no plano: $x^2 + y^2 < 1$	46
3.1	Trajétória da variável de estado $x_2(k)$	67
3.2	Trajétória de controle $u(k)$	67
3.3	Comparação das Trajetórias de Riccati e via LMI	68
4.1	Trajétória da variável de estado $x_1(k)$	90
4.2	Trajétória de controle	91

LISTA DE ABREVIATURAS E SIGLAS

ARE	Equação Algébrica de Riccati
IPM	Métodos de Ponto Interiores
KKT	Karush-Kunh-Tucker
LMI	Desigualdades Matriciais Lineares
LP	Programação Linear
LQR	Regulador Quadrático Linear
LTI	Linear Invariante no Tempo
QCQP	Programação Quadrática com Restrições Quadráticas
QP	Programação Quadrática
SOCP	Programação do Cone de Segunda Ordem
SDP	Programação Semidefinida
TRS	Subproblemas da Região de Confiança

Capítulo 1

Introdução e Fundamentos Teóricos

1.1 Introdução

Durante as últimas duas décadas, pesquisadores na área de controle propuseram diferentes técnicas para projetar o controle de sistemas lineares invariantes no tempo que são robustos e ótimos. Tais projetos de sistemas de controle necessitam de um modelo matemático que é utilizado para verificar o desempenho do controlador projetado e deve incluir todas as características importantes do processo. Considerando que nenhuma planta real é completamente linear e invariante no tempo, a propriedade de robustez do controlador torna possível que um sistema apresente um nível razoável de desempenho e estabilidade, mesmo quando operando na presença de incertezas. Apesar do sucesso da teoria de controle robusto lidando com uma grande classe de problemas de controle, pesquisadores têm procurado idéias novas e revolucionárias para substituir os métodos já existentes. Entre essas idéias revolucionárias, destacam-se uma classe poderosa de técnicas baseada em Desigualdades Matriciais Lineares (em inglês, Linear Matrix Inequalities (LMI)) e métodos de Otimização Convexa que podem ser resolvidos eficazmente pela Programação Semidefinida (em inglês, Semidefinite Programming (SDP)).

Nesse sentido, vários problemas da teoria de controle podem ser reformulados como problemas de Programação Semidefinida, ou seja, a minimização de uma função objetivo linear convexa restrita a um conjunto de LMIs [6, 42]. Da teoria da dualidade em otimização convexa, é possível obter condições para a viabilidade de LMIs e reformulação de problemas de otimização dual. Estas, por sua vez, podem ser interpretadas em termos da teoria de controle. Há considerável pesquisa sobre algoritmos e softwares com aplicações de SDP para a solução numérica de problemas de análise e síntese de controle [41, 51, 61].

Paralelamente, a Otimização H_∞ [9, 10, 19, 28, 68] tem se tornado muito popular na Teoria de Controle, devido ao fato dessa metodologia incorporar requerimentos de robustez, atenuação de perturbações e propriedades de desempenho dentro de um problema de otimização para projeto de controladores robustos. No entanto, o estudo de robustez, com respeito às incertezas existentes em sistemas de controle, colocou em evidência a necessidade de se buscar métodos mais abrangentes de solução, uma vez que estes problemas mais gerais, envolvendo robustez, não podiam ser resolvidos com o auxílio da equação de Riccati. O uso de métodos mais abrangentes, isto é, capazes de obter soluções globais, tinha como dificuldade principal o fato de que os problemas que se desejavam resolver não eram convexos. Esta dificuldade tem sido superada, por exemplo, com a aplicação de programação SDP, possibilitando a determinação de transformação de variáveis que permite converter os mais diversos problemas de controle em problemas de otimização convexa [40, 45, 52].

1.1.1 Motivação e Objetivos do Trabalho

A principal motivação para esta dissertação resulta da necessidade de desenvolver melhores técnicas numéricas eficientes e efetivas de projeto de controladores

via Otimização Convexa para atender aos requerimentos de estabilidade e desempenho robusto de sistemas dinâmicos discretos e lineares. Esse problema tem recebido bastante atenção de pesquisadores da área de controle. Há várias técnicas de projetos com aplicações limitadas. O objetivo deste trabalho é propor uma nova metodologia de otimização convexa baseada em LMIs para o projeto de controladores robustos com especificações de pior caso. Nessa nova metodologia, obtém-se, inicialmente, a solução do problema de controle em malha aberta via otimização multinível, através de um esquema de decomposição paramétrica e de transformações duais, em um conjunto de equações estáticas mais simples de resolver. A partir dessas equações, o problema de otimização é reformulado como um problema quadrático com restrições quadráticas (região de confiança) e resolvido como um problema de Programação Semidefinida (SDP) que garanta melhores propriedades de estabilidade e desempenho robusto. Dessa forma, é proposto um novo algoritmo para a determinação de uma matriz de ganho de realimentação de estado utilizado para o projeto de controle LQR e H_∞ discreto no domínio do tempo sem a necessidade do emprego das equações algébricas de Riccati (ARE).

1.1.2 Organização da Dissertação

Esta dissertação está organizada em 5 capítulos, a saber:

Capítulo 1: Introdução e Fundamentos Teóricos

Neste capítulo, é feita uma revisão de alguns fundamentos teóricos relacionados com otimização convexa, dualidade, regiões de confiança e métodos de pontos interiores; é comentada a solução do problema do regulador linear quadrático para sistemas via solução de Riccati. Procura, de forma concisa, apresentar ferramentas matemáticas que servirão de base para o desenvolvimento do trabalho.

Capítulo 2: Programação Semidefinida

Este capítulo, apresenta a definição e principais propriedades da Programação

Semidefinida, formulando-se o problema primal-dual bem como sua solução via métodos de pontos interiores. Apresenta, ainda, a definição e propriedades de LMIs que é a base necessária ao entendimento deste trabalho.

Capítulo 3: Controle Ótimo Via Programação Semidefinida

Este capítulo contém a proposta de uma nova metodologia para o projeto de controladores via otimização multinível, teoria da dualidade e Programação Semidefinida. A idéia básica envolve técnicas de decomposição e hierarquização do sistema dinâmico e a extensão do conceito da teoria da dualidade para estabelecer relações simétricas entre estruturas algébricas e topológicas, resultantes de transformações duais de um conjunto de equações dinâmicas complexas, em um conjunto de equações estáticas mais simples. A partir dessas equações estáticas, e o uso da Otimização Convexa, transformando um problema convexo quadrático com restrições quadráticas (QCQP) em um problema (SDP), propomos dois procedimentos numéricos para regularizar o problema de otimização multinível, com a finalidade de obter uma maior estabilidade numérica e gerar um controle em malha fechada com características melhores de estabilidade e desempenho. Alguns resultados e aplicações dessa nova metodologia de otimização e controle são apresentados e analisados.

Capítulo 4: Controle Robusto Via Programação Semidefinida

Neste capítulo, utilizamos a metodologia proposta no Capítulo 3, com a finalidade de obter um novo algoritmo para a síntese de controladores H_∞ . Para ilustrar a análise e síntese de controlador H_∞ proposto, resultados e aplicações são apresentados e analisados.

Capítulo 5: Conclusões

Este capítulo resume as contribuições desta dissertação e indica direções para pesquisas futuras.

1.2 Fundamentos Teóricos

Nesta seção, desenvolvemos uma breve revisão dos mecanismos básicos relacionados à Otimização Convexa, Regiões de Confiança e Métodos de Pontos Interiores com suas definições e propriedades.

1.2.1 Otimização Convexa

A Otimização representa um papel importante em muitas aplicações da Engenharia. Visa evidenciar condições algébricas que permitem caracterizar os mínimos ou máximos de uma função de várias variáveis de decisão, cujos valores estão dentro de uma determinada região do espaço multi-dimensional. Assim, as ferramentas matemáticas possíveis de ser empregadas dependem das hipóteses adotadas, a respeito da função objetivo e sua relação com as variáveis de decisão, e do tipo de conjunto de restrições.

Nesse contexto, nosso enfoque tem em vista os aspectos teóricos da otimização que efetivamente têm contribuído no desenvolvimento de algoritmos práticos para a Teoria de Controle. Assim a Otimização Convexa trata o problema de minimizar uma função convexa restrita a um conjunto convexo. Duas razões são importantes: primeiro, as hipóteses de convexidade em um problema de minimização implicam que as condições necessárias de otimalidade são também condições de suficiência. Segundo, para problemas convexos, o dual é sempre equivalente ao primal. Durante as últimas duas décadas, Otimização Convexa [41, 45], e particularmente, a Programação Semidefinida (SDP) [5] tem se tornado uma valiosa ferramenta numérica para projeto e análise de sistemas de controle. Vários problemas da Teoria de Controle podem ser reformulados como problemas SDP, ou seja, a minimização de uma função objetivo linear convexa restrita a um conjunto de desigualdades matriciais lineares (LMIs) [6, 42].

Matematicamente, podemos definir um problema de Otimização como: minimizar uma função $f(x)$, chamada função objetivo, onde x é chamada de *variável de decisão*, restrito ao conjunto C . Formalmente,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in C \subset \mathbb{R}^n \end{aligned} \tag{1.1}$$

O conjunto C é chamado *conjunto factível*.

A estrutura de $f(x)$ e do conjunto C é muito importante na utilização de algoritmos eficientes para resolver o problema de Otimização. A noção de convexidade representa um papel importante na Teoria de Otimização e reduz a complexidade computacional dos algoritmos.

Definição 1. Conjuntos Convexos

Um conjunto C é convexo se

$$\lambda x + (1 - \lambda)y \in C$$

é válida para todo $\lambda \in [0, 1]$, e para todo $x, y \in C$.

A interpretação geométrica dessa definição é que o segmento entre quaisquer dois pontos de C ainda está em C .

Definição 2. Função Convexa

Uma função $f : C \mapsto \mathbb{R}$ é dita convexa se C é um conjunto convexo e vale a seguinte desigualdade:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

para todo $\lambda \in [0, 1]$ e para todo $x, y \in C$.

Teorema 1. *Seja $f : C \mapsto \mathbb{R}$ uma função convexa. Se f possui um ponto de mínimo local em $x_0 \in C$ então $f(x_0)$ é também um mínimo global de f . Se f é estritamente convexa então x_0 é único.*

Demonstração. Seja f uma função convexa e suponha que f possui um ponto de mínimo local em $x_0 \in C$. Então, para todo $x \in C$ e $\lambda \in [0, 1]$,

$$f(x_0) \leq f((1 - \lambda)x_0 + \lambda x) = f(x_0 + \lambda(x - x_0)) \leq (1 - \lambda)f(x_0) + \lambda f(x) \quad (1.2)$$

Isso implica que

$$0 \leq \lambda(f(x) - f(x_0)) \quad (1.3)$$

isto é, $f(x_0) \leq f(x)$. Assim $f(x_0)$ é um mínimo global. Se f é estritamente convexa então a segunda desigualdade em (1.2) é estrita de modo que (1.3) torna-se estrito para todo $x \in C$. Portanto, x_0 é único. \square

Consideremos agora o problema geral de programação não linear:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) = 0 \\ & h(x) \leq 0 \end{aligned} \quad (1.4)$$

onde $f : R^n \mapsto R$, $g : R^n \mapsto R^m$ e $h : R^n \mapsto R^p$.

Definição 3. *Seja $x \in C = \{x \in R^n | g(x) = 0 \text{ e } h(x) \leq 0\}$, x é um ponto regular se o posto de $\{\nabla g_1(x), \dots, \nabla g_m(x)\} \sqcup \{\nabla h_1(x), \dots, \nabla h_p(x)\}$ é um conjunto linearmente independente.*

Teorema 2. *Seja x^* um minimizador local regular de (1.4). Seja $I = \{i \in \{1, \dots, p\} | h_i(x^*) = 0\}$. Se $\{\nabla g_1(x^*), \dots, \nabla g_m(x^*)\} \sqcup \{\nabla h_1(x^*), \dots, \nabla h_p(x^*)\}$ é um*

conjunto linearmente independente, então existem únicos $\lambda_1, \dots, \lambda_m \in R$ e $\mu_i \geq 0$ para todo $i \in I$ tais que:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i \in I} \mu_i \nabla h_i(x^*) = 0$$

Se x é um ponto regular e minimizador local para o problema (1.4), definindo $\mu_i = 0$ se $i \notin I$, então temos as condições gerais de KKT (Karush-Kuhn-Tucker) da seguinte forma:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i \in I} \mu_i \nabla h_i(x^*) = 0 \quad (1.5)$$

$$g(x) = 0 \quad (1.6)$$

$$\mu_i h_i(x) = 0, \quad i = 1, \dots, p \quad (1.7)$$

$$\mu_i \geq 0, \quad i = 1, \dots, p \quad (1.8)$$

$$h_i(x_i) \leq 0, \quad i = 1, \dots, p \quad (1.9)$$

As equações (1.5) – (1.7) formam um sistema não linear de $n + m + p$ equações nas incógnitas $x \in R^n, \lambda \in R^m$ e $\mu \in R^p$. As soluções deste sistema que satisfazem as equações (1.8) e (1.9) são os pontos estacionários de (1.4)

Demonstração. Vide referência [36]. □

Teorema 3. *Sejam $C \subset R^n$ aberto e convexo, $f : C \mapsto R$, $g \in C^1(C)$. Então f é convexa se, e somente se, $f(x) \geq f(y) + \nabla^T f(x)(y - x)$, para todo $x, y \in C$.*

Demonstração. Vide referência [36]. □

Teorema 4. *Se o problema geral de otimização (1.4) é um problema de programação convexa e em x^* valem as condições de KKT, então x^* é um minimizador global (a regularidade não é necessária).*

Demonstração. Definimos $C = \{x \in R^n | g(x) = 0, h(x) \leq 0\}$ e tomamos $x \in C, x \neq x^*$.

Se $\lambda \in R^m$ e $\mu \in R^p$ são os multiplicadores, então temos:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i \in I} \mu_i \nabla h_i(x^*) = 0 \quad (1.10)$$

$$g_i(x^*) = 0 \quad (1.11)$$

$$\mu_i h_i(x^*) = 0, \quad i = 1, \dots, p \quad (1.12)$$

$$\mu_i \geq 0, \quad i = 1, \dots, p \quad (1.13)$$

$$h_i(x^*) \leq 0, \quad i = 1, \dots, p \quad (1.14)$$

□

DUALIDADE

Considerando o problema geral de otimização não linear (1.4):

Definição 4. Chamamos o problema dual ao problema

$$\begin{aligned} \max \quad & \phi(x, \lambda, \mu) \\ \text{s.a} \quad & \nabla_x \phi(x, \lambda, \mu) = 0 \\ & \mu \geq 0 \end{aligned}$$

onde $\phi(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \mu_i h_i(x)$.

Reescrevendo o problema dual, temos:

$$\begin{aligned} \max \quad & f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \mu_i h_i(x) \\ \text{s.a} \quad & \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{i \in I} \mu_i \nabla h_i(x) = 0 \\ & \mu \geq 0 \end{aligned} \quad (1.15)$$

Teorema 5. Se o problema (1.4) é um problema de programação convexa, e em x^* valem as condições gerais de KKT com os multiplicadores correspondentes λ^* e μ^* ,

então (x^*, λ^*, μ^*) é solução do dual (1.15). Além disso, o valor da função objetivo primal e dual coincidem, isto é, $f(x^*) = \phi(x^*, \lambda^*, \mu^*)$.

Demonstração. Vide referência [36]. □

1.2.2 Programação Quadrática

A Otimização Quadrática compreende uma das áreas mais importantes da programação não linear. Um grande número destes problemas podem ser formulados como problemas da Programação Quadrática (QP), ou seja, minimizar uma função objetivo quadrática cujas restrições são um conjunto de restrições de igualdade ou desigualdade. A Programação Quadrática (QP) pode ser aplicada a problemas reais, incluindo entre eles projeto e análise de sistema de controle na engenharia.

Particularmente, problemas quadráticos com restrições quadráticas ocorrem freqüentemente na Teoria de Controle. Além disso, o problema quadrático é conhecido por ser NP-hard que o torna uma das classes mais interessantes e desafiantes dos problemas de Otimização. Portanto, está claro que o problema de Programação Quadrática tem grande importância, tanto do ponto de vista matemático, quanto do ponto de vista de aplicação, conforme pode ser visto no grande número de publicações nas quais são propostas soluções para estes problemas. Tradicionalmente, problemas quadráticos são tratados como uma subclasse do problema de programação não linear e os métodos para resolvê-los baseiam-se em técnicas de otimização local. Porém, nos últimos anos, há pesquisas para desenvolver algoritmos de otimização global para essa classe de problemas [16].

O Problema Geral de Programação Quadrática

O problema geral de programação quadrática consiste em minimizar uma função objetivo quadrática, restrito a um conjunto de desigualdades lineares, como

mostra abaixo:

$$(PQ) \begin{cases} \min Q(x) = c^T x + \frac{1}{2} x^T D x \\ s.a & Ax \leq b \\ & x \geq 0 \end{cases} \quad (1.16)$$

onde $c \in R^n$, $b \in R^m$, A é uma matriz $m \times n$ e D é uma matriz $n \times n$. Sem perda de generalidade, supomos D uma matriz simétrica.

Se a matriz D é semidefinida positiva ou definida positiva, o problema (PQ) torna-se um problema de programação convexa. Como qualquer ponto de mínimo local é equivalente ao ótimo global em problemas convexos, o problema (PQ) pode ser resolvido por qualquer algoritmo da programação quadrática convexa. Em particular, o problema quadrático convexo pertence à classe de problemas resolvidos em tempo polinomial, portanto existem algoritmos polinomiais que podem ser aplicados para resolver esses problemas. Por exemplo, um algoritmo polinomial (incluindo abordagem de pontos interiores) para programação quadrática convexa foi proposto por Nesterov [41], Monteiro [37] e Ye [63].

Problemas Quadráticos com Restrições Quadráticas

A classe mais geral de problemas quadráticos surge quando incluímos restrições quadráticas no problema (PQ) . Esses problemas podem ser formulados como segue:

$$(QCQP) \begin{cases} \min Q(x) = x^T D x + c^T x \\ s.a & x^T A_j x + B_j x \leq b_j, \quad j = 1, \dots, m \\ & x \geq 0 \end{cases} \quad (1.17)$$

A_j é uma matriz $n \times n$ correspondendo à m -ésima restrição quadrática e B_j é a j -ésima linha da matriz $B_{m \times n}$.

Um problema particular com restrições quadráticas que tem recebido atenção considerável é o problema que envolve a minimização de uma função quadrática

sobre uma esfera:

$$\begin{aligned} \min \quad & \frac{1}{2}x^T D x + c^T x \\ \text{s.a} \quad & x^T x \leq 1 \end{aligned} \tag{1.18}$$

Este problema surge freqüentemente como um subproblema em algoritmos de otimização geral. Freqüentemente, a função objetivo na programação não linear é aproximada localmente por uma função quadrática. Em tais casos, a aproximação é restringida a uma região pequena ao redor da iteração atual. Se a norma Euclidiana é usada para definir esta região, então usamos o problema (1.18). Tais métodos são chamados métodos de região de confiança. Por causa da presença da restrição esférica, é provável que a solução de (1.18) seja um número irracional, o que implica que não é possível calcular a solução exatamente. Porém, existem algoritmos em tempo polinomial que calculam com eficiência a solução aproximada para este problema dado uma precisão desejada. Soresen [54] propôs um algoritmo geral que usa modelo de regiões de confiança, enquanto Ye [64] e Karmarkar [27] propuseram algoritmos fortemente polinomiais para resolver (1.18). Um algoritmo é dito fortemente polinomial se for possível encontrar limitantes polinomiais para o número de operações aritméticas que dependam somente da estrutura do problema.

1.2.3 Regiões de Confiança

Em otimização irrestrita, lidamos com o problema de encontrar o mínimo de uma função $f : R^n \mapsto R$. Se a função f é de classe C^2 , podem ser aplicados muitos métodos para encontrar o mínimo. O método do gradiente e o método de Newton provavelmente estão entre os mais conhecidos. A idéia fundamental do método de Newton é aproximar a função f a ser minimizada ou maximizada por uma função quadrática, construída com base nos valores da primeira e segunda derivadas de f . Mais exatamente, seja $f : R^n \mapsto R$ de classe C^2 e seja x_0 um ponto suficiente próximo da

solução ótima desejada. A fórmula de Taylor de segunda ordem em x_0 fornece uma aproximação quadrática para f , dada por

$$\tilde{f}(x_0 + h) = f(x_0) + \nabla^T f(x_0)h + \frac{1}{2}h^T \nabla^2 f(x_0)h$$

A aproximação \tilde{f} tem mínimo para $h = -(\nabla^2 f(x_0))^{-1} \nabla f(x_0)$. Esse passo h gera uma nova aproximação $x_1 = x_0 + h$ para a solução ótima. Assim, o método de Newton gera a seqüência definida recursivamente por

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)^T$$

Embora seja muito eficiente, uma de suas desvantagens é que não possui convergência global e o desempenho do método é muito dependente da estimativa inicial. Em particular, a seqüência (x_k) pode convergir para um ponto de sela, um máximo local, ou um mínimo local.

Uma maneira de contornar essas dificuldades é minimizar a cada iteração o mesmo modelo quadrático como no método de Newton, mas em vez de considerar o espaço inteiro para a minimização, restringimo-nos à uma bola fechada chamada de região de confiança. Métodos de região de confiança (minimização de uma função objetivo quadrática sujeito a uma uma restrição quadrática) são técnicas eficientes e robustas para resolver problemas de otimização irrestrita. Esses métodos possuem propriedades teóricas de convergência mais fortes que os métodos de busca linear. A principal dificuldade é resolver eficazmente, em cada iteração, o subproblema de região de confiança (do inglês, trust region subproblems (TRS))

$$\begin{aligned} \min_p L(p) &= \frac{1}{2}p^T \tilde{H}p + p^T \tilde{b} + \tilde{c} \\ \text{s.a.} \quad &\|p\|_2 \leq \Delta, \quad \Delta > 0 \end{aligned} \tag{1.19}$$

onde, \tilde{H} é uma matriz simétrica real de ordem n , $\tilde{b} \in R^n, \tilde{c} \in R, \Delta$ é um escalar positivo e $p \in R^n$ é a variável.

Um subproblema de região de confiança tem muitas aplicações em diversas áreas, como por exemplo minimização de função, programação quadrática sequencial, regularização de problemas mal-condicionados e otimização discreta.

Dentre os artigos relacionados com métodos de região de confiança e com subproblemas de região de confiança, muitos deles citam os trabalhos de Levenberg [29] e Marquardt [34] para problemas de mínimos quadrados não lineares. Powell [48] foi o primeiro a estabelecer convergência dos métodos de região de confiança para otimização irrestrita. Moré [38, 39] faz um tratado sobre o desenvolvimento dos métodos de região de confiança. Fletcher [15] e Sorensen [54] provaram propriedades de convergência global para algoritmos de região de confiança usando informações de segunda ordem. Recentemente, Wolkowicz [59, 60] usou técnicas da programação semidefinida para resolver problemas (TRS).

Dependendo dos valores de \tilde{H} , \tilde{b} e Δ , diferentes modos de resolver o subproblema de região de confiança precisam ser considerados. Dois casos diferentes podem acontecer: o caso fácil e o difícil. O caso difícil e próximo a ele causam dificuldades numéricas na solução do problema.

Análise do subproblema da Região de Confiança

Teorema 6. *Seja o problema (TRS)*

$$\begin{aligned} \min_p L(p) &= \frac{1}{2} p^T \tilde{H} p + p^T \tilde{b} + \tilde{c} \\ \text{s.a.} \quad & \| p \|_2 \leq \Delta, \Delta > 0 \end{aligned} \tag{1.20}$$

Se o vetor $p^ \in B[0; \Delta]$ resolve o problema TRS então p^* é solução da equação:*

$$(\tilde{H} + \alpha I)p^* = -\tilde{b} \tag{1.21}$$

*com $\alpha \geq 0$, $\alpha(p^{*T} p^* - \Delta^2) = 0$ e $(\tilde{H} + \alpha I) \geq 0$.*

Demonstração. O problema (1.20) é equivalente a :

$$\begin{aligned} & \min_p L(p) & (1.22) \\ \text{s.a} \quad & p^* p \leq \Delta^2 \end{aligned}$$

Como p^* é solução de (1.20), então p^* satisfaz as condições de Kunh-Tucker [3] para (1.22), isto é, existe um $\alpha \geq 0$ tal que $\tilde{H}p + \tilde{b} + \alpha p^* = 0$ e $\alpha(p^{*T} p^* - \Delta^2) = 0$. Portanto, p^* e α verificam (1.21).

Para concluir que $(\tilde{H} + \alpha I) \geq 0$, considere inicialmente $p^* \neq 0$. Como p^* é solução de (1.20), p^* também é minimizador global de $L(p)$ sujeita a $\|p\| = \|p^*\|$. Então:

$$L(p) \geq L(p^*) \quad (1.23)$$

para todo p tal que $\|p\| = \|p^*\|$.

Substituindo (1.21) em (1.23), temos:

$$\frac{1}{2} p^T \tilde{H} p - p^T (\tilde{H} + \alpha I) p \geq \frac{1}{2} p^{*T} \tilde{H} p^* - p^{*T} (\tilde{H} + \alpha I) p^* \quad (1.24)$$

Rearranjando (1.24), segue que:

$$\frac{1}{2} (p - p^*)^T (\tilde{H} + \alpha I) (p - p^*) \geq 0 \quad (1.25)$$

para todo p tal que $\|p\| = \|p^*\| = \Delta$. Como $p^* \neq 0$ e o resultado da desigualdade resultante vale para todo p com $\|p\| = \Delta$, deduzimos que $(\tilde{H} + \alpha I) \geq 0$.

Se $p^* = 0$, por (1.21) temos $\tilde{b} = 0$. Então $p^* = 0$ é minimizador local de

$$\begin{aligned} \min_p L(p) &= \frac{1}{2} p^T \tilde{H} p + \tilde{c} & (1.26) \\ \text{s.a} \quad & \|p\|_2 \leq \Delta, \quad \Delta > 0 \end{aligned}$$

uma vez que $\tilde{H} \geq 0$ e $\nu^T (\tilde{H} + \alpha I) \nu \geq 0$ vale para todo $\nu \in R^{nT}$ com $\alpha = 0$. \square

O teorema seguinte fornece as condições suficientes que garantem que p^* é a solução do problema (TRS) (1.20).

Teorema 7. *Sejam $\alpha \in R$ e $p^* \in R^{nT}$ tais que :*

$$(\tilde{H} + \alpha I)p^* = -\tilde{b} \text{ com } (\tilde{H} + \alpha I) \geq 0. \quad (1.27)$$

1. *Se $\alpha = 0$ e $\|p^*\| \leq \Delta$ então p^* é a solução de (1.20)*

2. *Se $\|p^*\| = \Delta$ então p^* é a solução de*

$$\begin{aligned} \min_p L(p) &= \frac{1}{2}p^T \tilde{H}p + \tilde{c} \\ \text{s.a.} \quad &\|p\|_2 = \Delta \end{aligned}$$

3. *Se $\alpha \geq 0$ e $\|p^*\| = \Delta$ então p é a solução de (1.20). Além disso, se $(\tilde{H} + \alpha I) > 0$, então p^* é única em 1, 2 e 3.*

Demonstração. Se α e p^* satisfazem (1.27), p^* é minimizador da quadrática:

$$L(p) = \frac{1}{2}p^T(\tilde{H} + \alpha I)p + \tilde{b}^T p + \tilde{c}$$

Logo,

$$\frac{1}{2}p^T(\tilde{H} + \alpha I)p + \tilde{b}^T p + \tilde{c} \geq \frac{1}{2}p^{*T}(\tilde{H} + \alpha I)p^* + \tilde{b}^T p^* + \tilde{c} \quad (1.28)$$

para todo $p \in R^{nT}$. De (1.28), segue que

$$L(p) \geq L(p^*) + \frac{\alpha}{2}(p^{*T} p^* - p^T p) \quad (1.29)$$

para todo $p \in R^{nT}$.

As afirmações 1, 2 e 3 são conseqüências imediatas de (1.29). A unicidade segue de (1.28), pois se $(\tilde{H} + \alpha I) > 0$, a desigualdade é estrita para $p \neq p^*$. \square

Os teoremas acima mostram que, se há uma solução p^* para o problema (1.20), situada na fronteira da bola, ela deve satisfazer, com seu multiplicador correspondente α , as seguintes equações:

$$(\tilde{H} + \alpha I)p^* = -\tilde{b}, \quad \|p^*\| = \Delta \quad (1.30)$$

com $\alpha \geq 0$ e $(\tilde{H} + \alpha I) \geq 0$.

Considerando a decomposição espectral $\tilde{H} = V\Lambda V^T$ onde as colunas de V são os autovetores ortonormais de \tilde{H} e Λ é uma matriz diagonal cujos elementos são os autovalores de \tilde{H}

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad \lambda_1 \leq \dots \leq \lambda_n \quad (1.31)$$

Como $(\tilde{H} + \alpha I) = V(\Lambda + \alpha I)V^t$, a equação (1.30) é equivalente a:

$$(\Lambda + \alpha I)p = -V^T \tilde{b} \quad (1.32)$$

A condição $(\tilde{H} + \alpha I) \geq 0$ é equivalente a $\alpha \geq -\lambda_1$. Na região em que $\alpha \geq \max\{0, -\lambda_1\}$, o sistema $(\tilde{H} + \alpha I)p^* = -\tilde{b}$ tem como solução única $p^* = -(\tilde{H} + \alpha I)^{-1}\tilde{b}$, pois $(\tilde{H} + \alpha I)$ é não singular. Portanto, encontrar α satisfazendo (1.30) é equivalente a resolver

$$\|(\tilde{H} + \alpha I)^{-1}\tilde{b}\|^2 = \Delta^2, \quad \alpha \geq -\lambda_1, \alpha \geq 0 \quad (1.33)$$

Casos a considerar

Denotaremos $\mathcal{N}(\cdot)$ o espaço nulo e $\mathcal{R}(\cdot)$ o espaço imagem. Vamos definir o seguinte:

1. **Caso Fácil:** If \tilde{b} não é perpendicular a $\mathcal{N}(\tilde{H} + \lambda_1(\tilde{H})I)$ (equivalentemente $\tilde{b} \notin \mathcal{R}(\tilde{H} + \lambda_1(\tilde{H})I)$), então temos o caso fácil; isso implica que $(\tilde{H} + \alpha^* I) \succ 0$. Em particular, $(\tilde{H} + \alpha^* I)$ é não-singular e podemos então trabalhar com (1.33). O ponto ótimo $p^* = p(\alpha^*)$ é único.
2. **Caso Difícil:** If \tilde{b} é perpendicular a $\mathcal{N}(\tilde{H} + \lambda_1(\tilde{H})I)$ (equivalentemente $\tilde{b} \in \mathcal{R}(\tilde{H} + \lambda_1(\tilde{H})I)$), duas possibilidades podem ocorrer:
 - (a) **Caso Difícil (caso 1):** If $\alpha^* > -\lambda_1(\tilde{H})$, nenhuma dificuldade óbvia acontece. Podemos trabalhar com (1.33). O ponto ótimo $p^* = p(\alpha^*)$ é único.
 - (b) **Caso Difícil (caso 2):** If $\alpha^* = \lambda_1(\tilde{H})$, então existem duas possibilidades:

- i. Quando $\|(\tilde{H} + \alpha^* I)^\dagger \tilde{b}\| = \Delta$ ou $\alpha^* = 0$, então o par $p^* = (p(\alpha^*), \alpha^*)$ satisfaz as condições de otimalidade. Portanto, podemos trabalhar com (1.33) e usar a matriz inversa generalizada de Moore-Penrose para o caso singular.
- ii. Quando $p(\alpha^*) = \|(\tilde{H} + \alpha^* I)^\dagger \tilde{b}\| < \Delta, \alpha^* > 0$, quase sempre falha a condição de otimalidade. Contudo, podemos escolher um autovetor $z \in \mathcal{N}(\tilde{H} + \alpha^* I)$ e calcular τ apropriado, de modo que $\Delta = \|p(\alpha^*) + \tau z\|$. (Observe que $p(\alpha^*) \perp z$, i.e. $\Delta^2 = \|p(\alpha^*)\|^2 + \|\tau z\|^2$). Este é o único caso onde o ótimo não é único. A variedade de soluções ótimas é dada pela interseção do *autoespaço generalizado* $\{p : (\tilde{H} + \alpha^* I)p - \tilde{b} = 0\}$ com a fronteira da bola, $\{p : \|p\| = \Delta\}$.

Tabela 1.1: Os três casos diferentes para o subproblema da região de confiança.

1. Caso fácil	2.(a) Caso difícil (caso 1)	2.(b) Caso difícil (caso 2)
$b \notin \mathcal{N}(\tilde{H} + \lambda_1(\tilde{H})I)$	$\tilde{b} \perp \mathcal{N}(\tilde{H} + \lambda_1(\tilde{H})I)$	$\tilde{b} \perp \mathcal{N}(\tilde{H} + \lambda_1(\tilde{H})I)$
(implica $\alpha^* > -\lambda_1(\tilde{H})$)	e $\alpha^* > -\lambda_1(\tilde{H})$	e $\alpha^* = \lambda_1(\tilde{H})$
		(i) $\ (\tilde{H} + \alpha^* I)^\dagger \tilde{b}\ = \Delta$ ou $\alpha^* = 0$ (ii) $\ (\tilde{H} + \alpha^* I)^\dagger \tilde{b}\ < \Delta, \alpha^* < 0$

Neste trabalho, usamos o caso fácil do problema TRS, como uma maneira de regularizar o problema de minimizar uma função quadrática irrestrita:

$$\min_p L(p) = \frac{1}{2} p^T \tilde{H} p + p^T \tilde{b} + \tilde{c} \quad (1.34)$$

isto é, resolver o problema (1.34) quando \tilde{H} é mal-condicionada. Quando isso acontece, a solução do problema (1.34) não é exata, por ser sensível aos erros dos dados ou ao arredondamento.

1.2.4 Métodos de Pontos Interiores

Desde que possamos resolver problemas de minimização irrestrita mais facilmente que problemas de minimização restrita, é natural tentar converter problemas restritos em problema irrestritos. Dentre as várias técnicas que transformam um problema restrito em problema irrestrito, ressaltamos o método da função penalidade bastante conhecido na literatura [3, 11, 32] .

Em geral, uma função de penalidade deve gerar uma penalidade positiva para pontos não factíveis e não penalizar pontos factíveis. Estes métodos podem ser divididos em:

- Métodos de penalidade interna, de barreira ou de pontos interiores, os quais geram uma seqüência de pontos factíveis, cujo ponto limite é a solução ótima do problema original.
- Métodos de penalidade externa, de pontos exteriores ou simplesmente métodos de penalidades. Esses métodos geram uma seqüência de pontos não factíveis, cujo ponto limite é a solução ótima do problema original.

Consideremos o problema a seguir:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in S \end{aligned} \tag{1.35}$$

onde f é uma função contínua em R^n e S é um conjunto de restrições em R^n com interior não vazio, sabendo-se que o interior de qualquer conjunto $S \subset R^n$ é o conjunto de pontos $x \in S$ os quais são o centro de alguma esfera contida em S .

Os métodos de barreira funcionam estabelecendo uma barreira na fronteira da região factível evitando que um procedimento de busca abandone a região. Uma função barreira é uma função $B(x)$ definida no interior de S tal que:

1. $B(x)$ é contínua;
2. $B(x) \geq 0$;
3. $B(x) \rightarrow \infty$ quando x tende à fronteira de S .

Métodos de pontos interiores operam sempre no interior do conjunto S , usando uma função $B(x)$ que cresce indefinidamente próxima à fronteira de S e uma seqüência decrescente de multiplicadores λ .

Equivalente ao problema (1.35), consideremos o problema de otimização irrestrito:

$$\begin{aligned}
 \min \quad & f(x) + \lambda B(x) & (1.36) \\
 \text{s.a} \quad & x \in \text{int}S \\
 & \lambda > 0
 \end{aligned}$$

A vantagem desse problema consiste em ele ser de simples implementação, por usar métodos de otimização sem restrições. No entanto, tipicamente é de convergência lenta.

Método das Barreiras

O procedimento para resolver o problema (1.35) pelo método das barreiras é como o que segue: Seja (λ_k) , $k = 1, 2, \dots$ uma seqüência dos parâmetros tal que para cada k ,

1. $\lambda_k \geq 0$
2. $\lambda_{k+1} < \lambda_k$
3. $\lim_{k \rightarrow \infty} \lambda_k = 0$

A seguir, definimos uma função $B : R^n \rightarrow R$, denominada função de barreira tal que

$$\begin{cases} B(x) = 0 & \text{se } x \in S \\ B(x) > 0 & \text{se } x \notin S \end{cases} \quad (1.37)$$

Consideremos a função

$$q(\lambda, x) := f(x) + \lambda B(x) \quad (1.38)$$

Para cada k , resolve-se o problema

$$q(\lambda_k, x_k) = \min_{x \in \mathbb{R}^n} q(\lambda_k, x) \quad (1.39)$$

Desta forma, para λ_k , calcule x_k tal que (1.39) seja satisfeita. Se x_k for um ponto ótimo para o problema (1.35), então pare; caso contrário, continue com λ_{k+1} .

Lema 1.

$$q(\lambda_k, x_k) \geq q(\lambda_{k+1}, x_{k+1}) \quad (1.40)$$

$$B(x_k) \leq B(x_{k+1}) \quad (1.41)$$

$$f(x_k) \geq f(x_{k+1}) \quad (1.42)$$

Demonstração. Vide referências [3, 11, 32]. □

Lema 2. *Seja x^* uma solução do problema (1.35). Então, para cada k , $f(x^*) \geq q(\lambda_k, x_k) \geq f(x_k)$.*

Demonstração. Vide referências [3, 11, 32]. □

Teorema 8. *Seja (x_k) uma seqüência gerada pelo método das barreiras. Então, qualquer ponto limite da seqüência é uma solução de (1.35).*

Demonstração. Suponha que a subseqüência (x_{k_i}) é uma subseqüência convergente da seqüência (x_k) com limite a . Então, pela continuidade de f , temos

$$\lim_{i \rightarrow \infty} f(x_{k_i}) = f(a) \quad (1.43)$$

Seja f^* o valor ótimo associado ao problema (1.35). Então, pelos Lemas 1 e 2, a seqüência de valores $\{q(\lambda_k, x_k)\}$ é não-decrescente e limitada inferiormente por f^* .

Desse modo,

$$\lim_{i \rightarrow \infty} q(\lambda_k, x_k) = \lim_{k \rightarrow \infty} q(\lambda_k, x_k) = f^* \quad (1.44)$$

Subtraindo (1.43) de (1.44), obtemos

$$\lim_{k \rightarrow \infty} \lambda_k B(x_k) = f^* - f(a) \quad (1.45)$$

Como $B(x_k) \geq 0$ e $\lambda_k \rightarrow \infty$, a equação (1.45) será válida somente se $\lim_{i \rightarrow \infty} B(x_k) = 0$. Utilizando a continuidade de $B(x)$, isso implica $\lim_{i \rightarrow \infty} B(x_k) = B(a) = 0$. Portanto a é factível para (1.35), isto é, $a \in S$.

Para mostrar que a é ótimo, observamos pelo Lema 1 que, $f(x_k) \leq f^*$; portanto, $f(a) = \lim_{k \rightarrow \infty} f(x_{ki}) \leq f^*$. Assim provamos que $a \in S$ e $f(a) \leq f^*$. Conseqüentemente, o ponto a deve ser uma solução. \square

1.2.5 Sistemas Dinâmicos Lineares Discretos Invariantes no Tempo

Nesta seção, alguns resultados básicos são apresentados, para um sistema dinâmico linear discreto invariante no tempo descrito pelas seguintes equações a diferenças:

$$x(k+1) = Ax(k) + Bu(k); \quad x(0) = x_0 \quad (1.46)$$

$$y(k) = Cx(k) + Du(k) \quad (1.47)$$

onde $x(k) \in R^n$ é denominado vetor de estado do sistema, $u(k) \in R^p$ é chamado de vetor de controle (entrada) e $y(k) \in R^q$ é o vetor de saída do sistema. As matrizes reais constantes A, B, C e D têm dimensões apropriadas.

A matriz de transferência correspondente que relaciona a entrada $u(k)$ e a saída $y(k)$ é definida como:

$$Y(z) = G(z)U(z)$$

onde $U(z)$ e $Y(z)$ são, respectivamente, as transformada- z de $u(k)$ e $y(k)$ com condição inicial nula. Portanto, temos:

$$G(z) = C(zI - A)^{-1}B + D \quad (1.48)$$

Teorema 9. *Uma equação de estado linear invariante no tempo é dita controlável, se e somente se qualquer uma das seguintes condições equivalentes é satisfeita:*

1. *A matriz de controlabilidade :*

$$\Lambda = \begin{bmatrix} A^{-n}B & A^{-n+1}B & \dots & A^{-1}B \end{bmatrix}$$

tem posto n .

2. *A matriz complexa $\begin{bmatrix} A - \lambda I & B \end{bmatrix}$ tem posto n para todo $\lambda \in C$.*

3. *Os autovalores de $A+BK$ podem ser escolhidos livremente, se houver uma matriz K adequada.*

Demonstração. Vide referência [23].

□

Teorema 10. *Um sistema linear invariante no tempo é dito observável, se e somente se qualquer uma das seguintes condições equivalentes é satisfeita:*

1. *A matriz de observabilidade*

$$O = \begin{bmatrix} C & CA & \dots & CA^{n-1} \end{bmatrix}$$

tem posto n .

2. *A matriz complexa $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ tem posto n para todo $\lambda \in C$.*

3. *Os autovalores de $A+LC$ podem ser escolhidas livremente, se houver uma matriz L adequada.*

Demonstração. Vide referência [23].

□

1.2.6 Regulador Linear Quadrático

O controle linear quadrático (LQ) é uma técnica de controle moderno que surgiu aproximadamente nos anos 60. Nessa técnica, existem alguns casos especiais como: LQR (Linear Quadratic Regulator); LQG (Linear Quadratic Gaussian) e

LQG/LTR (Loop Transfer Recovery) que é o LQG com recuperação da malha de transferência.

O problema LQR , estudado nesta seção, é dado por:

$$\begin{aligned} \min J &= \frac{1}{2} \|x(T)\|_S^2 + \frac{1}{2} \sum_{k=0}^{T-1} (\|x(k)\|_Q^2 + \|u(k)\|_R^2) \\ \text{s.a} \quad &x(k+1) = Ax(k) + Bu(k), \quad x(0) = \xi \end{aligned} \quad (1.49)$$

e $Q \geq 0$, $S \geq 0$ e $R > 0$, onde Q, S e R são matrizes simétricas com dimensões apropriadas. O Hamiltoniano para o problema LQR é definido por:

$$H^k = \frac{1}{2} [\|x(k)\|_Q^2 + \|u(k)\|_R^2] + p^{*T}(k)(Ax(k) + Bu(k))$$

As condições de otimalidade fornecem:

$$x(k+1) = \frac{\partial H^k}{\partial p^*(k)} = Ax(k) + Bu(k) \quad (1.50)$$

$$p^*(k-1) = \frac{\partial H^k}{\partial x(k)} = Qx(k) + A^T p^*(k), \quad k = 1, \dots, T-1 \quad (1.51)$$

$$0 = \frac{\partial H^k}{\partial u(k)} = Ru(k) + B^T p^*(k), \quad k = 0, \dots, T-1 \quad (1.52)$$

$$p^*(T-1) = Qx(T) \quad (1.53)$$

A solução ótima do problema pode ser encontrada de duas maneiras: em malha aberta, obtendo uma seqüência ótima $u(k)$, ou em malha fechada, utilizando controle do tipo $u(k) = Kx(k)$.

Solução em Malha Aberta

Na solução em malha aberta, podemos utilizar diversas técnicas baseadas em programação matemática. Neste trabalho, vamos considerar um método de otimização multinível baseado em dualidade, dado que nosso problema é convexo e a função dual $\phi(\lambda(k))$ é côncava.

O problema dual é definido como:

$$\max \phi(\lambda(k)) = \max_{\lambda} \min_{x, u} L(x(k), u(k), \lambda(k))$$

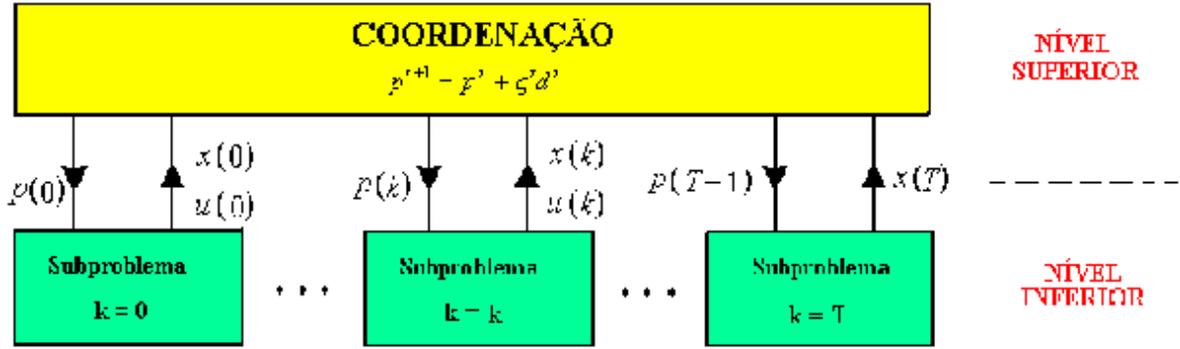


Figura 1.1: Estrutura do Controle Ótimo Multinível

onde $L(x(k), u(k), \lambda(k))$ é o lagrangeano, que é dado por:

$$L(x, u, p^*, k) = \frac{1}{2} \|x(T)\|_S^2 + \sum_{k=0}^{T-1} \{\mathcal{H}^k(x(k), u(k), p^*) - p^{*T}(k-1)x(k)\} \quad (1.54)$$

A Figura (1.1) mostra, esquematicamente, a estrutura hierárquica em dois níveis.

A tabela (1.2) descreve sucintamente o algoritmo de Otimização Multinível:

Tabela 1.2: Algoritmo de Otimização Multinível

Passo 1	Fazer $k = 0$ e escolher $p^T = (p^T(0), p^T(1), \dots, p^T(T-1))^T$
Passo 2	Resolver para $k = 0, 1, \dots, (T-1)$ as equações (1.50) – (1.53)
Passo 3	Encontrar $\nabla L(p^T)$ e verificar se $\ \nabla L(p^T)\ < \epsilon$
	Caso sim, fim .
	Caso não, ir para 4.
Passo 4	Fazer $p^{T+1} = p^T + \tau^T d^T$, $T = T + 1$, e voltar para 2
	onde $d^1 = \nabla L(p^1)$.
	$d^T = \nabla L(p^T) + \Lambda^{T-1} d^{T-1}$ é a direção de descida
	$\Lambda^{T-1} = \frac{\sum_{k \in K} \nabla L(p^T(k))^T \nabla L(p^T(k))}{\nabla L(p^{T-1}(k))^T \nabla L(p^{T-1}(k))}$ é o tamanho do passo .

Solução em Malha Fechada

O controle LQR em malha fechada é uma técnica no domínio do tempo que possui as seguintes características: ampla margem de estabilidade, margem de fase de 60° e margem de ganho infinito. Possui, portanto, a propriedade de estabilidade robusta. Contudo ele requer que o estado seja completamente acessível para medida, o que normalmente na prática não é possível. Esta solução é bem conhecida e é dada na Tabela (1.3):

Tabela 1.3: Regulador Linear Quadrático

Modelo do sistema	$x(k+1) = Ax(k) + Bu(k)$
Índice de desempenho	$J = \frac{1}{2} \ x(T)\ _S^2 + \frac{1}{2} \sum_{k=0}^{T-1} (\ x(k)\ _Q^2 + \ u(k)\ _R^2)$
Controle de realimentação : (Equação de Riccati)	$P = A^T P A - A^T P B (B^T S B + R)^{-1} B^T P A + Q$
Ganho de controlador	$K = (B^T S B + R)^{-1} B^T P A$
Controle a ser aplicado	$u(k) = K x(k)$
Índice de desempenho otimizado	$J(i) = \frac{1}{2} x^T(0) P x(0)$

1.3 Comentário Final

Neste capítulo, procuramos inicialmente revisar algumas definições e teoremas de otimização convexa, dualidade, regiões de confiança e métodos de pontos interiores que juntos são de importância fundamental para o desenvolvimento dos controladores que serão propostos na seqüência do trabalho. Foi também comentada a solução do problema do regulador em malha aberta e em malha fechada.

Capítulo 2

Programação Semidefinida

2.1 Conceito de SDP

O problema de Programação Semidefinida (SDP) é um problema de Otimização Convexa que consiste em minimizar uma função linear sob o cone das matrizes semidefinidas positivas [5, 40, 45].

Existem boas razões para estudarmos problemas SDPs. Particularmente, problemas cujas restrições são matrizes semidefinidas positivas surgem muitas vezes em importantes aplicações de áreas distintas, dentre elas: Análise Convexa [52], Programação Quadrática (QP) [16] e Métodos de Pontos Interiores [41, 51, 61] principalmente em análise e síntese de sistemas da teoria de controle [5, 6].

Além disso, muitos problemas de otimização convexa podem ser reformulados como problemas SDPs. Porém, talvez a razão mais importante seja a forma eficiente pela qual o problema pode ser resolvido, tanto na teoria quanto na prática.

A Programação SDP é uma extensão da programação linear (em inglês, Linear Programming (LP)) onde as variáveis vetoriais são substituídas por matrizes

e as restrições relativas a cada elemento não negativo ($x \in R_+^n$) são substituídas por matrizes semidefinidas positivas. Essa generalização engloba várias propriedades importantes da programação linear: convexidade do conjunto factível, uma rica teoria de dualidade (embora não tão forte quanto na programação linear) e admite solução eficiente baseada em métodos de pontos interiores [13, 14].

2.2 Cone das Matrizes Semidefinidas Positivas

O conjunto das matrizes M_n reais e quadradas de ordem n pode ser interpretado como um espaço vetorial em R^{n^2} . Neste trabalho, usamos S_n como o espaço vetorial das matrizes simétricas de dimensão $n(n+1)/2$.

Como o conjunto das matrizes semidefinidas positivas é um espaço vetorial, ele é um conjunto convexo.

Cones

- Um cone é um conjunto K tal que se $x \in K$ então $\forall \alpha \geq 0, \alpha x \in K$.
- Um cone é dito convexo se dados $x, y \in K$ implica que $x + y \in K$.
- Um cone é pontiagudo se $K \cap (-K) = \{0\}$.

Teorema 11. *O conjunto das matrizes semidefinidas positivas S_n^+ é um cone pontiagudo.*

Demonstração. $S_n^+ = \{X \in R^{n \times n} | X = X^T \geq 0\}$ é um cone, se $\alpha_1, \alpha_2 \geq 0$ e $A = A^T \geq 0$ e $B = B^T \geq 0$ então $\alpha_1 A + \alpha_2 B \geq 0$.

A convexidade pode ser vista diretamente da definição de matriz semidefinida positiva: Para qualquer $x \in R^n$ temos $x^T(\alpha_1 A + \alpha_2 B)x = \alpha_1 x^T A x + \alpha_2 x^T B x \geq 0$

se $A \geq 0, B \geq 0$ e $\alpha_1, \alpha_2 \geq 0$.

S_n^+ é um cone pontiagudo: $X \in S_n^+, -X \in S_n^+ \Rightarrow X = O_{n \times n}$. □

Definição 5. Um cone K é chamado um cone polar de K se $K^* = \{y : x^T y \geq 0, \forall x \in K\}$, isto é, $K = K^*$.

Exemplo 1. O cone das matrizes semidefinidas positivas S_n^+ é um cone polar, isto é, para $X = X^T, Y = Y^T \in R^{n \times n}$, $Tr XY \geq 0 \forall X \geq 0 \Leftrightarrow Y \geq 0$.

Suponha que Y não é uma matriz semidefinida positiva. Então existe $q \in R^n$ com $q^T Y q = Tr qq^T Y < 0$. Portanto a matriz semidefinida positiva $X = qq^T$ satisfaz $Tr XY < 0$. Logo Y não está no dual do cone das matrizes semidefinidas positivas.

Suponha agora, $Y = Y^T \geq 0$ e seja $X = X^T \geq 0$ qualquer matriz semidefinida positiva. Podemos expressar X em termos da decomposição de autovalores como $X = \sum_i \lambda_i q_i q_i^T$, onde os autovalores $\lambda_i \geq 0$. Temos $Tr YX = Tr Y \sum_i \lambda_i q_i q_i^T = \sum_i \lambda_i q_i^T Y q_i$ usando o fato que $Tr AB = Tr BA$. Como Y é semidefinida positiva, cada termo no somatório é não negativo, assim $Tr YX \geq 0$. Isto prova que Y está no dual do cone das matrizes semidefinidas positivas.

2.3 Problemas SDPs

O problema primal SDP é:

$$(P) \begin{cases} \min & C \bullet X \\ \text{s.a} & A_i \bullet X = b_i \\ & X \geq 0 \end{cases} \quad (2.1)$$

onde: X, C e A_i são todas matrizes quadradas reais e simétricas de ordem n .

$C \bullet X$ é o produto interno satisfazendo

$$C \bullet X := Tr (C^T X) = \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \quad (2.2)$$

A norma proveniente deste produto interno é a norma de Frobenius

$$\| X \| = \sqrt{\text{Tr } X^2}.$$

$A \geq 0$ significa que a matriz A é semidefinida positiva, isto é, para todo x , $x^T A x \geq 0$.

Em geral, $A \geq B$ significa que $A - B$ é uma matriz semidefinida positiva.

Se X satisfaz as restrições $A_i \bullet X = b$ e $X \geq 0$, então dizemos que a matriz X é factível, o conjunto de todas as matrizes factíveis chama-se conjunto factível.

2.4 Dualidade em SDP

O problema dual SDP é:

$$(D) \begin{cases} \max & b^T y \\ \text{s.a} & \sum_{i=1}^m y_i A_i + S = C \\ & S \geq 0 \end{cases} \quad (2.3)$$

onde $y \in R^m$ e $S \in R^{n \times n}$ são as variáveis.

O problema de programação semidefinida (2.1) é chamado de problema primal e os dois problemas juntos são chamados de problema primal-dual. Dizemos que (X, y, S) é uma solução factível do problema primal-dual, se X é uma solução factível satisfazendo $X \geq 0$ do problema primal (2.1) e (y, S) é uma solução factível satisfazendo $S \geq 0$ do problema dual (2.3).

Teorema 12. *Sejam $X \geq 0$ e $S \geq 0$. Então $X \bullet S = 0$ se e somente se $XS = 0$.*

Demonstração. \Rightarrow Observe que :

$$0 = X \bullet S = \text{Tr}(XS) = \text{Tr}(X^{1/2} S X^{1/2})$$

Como $X^{1/2} S X^{1/2} \geq 0$ e seu traço é zero, todos os seus autovalores são nulos assim

$$(X^{1/2} S^{1/2})^T (X^{1/2} S^{1/2}) = X^{1/2} S X^{1/2} = 0$$

e portanto $XS = 0$.

\Leftarrow Obviamente se $XS = 0$ então $\text{Tr}(XS) = X \bullet S = 0$. □

Teorema 13 (Dualidade Fraca). *Se X é uma solução do problema primal e (y, S) é uma solução do problema dual, então $C \bullet X \geq b^T y$.*

Demonstração.

$$\begin{aligned}
C \bullet X - b^T y &= C \bullet X - \sum_{i=1}^m b_i y_i \\
&= C \bullet X - \sum_{i=1}^m (A_i \bullet X) y_i \\
&= C \bullet X - \sum_{i=1}^m (y_i A_i) \bullet X \\
&= (C - \sum_{i=1}^m y_i A_i) \bullet X \\
&= S \bullet X \\
&\geq 0
\end{aligned}$$

□

Teorema 14 (Teorema da Separação). *Seja $A^T y \in K^*$, $b \notin K$. Então existe um espaço linear $\{x : y^T x = 0\}$ tal que $b^T y < 0$ e $y^T x \geq 0$ para todo $x \in K$.*

Demonstração. Vide referências [3, 11, 32].

□

Lema 3 (Lema de Farkas generalizado). *Se $A(K) = \{Ax : x \in K\}$ é fechado, então ou:*

1. $\exists x \in R^n$ tal que $Ax = b$ e $x \in K$ ou

2. $\exists y \in R^n$ tal que $A^T y \in K^*$ e $b^T y < 0$.

Demonstração. Se (1) é verdadeiro então $\forall y \in R^n, b^T y = x^T A^T y \geq 0$ pela definição de polaridade, e assim (2) não pode ser verdadeiro.

Se (1) é falso, então x tal que $Ax = b$ e $x \in K$, assim, $b \notin A(K)$. Pelo teorema da separação, existe um y tal que $b^T y < 0$ e $y^T z \geq 0$ para todo $z \in A(K)$, assim $y^T Ax \geq 0$ para todo $x \in K$, assim $A^T y \in K^*$ e (2) não pode ser verdadeira. □

Teorema 15 (Dualidade Forte). *Seja A uma matriz de posto completo. Suponha que $A(K) = \{Ax : x \in K\}$ é fechado. Se z_1 e z_2 são soluções ótimas e finitas dos problemas SDPs primal e dual, respectivamente, então $z_1 = z_2$.*

Demonstração. O teorema da dualidade fraca implica que $z_1 \geq z_2$. Consideremos $z_1 > z_2$. Então o sistema $c^T x = z_2$, $Ax = b$ e $x \in K$ pode não ser factível. Pelo lema de Farkas generalizado, deve existir um par (y_0, y) tal que $A^T y + c^T y_0 \in K^*$ e $b^T y + z_2 y_0 < 0$. Existem três casos:

1. $y_0 = 0$: Assim $A^T y \in K^*$ e $b^T y < 0$. Portanto $Ax = b$ não é factível para $x \in K$ o que significa que o problema primal não é factível e $z_1 = \infty$. Isso contradiz a hipótese que z_1 é finito.
2. $y_0 > 0$: Dividindo por y_0 , obtemos $c - A^T(\frac{-y}{y_0}) \in K^*$ e $z_2 - b^T(\frac{-y}{y_0}) < 0$ o que significa que $-\frac{y}{y_0}$ é factível dual e sua função objetivo é maior que z_2 . Isso é uma contradição à hipótese que z_2 é máximo.
3. $y_0 < 0$: Dividindo por $-y_0$ e multiplicando pela solução ótima primal x^* obtemos:

$$\begin{aligned} -x^* c + x^* A^T\left(\frac{-y}{y_0}\right) &\geq 0 \\ \Leftrightarrow -z_1 + b^T\left(\frac{-y}{y_0}\right) &\geq 0 \\ \Leftrightarrow b^T\left(\frac{-y}{y_0}\right) &\geq z_1 \\ \Rightarrow z_2 &\geq z_1 \end{aligned}$$

Isso é uma contradição para a hipótese $z_1 > z_2$.

□

2.5 Solução dos Problemas SDPs

A publicação em 1984 de um artigo de Karmarkar [26] foi o fato mais importante na programação linear (PL), depois da descoberta do método Simplex.

Esse novo algoritmo impulsionou uma grande quantidade de artigos na área de métodos de pontos interiores para programação (LP). Porém, toda essa pesquisa foi concentrada em algoritmos para programação linear e programação quadrática (convexa). Por conseguinte, a primeira extensão dos métodos de pontos interiores para a Programação Semidefinida foi feita por Nesterov e Nemirovskii [41] e, independentemente, por Alizadeh [1].

Depois do trabalho de Karmarkar, que na realidade é um algoritmo de barreiras, os métodos de penalidade interna, também conhecidos como métodos de barreira, tiveram um grande desenvolvimento nos últimos 20 anos, devido à convergência polinomial dos algoritmos de pontos interiores [22, 26, 41].

A função barreira logarítmica foi utilizada pela primeira vez em otimização por Frisch [17]. O método da função de barreira logarítmica é estudada no livro clássico de Fiacco e Mc Cormick [11].

Ultimamente os métodos primal-dual têm recebido muita atenção, devido ao sucesso na solução de problemas de grande porte.

2.5.1 Métodos da Barreira Logarítmica

Os métodos primais de barreira logarítmica usam o método de Newton para resolver uma seqüência de problemas da forma

$$\min \{ Tr (CX) - \mu \logdet(X) : Tr (A_i X) = b_i \ (i = 1, \dots, m) \} \quad (2.4)$$

Como na Programação Linear, os métodos primal-dual tornaram-se muito mais populares. Esses métodos minimizam o gap de dualidade onde o parâmetro μ é seqüencialmente decrescente a zero.

$$Tr (CX) - b^T y = Tr (XS)$$

e empregam a função barreira primal-dual combinada

$$f_{pd} := -(\logdet(X) + \logdet(S)) = -\logdet(XS) \quad (2.5)$$

Isso significa que a seqüência de problemas da seguinte forma são resolvidos:

$$\min \{ Tr (XS) - \mu \logdet(XS) : Tr (A_i X) = b_i, \sum_{i=1}^m y_i A_i + S = C \} \quad (2.6)$$

As condições de otimalidade de primeira ordem são:

$$\left\{ \begin{array}{l} Tr (A_i X) = b_i, \ i = 1, \dots, m \\ \sum_{i=1}^m y_i A_i + S = C \\ XS = \mu I \\ X > 0, \ S > 0 \end{array} \right. \quad (2.7)$$

Esse sistema tem um único par solução definida positiva denotada por $X(\mu) > 0$ e $S(\mu) > 0$. Métodos de barreira logarítmica primal-dual resolve o sistema (2.7), através de uma redução de μ .

O fundamental é obter passos primal e dual ΔX e ΔS , respectivamente, que satisfazem $X + \Delta X \geq 0$, $S + \Delta S \geq 0$ e

$$\begin{aligned} \text{Tr}(A_i \Delta X) &= 0, \quad i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0 \\ (X + \Delta X)(S + \Delta S) &= \mu I \end{aligned} \tag{2.8}$$

Dizemos assim que a última equação não é linear, e os métodos primal-dual diferem pela maneira como ela é linearizada. Além disso, devemos considerar que as matrizes ΔX e ΔS são simétricas.

2.5.2 Métodos Seguidor de Trajetória Prima-Dual

Os métodos de pontos interiores prima-dual são um dos métodos mais utilizados na programação linear [61]. Baseiam-se no seguinte resultado: Considere que os problemas primal (2.1) e dual (2.3) possuem soluções estritamente factíveis, então, pelo teorema da dualidade forte, uma solução ótima do SDP primal-dual satisfaz o seguinte sistema de equações:

$$\begin{cases} A_i \bullet X = b \\ \sum_{i=1}^m y_i A_i + S = c \\ XS = 0 \\ X > 0, S > 0 \end{cases} \tag{2.9}$$

A trajetória central C é um conjunto de soluções factíveis que representam um papel importante na teoria dos algoritmos primal-dual. A trajetória é parametrizada por um escalar $\mu > 0$ e cada ponto $(X(\mu), y(\mu), S(\mu)) \in C$ resolve o seguinte sistema:

$$\begin{cases} A_i \bullet X = b \\ \sum_{i=1}^m y_i A_i + S = c \\ XS = \mu I \\ X > 0, S > 0 \end{cases} \tag{2.10}$$

O mecanismo para seguir a trajetória central é: Dada uma iteração atual estritamente factível (X, y, S) , um parâmetro $\mu > 0$ é escolhido e o método de Newton é aplicado ao sistema (2.10) para determinar a próxima iteração. Observe que o cálculo da direção de Newton para (2.10) não é trivial, pois a equação $XS = \mu I$ não é linear e além disso, precisamos encontrar uma matriz simétrica X como solução. A tabela abaixo descreve sucintamente o algoritmo de pontos interiores primal-dual para SDP:

Tabela 2.1: Algoritmo de Pontos Interiores Primal-Dual para SDP

Passo 1	Escolha um critério de parada
Passo 2	Seja (X^0, y^0, S^0) um ponto inicial tal que $X^0 > 0$ e $S^0 > 0$
Passo 3	Faça $(X, y, S) = (X^0, y^0, S^0)$
Passo 4	Enquanto a iteração atual (X, y, S) não satisfaz o critério de parada
Passo 5	Escolha uma direção de busca (dX, dy, dS)
Passo 6	Escolha um tamanho do passo α_p e um tamanho do passo α_d tal que $X + \alpha_p dX > 0$ e $S + \alpha_d dS > 0$
Passo 7	$X = X + \alpha_p dX$, $(y, S) = (y, S) + \alpha_d(dy, dS)$.

2.6 Um breve histórico das LMIs na Teoria de Controle

A história das LMIs, na análise de sistemas dinâmicos, começou em 1890, quando Lyapunov mostrou que o sistema de equações diferenciais ordinárias

$$\dot{x}(t) = Ax(t) \tag{2.11}$$

é assintoticamente estável, isto é, todas suas trajetórias convergem para zero, se e somente se, existe uma matriz definida positiva P que deve satisfazer a seguinte desigualdade:

$$A^T P + P A < 0 \tag{2.12}$$

A desigualdade $P > 0$ bem como (2.12) constituem um tipo especial de LMI. Lyapunov também mostrou que essa LMI pode ser resolvida analiticamente por um conjunto de equações lineares algébricas, tomando-se qualquer $Q = Q^T > 0$ e, então resolvendo o sistema linear

$$A^T P + P A = -Q$$

para a matriz P , que é garantida ser definida positiva, se o sistema (2.11) for estável.

Nos anos 40 Lure, Postnikov e outros [33] aplicaram os métodos de Lyapunov a alguns problemas práticos na engenharia de controle. Especificamente, foi estudado o problema de estabilidade de um sistema de controle com uma não-linearidade no atuador. As desigualdades polinomiais geradas no trabalho de Lure apresentavam a forma de LMI e foram resolvidas analiticamente, limitando sua aplicação a sistemas de primeira, segunda ou terceira ordem.

Posteriormente, nos anos 60, Yakubovich [65, 66, 67], Popov [47], Kalman [24, 25] e outros trataram a solução das LMIs que surgiram no problema de Lure na forma gráfica, usando o que chamamos agora de Lema de Kalman-Yakubovich-Popov. Isso resultou no célebre critério de Popov, no critério do círculo, no critério de Tsykin e nas muitas outras variações. Embora esses critérios pudessem ser aplicados a sistemas de uma ordem mais elevada, eles não foram estendidos a sistemas com mais de uma não-linearidade.

Na década de 60, diversos pesquisadores trataram a solução das LMIs que apareceram no problema de Lure na forma gráfica, estendendo, desse modo, a análise de estabilidade para sistemas de ordem mais elevada.

Em 1971, Willems [58] mostrou que a seguinte LMI:

$$\begin{bmatrix} A^T P + P A + Q & P B + C^T \\ B^T P + C & R \end{bmatrix} \geq 0$$

poderia ser resolvida não somente por meio de gráficos, mas também resolvendo a seguinte equação algébrica de Riccati (ARE):

$$A^T P + P A - (P B + C^T) R^{-1} (B^T P + C) + Q = 0$$

Em 1971 Willems [58] comentou o seguinte: "A importância básica da LMI apresentada em (2.12) parece ser muito desprezada. Seria interessante verificar se esta abordagem pode ou não, ser explorada por algoritmos computacionais."

A sugestão acima dada por Willems fez prever o capítulo seguinte da história das LMIs. A descoberta principal seguinte era a simples observação que as LMIs que aparecem em sistemas da teoria de controle poderiam ser formuladas como um problema de Otimização Convexa, conseqüentemente essas LMIs possuem solução computacional. Essa observação simples sublinha o fato que, embora não possamos resolver analiticamente muitas dessas LMIs, podemos resolvê-las numericamente de um modo confiável. Tal observação foi feita pela primeira vez, por Pyatnitskii e Skodinskii [46], pesquisadores que reduziram as LMIs surgidas no problema de Lure a um problema de Otimização Convexa. Esse problema foi resolvido usando o algoritmo do elipsóide.

Esses teóricos foram os primeiros a formular a busca da função de Lyapunov como um problema de otimização convexa. Apesar de o algoritmo de elipsóide resolver o problema de Otimização Convexa em tempo polinomial (isto é, a complexidade do problema aumenta de modo polinomial, quando o tamanho do problema aumenta), este algoritmo não era eficiente na prática.

O capítulo final na história das LMIs é bastante importante, pois elas ganharam grande popularidade na Teoria de Controle, devido ao desenvolvimento de poderosos e eficientes métodos de pontos interiores. O início desta popularidade se deu em 1984, quando N. Karmarkar [26] introduziu um novo algoritmo de programação linear que resolvia problemas lineares em tempo polinomial e, em contraste com o método do elipsóide, é também muito eficiente na prática. Pensava-se que o método Simplex era o melhor método para resolver programas lineares durante anos. Embora fosse como algoritmo com complexidade combinatória, era muito mais eficiente que o algoritmo de elipsóide. O algoritmo de Karmarkar trouxe um grande impacto na literatura da programação matemática, que teve seu espaço até mesmo na primeira página do New York Times. O trabalho de Karmarkar induziu uma quantidade enorme de artigos na área de algoritmos de ponto interiores para programação linear.

Em 1988, Nesterov e Nemirovskii [41] desenvolveram métodos de pontos interiores aplicáveis diretamente a problemas convexos que envolvem LMIs. Embora exista muito a ser pesquisado nesta área, vários algoritmos de pontos interiores para problemas LMIs vêm sendo implementados, mostrando-se extremamente eficientes.

2.7 O que são LMIs?

Uma *desigualdade matricial linear* é uma expressão da forma

$$F(x) := F_0 + x_1 F_1 + \cdots + x_m F_m \geq 0 \quad (2.13)$$

onde

- $x = (x_1, \dots, x_m)$ é um vetor de componentes reais.
- F_0, \dots, F_m são matrizes simétricas, isto é, $F_i = F_i^T \in R^{n \times n}, i = 0, \dots, m$ para algum $n \in Z_+$.
- a desigualdade $F(x) \geq 0$ em (2.13) significa que F é uma matriz semidefinida positiva, isto é, $x^T F(x) x > 0$ para todo $x \in R^n, x \neq 0$.

Definição 6. *Desigualdade matricial linear*

Uma *desigualdade matricial linear (LMI)* é uma desigualdade

$$F(x) \geq 0 \quad (2.14)$$

onde F é uma função afim de um espaço vetorial de dimensão finita V no conjunto das matrizes reais simétricas $S^n = \{M | M = M^T \in R^{n \times n}, n > 0\}$, isto é,

$$F(\alpha x + (1 - \alpha)y) = \alpha F(x) + (1 - \alpha)F(y), \forall x, y \in V, \alpha \in R.$$

Um nome mais apropriado para este tipo de desigualdade seria desigualdades matriciais afins. Contudo, adotamos aqui o nome tradicional.

Uma propriedade importante das LMIs é que o conjunto factível é convexo.

Teorema 16. *O conjunto*

$$C = \{x \in R^n | F(x) \geq 0\}$$

é convexo.

Demonstração. Sejam x e y dois vetores tais que $F(x) \geq 0$ e $F(y) \geq 0$ e seja $\lambda \in [0, 1]$.

Então

$$\begin{aligned}
 F(\lambda x + (1 - \lambda)y) &= F_0 + \sum_{i=1}^m (\lambda x_i + (1 - \lambda)y_i)F_i \\
 &= \lambda F_0 + (1 - \lambda)F_0 + \lambda \sum_{i=1}^m x_i F_i + (1 - \lambda) \sum_{i=1}^m y_i F_i \\
 &= \lambda F(x) + (1 - \lambda)F(y) \\
 &\geq 0.
 \end{aligned}$$

□

LMIs podem representar muitas restrições convexas, por exemplo, desigualdades lineares, desigualdades quadráticas convexas e desigualdades de norma matricial. Adquirimos um modo diferente para caracterizar LMIs, reconhecendo o fato de que uma matriz simétrica é semidefinida positiva, se e somente se seus menores principais são não negativos, (Teorema (21), Apêndice A). Como os menores principais são polinômios da variável x , (2.14) é equivalente a um conjunto de desigualdades polinomiais. Isso é mostrado no exemplo seguinte.

Exemplo 2. *Considere a LMI*

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + x \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix} \geq 0$$

\iff

$$\begin{bmatrix} x - 1 & x - y \\ x - y & y - 1 \end{bmatrix} \geq 0$$

isto é equivalente a

$$x - 1 \geq 0$$

$$y - 1 \geq 0$$

$$\begin{vmatrix} x-1 & x-y \\ x-y & y-1 \end{vmatrix} = -x^2 - y^2 + 3xy - x - y + 1 \geq 0.$$

Matrizes como variáveis

Muitas vezes é conveniente especificar LMIs usando as matrizes como variáveis. Contudo que as variáveis apareçam como combinação afim, sempre é possível transformá-las na forma (2.13). Esse passo pode ser muito tedioso, é feito, então, através de softwares.

Exemplo 3. *Considere a LMI*

$$(A + I) \geq 0$$

onde A é a variável matricial simétrica. Podemos reescrevê-la como

$$\begin{aligned} & \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \geq 0. \\ & \iff \\ & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \geq 0. \end{aligned}$$

Desse modo, é possível escrever a LMI com variável matricial na forma padrão (2.13) cujos componentes básicos são dados por

$$F_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, F_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, F_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Várias Restrições LMI

Se temos várias LMIs, sempre é possível reescrevê-las como uma única LMI, usando a propriedade: uma matriz bloco diagonal é uma matriz semidefinida positiva, se e somente se cada bloco é uma matriz semidefinida positiva. Por exemplo, as restrições, $F(x) \geq 0$ e $G(x) \geq 0$ podem ser reescritas como:

$$\begin{bmatrix} F(x) & 0 \\ 0 & G(x) \end{bmatrix} \geq 0.$$

2.7.1 Complemento de Schur

Definição 7. Diz-se que uma matriz B é congruente com uma matriz A , se existe uma matriz P não-singular tal que $B = P^T A P$.

Teorema 17. Se A e B são congruentes então $A > 0$ se e somente se $B > 0$.

Demonstração. Se $A > 0 \Leftrightarrow x^T A x > 0, \forall x \in R^n, x \neq 0$.

Como A e B são congruentes, existe P não-singular tal que $B = P^T A P$. Portanto, usando o fato que P é não-singular, para todo $x \neq 0$, o vetor $y := P^{-1}x \neq 0$ e $A > 0 \Leftrightarrow x^T A x = y^T P^T A P y = y^T B y > 0 \Leftrightarrow B > 0$. \square

Uma ferramenta útil para reformularmos LMIs são os complementos de Schur.

Seja $M := \begin{bmatrix} S & G^T \\ G & R \end{bmatrix}$ onde R é não-singular, então o termo $S - G^T R^{-1} G$

é chamado o *Complemento de Schur* de S em M . O Complemento de Schur surge em muitos contextos diferentes, e nesta dissertação ele é usado para representar algumas restrições não lineares convexas como LMIs. Essa é a consequência do teorema seguinte.

Teorema 18. *Sejam S e R matrizes simétricas e R inversível. Então*

$$\begin{bmatrix} S & G^T \\ G & R \end{bmatrix} \geq 0 \quad (2.15)$$

se e somente se

$$R > 0, \quad S - G^T R^{-1} G \geq 0 \quad (2.16)$$

Demonstração. A Matriz $\begin{bmatrix} S & G^T \\ G & R \end{bmatrix}$ é congruente a $\begin{bmatrix} S - G^T R^{-1} G & 0 \\ 0 & R \end{bmatrix}$

Com efeito:

$$\begin{aligned} & \begin{bmatrix} S & G^T \\ G & R \end{bmatrix} \geq 0 \\ & \iff \\ & \begin{bmatrix} I & -G^T R^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} S & G^T \\ G & R \end{bmatrix} \begin{bmatrix} I & 0 \\ -R^{-1} G & I \end{bmatrix} \geq 0 \\ & \iff \\ & \begin{bmatrix} S - G^T R^{-1} G & 0 \\ 0 & R \end{bmatrix} \geq 0 \\ & \iff \\ & \begin{matrix} S - G^T R^{-1} G & \geq & 0 \\ R & \geq & 0 \end{matrix} \end{aligned}$$

Como R é inversível, então $R > 0$. □

Exemplo 4. *Uma restrição da norma Euclideana de uma variável $x \in R^n$ pode ser colocada no formato LMI usando o complemento de Schur. Considere a restrição $x^T x \leq \gamma^2$, onde γ é a restrição em x . Isto é equivalente às desigualdades em (2.16) se $G = x, S = \gamma^2$ e $R = I_n$. A LMI em x torna-se*

$$\begin{bmatrix} \gamma^2 & x^T \\ x & I_n \end{bmatrix} \geq 0$$

Comentário 1. Note que é essencial que as desigualdades sejam da forma (2.16). Se mudamos o sinal da desigualdade em (2.16) para

$$R < 0, \quad (S - G^T R^{-1} G) \geq 0 \quad (2.17)$$

não podemos reescrevê-las como uma LMI. Isso devido ao fato de o conjunto factível (2.17) não ser sempre convexo e o conjunto factível de uma LMI é sempre um conjunto convexo. O exemplo seguinte mostra um caso onde o conjunto factível de (2.17) não é convexo.

Exemplo 5. Se $S = -1$ e $R = -1$, então G é uma variável escalar. A desigualdade em (2.17) torna-se,

$$G^2 \geq 1,$$

cujos conjunto factível $\{G \mid G \leq -1 \text{ ou } G \geq 1\}$ não é conjunto convexo.

Exemplo 6. Exemplos de LMIs no Plano

- $y > x \iff y - x > 0$
- $y > x^2 \iff \begin{bmatrix} 1 & x \\ x & y \end{bmatrix} > 0$
- $x^2 + y^2 < 1 \iff \begin{bmatrix} 1 & x & y \\ x & 1 & 0 \\ y & 0 & 1 \end{bmatrix} > 0$

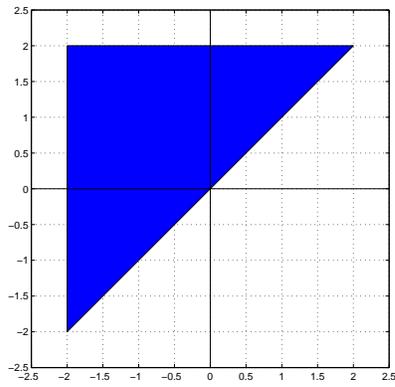


Figura 2.1: LMI no plano: $y - x > 0$

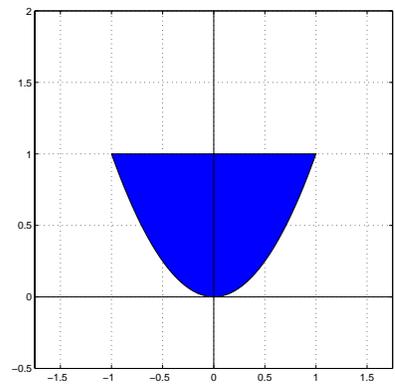


Figura 2.2: LMI no plano: $y > x^2$

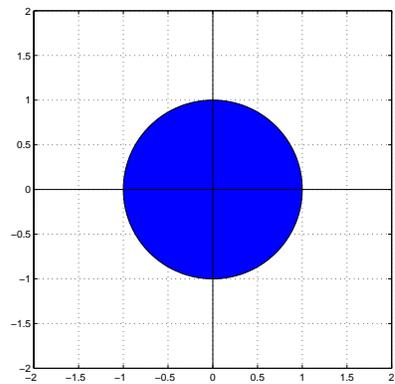


Figura 2.3: LMI no plano: $x^2 + y^2 < 1$

2.7.2 Problemas da Programação Linear Expressos em LMIs

Um problema de programação linear é da forma

$$\min c^T x \quad \text{sujeito a} \quad Ax \leq b$$

onde A é uma matriz $n \times m$. Definindo

$$F_0 = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{bmatrix} \quad \text{e} \quad F_i = \begin{bmatrix} -A_{1,i} & 0 & \cdots & 0 \\ 0 & -A_{2,i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -A_{n,i} \end{bmatrix}$$

é possível expressar as restrições por uma LMI da forma

$$F(x) \geq 0$$

2.7.3 Problemas da Programação Quadrática Expressos em LMIs

É também possível expressar um problema de programação quadrática como uma LMI. Senão vejamos: Considere o problema

$$\min \frac{1}{2} x^T H x + c^T x \quad \text{sujeito a} \quad Ax \leq b$$

onde H é simétrica e definida positiva. Isto é equivalente ao problema

$$\min t \quad \text{s.a.} \quad t \geq \frac{1}{2} x^T H x + c^T x \quad \text{e} \quad Ax \leq b$$

Uma restrição quadrática convexa $(Ax + b)^T (Ax + b) - c^T x - d \leq 0$ com $x \in R^k$, pode ser escrita como

$$\begin{bmatrix} I & Ax + b \\ (Ax + b)^T & c^T x + d \end{bmatrix} \geq 0$$

Esta restrição pode ser expressa como uma LMI: $F(x) = F_0 + x_1 F_1 + \cdots + x_k F_k \geq 0$,

cujos componentes básicos são dados por

$$F_0 = \begin{bmatrix} I & b \\ b^T & d \end{bmatrix}, F_i = \begin{bmatrix} 0 & a_i \\ a_i^T & c_i \end{bmatrix}, i = 1, \dots, k,$$

onde $A = [a_1 \cdots a_k]$. Portanto, um programa convexo quadrático com restrições quadráticas (QCQP)

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.a} \quad & f_i(x) \leq 0 \quad i = 1, \dots, L. \end{aligned} \tag{2.18}$$

onde cada f_i é uma função quadrática convexa $f_i(x) = (A_i x + b)^T (A_i x + b) - c_i^T x - d_i$, pode ser escrito como um SDP (em x e t):

$$\begin{aligned} \min \quad & t \\ \text{s.a} \quad & \begin{bmatrix} I & A_0 x + b_0 \\ (A_0 x + b_0)^T & c_0^T x + d_0 + t \end{bmatrix} \geq 0 \\ & \begin{bmatrix} I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \geq 0 \end{aligned}$$

que é um programa semidefinido com variáveis $x \in R^k$ e $t \in R$. Este programa semidefinido tem dimensões $m = k + 1$ e $n = n_0 + \dots + n_L$, onde $A_i \in R^{n_i \times k}$.

Comentário 2. *É interessante observar que problemas QCQPs podem ser escritos como problemas SDP, mas essa formulação é menos eficiente, especialmente quando as matrizes A_i têm posto muito grande. Um método de pontos interiores mais eficiente para QCQPs usa a formulação de Nesterov e Nemirovski [41] como um problema do cone de segunda ordem.*

2.8 Comentário Final

Neste capítulo, apresentamos a definição e principais propriedades da programação SDP. Em particular, estudamos uma estrutura matemática chamada LMI devido ao grande número de problemas que surgem na Teoria de Controle. Além da simplicidade e consistência algébrica, uma das principais vantagens dessa metodologia é a disponibilidade de algoritmos computacionais eficientes, derivados dos métodos de pontos interiores para programação convexa. Atualmente, as LMIs estão se tornando ferramentas básicas na Teoria de Controle, assim como as equações de Riccati se tornaram ferramentas básicas nos anos sessenta.

Capítulo 3

Controle Ótimo Via Programação Semidefinida

3.1 Formulação do Problema de Controle Ótimo de Sistemas Dinâmicos Discretos

Considere um sistema dinâmico discreto, para $k \in K = \{0, \dots, T - 1\}$ dado por:

$$x(k + 1) = Ax(k) + Bu(k), \quad x(0) = \xi \quad (3.1)$$

$$y(k) = Cx(k) \quad (3.2)$$

onde $x \in R^n$, $u \in R^m$ e $y \in R^l$ são, respectivamente, os vetores de estado, de controle e de saída. A , B e C são matrizes de dimensões $n \times n$, $n \times m$ e $l \times n$, respectivamente.

O índice de desempenho associado ao sistema (3.1) - (3.2) é representado por:

$$J = \frac{1}{2} \|x(T)\|_S^2 + \frac{1}{2} \sum_{k=0}^{T-1} (\|x(k)\|_Q^2 + \|u(k)\|_R^2) \quad (3.3)$$

onde S e Q são matrizes semidefinidas positivas com dimensões $n \times n$, R é uma matriz definida positiva com dimensão $m \times m$ e $\|x(k)\|_Q^2 := x^T(k)Qx(k)$.

O problema de controle ótimo é gerar uma seqüência de vetores de controle $u(k)$ tal que (3.3) é minimizado sujeito a (3.1), (3.2) e a um conjunto de restrições de desigualdades:

$$\begin{cases} x_{min} \leq x(k) \leq x_{max}, & k = 1, \dots, T \\ u_{min} \leq u(k) \leq u_{max}, & k = 0, \dots, T - 1 \end{cases} \quad (3.4)$$

3.2 Metodologia de Otimização Multinível

A idéia da metodologia de otimização multinível proposta inicia-se com a modificação do problema descrito na seção anterior, através da restrição do espaço de controle e da alteração da função objetivo (3.3) para uma função objetivo aumentada, via método de Lagrange, transformando o problema original pela decomposição em $T + 1$ problemas de otimização paramétrica irrestritos para o primeiro nível da estrutura hierárquica. Consideremos a seguinte função objetivo aumentada:

$$L(x, u, p^*, k) = \frac{1}{2} \|x(T)\|_S^2 + \sum_{k=0}^{T-1} \{H^k(x(k), u(k), p^*) - p^{*T}(k-1)x(k)\} \quad (3.5)$$

onde H^k é a função Hamiltoniana é definida por:

$$H^k = \frac{1}{2} [\|x(k)\|_Q^2 + \|u(k)\|_R^2] + p^{*T}(k)(Ax(k) + Bu(k)) \quad (3.6)$$

e $p^*(k)$ é a seqüência de variáveis duais fixadas, com $p(-1) = 0$ e $p(k) = 0$ para $k > T$.

De (3.5) e (3.6), temos para o primeiro nível os seguintes $T + 1$ problemas de controle ótimo:

1. Para $k = 0$,

$$\begin{aligned} \min_{u(0)} \quad & H^k(x(0), u(0), p^*(0)) \\ \text{s.a} \quad & x(0) = \xi \\ & u_{\min} \leq u(0) \leq u_{\max} \end{aligned} \quad (3.7)$$

2. Para $k = 1, \dots, T - 1$,

$$\begin{aligned} \min_{u(k), x(k)} \quad & H^k(x(k), u(k), p^*(k)) - p^*(k-1)x(k) \\ \text{s.a} \quad & x_{\min} \leq x(k) \leq x_{\max} \\ & u_{\min} \leq u(k) \leq u_{\max} \end{aligned} \quad (3.8)$$

3. Para $k = T$,

$$\begin{aligned} \min_{x(T)} \quad & x^T(T)Qx(T) - p^{*T}(T-1)x(T) \\ \text{s.a} \quad & x_{\min} \leq x(T) \leq x_{\max} \end{aligned} \quad (3.9)$$

Existem vários métodos para resolver os subproblemas no nível inferior. Basicamente, esses métodos iniciam-se com a solução das seguintes condições de otimalidade:

$$x(k+1) = \frac{\partial H^k}{\partial p^*(k)} = Ax(k) + Bu(k) \quad (3.10)$$

$$p^*(k-1) = \frac{\partial H^k}{\partial x(k)} = Qx(k) + A^T p^*(k), \quad k = 1, \dots, T-1 \quad (3.11)$$

$$0 = \frac{\partial H^k}{\partial u(k)} = Ru(k) + B^T p^*(k), \quad k = 0, \dots, T-1 \quad (3.12)$$

$$p^*(T-1) = Qx(T) \quad (3.13)$$

Como o problema é convexo, essas condições são necessárias e suficientes.

Com esse problema de decomposição paramétrica e, devido à inexistência de procedimentos de recorrência, os cálculos de $x(k)$ e $u(k)$ em (3.11) e (3.12) dependem

somente de $p^*(k)$, previamente fixado pelo nível superior (teorema da dualidade). O procedimento computacional de cálculo de $u(k)$ e $x(k)$, $k \in K$ pode ser calculado, a partir das seguintes equações:

$$u(k) = \begin{cases} u_{\max} & \text{se } -R^{-1}B^T p(k) \geq u_{\max} \\ -R^{-1}B^T p(k) & \text{se } u_{\min} < -R^{-1}B^T p(k) < u_{\max} \\ u_{\min} & \text{se } -R^{-1}B^T p(k) \leq u_{\min} \end{cases} \quad (3.14)$$

$$x(k) = \begin{cases} x_{\max} & \text{se } Q^{-1}\{p(k-1) - A^T p(k)\} \geq x_{\max} \\ Q^{-1}\{p(k-1) - A^T p(k)\} & \text{se } x_{\min} < Q^{-1}\{p(k-1) - A^T p(k)\} < x_{\max} \\ x_{\min} & \text{se } Q^{-1}\{p(k-1) - A^T p(k)\} \leq x_{\min} \end{cases} \quad (3.15)$$

$$x(T) = Q^{-1}p(T-1) \quad (3.16)$$

3.3 Estratégia de Coordenação

Um aspecto importante dos problemas de Otimização Convexa restritos é que tais problemas podem ser transformados em problemas duais os quais, em muitos casos, são mais fáceis de resolver. Para considerar o problema de otimização dual do nível superior para determinar $p^*(k)$, maximizamos a seguinte função dual $L(p(k))$:

$$\begin{aligned} \max_{p(k)} L(p(k)) = \max_{p(k)} \min_{u(k), x(k)} \sum_{k=0}^{T-1} & \left[\frac{1}{2} (\|x(k)\|_Q^2 + \|u(k)\|_R^2) \right. \\ & \left. + p^T(k)(Ax(k) + Bu(k)) - p^T(k-1)x(k) \right] + \frac{1}{2} \|x(T)\|_Q^2 - p^T(T-1)x(T) \end{aligned} \quad (3.17)$$

$$\begin{aligned}
s.a \quad p(k-1) &= \frac{\partial H^k}{\partial x^*(k)}, \\
\frac{\partial H^k}{\partial u^*(k)} &= 0, \\
x(0) &= \xi, \\
p(T-1) &= Qx(T)
\end{aligned} \tag{3.18}$$

A solução do problema dual pode ser obtida numericamente por técnicas do tipo gradiente. O algoritmo no nível superior é simples, no entanto envolve uma grande quantidade de cálculos, implicando em interações intensas entre a coordenação e os subproblemas no nível inferior. A Tabela (1.2) descreve sucintamente o algoritmo de otimização multinível.

3.4 Reformulação do Algoritmo do Nível Superior

Para reduzir os custos computacionais do processo de coordenação, propomos, dada a concavidade estrita da função dual $L(p(k))$, reformular o problema dual dinâmico para o seguinte problema de otimização estática:

Proposição 1. *Dada a concavidade da função dual $L(p(k))$, o problema dual dinâmico pode ser resolvido através do seguinte problema SDP:*

$$\begin{aligned}
\min \quad & t \\
s.a \quad & \begin{bmatrix} \tilde{H} & \frac{1}{\sqrt{2}}\tilde{H}p \\ \frac{1}{\sqrt{2}}p^T\tilde{H} & -p^T\tilde{b} - \tilde{c} + t \end{bmatrix} \geq 0
\end{aligned} \tag{3.19}$$

onde \tilde{H} é simétrica definida positiva com dimensão $nT \times nT$, \tilde{b} é um vetor $nT \times 1$, \tilde{c}

e t são escalares e p é um vetor $\in R^{nT}$ dado por

$$p = \begin{bmatrix} p^T(0) & \cdots & p^T(T-1) \end{bmatrix}^T.$$

Demonstração. Para a obtenção do problema (3.19), iniciamos com a substituição das restrições dadas pelas equações (3.18) na função $L(p(k))$. Reescrevendo essas restrições como:

$$\begin{aligned} x(0) &= \xi \\ x(k) &= -Q^{-1}(-p(k-1) + A^T p(k)) \\ x^*(T) &= Q^{-1}p(T-1) \\ u^*(k) &= -R^{-1}B^T p(k) \end{aligned}$$

e substituindo-as em (3.17), obtemos, após a realização de algumas manipulações algébricas:

$$\begin{aligned} L(p(k)) &= -\frac{1}{2}p^T(T-1)Q^{-1}p(T-1) + \frac{1}{2}\xi^T Q\xi - \frac{1}{2}p^T(0)[BR^{-1}B]p(0) + p^T(0)(A\xi) \\ &\quad + \sum_{k=1}^{T-1} \left\{ -\frac{1}{2}p^T(k-1)Q^{-1}p(k-1) + \frac{1}{2}p^T(k)[AQ^{-1}]p(k-1) \right. \\ &\quad \left. - \frac{1}{2}p^T(k)[AQ^{-1}A^T + BR^{-1}B]p(k) \right. \\ &\quad \left. + \frac{1}{2}p^*(k-1)Q^{-1}A^T p(k) \right\} \end{aligned}$$

Desenvolvendo o somatório para $k \in K$, obtemos $L(p) = \frac{1}{2}p^T \tilde{H} p + p^T \tilde{b} + \tilde{c}$, onde \tilde{H} é uma matriz bloco tridiagonal, simétrica e definida positiva com dimensão $nT \times nT$ e blocos matriciais com dimensões $n \times n$. Essa matriz é dada por:

$$\tilde{H} = \begin{bmatrix} V & -Q^{-1}A^T & 0 & \cdots & 0 \\ -AQ^{-1} & V + AQ^{-1}A^T & -Q^{-1}A^T & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & \cdots & -Q^{-1}A^T \\ 0 & 0 & 0 & -A^T Q^{-1} & V + AQ^{-1}A^T \end{bmatrix}$$

onde $V = BR^{-1}B^T + Q^{-1}$, $\tilde{b} = \begin{bmatrix} -A\xi & 0 & \cdots & 0 \end{bmatrix}^T$, $p = \begin{bmatrix} p^T(0) & \cdots & p^T(T-1) \end{bmatrix}^T$ e $\tilde{c} = -\frac{1}{2}\xi^T Q\xi$.

Como o problema $\max L(p)$ é equivalente a $\min -L(p)$, temos o problema convexo quadrático equivalente a:

$$\min t \quad s.a \quad \frac{1}{2}p^T \tilde{H}p + p^T \tilde{b} + \tilde{c} \leq t \quad (3.20)$$

A partir do complemento de Schur, representamos a restrição quadrática convexa do problema (3.20) como uma LMI: $F(p) = F_0 + p_1 F_1 + \dots + p_k F_k \geq 0$, cujos componentes básicos são dados por

$$F_0 = \begin{bmatrix} \tilde{H} & 0 \\ 0 & -\tilde{c} + t \end{bmatrix}, \quad F_i = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}\tilde{H}_i \\ \frac{1}{\sqrt{2}}\tilde{H}_i^T & -\tilde{b}_i \end{bmatrix}, i = 1, \dots, k,$$

O que prova a Proposição 1. □

O método de Newton [3, 32] é eficiente para resolver problemas de otimização paramétrica, porém requer a inversa da matriz Hessiana \tilde{H} com dimensão $nT \times nT$, o que pode resultar em instabilidade numérica. Já os métodos do tipo gradiente conjugado são mais usados para a coordenação proposta [3], apesar de o número de iterações desses métodos ser igual à dimensão de \tilde{H} , ou seja, nT . Além disso, o método do gradiente conjugado pode ser intoleravelmente lento em problemas de grande porte, se o número de condição da matriz é grande [3]. No entanto, temos verificado que a dimensão da matriz \tilde{H} , proveniente de discretização, aumenta quando o número de pontos cresce, tendo como consequência o aumento do número de condição de \tilde{H} .

Comentário : *É interessante observar que problemas QCQPs podem ser escritos como problemas SDPs, os quais podem ser resolvidos eficazmente via métodos de pontos interiores. Tais métodos têm complexidade polinomial para o pior caso em relação ao tamanho do problema. No entanto, esta formulação SDP é menos eficiente, especialmente quando as matrizes \tilde{H}_i têm posto muito grande [5].*

3.5 Modificação da Estratégia de Coordenação

Nesta seção, propomos uma estratégia para regularizar o problema definido em (3.19) através dos seguintes procedimentos:

Procedimento 1 :

Construir um problema equivalente ao original de tal forma que a matriz \tilde{H} tenha quase dominância bloco diagonal [30].

Seja o problema modificado:

$$\begin{aligned} \min \quad & t & (3.21) \\ \text{s.a} \quad & \begin{bmatrix} \tilde{H} & \frac{1}{\sqrt{2}}\tilde{H}p \\ \frac{1}{\sqrt{2}}p^T\tilde{H} & -p^T\tilde{b} - \tilde{c} + t \end{bmatrix} \geq 0 \\ & \begin{bmatrix} I & p \\ p^T & \Delta^2 \end{bmatrix} \geq 0 \end{aligned}$$

Esse problema é equivalente ao (TRS) (1.19) onde $\|(\tilde{H} + \alpha I)^{-1}p\| = \Delta$. Contrariamente ao (3.19), o problema modificado, que é um programa (QCQP), sempre tem solução, uma vez que as quadráticas são funções contínuas e a região factível em (3.21) é uma bola fechada, portanto, um conjunto compacto em R^{nT} .

Alguns autores utilizam (3.21) como uma forma de regularizar o problema de minimização de uma quadrática sem restrições, quando \tilde{H} é muito mal condicionada e a solução de (3.19) não é exata, por ser extremamente sensível aos erros de dados ou ao arredondamento. Por outro lado, o problema (3.21) é bem condicionado se Δ não é grande. Portanto, substituir (3.19) por (3.21) representa um certo sacrifício em termos do erro no resíduo $\|\nabla L(p)\|$, mas normalmente compensado por uma maior estabilidade numérica. A estrutura especial do problema (3.21) proporciona caracterizações dos minimizadores muito mais poderosas que no caso geral de minimização irrestrita.

Procedimento 2 :

A partir da decomposição temporal, é possível resolver sucessivamente o problema em malha aberta por intervalos, cujo objetivo é melhorar o condicionamento de \tilde{H} .

Através dos experimentos realizados, verificamos que, para T pequeno e $x(0)$ próximo da solução, o erro do resíduo $\| \nabla L(p) \|$ é reduzido, tornando a matriz

$$\begin{bmatrix} \tilde{H} & \frac{1}{\sqrt{2}}\tilde{H}p \\ \frac{1}{\sqrt{2}}p^T\tilde{H} & -p^T\tilde{b} - \tilde{c} + t \end{bmatrix}$$

melhor condicionada. Em geral, quando a matriz \tilde{H} aumenta de dimensão em função de T , seu número de condição também aumenta. Nesse sentido, é possível uma redução significativa da dimensão da matriz \tilde{H} , de nT para nt , onde $t \ll T$, diminuindo seu número de condição.

3.6 Uma Proposta para Controle em Malha Fechada

A principal desvantagem do algoritmo apresentado nas seções anteriores deve-se ao fato de que as ações obtidas são em malha aberta. Nesse caso, há a necessidade de se recalcular os controles toda vez que o estado inicial alterar e de armazenar as trajetórias de controle antes da implantação em malha aberta. Assim, particularmente, uma matriz constante de realimentação de estado para o controle

$$u(k) = Gx(k) \tag{3.22}$$

é claramente desejável. A matriz G de dimensão $m \times n$ deve ser obtida de tal forma que os autovalores da matriz

$$A_f = A + BG \tag{3.23}$$

estejam dentro de uma região específica.

Muitos métodos têm sido desenvolvidos para a solução deste problema [32]. No entanto, muitas dessas metodologias, que são aplicáveis para sistemas dinâmicos de pequeno porte, encontram dificuldades numéricas, normalmente críticas, quando os sistemas são de grande porte. Uma das razões para essas dificuldades é que pouca atenção tem sido dada à estabilidade numérica dos algoritmos e ao bom condicionamento das matrizes computadas.

Propomos, portanto, um novo algoritmo para o projeto de sistemas de controle em malha fechada através de soluções bem condicionadas e numericamente estáveis. A partir do subproblema de coordenação, a matriz G pode ser obtida para $T \rightarrow \infty$ através da maximização da função dual $L(p)$, das seqüências $x(k)$ e $u(k)$ obtidas das equações (3.11) e (3.12) e do emprego dos procedimentos 1 e 2, descritos na seção anterior.

A matriz G resulta da aplicação do procedimento 2, onde controles em malha fechada podem ser gerados, resolvendo-se sucessivamente o problema de malha aberta a cada dois instantes de tempo para todo $k \in K$. Considerando, sem perda de generalidade, o caso para $k = 0, 1$, temos o seguinte problema SDP:

$$\begin{aligned}
& \min \quad t \\
& \text{s.a} \quad M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \geq 0 \\
& \quad \left[\begin{array}{c} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} p(0) & p(1) \end{bmatrix} \end{array} \quad \begin{bmatrix} p(0) \\ p(1) \\ \Delta^2 \end{bmatrix} \right] \geq 0
\end{aligned}$$

onde

$$M_{11} = \begin{bmatrix} BR^{-1}B^T + Q^{-1} & -Q^{-1}A^T \\ -AQ^{-1} & BR^{-1}B^T + Q^{-1} + AQ^{-1}A^T \end{bmatrix}$$

$$M_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} BR^{-1}B^T + Q^{-1} & -Q^{-1}A^T \\ -AQ^{-1} & BR^{-1}B^T + Q^{-1} + AQ^{-1}A^T \end{bmatrix} \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$$

$$M_{21} = \frac{1}{\sqrt{2}} \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} BR^{-1}B^T + Q^{-1} & -Q^{-1}A^T \\ -AQ^{-1} & BR^{-1}B^T + Q^{-1} + AQ^{-1}A^T \end{bmatrix}$$

$$M_{22} = \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} A\xi \\ 0 \end{bmatrix} + \frac{1}{2}\xi^T Q \xi + t$$

Notemos que, quanto menor Δ , a matriz M tende para:

$$\begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix}$$

Portanto, com o ajuste de Δ , é possível garantir a dominância bloco diagonal da matriz M e alocar seus autovalores dentro de uma região desejada.

Os valores de $p(0)$ e $p(1)$ são obtidos resolvendo-se o problema SDP via métodos de pontos interiores (IPM). Substituindo $p(0)$ e $p(1)$ em (3.11) para $k = 1$, temos:

$$x(1) = Q^{-1}(p(0) - A^T p(1)) \quad (3.24)$$

Desta forma, podemos definir a matriz de ganho G a partir da equação (3.24).

Para isso escolhemos arbitrariamente um conjunto de vetores $\{x_0^1, x_0^2, \dots, x_0^n\}$, linearmente independente que representa os n estados iniciais, de modo que a matriz

$$X = \begin{bmatrix} x_0^1 & x_0^2 & \dots & x_0^n \end{bmatrix} \quad (3.25)$$

seja não-singular. Então, para $\bar{X} = \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^n \end{bmatrix}$ obtidas de (3.24) temos:

$$\bar{X} = A_f X \text{ ou } A_f = \bar{X} X^{-1} \quad (3.26)$$

A inversão de X não apresenta nenhum problema, até mesmo quando sistemas de grande porte são considerados, pois a inversão é feita off-line.

Comentário 3. *Observemos que a habilidade para inverter a matriz X depende basicamente da independência linear dos estados iniciais escolhidos. Em casos práticos, o projetista pode resolver o problema off-line n vezes sucessivamente para os estados iniciais:*

$$x(0) = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} \dots \\ \dots \\ \vdots \\ \dots \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \end{bmatrix}$$

Então $A_f = \bar{X}$ e não é necessário inverter X . Isso requer mais cálculo claro. Esse cálculo é realizado off-line de forma descentralizada.

Está claro que a matriz A_f depende de Δ . Portanto, podemos, por meio de Δ , alocar os autovalores da matriz A_f numa determinada região.

O próximo teorema determina as condições para a existência de G .

Teorema 19. *Dado*

$$A_f = \begin{bmatrix} x_1^1 & x_1^2 \end{bmatrix} \begin{bmatrix} x_0^1 & x_0^2 \end{bmatrix}^{-1}$$

existe G , a solução de (3.23), se e somente se

$$U_1^T(BG) = 0 \quad (3.27)$$

onde

$$B = \begin{bmatrix} U_0 & U_1 \end{bmatrix} \begin{bmatrix} Z \\ 0 \end{bmatrix} \quad (3.28)$$

com a matriz $U = \begin{bmatrix} U_0 & U_1 \end{bmatrix}$ ortogonal e Z não singular. Então, G é dado explicitamente por:

$$G = Z^{-1}U_0(A_f - A) \quad (3.29)$$

Demonstração. Consideremos que B tem posto completo. Isso implica na existência da decomposição em (3.28). De (3.23), G deve satisfazer

$$BG = A_f - A \quad (3.30)$$

e pré-multiplicando por U^T , temos as seguintes equações:

$$ZG = U_0^T(A_f - A) \quad (3.31)$$

$$0 = U_1^T(A_f - A) \quad (3.32)$$

das quais (3.27) e (3.29) seguem diretamente. \square

Comentário 4. A equação (3.30) implica que G existe se e somente se $\Re\{A_f - A\} \subset \Re U_0$, onde $\Re\{.\}$ é o espaço imagem. Para a decomposição de B em (3.28), podemos utilizar, por exemplo, a decomposição em valores singulares (SVD). Um outro método de fatorização que pode ser utilizado é a decomposição QR.

3.7 Análise e Resultados

Apresentamos alguns exemplos da aplicação da nova metodologia de otimização para controle ótimo. Os resultados obtidos foram comparados com o controle ótimo LQR via ARE .

Tabela 3.1: Algoritmo para o Regulador Linear Quadrático

Sistema:	$x(k+1) = Ax(k) + Bu(k)$
Índice de desempenho	$J = \frac{1}{2} \ x(T)\ _S^2 + \frac{1}{2} \sum_{k=0}^{T-1} (\ x(k)\ _Q^2 + \ u(k)\ _R^2)$
Controle :	
<u>Solução Clássica</u>	$S = A^T S A - A^T S B (B^T S B + R)^{-1} B^T S A + Q$ $G = (B^T S B + R)^{-1} B^T S A$ $u(k) = Gx(k)$
<u>Nova Proposta</u>	$BG = A_f - A$ $A_f = \begin{bmatrix} x_1^1 & x_1^2 \end{bmatrix} \begin{bmatrix} x_0^1 & x_0^2 \end{bmatrix}^{-1}$

Exemplo 7. Consideremos o seguinte sistema térmico dinâmico bastante estudado em [23], dado por:

$$x(k+1) = \begin{bmatrix} 0.6277 & 0.3597 \\ 0.0899 & 0.8526 \end{bmatrix} x(k) + \begin{bmatrix} 0.0251 \\ 0.1150 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} 1 & 1 \end{bmatrix} x(k)$$

onde as matrizes A e B foram obtidas a partir de um modelo contínuo, cujo período de amostragem foi $T = 0.25$.

A função objetivo a ser minimizada é:

$$J = \frac{1}{2} \sum_{k=0}^T (\|x\|_Q + \|u\|_R)$$

cujas matrizes de ponderação são dadas por:

$$Q = \begin{bmatrix} 50 & 0 \\ 0 & 10 \end{bmatrix}, \quad R = 1$$

Comentário 5. A vantagem de formular um problema com restrições LMI é a disponibilidade de algoritmos que calculam soluções factíveis para estes problemas. Existem muitas ferramentas para resolver as LMIs. Neste trabalho, empregamos o software LMISol [20] para a obtenção da variável de co-estado p .

Nas Tabelas (3.2) e (3.3), mostramos os valores das variáveis p e t , escolhendo-se uma condição inicial arbitrária (x_0) e variando o raio da bola fechada Δ . Nesse sentido, é interessante que haja uma faixa de variação de Δ que garanta o bom condicionamento da matriz M de tal forma que uma melhor determinação de G possa ser obtida.

Nas Figuras (3.1) e (3.2) são apresentadas as trajetórias de estado $x(2)$ e as trajetórias de controle, respectivamente, para vários valores de Δ usando a solução via ARE (curva vermelha) e via nova proposta. Na Figura (3.3), observamos o bom desempenho do sistema dinâmico obtido através da nova proposta, quando comparado com a solução via ARE. Conforme podemos verificar, a solução obtida através da nova proposta ($\Delta^2 = 1$ e $\Delta^2 = 0.5$) é melhor do que aquela obtida através do controle ótimo via ARE, uma vez que o sistema apresenta um ganho maior obtendo assim uma resposta bem mais rápida. Verificamos, portanto, que o projetista tem a seu dispor uma família de soluções estáveis, sendo que a escolha de uma destas dependerá das restrições físicas sobre as variáveis de estado e de controle.

Tabela 3.2: Valores de p e t

ponto inicial	Δ^2	iterações	t	p	Δ^2	iterações	t	p
$[1 \ 0]'$	1	6	-25.62	-0.9919 -0.1231 -0.0254 -0.0141	2	4	-25.86	-1.3857 -0.1619 -0.0479 -0.0207
$[1 \ 0]'$	0.5	4	-25.44	-0.6991 -0.0905 -0.0143 -0.0077	0.1	4	-25.19	-0.3125 -0.0427 -0.0035 -0.0019
$[1 \ 0]'$	10	8	-26.89	-3.1372 -0.3217 -0.2194 -0.0690	16	8	-27.36	-3.9654 -0.3870 -0.3382 -0.0864
$[1 \ 0]'$	100	6	-30.20	-9.4476 -0.7815 -1.6164 -0.1143	625	12	-35.53	-23.7212 -2.3093 -7.5462 0.1134

Tabela 3.3: Valores de p e t

ponto inicial	Δ^2	iterações	t	p	Δ^2	iterações	t	p
[0 1]'	1	6	-5.87	-0.4174	2	7	-6.21	-0.6056
				-0.9047				-1.2676
				-0.0401				-0.0789
				-0.0742				-0.1415
[0 1]'	0.5	5	-5.63	-0.2886	0.1	5	-5.28	-0.1259
				-0.6439				-0.2899
				-0.0196				-0.0041
				-0.0392				-0.0082
[0 1]'	10	8	-7.52	-1.4790	16	8	-8.09	-1.9388
				-2.7163				-3.3633
				-0.3714				-0.5768
				-0.5438				-0.7714
[0 1]'	100	10	-11.34	-5.8485	625	12	-15.64	-17.6082
				-7.2016				-12.8970
				-2.9616				-11.5836
				-2.2641				-3.7919

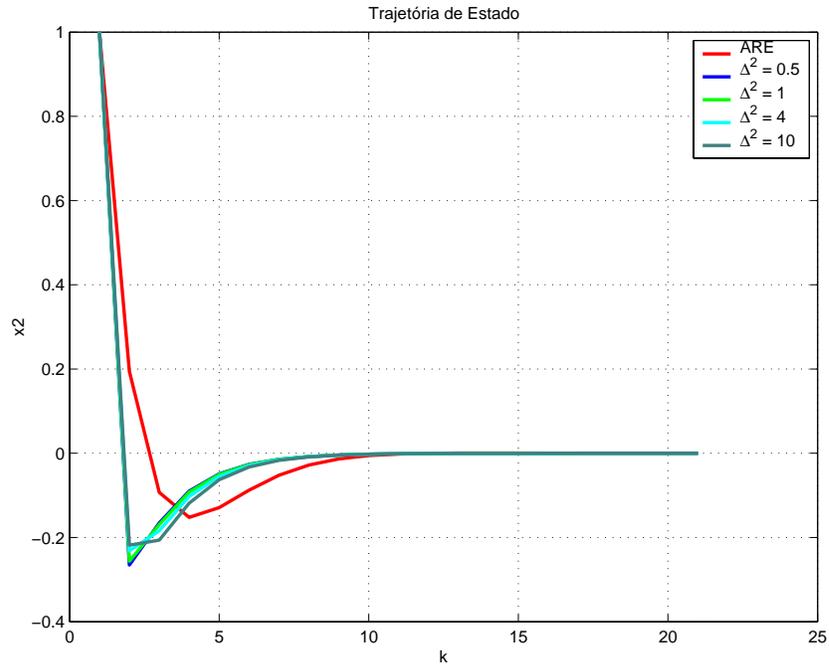


Figura 3.1: Trajetória da variável de estado $x_2(k)$

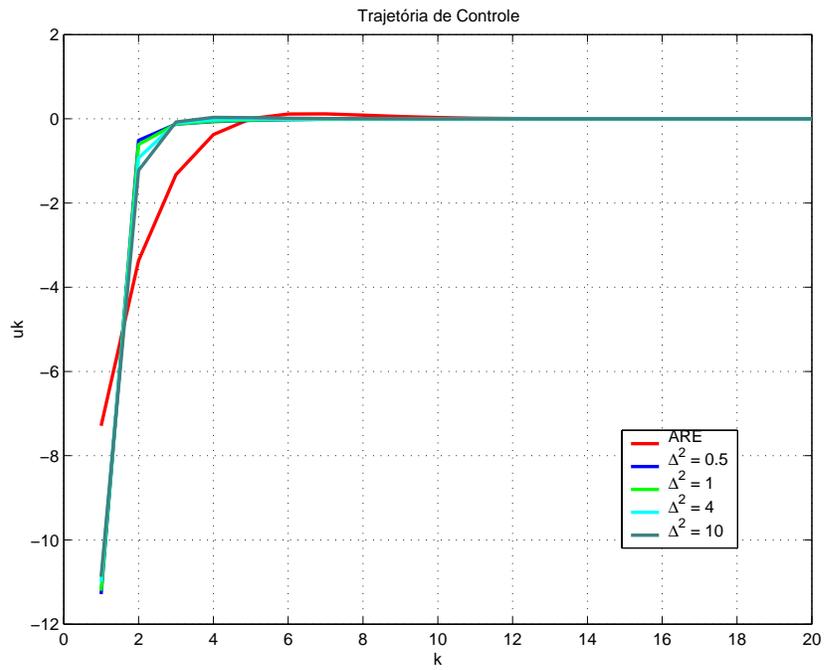


Figura 3.2: Trajetória de controle $u(k)$

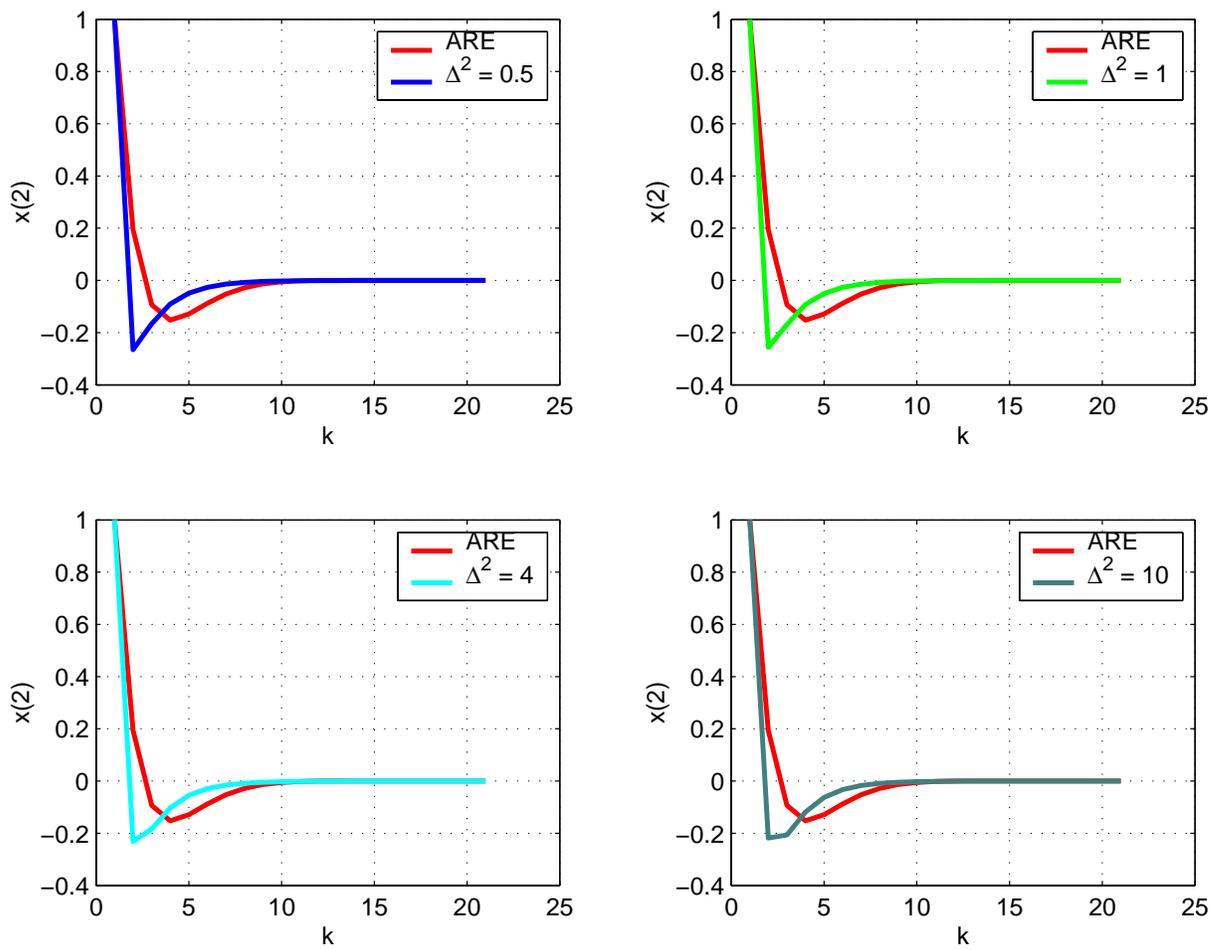


Figura 3.3: Comparação das Trajetórias de Riccati e via LMI

O comportamento da norma do ganho do controlador G , bem como os autovalores máximos e mínimos da matriz $A + BG$ via nova metodologia são mostrados na Tabela (3.4).

Tabela 3.4: Comportamento dos ganhos e autovalores do controlador

Δ^2	$\begin{pmatrix} \lambda_{min} \\ \lambda_{max} \end{pmatrix}$	$ G $	Δ^2	$\begin{pmatrix} \lambda_{min} \\ \lambda_{max} \end{pmatrix}$	$ G $
1	$\begin{pmatrix} 0.0385 \\ 0.5398 \end{pmatrix}$	7.6692	0.5	$\begin{pmatrix} 0.0270 \\ 0.5414 \end{pmatrix}$	7.7525
4	$\begin{pmatrix} 0.0796 \\ 0.5317 \end{pmatrix}$	7.3930	10	$\begin{pmatrix} 0.1269 \\ 0.5173 \end{pmatrix}$	7.1240
16	$\begin{pmatrix} 0.1597 \\ 0.5032 \end{pmatrix}$	6.9771	100	$\begin{pmatrix} 0.3275 - 0.0869i \\ 0.3275 + 0.0869i \end{pmatrix}$	7.2357

3.8 Comentário Final

Neste capítulo, apresentamos uma nova proposta de otimização para projeto de controladores para sistemas discretos no tempo. Esses controladores podem ser eficientemente obtidos com boas condições de desempenho e estabilidade, sem o emprego de soluções simétricas e definidas positivas das equações de Riccati. A nova metodologia proposta apresenta um bom potencial para tratar problemas envolvendo incertezas, ou seja, problemas de controle robusto. Foi verificado que os procedimentos 1 e 2 podem ser empregados tanto para regularizar problemas de otimização mal condicionados, como para determinar a matriz de ganho de realimentação pela proposta apresentada. Desta forma, o projetista terá ao seu dispor a escolha da melhor trajetória para os controles e para o estado, de acordo com as restrições físicas do projeto, graças à obtenção de uma família de ganhos de realimentação.

Capítulo 4

Controle Robusto Via Programação Semidefinida

4.1 Introdução

Durante a década de 80, um novo paradigma na teoria de controle de sistemas dinâmicos multivariáveis emergiu, o controle H_∞ . Esse problema de controle foi formulado inicialmente por G. Zames [68], como essencialmente um método de otimização no domínio da frequência para projetos de sistemas de controle robusto. O aspecto de robustez tornou-se o foco principal na pesquisa da comunidade de controle. Com isso, muitas técnicas para projeto de sistemas de controle multivariável surgiram, entre as quais: controle H_∞ [9, 68], síntese μ [28], QFT (Quantitative Feedback Theory) por I. Horowitz [68] e métodos baseados na teoria de Kharitonov para incerteza estruturada [10, 68]. Todas essas técnicas estão ainda sendo desenvolvidas e refinadas.

O termo H_∞ se refere ao espaço das funções de transferências próprias e estáveis. As funções de transferência próprias são as que apresentam o grau do denominador maior ou igual ao grau do numerador e as estáveis são as que apresentam pólos

estritamente no lado esquerdo do semiplano, no caso contínuo ou dentro do círculo unitário, no caso discreto. Para comparar as funções de transferências, a Teoria de Controle H_∞ emprega a norma H_∞ como medida de tamanho. No apêndice A, é feita uma revisão sucinta sobre esta norma.

Avanços significativos na Teoria de Controle H_∞ para sistemas dinâmicos discretos têm ocorrido durante a última década, [9, 10, 28]. A abordagem por espaço de estado para o controle H_∞ de sistemas lineares é desenvolvida com base na relação entre norma H_∞ e as equações de Riccati. Os resultados obtidos para sistemas contínuos são estendidos para sistemas discretos.

Muitos pesquisadores têm utilizado problemas de controle robusto H_∞ para sistemas discretos e estabelecido condições suficientes em termos de equações algébricas de Riccati modificadas.

Neste capítulo, foi desenvolvido um novo procedimento de solução para o problema de controle H_∞ via metodologia proposta no Capítulo 3. A partir do índice de desempenho derivado da norma H_∞ , obtivemos uma lei de controle em malha fechada, sem a utilização da *ARE* correspondente.

4.2 Formulação do Problema

Consideremos o seguinte sistema dinâmico, discreto no tempo, com realimentação de estado:

$$x(k+1) = Ax(k) + B_1w(k) + B_2u(k) \quad (4.1)$$

$$z(k) = C_1x(k) + D_{12}u(k) \quad (4.2)$$

onde $x(k) \in R^n$ é o estado, $w(k) \in R^m$ é a perturbação e $z(k) \in R^l$ é a variável controlada. Supondo uma lei de controle de realimentação de estados $u(k) \in R^m$,

$$u(k) = Gx(k) \quad (4.3)$$

o sistema em malha fechada torna-se:

$$x(k+1) = (A + B_2G)x(k) + B_1w(k) \quad (4.4)$$

$$z(k) = (C_1 + D_{12}G)x(k) \quad (4.5)$$

e a função de transferência de w para z é dada por:

$$T_{zw}(z) = (C_1 + D_{12}G)(zI - (A + B_2G))^{-1}B_1 \quad (4.6)$$

Para considerar a atenuação da perturbação pior possível para o sistema, a norma H_∞ da função de transferência em malha fechada da perturbação para a variável controlada $z(k)$, $\|T_{zw}\|_\infty$ é minimizada [9], isto é:

$$\min \|T_{zw}\|_\infty = \min_{z \in L_2} \max \frac{\|z\|_2}{\|w\|_2} \quad (4.7)$$

Dizemos que $w(k)$ é a pior possível no sentido de que a mesma aceita o máximo ganho de energia da entrada de $w(k)$ para a variável controlada $z(k)$.

Supondo o sistema $z = T_{zw}w$ estável, onde L_2 é o espaço formado pelas funções quadraticamente integráveis, a norma $\|T_{zw}\|_\infty$, limitada por um número γ , é reescrita no domínio do tempo da seguinte forma [32]:

$$\max_{z \in L_2} \frac{\|z\|_2}{\|w\|_2} < \gamma^2 \Leftrightarrow \|z\|_2^2 - \gamma^2 \|w\|_2^2 \leq \epsilon \|w\|_2^2, \quad \epsilon > 0 \quad (4.8)$$

Portanto, podemos definir a função objetivo a ser minimizada como:

$$J(u(k), x(k), k) = \sum_{k=0}^{\infty} (\|z(k)\|_2^2 - \gamma^2 \|w\|_2^2)$$

ou equivalentemente:

$$J(u(k), x(k), k) = \sum_{k=0}^{\infty} (z^T(k)z(k) - \gamma^2 w^T(k)w(k)) \quad (4.9)$$

A minimização da função objetivo (4.9) implica na escolha de uma entrada de controle, levando-se em consideração que a perturbação é a pior possível para o sistema.

Chamando de \mathcal{G} o conjunto dos ganhos estabilizantes, o problema de controle H_{∞} no horizonte infinito pode ser escrito na forma:

$$\{G \in \mathcal{G} : \|T_{zw}\|_{\infty} \leq \gamma\}$$

4.3 Projeto de Controladores H_{∞} via ARE

Aplicando as condições de otimalidade e definindo o Hamiltoniano como

$$H(x(k), w(k), u(k), k) = L(x(k), w(k), k) + p^T(k+1)(Ax(k) + B_1w(k) + B_2u(k))) \quad (4.10)$$

temos:

$$x(k+1) = \frac{\partial H^k}{\partial p^*(k+1)} = Ax(k) + B_1w(k) + B_2u(k) \quad (4.11)$$

$$p(k) = \frac{\partial H^k}{\partial x(k)} = C_1^T C_1 x(k) + A^T p^*(k+1), \quad k = 1, \dots, T-1 \quad (4.12)$$

$$0 = \frac{\partial H^k}{\partial u(k)} = D_{12}^T D_{12} u(k) + B_2^T p^*(k+1), \quad k = 0, \dots, T-1 \quad (4.13)$$

$$0 = \frac{\partial H^k}{\partial w(k)} = \gamma^2 w(k) + B_1^T p^*(k+1), \quad k = 0, \dots, T-1 \quad (4.14)$$

$$p^*(T) = S(T)x(T) \quad (4.15)$$

onde é considerado que $C_1^T D_{12} = 0$, $D_{12}^T D_{12} = I$, ou seja, não há ponderação cruzada entre o estado e a matriz de ponderação do controle é normalizada em (4.9). Consideramos, também, que o par (A, B_2) é controlável.

Das equações (4.13) e (4.14), obtemos:

$$w^*(k) = \gamma^{-2} B_1^T p(k+1) \quad (4.16)$$

$$u^*(k) = -B_2^T p(k+1) \quad (4.17)$$

Substituindo-se (4.16) em (4.11), temos:

$$x(k+1) = Ax(k) + (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) p(k+1) \quad (4.18)$$

Considerando que $p^*(k) = S(k)x(k)$, e substituindo essa equação em (4.12) e (4.18), resultam:

$$S(k)x(k) = C_1^T C_1 x(k) + A^T S(k+1)x(k+1) \quad (4.19)$$

$$x(k+1) = [I - (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) S(k+1)]^{-1} Ax(k) \quad (4.20)$$

Substituindo (4.20) em (4.19) obtemos:

$$S(k) = A^T S(k+1) [I - (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) S(k+1)]^{-1} A + C_1^T C_1 \quad (4.21)$$

Essa é a equação de Riccati para o problema de controle H_∞ com realimentação de estado.

Da equação (4.18), temos:

$$x(k) = A^{-1} x(k+1) - A^{-1} (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) p(k+1) \quad (4.22)$$

Substituindo (4.22) em (4.12), obtemos;

$$p(k) = C_1^T C_1 A^{-1} x(k+1) + [A^T - C_1^T C_1 A^{-1} (\gamma^{-2} B_1 B_1^T - B_2 B_2^T)] p(k+1) \quad (4.23)$$

Assim, de (4.22) e (4.23), obtemos:

$$\begin{bmatrix} x(k) \\ p(k) \end{bmatrix} = \begin{bmatrix} A^{-1} & -A^{-1} (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) \\ C_1^T C_1 A^{-1} & A^T - C_1^T C_1 A^{-1} (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) \end{bmatrix} \begin{bmatrix} x(k+1) \\ p(k+1) \end{bmatrix} \quad (4.24)$$

onde, do sistema (4.24), definimos a matriz Hamiltoniana H_a como:

$$H_a = \begin{bmatrix} A^{-1} & -A^{-1}(\gamma^{-2}B_1B_1^T - B_2B_2^T) \\ C_1^TC_1A^{-1} & A^T - C_1^TA^{-1}(\gamma^{-2}B_1B_1^T - B_2B_2^T) \end{bmatrix}$$

Realizando algumas manipulações algébricas em (4.16) e (4.17) e considerando a equação de estado do sistema sem levar em conta a perturbação, o controle é dado por:

$$u(k) = Gx(k) = -(I + B_2^T SB_2)^{-1} B_2^T SAx(k) \quad (4.25)$$

Para a perturbação no pior caso, ou seja, a perturbação que maximiza a norma da matriz função de transferência, o controle é dado por:

$$w^* = G_w x(k) = \gamma^{-2} B_1^T S(A + B_2 G)x(k) \quad (4.26)$$

onde $G = -(I + B_2^T SB_2)^{-1} B_2^T SA$ e $G_w = \gamma^{-2} B_1^T S(A + B_2 G)$ são os ganhos de realimentação de estado e da perturbação.

Algoritmo Para a Solução do Problema de Controle H_∞

A solução deste problema envolve um processo iterativo [9], conforme descrito no algoritmo da tabela a seguir.

Tabela 4.1: Algoritmo para Solução do Controle H_∞

Passo 1	Selecionar um número γ positivo
Passo 2	Testar se $\ T_{zw}\ < \gamma$ calculando os autovalores de H_a
Passo 3	Se não existirem autovalores na fronteira da circunferência unitária, diminuir γ e voltar para o passo 2. Caso contrário, aumentar γ . Este processo deve continuar até obter-se o menor γ que satisfaça a condição de que não existam autovalores sobre o círculo unitário.

4.4 Projeto de Controladores H_∞ via Otimização Convexa

Adotando a mesma terminologia utilizada no capítulo 3, consideremos o sistema dinâmico discreto, representado pelas equações: (4.1) e (4.2).

A função objetivo associada ao sistema dinâmico é dada por:

$$J = \frac{1}{2} \| z(T) \|^2 + \frac{1}{2} \sum_{k=0}^{T-1} (\| z(k) \|^2 - \gamma^2 \| w(k) \|^2) \quad (4.27)$$

O problema de controle robusto é gerar uma seqüência de vetores de controle $u(k)$ tal que (4.27) seja minimizada sujeita a (4.1) e (4.2).

O problema de minimização da função (4.27) sujeito ao sistema dinâmico será modificado através do espaço de controle e da alteração da função objetivo aumentada via método de Lagrange, conforme descrito no Capítulo 3. Considerando as restrições (4.1) e (4.2), o Lagrangeano é dado por:

$$L(z, w, p^*, k) = \sum_{k=0}^{T-1} \{ H^k(z(k), u(k), p^*(k)) - p^{*T}(k-1)x(k) \} \quad (4.28)$$

onde H^k é a função Hamiltoniana definida por:

$$H^k = \frac{1}{2} [\| z(k) \|^2 - \gamma^2 \| w(k) \|^2] + p^{*T}(k)(Ax(k) + B_1w(k) + B_2u(k)) \quad (4.29)$$

onde $p^*(k)$ é a seqüência de variáveis duais fixadas, com $p(-1) = 0$ e $p(k) = 0$ para $k > T$.

Portanto, temos para o primeiro nível, os seguintes $T + 1$ problemas de controle ótimo:

1. Para $k = 0$,

$$\begin{aligned} \min_{u(0)} \quad & H^k(z(0), u(0), p^*(0)) \\ \text{s.a} \quad & x(0) = \xi \end{aligned} \quad (4.30)$$

2. Para $k = 1, \dots, T - 1$,

$$\min_{u(k), x(k)} H^k(z(k), u(k), p^*(k)) - p^*(k-1)x(k) \quad (4.31)$$

3. Para $k = T$,

$$\min_{x(T)} z^T(T)z(T) - p^{*T}(T-1)x(T) \quad (4.32)$$

Aplicando as condições de otimalidade:

$$p^*(k-1) = \frac{\partial H^k}{\partial z(k)} = C_1^T C_1 x(k) + A^T p^*(k), \quad k = 1, \dots, T-1 \quad (4.33)$$

$$0 = \frac{\partial H^k}{\partial u(k)} = u(k) + B_2^T p^*(k), \quad k = 0, \dots, T-1 \quad (4.34)$$

$$0 = \frac{\partial H^k}{\partial w(k)} = -\gamma^2 w(k) + B_1^T p^*(k), \quad k = 0, \dots, T-1 \quad (4.35)$$

onde,

$$p^*(T-1) = Qx(T) \quad (4.36)$$

Devido à inexistência de recorrência, obtemos para o cálculo de $u(k)$ e $x(k)$ as seguintes equações:

$$u(k) = -B_2^T p(k) \quad (4.37)$$

$$x(k) = Q^{-1}\{p(k-1) - A^T p(k)\} \quad (4.38)$$

$$x(T) = Q^{-1}p(T-1) \quad (4.39)$$

Transformando este problema de otimização, convexa em $x(k)$ e $u(k)$ e cônica em $w(k)$, com restrições em um problema dual do nível superior para determinar $p^*(k)$, temos a seguinte função dual $L(p(k))$ a ser maximizada:

$$\begin{aligned} \max_{p(k)} L(p(k)) = \max_{p(k)} \min_{u(k), x(k)} \sum_{k=0}^{T-1} & \left[\frac{1}{2} (\|x(k)\|_Q^2 + \|u(k)\|^2 - \gamma^2 \|w(k)\|^2 \right. \\ & \left. + p^T(k)(Ax(k) + B_1w(k) + B_2u(k)) - p^T(k-1)x(k) \right] \\ & + \frac{1}{2} \|x(T)\|_Q^2 - p^T(T-1)x(T) \end{aligned} \quad (4.40)$$

sujeito às restrições (4.33), (4.34), (4.35) e (4.36).

Proposição 2. *O problema (4.40) pode ser reformulado, supondo a concavidade da função dual $L(p(k))$, no seguinte problema SDP:*

$$\begin{aligned} \min \quad & t \\ \text{s.a} \quad & \begin{bmatrix} \tilde{H} & \frac{1}{\sqrt{2}}\tilde{H}p \\ \frac{1}{\sqrt{2}}p^T\tilde{H} & -p^T\tilde{b} - \tilde{c} + t \end{bmatrix} \geq 0 \end{aligned} \quad (4.41)$$

onde \tilde{H} é uma matriz bloco tridiagonal, simétrica e definida positiva com dimensão $nT \times nT$, \tilde{b} é um vetor $nT \times 1$, \tilde{c} é um escalar e p é um vetor $\in \mathbb{R}^{nT}$ dado por $p = \begin{bmatrix} p^T(0) & \cdots & p^T(T-1) \end{bmatrix}^T$.

Demonstração. Para isso, basta substituir as restrições dadas em (4.33), (4.34), (4.35) e (4.36), na função $L(p(k))$, de forma a obter:

$$\begin{aligned} L(p(k)) = & -\frac{1}{2}p^T(T-1)Q^{-1}p(T-1) + \frac{1}{2}\xi^T Q \xi - \frac{1}{2}p^T(0)[B_2B_2^T - \gamma^{-2}B_1B_1^T]p(0) \\ & + p^T(0)(A\xi) + \frac{1}{2} \sum_{k=1}^{T-1} \{p^T(k-1)Q^{-1}p(k-1) + p^T(k)[AQ^{-1}]p(k-1)\} \\ & + \frac{1}{2} \sum_{k=1}^{T-1} \{-p^T(k)[AQ^{-1}A^T + B_2B_2^T - \gamma^{-2}B_1B_1^T] + p^T(k-1)Q^{-1}A^T p(k)\} \end{aligned}$$

Desenvolvendo o somatório para $k \in K$, obtemos $L(p) = \frac{1}{2}p^T\tilde{H}p + p^T\tilde{b} + \tilde{c}$, onde \tilde{H} é dada por:

$$\tilde{H} = \begin{bmatrix} V & -Q^{-1}A^T & 0 & \cdots & 0 \\ -AQ^{-1} & V + AQ^{-1}A^T & -Q^{-1}A^T & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & \ddots & -Q^{-1}A^T \\ 0 & 0 & 0 & -A^TQ^{-1} & V + AQ^{-1}A^T \end{bmatrix}$$

$$V = B_2B_2^T - \gamma^{-2}B_1B_1^T + Q^{-1}, \quad \tilde{b} = - \begin{bmatrix} A\xi & 0 & \cdots & 0 \end{bmatrix}^T \text{ e } \tilde{c} = -\frac{1}{2}\xi^TQ\xi.$$

Como o problema $\max L(p(k))$ é equivalente a $\min -L(p(k))$, temos:

$$\min_p L(p) = \frac{1}{2}p^T \tilde{H}p + p^T \tilde{b} + \tilde{c} \quad (4.42)$$

O Programa Convexo Quadrático (4.42) é equivalente ao programa SDP

$$\min t \quad s.a \quad \frac{1}{2}p^T \tilde{H}p + p^T \tilde{b} + \tilde{c} \leq t \quad (4.43)$$

A partir do complemento de Schur, representamos a restrição quadrática convexa do problema (4.41) como uma LMI: $F(p) = F_0 + p_1F_1 + \dots + p_kF_k \geq 0$, cujos componentes básicos são dados por:

$$F_0 = \begin{bmatrix} \tilde{H} & 0 \\ 0 & -\tilde{c} + t \end{bmatrix}, \quad F_i = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}\tilde{H}_i \\ \frac{1}{\sqrt{2}}\tilde{H}_i^T & -\tilde{b}_i \end{bmatrix}, i = 1, \dots, k, \quad \square$$

4.5 Modificação da Estratégia de Coordenação

Resta-nos, agora, regularizar o problema proposto em (4.42) de tal forma que a matriz \tilde{H} tenha dominância quase diagonal. Portanto, seguindo o mesmo método realizado no Capítulo 3, propomos dois procedimentos:

Procedimento 1:

Seja o problema modificado:

$$\begin{aligned}
& \min && t \\
& \text{s.a} && \begin{bmatrix} \tilde{H} & \frac{1}{\sqrt{2}}\tilde{H}p \\ \frac{1}{\sqrt{2}}p^T\tilde{H} & -p^Tb - \tilde{c} + t \end{bmatrix} \geq 0 \\
& && \begin{bmatrix} I & p \\ p^T & \Delta^2 \end{bmatrix} \geq 0.
\end{aligned} \tag{4.44}$$

Para resolver o problema (4.44), basta aplicar qualquer algoritmo que resolva o problema SDP. No nosso caso, usamos o software *LMISOL*.

Procedimento 2:

Este é o procedimento através do qual obtemos uma matriz \tilde{H} com um bom condicionamento, pois o problema é resolvido sucessivamente para cada dois intervalos de tempo. Também é aplicado para a obtenção de um ganho de realimentação constante para a solução do problema em malha fechada. Nesse caso, de (4.40), para $k = 0, 1$, temos:

$$\begin{aligned}
L(p(k)) = \frac{1}{2} \|x(k)\|_Q^2 - p^T x(k) & \sum_{k=0}^{T-1} \left[\frac{1}{2} (\|x(k)\|_Q^2 + \|u(k)\|^2 - \gamma^2 \|w(k)\|^2 \right. \\
& \left. + p^T(k)(Ax(k) + B_1w(k) + B_2u(k)) - p^T(k-1)x(k) \right] \tag{4.45}
\end{aligned}$$

onde $T = 2$, $p = [p(0), p(1)]^T$, $\tilde{b} = -[b(0), b(1)]^T$, \tilde{c} é uma constante e \tilde{H} é uma matriz semidefinida positiva.

Desenvolvendo o somatório e substituindo as restrições (4.33), (4.34), (4.35) e (4.36),

em (4.45) obtemos:

$$\begin{aligned}
L(p(k)) &= \frac{1}{2}[x^T(2)Qx(2) - p^T(1)x(2)] \\
&+ \frac{1}{2}[(x^T(0)Qx(0) + u^T(0)u(0)) - \gamma^2 w^T(0)w(0)] \\
&\quad + \frac{1}{2}[p^T(0)(Ax(0) + B_1w(0) + B_2u(0))] \\
&+ \frac{1}{2}[(x^T(1)Qx(1) + u^T(1)u(1)) - \gamma^2 w^T(1)w(1)] \\
&+ \frac{1}{2}[p^T(1)(Ax(1) + B_1w(1) + B_2u(1)) - p^T(0)x(1)]
\end{aligned} \tag{4.46}$$

Após algumas manipulações, verificamos que:

$$\begin{aligned}
L(p) &= \frac{1}{2}p^T(1)[-Q^{-1} - AQ^{-1}A^T - B_2B_2^T + \gamma^{-2}B_1B_1^T - Q^{-1}]p(1) \\
&\quad + \frac{1}{2}p^T(0)[-B_2B_2^T + \gamma^{-2}B_1B_1^T - Q^{-1}]p(0) \\
&\quad + \frac{1}{2}p^T(0)Q^{-1}A^T p(1) + p^T(1)AQ^{-1}p(0) \\
&\quad + \frac{1}{2}\xi^T Q\xi + p^T(0)A\xi
\end{aligned} \tag{4.47}$$

De (4.42) e (4.47), encontramos o seguinte resultado:

$$\tilde{H} = \begin{bmatrix} B_2B_2^T - \gamma^{-2}B_1B_1^T + Q^{-1} & -Q^{-1}A^T \\ -AQ^{-1} & B_2B_2^T - \gamma^{-2}B_1B_1^T + Q^{-1} + AQ^{-1}A^T \end{bmatrix}$$

$$\tilde{b} = \begin{bmatrix} -A\xi \\ 0 \end{bmatrix} \quad \text{e} \quad \tilde{c} = -\frac{1}{2}\xi^T Q\xi$$

Análogo ao Capítulo 3, temos o seguinte problema SDP:

$$\begin{aligned} \min \quad & t \\ \text{s.a} \quad & M = \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \geq 0 \\ & \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} p(0) \\ p(1) \end{bmatrix} \\ \begin{bmatrix} p(0) & p(1) \end{bmatrix} & \Delta^2 \end{bmatrix} \geq 0 \end{aligned}$$

onde

$$M_1 = \begin{bmatrix} B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} & -Q^{-1} A^T \\ -A Q^{-1} & B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} + A Q^{-1} A^T \end{bmatrix}$$

$$M_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} & -Q^{-1} A^T \\ -A Q^{-1} & B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} + A Q^{-1} A^T \end{bmatrix} \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$$

$$M_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} & -Q^{-1} A^T \\ -A Q^{-1} & B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} + A Q^{-1} A^T \end{bmatrix}$$

$$M_4 = \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} A\xi \\ 0 \end{bmatrix} + \frac{1}{2} \xi^T Q \xi + t$$

Notemos que, quanto menor Δ , a matriz M tende para:

$$\begin{bmatrix} M_1 & 0 \\ 0 & M_4 \end{bmatrix}$$

Portanto, com o ajuste de Δ , é possível garantir a dominância bloco diagonal da matriz M e alocar seus autovalores dentro de uma região desejada.

4.6 Proposta de Controle H_∞

Após as considerações feitas até aqui, propomos um novo algoritmo para o projeto de sistema de controle em malha fechada, partindo-se do índice de desempenho do problema de controle H_∞ , através de soluções bem condicionadas e numericamente estáveis. A partir do subproblema de coordenação, a matriz G pode ser obtida para $T \rightarrow \infty$ através da maximização da função dual $L(p)$, das seqüências $x(k)$ e $u(k)$ obtidas das equações (4.12), (4.13) e do emprego dos procedimentos 1 e 2, descritos na seção anterior.

Portanto, ajustando-se Δ , é possível alocar os autovalores de $(\tilde{H} + \alpha I)^{-1}$ dentro de uma região desejada. Dessa forma, conforme exposto no Capítulo 3, obtemos:

$$A_f = \begin{bmatrix} x_1^1 & x_1^2 \end{bmatrix} \begin{bmatrix} x_0^1 & x_0^2 \end{bmatrix}^{-1} \quad (4.48)$$

Para o cálculo do ganho G , sem levar em conta a perturbação, temos:

$$u^*(k) = Gx(k) \quad (4.49)$$

onde o ganho é encontrado resolvendo-se:

$$B_2G = A_f - A \quad (4.50)$$

A matriz função de transferência T_{zw} do sistema em malha fechada é:

$$T_{zw}(z) = (C1 + D_{12}G)(zI - A_f)^{-1}B_1 \quad (4.51)$$

Proposta de Cálculo do Ganho de Perturbação no pior Caso

Para o cálculo de G_w , vamos considerar o problema TRS formulado como no Capítulo 1, em que o sistema $(\tilde{H} + \alpha I)p^* = -\tilde{b}$ tem solução única. Nesse caso, para $k = 0, 1$, temos o sistema de equações:

$$\begin{bmatrix} \tilde{H} + \alpha I_{2 \times 2} \end{bmatrix} \begin{bmatrix} p(0) \\ p(1) \end{bmatrix} = \begin{bmatrix} A\xi \\ 0 \end{bmatrix} \quad (4.52)$$

Como visto no Capítulo 3, quanto menor Δ , a matriz $(\tilde{H} + \alpha I)$ tende para:

$$\begin{bmatrix} B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} + \alpha I_{2 \times 2} & O_{2 \times 2} \\ O_{2 \times 2} & B_2 B_2^T - \gamma^{-2} B_1 B_1^T + Q^{-1} + A Q^{-1} A^T + \alpha I_{2 \times 2} \end{bmatrix}$$

Portanto, com o ajuste de Δ , é possível alocar os autovalores de $\overline{H} = (\tilde{H} + \alpha I)^{-1}$ dentro de uma região desejada. Isso é equivalente a alocar os autovalores da matriz M . Os valores de $p(0)$ e $p(1)$ são obtidos como:

$$p(0) = \overline{H}_{11} A x(0) \quad (4.53)$$

$$p(1) = \overline{H}_{21} A x(0) \quad (4.54)$$

Para calcular o ganho G_w que gera a perturbação no pior caso,

$$w^*(k) = G_w x(k) \quad (4.55)$$

empregamos a condição de otimalidade (4.12) para $k = 0$:

$$p(0) = Q x(1) + A^T p(1) \quad (4.56)$$

Desprezando a perturbação, temos da equação (4.11) que :

$$x(1) = A x(0) + B_2 u(0) = (A + B_2 G) x(0) \quad (4.57)$$

Então, substituindo as equações (4.54) e (4.57) em (4.56) encontramos:

$$p(0) = [Q(A + B_2 G) + A^T \overline{H}_{21} A] x(0) \quad (4.58)$$

Da condição de otimalidade (4.35) temos:

$$w(0) = \gamma^{-2} B_1^T p(0) \quad (4.59)$$

Assim, o ganho da perturbação é dado por:

$$G_w = \gamma^{-2} B_1^T (Q A_f + A^T \overline{H}_{21}) A \quad (4.60)$$

4.7 Aplicações e Resultados

Apresentamos, a seguir, alguns exemplos da aplicação da nova metodologia de otimização para controle H_∞ . Os resultados obtidos foram comparados com o controle H_∞ com realimentação de estado via *ARE* não padrão.

Tabela 4.2: Algoritmo para o controle H_∞ com realimentação de estado

Sistema:	$x(k+1) = Ax(k) + B_1w(k) + B_2u(k)$
Índice de desempenho	$J(u(k), x(k), k) = \sum_{k=0}^{\infty} (\ z(k)\ _2^2 - \gamma^2 \ w\ _2^2)$
Controle :	
<u>Solução Clássica</u>	$S(k) = A^T S(k+1)[I - (\gamma^{-2}B_1B_1^T - B_2B_2^T)S(k+1)]^{-1}A + C_1^T C_1$ $G = -(I + B_2^T S B_2)^{-1}B_2^T S A$ $G_w = \gamma^{-2}B_1^T S(A + B_2G)$ $u(k) = Gx(k)$ $w(k) = G_w x(k)$
<u>Nova Proposta</u>	$A_f = \begin{bmatrix} x_1^1 & x_1^2 \end{bmatrix} \begin{bmatrix} x_0^1 & x_0^2 \end{bmatrix}^{-1}$ $B_2G = A_f - A$ $G_w = \gamma^{-2}B_1^T(QA_f + A^T\bar{H}_{21})A$

Exemplo 8. Consideremos o seguinte sistema térmico dinâmico estudado em [23], dado por:

$$x(k+1) = \begin{bmatrix} 0.6277 & 0.3597 \\ 0.0899 & 0.8526 \end{bmatrix} x(k) + \begin{bmatrix} 0.0126 \\ 0.0575 \end{bmatrix} w(k) + \begin{bmatrix} 0.0251 \\ 0.1150 \end{bmatrix} u(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k)$$

onde $z(k)$ é a variável controlada e a saída medida não aparece no modelo, visto que se trata de um problema de informação completa.

A função objetivo a ser minimizada é:

$$J(u(k), x(k), k) = \sum_{k=0}^{\infty} (\|z(k)\|_2^2 - \gamma^2 \|w\|_2^2)$$

onde

$$Q = C_1^T C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Comentário 6. Neste trabalho, empregamos o software *LMISol* [20] para a obtenção da variável de co-estado p .

Nas Tabelas (4.3) e (4.4), mostramos alguns resultados de p e t , escolhendo uma condição inicial arbitrária, um valor para γ e variando o raio da bola fechada Δ . Observamos que o programa não convergirá se a matriz $\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}$ for indefinida. Para torná-la definida positiva, ajustamos o valor do raio da bola Δ e o valor de γ . De forma análoga ao Capítulo 3, é interessante que haja uma faixa de variação de Δ que garanta o bom condicionamento da matriz M de tal forma que uma melhor determinação da matriz de ganho de realimentação G possa ser obtida.

Na Tabela (4.5), a norma das matrizes de ganho G e G_w são obtidas para $\gamma = 0.51$ (pior caso) e para valores crescentes de Δ , bem como os autovalores máximos e mínimos da matriz A_f . O ajuste de $\gamma = 0.51$ foi obtido seguindo o algoritmo descrito na Tabela (4.1). Com o ajuste de Δ , é possível gerar uma família de ganhos estabilizantes para esse pior caso, no qual os autovalores podem ser alocados numa determinada região que satisfaça os requerimentos de estabilidade e desempenho robusto.

Tabela 4.3: Valores de p e t

γ	$x(0)$	Δ^2	iterações	t	p	Δ^2	iterações	t	p
0.51	$[1 \ 0]'$	1	8	-0.8291	-0.7171 -0.5675 -0.2796 -0.2920	4	10	-0.9585	-0.8710 -1.5309 -0.3421 -0.8834
0.51	$[1 \ 0]'$	0.5	8	-0.7757	-0.5931 -0.2920 -0.2144 -0.1301	0.1	7	-0.6597	-0.2991 -0.0720 -0.0706 -0.0184
0.51	$[1 \ 0]'$	10	10	-1.0884	-0.9582 -2.5715 -0.3606 -1.5292	25	11	-1.2844	-1.0648 -4.1673 -0.3716 -2.5219
0.51	$[1 \ 0]'$	100	12	-1.7992	-1.3184 -8.4413 -0.3807 -5.1627	400	13	-2.7829	-1.8001 -16.9397 -0.3846 -10.4714
0.51	$[1 \ 0]'$	2.63	10	-0.9137	-0.8315 -1.1801 -0.3295 -0.6666	2.8	10	-0.9197	-0.8374 -1.2269 -0.3315 -0.6955

Tabela 4.4: Valores de p e t

γ	$x(0)$	Δ^2	iterações	t	p	Δ^2	iterações	t	p
0.51	$[0 \ 1]'$	1	5	-0.8906	-0.2126	4	10	-1.7580	-0.2986
					-0.9749				-1.9744
					-0.0595				-0.1012
					-0.0179				-0.0480
0.51	$[0 \ 1]'$	0.5	7	-0.6339	-0.1717	0.1	6	-0.2872	-0.0963
					-0.6846				-0.3008
					-0.0407				-0.0142
					-0.0105				-0.0028
0.51	$[0 \ 1]'$	10	10	-2.7571	-0.3532	25	10	-4.3287	-0.4020
					-3.1386				-4.9784
					-0.1302				-0.1558
					-0.0872				-0.1524
0.51	$[0 \ 1]'$	100	11	-8.5837	-0.4653	400	12	-17.0433	-0.5258
					-9.9817				-19.9796
					-0.1847				-0.2020
					-0.3358				-0.7064
0.51	$[0 \ 1]'$	2.63	7	-1.4325	-0.2720	2.8	7	-1.4753	-0.2760
					-1.5976				-1.6470
					-0.0874				-0.0894
					-0.0361				-0.0376

Tabela 4.5: Resultados numéricos do projeto via nova proposta com $\gamma = 0.51$

Δ^2	$\begin{pmatrix} \lambda_{min} \\ \lambda_{max} \end{pmatrix}$	G	$\ G\ _\infty$	G_w	$\ G_w\ _\infty$
0.5	$\begin{pmatrix} 0.0976 \\ 0.4827 \end{pmatrix}$	$\begin{bmatrix} -2.5981 & -7.2621 \end{bmatrix}$	7.2621	$\begin{bmatrix} 13.6174 & 13.5289 \end{bmatrix}$	13.6174
1	$\begin{pmatrix} 0.1300 \\ 0.4827 \end{pmatrix}$	$\begin{bmatrix} -2.4835 & -7.0049 \end{bmatrix}$	7.0049	$\begin{bmatrix} 13.6199 & 13.5359 \end{bmatrix}$	13.6199
4	$\begin{pmatrix} 0.2104 \\ 0.5380 \end{pmatrix}$	$\begin{bmatrix} -1.5738 & -6.0228 \end{bmatrix}$	6.0228	$\begin{bmatrix} 13.6376 & 13.5672 \end{bmatrix}$	13.6376
10	$\begin{pmatrix} 0.2741 \\ 0.6434 \end{pmatrix}$	$\begin{bmatrix} -0.4474 & -4.7959 \end{bmatrix}$	4.7959	$\begin{bmatrix} 13.6595 & 13.6061 \end{bmatrix}$	13.6595
25	$\begin{pmatrix} 0.3233 \\ 0.8695 \end{pmatrix}$	$\begin{bmatrix} 1.3294 & -2.7888 \end{bmatrix}$	2.7888	$\begin{bmatrix} 13.6943 & 13.6692 \end{bmatrix}$	13.6943

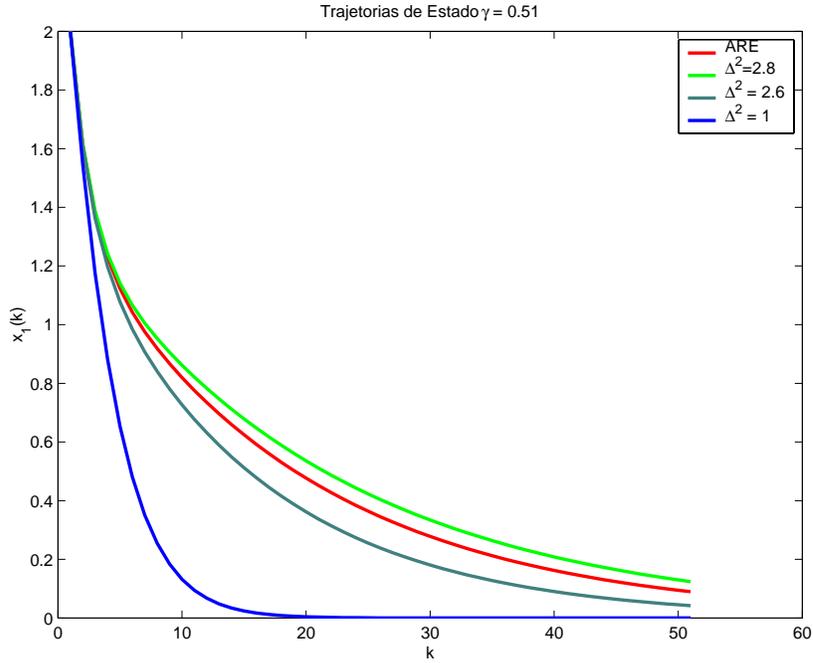


Figura 4.1: Trajetória da variável de estado $x_1(k)$

Através da Tabela (4.5), verificamos que a nova metodologia de controle H_∞ via LMI tem uma certa flexibilidade, no sentido de podermos obter soluções eficientes através do parâmetro Δ .

Na Figura (4.1) são apresentadas as trajetórias de estado para vários valores de Δ , e comparando-as com as trajetórias obtidas através da síntese H_∞ via ARE. Nesse exemplo, fica claro que é possível, com o ajuste de Δ , obter melhores respostas para as trajetórias de estado. Observamos que a solução obtida através da nova proposta ($\Delta^2 = 1$ e $\Delta^2 = 2.6$) é melhor do que aquela obtida através do controle ótimo via ARE, uma vez que o sistema apresenta um ganho maior obtendo assim uma resposta mais rápida.

Na Figura (4.2) são apresentadas as trajetórias de controle para vários valores de Δ , observamos o bom desempenho do sistema dinâmico obtido através da nova proposta quando comparado com a solução de Riccati. De fato, à medida que o

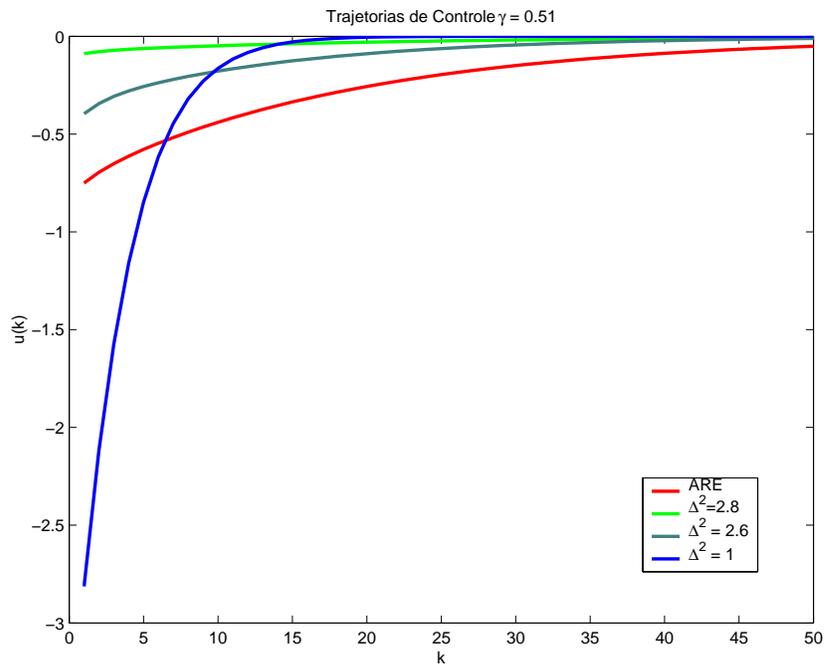


Figura 4.2: Trajetória de controle

parâmetro de ajuste Δ for aumentando, o esforço de controle, obtido pela nova proposta, faz com que o sistema atinja mais rapidamente o valor de regime desejado.

4.8 Comentário Final

Com o intuito de dar uma contribuição ao estudo da síntese de controle H_∞ de sistemas dinâmicos discretos e lineares, foi desenvolvida uma nova proposta, através do emprego da metodologia de Otimização Convexa Multinível e Programação semidefinida. Diferente das abordagens tradicionais, a matriz de ganho de realimentação foi obtida sem a necessidade do uso das equações de Riccati acopladas, o que resulta em um projeto de controladores muito mais simples, cuja metodologia pode assegurar boas condições de estabilidade e desempenho robusto, devido ao parâmetro de ajuste Δ .

Capítulo 5

Conclusão

Este trabalho apresenta o desenvolvimento de novos algoritmos usando LMI para sínteses de ganhos estabilizantes de controladores LQR e H_∞ . Foi apresentada uma técnica iterativa simples via Programação Semidefinida obtida a partir de métodos de otimização convexa multinível, que apresenta vantagens significativas do ponto de vista:

1. das condições de estabilidade numérica (melhor condicionamento das matrizes computadas).
2. do emprego de métodos de pontos interiores para Programação Semidefinida. Esses métodos são eficazes e eficientes e possuem complexidade polinomial.
3. de especificações distintas na síntese de controladores, para explorar limites de desempenho e estabilidade.
4. das restrições de convexidade impostas pelo problema SDP, em relação à não garantia de convexidade de técnicas tradicionais baseadas na solução via ARE.

Na metodologia proposta, empregamos a técnica LMI e a formulação via TRS, para redução da ordem do problema de otimização, garantindo convexidade e

gerando diferentes ganhos de malha fechada para diferentes restrições H_∞ . Tal abordagem tornou possível atingirmos especificações de desempenho e estabilidade aceitáveis para uma ampla faixa de ganhos. Com esta nova metodologia, as equações matriciais são bem condicionadas, há uma simplicidade algébrica de cálculo off-line e é possível dispormos de uma variedade de métodos numéricos com boas propriedades de estabilidade e precisão numérica. Estas características, normalmente críticas em outras metodologias de projeto de controle, asseguram a determinação de uma família de matrizes de ganho de realimentação, garantindo estabilidade e desempenho robustos.

Para isso, basta definir um γ (controle H_∞ subótimo), tal que a matriz Hamiltoniana não apresente autovalores sobre a circunferência de raio unitário. A seguir, através do ajuste no raio da bola Δ , obtemos uma matriz de ganho de realimentação desde que os pólos em malha fechada estejam dentro de uma região desejada, com o cuidado de computar-se a norma H_∞ e verificar se esta é menor que γ .

5.1 Trabalhos Futuros

Como perspectivas futuras é possível destacar:

- Projetar controladores via metodologia multinível e algoritmos de região de confiança.
- Projetar controladores via Programação do cone de segunda ordem (do inglês, Programming Cone Second-order (SOCP)).
- Desenvolver algoritmos de projeto de controle robusto, utilizando-se esta nova metodologia, para o problema de controle H_∞ com realimentação de saída.
- Estender as metodologias de otimização para o problema (LQR) e H_∞ de sistemas discretos com atraso.

Referências Bibliográficas

- [1] ALIZADEH, F. **Interior Point Methods in Semidefinite Programming With Applications To Combinatorial Optimization**. SIAM J. Optim. 5(1), 13-51. 1995.
- [2] ALIZADEH, F.; I. Adler. **Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems**. Technical Report RRR-111-95, RUTCOR, Rutgers University, 1995.
- [3] BAZARAA, M.S. **Nonlinear Programming: Theory and Algorithms**. John Wiley & Sons, 1993.
- [4] BELLMAN R.; FAN K. **On systems of linear inequalities in Hermitian matrix variables, in Convexity**. V. L. Klee, ed., vol. 7 of Proceedings of Symposia in Pure Mathematics, American Mathematical Society, 1963, pp. 1-11.
- [5] BOYD, S.; VANDENBERGHE L. **Semidefinite Programming**. SIAM Review 38, 49-95. 1996.
- [6] BOYD, S.P.,L.E. Ghaoui, E. Feron.; V. Balakrishnan. **Linear Matrix Inequalities in System and Control Theory**, SIAM Studies in Applied Mathematics. SIAM, Philadelphia, 1994.
- [7] COLANERI, Patrizio.; LOCATELLI Arturo. **Controllo in $RH_2 - RH_\infty$** . Pitagora Editrice Bologna, Bologna, 1993.

- [8] DENNIS, J. E. Jr.; SCHNABEL, R. B., **Numerical Methods for Unconstrained Optimization and Nonlinear Equations**, Prentice Hall Englewood Cliffs, New Jersey., 1983.
- [9] DOYLE, J. C.; GLOVER, K.; KHARGONEKAR, P.P.; FRANCIS, B.A., **State-Space Solutions to Standard H_2 e H_∞ Control Problems**. IEEE Transactions on Automatic Control, vol. AC-34, pg 831-847, 1989.
- [10] FRANCIS, B.A., **A Course in H_∞ Theory**. Lectures Notes in Control and Information sciences, vol88, 1986.
- [11] FIACCO, A. V.; McCORMICK, G. P. **Nonlinear Programming: Sequential Unconstrained Minimization Techniques**, Wiley, New York, 1968. Reprinted by SIAM Publications, 1990.
- [12] FILHO, José Tarcisio Costa. **Uma Contribuição ao Controle Ótimo de Alto Desempenho**. Tese submetida Departamento de Engenharia de Eletricidade da UFMA para o concurso de Professor Titular na área de Automação e Controle. 1997.
- [13] FLETCHER R.; **A nonlinear programming problem in statistics**. SIAM J. Sci. Stat. Comput., 2:257-267, 1981.
- [14] FLETCHER R. **Semi-definite matrix constraints in optimization**. SIAM J. Control Optim., 23:493-513, 1985.
- [15] FLETCHER R. **Practical Methods of Optimization**, 2 ed., New York, John Wiley Sons, 1987.
- [16] FLOUDAS, A. C.; VISWESWARAN, V. **Quadratic Programming**, 1995.

- [17] FRISCH, K. R. **The logarithmic potential method for convex programming**. Unpublished manuscript, Institute of Economics, University of Oslo, oslo, Norway, May 1955.
- [18] GAHINET, P.; NEMIROVSKI A.; CHILALI M. **LMI Control Toolbox**, The Mathworks Inc, Natick MA, 1995.
- [19] GEROMEL, J. C.; COLANERI, Patrizio; LOCATELLI Arturo. **Control Theory and Design - An RH_2 and RH_∞ viewpoint**. Academic Press, Califórnia, 1997.
- [20] GEROMEL, J.C.; CARVALHO Mauricio. **LMISol Toolbox: users' guide**, versão 1.0 alpha disponível em <http://www.dt.fee.unicamp.br/mauricio/>, (1997).
- [21] GILL, Philip E.; Walter Murray.; Margaret H. Wright **Practical Optimization**. New York: Academic Press 1981.
- [22] HERTOOG, D. D. **Interior Point Approach to Linear, Quadratic and Convex Programming**. Mathematics and its Applications 277. Kluwer Academic Publishers, Dordrecht, 1992.
- [23] JACQUOT, R.G. **Modern Digital Control Systems**, Marcel Dekker, Inc., New York (1994).
- [24] KALMAN, R. E., **Lyapunov functions for the problem of Lure in automatic control**. Proc. Nat. Acad. Sci., USA, 1963.
- [25] KALMAN, R. E., **On a new characterization of linear passive systems**. In proc. First Annual Allerton Conf. on Communication, Control and Computing, pages 456-470, 1963.
- [26] KARMAKAR, N., **A new polynomial-time algorithm for linear programming**. Combinatoria, 1984.

- [27] KARMARKAR, N., **An interior-point approach to NP-complete problems.** Contemporary Mathematics, 114:297-308, 1990.
- [28] KHARGONEKAR, P.; PETERSEN I.R.; ROTEA A., **H_∞ Optimal Control with State-Feedback**, IEEE Transactions on Automatic Control, vol 33, 1988.
- [29] LEVENBERG, K., **A method for the solution of certain nonlinear problems in least squares.** Quart. Appl. Math, 2:164-168, 1944.
- [30] LEWIS, F.L.; SYRMOS, V.L., **Optimal Control**, New York, John Wiley and Sons, Inc, 1995.
- [31] LIMA, Elon L., **Curso de Análise, Volume 2.** Rio de Janeiro, IMPA, 1981.
- [32] LUENBERGER, D. G., **Linear and Nonlinear Programming**, Addison- Wiley, Reading 1986.
- [33] LURE, A. I., **Some Nonlinear Problems in the Theory of Automatic Control.** H. M. Stationery Off., London, 1957. In Russian, 1951.
- [34] MARQUARDT, D. W. **An algorithm for least-squares estimation of nonlinear parameters.** SIAM Journal on Applied Mathematics, 11(2): 431-441, 1963.
- [35] MANGASARIAN, Olvi L. **Nonlinear Programming**, SIAM Studies In Applied Mathematics. SIAM, Philadelphia, Pennsylvania, 1994.
- [36] MARTINEZ, J.M., **Local Minimizers of Quadratic Functions On Euclidean Balls and Spheres.** SIAM Journal on Optimization 4, 1994.
- [37] MONTEIRO, R.D.C.; ADLER I. **Interior path following primal-dual algorithms. II: Convex quadratic programming.** Mathematical Programming, 44(1):43-66, 1989.

- [38] MORE[́], J.J. **Recent developments in algorithms and software for trust region methods**. In A. Bachem, M. Grötschel, and B. Korte, editors, *Mathematical Programming, The State of the Art*. Springer Verlag, 1983.
- [39] MORE[́], J.J.; SORENSEN, D.C. **Computing a trust region step**. *SIAM Journal on Sci. Statist. Comput.*, 4:553-572, 1983.
- [40] NEMIROVSKII, A.; BEN-TAL Aharon. **Lectures On Modern Convex Optimization: analysis, algorithms and engineering applications**, MPS-SIAM Series on Optimization, Philadelphia, Pennsylvania, 2001.
- [41] NESTEROV, Y. NEMIROVSKII, A. **Interior-Point Polynomial Algorithm in Convex Programming**. *SIAM Studies In Applied Mathematics*. SIAM, Philadelphia, Pennsylvania, 1994.
- [42] NICULESCU, S.I.; GHAOUI L. E. **Advances in Linear Matrix Inequalities Methods in Control**. *SIAM Studies In Applied Mathematics*. SIAM, Philadelphia, Pennsylvania, 2000.
- [43] NOCEDAL, J.; WRIGHT S.J. **Numerical Optimization**. *Springer Series in Operations Research*. Springer- Verlag, 1999.
- [44] OGATA, Katsuliko. **Discrete-Time Control Systems**. Prentice Hall. New Jersey, 1995.
- [45] PARDALOS, P. M.; WOLKOWICZ Henry. **Topics In Semidefinite and Interior Point Methods**. American Mathematical Society, 1991.
- [46] PYATNITSKII, E.S .; V. I SKORODINSKII. **A criterion of absolute stability of non-linear sampled-data control systems in the form of numerical procedures**. *Automation and Remote Control*, 47(9):1190-1198, August 1986.

- [47] POPOV, V. M.; **Absolute stability of nonlinear systems of automatic control.** Automation and Remote Control, 1962.
- [48] POWELL, M.J.D. **Convergence properties of a class of minimization algorithms.** In O. L. Mangansarian, R. R. Meyer, and S. M. Robinson, editors, Nonlinear Programming 2, pages 1-27. Academic Press, New York, 1975.
- [49] RAMANA M.; TUNCEL L.; WOLKOWICZ H. **Strong duality for semidefinite programming,** SIAM J. on Optimization, vol. 7, Aug. 1997.
- [50] RENDL F.; WOLKOWICZ H. **A semidefinite framework for trust region subproblems with applications to large scale minimization.** Waterloo, Canadá, 1995.
- [51] RENEGAR, James. **A Mathematical View Of Interior-Point Methods In Convex Optimization.** Philadelphia: Society for Industrial and Applied Mathematics, 2001.
- [52] ROCKAFELLAR, R. T. **Convex Analysis.** Princeton University Press, 1970.
- [53] SOUZA, Valeska M.; COSTA FILHO , J. Tarcisio. **Projeto de Controlador Robusto via Programação Semidefinida.** Advances in logic, artificial intelligence and robotics, Vol. II, Proceedings of the 3º Congress of Logic Applied to Technology- LAPTEC' 2002, São Paulo, Brazil.
- [54] SORENSEN D.C.; **Newton's Method With A Model Trust Region Modification.** SIAM J. Numer. Anal., 19:409-426, 1982.
- [55] STRANG Gilbert.; **Linear Algebra Its Applications.** Academic Press, Inc. Florida, 1982.
- [56] VANANTWERP, J.G.; BRAATZ R.D. **A Tutorial on Linear and Bilinear Matrix Inequalities.** Journal of Process Control, 10, 363-385, 2000.

- [57] VANDENBERGHE L.; BALAKRISHNAN V. **Connections between duality in control theory and convex optimization**, in proceedings of the American Control Conference, pp. 4030-4035, 1995.
- [58] WILLEMS, J. C., **Least squares stationary optimal control and the algebraic Riccati equation**. IEEE Trans. Aut. Control, AC-16(6):621-634, December 1971.
- [59] WOLKOWICZ H. **Some Applications of optimization in matrix theory**, Linear Algebra and Appl., 40 (1981), pp.239-258.
- [60] WOLKOWICZ H.; RENDL F. **A semidefinite framework for trust region subproblems with applications to large scale minimization**. Mathematical Programming Series B, 77(2):273-299,1997.
- [61] WRIGHT, Stephen J., **Primal-Dual Interior Point Methods**, SIAM Studies In Applied Mathematics. SIAM, Philadelphia, Pennsylvania, 1997.
- [62] WRIGHT, Margaret. **The Interior Point Revolution in Constrained Optimization**.
- [63] YE, Y., **An extension of Karmarkar's algorithm and the trust region method for quadratic programming**. pages 49-63, 1989.
- [64] YE, Y., **On Affine Scaling Algorithms for Nonconvex Quadratic Programming**. Mathematical Programming, 56:285-300, 1992.
- [65] YAKUBOVICH, V. A., **The solution of certain matrix inequalities in automatic control theory**. Soviet Math. Dokl., 3:620-623, 1962. In Russian, 1961.
- [66] YAKUBOVICH, V. A., **Solution of certain matrix inequalities encountered in nonlinear control theory**. Soviet Math. Dokl., 5:652-656, 1964.

- [67] YAKUBOVICH, V. A., **The method of matrix inequalities in the stability theory of nonlinear control systems, I, II, III.** Automation and Remote Control, 1967.
- [68] ZAMES, G., **Feedback and Optimal Sensitivity: Model Reference Transformations; Multiplicative Seminorms and Approximate Inverse.** IEEE Trans. Aut. Control, AC-26, 301-320, 1981.

Apêndice A

Normas e Valores Singulares

A.1 Normas de Vetores e Matrizes

A norma- p de x é definida como:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p \leq \infty \quad (\text{A.1})$$

Em particular, quando $p = 1, 2$ e ∞ , temos:

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^n |x_i| \\ \|x\|_2 &= \sqrt{\sum_{i=1}^n |x_i|^2} \\ \|x\|_\infty &= \max |x_i| \end{aligned}$$

Claramente, norma é uma abstração e extensão usual do conceito de comprimento no espaço Euclidiano tridimensional. Similarmente podemos introduzir alguns tipos de medidas de uma matriz.

Seja uma matriz $A \in C^{m \times n}$, então a norma da matriz A , induzida pela $\|\cdot\|_p$ é definida como:

$$\| A \|_p = \sup_{\|x\| \neq 0} \frac{\| Ax \|_p}{\| x \|_p} \quad (\text{A.2})$$

Em particular, para $p = 1, 2$ e ∞ , a norma induzida pode ser computada como:

$$\begin{aligned} \| A \|_1 &= \max \sum_{i=1}^m | a_{ij} | \\ \| A \|_2 &= \sqrt{\lambda_M(A^* A)} \\ \| A \|_\infty &= \max \sum_{i=1}^m | a_{ij} | \end{aligned}$$

onde λ_M é o maior autovalor da matriz indicada e A^* é a transposta conjugada de A . Do ponto de vista teórico, para um sistema, as normas induzidas têm a interpretação de ganhos de amplificação entrada/saída.

Uma outra norma de matriz, freqüentemente usada, é a norma de Frobenius. Ela é definida como:

$$\| A \|_F = \sqrt{\text{Tr}(A^* A)} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n | a_{ij} |^2} \quad (\text{A.3})$$

A.2 Valores Singulares

Os valores singulares de uma matriz A são definidos por:

$$\sigma_i(A) = \sqrt{\lambda_i(A^* A)}, \quad i = 1, 2, \dots, n \quad (\text{A.4})$$

onde $\sigma_i(A)$ representa o i -ésimo valor singular de A . $\lambda_i(A^* A)$ representa o i -ésimo autovalor da matriz $A^* A$.

O máximo valor singular σ_M fornece a norma espectral de uma matriz, isto é, seja A uma matriz, então:

$$\| A \| = \sqrt{\lambda_M(A^* A)} \quad (\text{A.5})$$

Uma matriz A é dita grande, se seu menor valor singular σ_m for grande; pequena, se seu maior valor singular σ_M for muito pequeno, isto é:

Matriz grande: $\sigma_m \gg 1$

Matriz pequena: $\sigma_M \ll 1$

A.3 Norma de Matrizes de Transferência

Há dois tipos de normas de matrizes de transferência: as dependentes da frequência, e as que não dependem da frequência.

Seja $G(z) \in C^{n \times m}$, uma matriz de transferência, sua norma em função da frequência é dada por

$$\|G(z)\| = \sigma_M[G(z)] \quad (\text{A.6})$$

Das normas que não dependem da frequência, as usuais são as normas H_2 e H_∞ . Seja $G(z) \in C^{n \times m}$ uma matriz de transferência estável, dada por:

$$G(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

onde $A \in C^{n \times m}$, $B \in C^{n \times p}$, $C \in C^{n \times n}$, D são matrizes do sistema cuja função de transferência é dada por:

$$G(z) = C(zI - A)^{-1}B + D$$

A norma H_2 é definida como:

$$\|G(z)\|_2^2 = \text{Tr}(DD^T + CL_c C^T) = \text{Tr}(D^T D + B^T L_o B) \quad (\text{A.7})$$

onde L_c e L_o são os Gramianos de controlabilidade e observabilidade, dados pelas soluções das seguintes equações lineares:

$$AL_cA^T - L_c + BB^T = 0 \quad (\text{A.8})$$

$$A^TL_oA - L_o + C^TC = 0 \quad (\text{A.9})$$

A norma H_∞ de $G(z)$ é definida como:

$$\|G(z)\|_{H_\infty} = \sup \sigma_M[G(z)] \quad (\text{A.10})$$

Apêndice B

Topologia do R^n

Neste apêndice, colecionamos algumas propriedades topológicas do R^n que podem ser encontradas em grandes detalhes nos livros de Análise e de Topologia.

B.1 Conjuntos Abertos e Fechados

Definição 8 (Bola Aberta). *A bola aberta de centro em um ponto $a \in R^n$ e raio $r > 0$ é o conjunto dos pontos $x \in R^n$, cuja distância ao ponto a é menor do que r . Ou seja,*

$$B(a; r) = \{x \in R^n; \|x - a\| < r\}.$$

Definição 9. Bola Fechada

A bola fechada de centro em um ponto $a \in R^n$ e raio $r > 0$ é o conjunto dos pontos $x \in R^n$, cuja distância ao ponto a é menor ou igual a r . Ou seja,

$$B(a; r) = \{x \in R^n; \|x - a\| \leq r\}.$$

Definição 10. Ponto Interior

Seja D um subconjunto de R^n . Um ponto $a \in D$ chama-se um ponto interior a D quando é centro de alguma bola inteiramente contida em D .

Definição 11. Interior de um conjunto

O interior de D é o conjunto $\text{int } D$, formado pelos pontos interiores a D . Quando $a \in \text{int } V$, dizemos que o conjunto V é uma vizinhança do ponto a .

Definição 12. Conjunto Aberto

Um conjunto $D \subset \mathbb{R}^n$ chama-se aberto quando todos os seus pontos são interiores.

Definição 13. Ponto Aderente

Um ponto $a \in \mathbb{R}^n$ é dito aderente a um conjunto $D \subset \mathbb{R}^n$, se para cada $r > 0$,

$$B(a; r) \cap D \neq \emptyset$$

Outras maneiras equivalentes de dizer que a é aderente a D são:

1. para todo conjunto aberto A contendo a , tem-se $A \cap D \neq \emptyset$;
2. toda vizinhança de a tem pontos em comum com o conjunto D .

Definição 14. Fecho de um conjunto

O fecho (ou aderência) de um conjunto $D \subset \mathbb{R}^n$ é o conjunto \overline{D} dos pontos de \mathbb{R}^n que são aderentes a D . Portanto, escrever $a \in \overline{D}$ é o mesmo que afirmar que o ponto a é aderente a D .

Definição 15. Ponto de acumulação

Seja D um subconjunto de \mathbb{R}^n . Um ponto $a \in \mathbb{R}^n$ chama-se ponto de acumulação do conjunto D , quando toda bola aberta de centro a contém algum ponto de D , diferente do ponto a . Em linguagem simbólica.

$$\forall r > 0, \exists y \in B(a; r) \cap D, y \neq a.$$

Definição 16. Conjunto Fechado

Um conjunto $F \subset \mathbb{R}^n$ é dito fechado quando contém todos seus pontos aderentes.

B.2 Seqüências

Definição 17. *Seqüência*

Uma seqüência em R^n é uma aplicação $x : N \rightarrow R^n$, definida no conjunto N dos números naturais.

Usamos as seguintes notações (x_n) ou $(x_1, x_2, \dots, x_n, \dots)$.

Definição 18. *Limite de uma Seqüência*

O ponto $a \in R^n$ é dito o limite da seqüência de pontos $x_n \in R^n$, quando para todo $\epsilon > 0$ dado, existe um número $n_0 \in N$ tal que $n > n_0 \Rightarrow |x_n - a| < \epsilon$.

Definição 19. *Subseqüência*

Uma subseqüência de (x_n) é restrição da seqüência a um subconjunto infinito $N' = \{n_1 < n_2 < \dots < n_i < \dots\} \subset N$.

Definição 20. *Seqüência crescente*

Uma seqüência (x_n) de números reais é dita crescente quando temos

$x_1 < x_2 < \dots < x_n < \dots$, isto é, $x_n < x_{n+1}$ para todo $n \in N$.

Quando vale apenas $x_n \leq x_{n+1}$, a seqüência é dita não-decrescente.

Analogamente se define seqüência decrescente e não-crescente. Uma seqüência de um desses quatro tipos é chamada monótona.

Definição 21. *Seqüência Limitada*

Uma seqüência (x_n) é dita limitada, se existe um número real $c > 0$ tal que $|x_n| \leq c$ para todo $n \in N$.

Definição 22. ponto limite Um ponto a é um ponto limite da seqüência (x_n) , se existe uma subseqüência (x_{n_i}) que converge para a .

Teorema 20. Toda seqüência monótona e limitada de números reais é convergente.

Demonstração. Consideremos a seqüência não-decrescente $x_1 \leq x_2 \leq \dots \leq x_n \leq \dots$. Tomemos $a = \sup_{n \in \mathbb{N}} x_n$. Afirmação: $a = \lim x_n$. Com efeito, dado arbitrariamente $\varepsilon > 0$, o número $a - \varepsilon$, sendo menor do que a , não pode ser cota superior do conjunto dos valores x_n . Logo existe $n_0 \in \mathbb{N}$ tal que $a - \varepsilon < x_{n_0} \leq a$. Então $n > n_0 \Rightarrow a - \varepsilon < x_{n_0} \leq x_n \leq a < a + \varepsilon \Rightarrow a - \varepsilon < x_n < a + \varepsilon$. Isto é, $x_n \rightarrow a$. □

B.3 Conjuntos Compactos

Definição 23. *Conjunto Limitado*

Um conjunto $K \subset \mathbb{R}^n$ é dito limitado se, e somente se, estiver contido em alguma bola.

Definição 24. *Conjunto Compacto*

Um conjunto $K \subset \mathbb{R}^n$ é dito compacto se satisfaz as seguintes condições equivalentes:

1. K é fechado e limitado;
2. (Borel-Lebesgue) Toda cobertura aberta de $K \subset \mathbb{R}^n \cup_{\lambda \in L} A_\lambda$ possui subcobertura finita $K \subset A_{\lambda_1} \sqcup \dots \sqcup A_{\lambda_i}$;
3. (Bolzano-Weierstrass) Toda seqüência de pontos de K possui uma subseqüência que converge para um ponto de K .

Apêndice C

Álgebra Linear

C.1 Menores Principais

Seja A uma matriz quadrada real de ordem n . Os menores principais são os determinantes de todas as submatrizes de A , cujas diagonais coincidem ou fazem parte da diagonal principal de A .

Exemplo 9. *Consideremos a matriz*

$$A := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Então a_{11}, a_{22}, a_{33} ,

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

e $\det A$ são os menores principais da matriz A .

C.2 Matrizes Definidas

Definição 25. *Uma matriz $A \in R^{n \times n}$ é dita:*

- *definida positiva, se $x^T Ax > 0, \forall x \in R^n, x \neq 0$*
- *definida negativa, se $x^T Ax < 0, \forall x \in R^n, x \neq 0$*
- *semidefinida positiva, se $x^T Ax \geq 0, \forall x \in R^n$*
- *semidefinida negativa, se $x^T Ax \leq 0, \forall x \in R^n$*

Se A não satisfaz nenhuma das alternativas acima, é dita indefinida.

Existem vários modos para verificarmos se uma matriz é definida.

Um deles é calcular os autovalores de acordo com os teoremas abaixo:

Teorema 21. *Uma matriz simétrica é definida positiva (negativa) se, e somente se, todos os seus autovalores são positivos (negativos).*

Uma matriz simétrica é semidefinida positiva (negativa) se, e somente se, todos os seus autovalores são não negativos (não positivos).

Teorema 22. *Uma matriz simétrica é definida positiva se, e somente se, todos os seus menores principais são positivos e semidefinida positiva se, e somente se, todos seus menores principais são negativos.*

C.3 Desigualdades Matriciais

Definição 26. *Sejam A e B duas matrizes simétricas então as relações de desigualdade $>, <, \geq,$ e \leq são definidas como*

- *$A > B$ se, e somente se, $A - B$ é definida positiva*
- *$A < B$ se, e somente se, $A - B$ é definida negativa*
- *$A \geq B$ se, e somente se, $A - B$ é semidefinida positiva*
- *$A \leq B$ se, e somente se, $A - B$ é semidefinida negativa.*

C.4 Inversas de Matrizes de Posto Completo

Seja $A \in R^{m \times n}$. O posto $P(A)$ de uma transformação linear A é a dimensão de seu espaço imagem, igual ao número de colunas linearmente independentes (e também igual ao número de linhas LI). Se $P(A) = \min(m, n)$ então A é de posto completo.

C.4.1 Inversa de Uma Matriz Quadrada

Seja $A \in R^{n \times n}$ e $P(A) = n$. Então existe uma única matriz A^{-1} , chamada a inversa de A , tal que

$$AA^{-1} = A^{-1}A = I_n$$

C.4.2 Inversa de Uma Matriz Retangular

Inversa à Direita

Seja $A \in R^{m \times n}$ e $P(A) = m$. Uma matriz $X \in R^{n \times m}$ é dita uma inversa à direita de A se $AX = I_m$. Nesse caso, uma das inversas à direita é dada por:

$$X = A^T(AA^T)^{-1}$$

Inversa à Esquerda

Reciprocamente, se $P(A) = n$. Uma matriz $Y \in R^{n \times m}$ é chamada uma inversa à esquerda de A se $YA = I_n$. Portanto, uma das inversas à esquerda é

$$Y = (A^T A)^{-1} A^T$$

Observemos que inversas à direita (ou à esquerda) existem somente quando A é de posto completo. Em ambos os casos a inversa não é única.

Exemplo 10. Qualquer matriz da forma $\begin{bmatrix} I \\ * \end{bmatrix}$ é uma inversa à direita de $\begin{bmatrix} I & 0 \end{bmatrix}$.

C.5 Inversa Generalizada de Moore-Pensore

Geralmente, se uma matriz A não possui posto completo então A não é inversível, será útil usar a chamada pseudo-inversa, também conhecida como inversa de Moore-Pensore. Essa pseudoinversa denotada por A^\dagger é a única matriz que satisfaz as seguintes condições:

1. $AA^\dagger A = A$
2. $A^\dagger AA^\dagger = A^\dagger$
3. $(AA^\dagger)^T = AA^\dagger$
4. $(A^\dagger A)^T = A^\dagger A$

Uma maneira de calcular A^\dagger é escrever $A = BC$, de modo que B tenha posto coluna completo e C tenha posto linha completo. Então

$$A^\dagger = C^T(CC^T)^{-1}(B^T B)^{-1}B^T$$

Outro modo de calcular A^\dagger é usar a decomposição em valor singular (SVD). Suponhamos que A tenha uma decomposição em valor singular $A = U\Sigma V^T$

$$\text{com } \Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}, \Sigma_i > 0.$$

$$\text{Então } A^\dagger = V\Sigma^\dagger U^T \text{ com } \Sigma^\dagger = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

C.6 Número de Condição

O número de condição é um dado muitas vezes difícil de se obter, mas que tem papel importante quando temos vários sistemas para resolver e a diferença entre seus elementos é pequena. Neste caso, através do número de condição podemos saber

se a solução varia muito ou não dependendo da oscilação dos elementos.

Vamos considerar o efeito de pequenas perturbações nos elementos da matriz $(\tilde{H} + \alpha I) > 0$ e do vetor \tilde{b} na obtenção da solução de p no sistema

$$(\tilde{H} + \alpha I)p = -\tilde{b} \tag{C.1}$$

A questão que fazemos é: Qual deve ser o número de condição antes que o sistema (C.1) seja considerado mal condicionado ?

Conhecer o condicionamento de um sistema é muito significativo, uma vez que, se ele é mal condicionado, não há um método com razoável estabilidade numérica para resolver o sistema (C.1) de forma a obter uma solução precisa. Nesse sentido, consideremos os seguintes casos:

Caso 1

Consideremos o sistema (C.1) onde $\Upsilon = (\tilde{H} + \alpha I) > 0$ e que o sistema tem solução única $p = -\Upsilon^{-1}\tilde{b}$. Se a matriz Υ permanecer a mesma mas o vetor \tilde{b} é alterado para $(\tilde{b} + \delta\tilde{b})$ a solução p também terá uma mudança δp que deve satisfazer a equação

$$\Upsilon(p + \delta p) = \tilde{b} + \delta\tilde{b}$$

desse modo temos:

$$\delta p = \Upsilon^{-1}\delta\tilde{b}$$

aplicando norma:

$$\|\tilde{b}\| \leq \|\Upsilon\| \|p\| \tag{C.2}$$

$$\|\delta p\| \leq \|\Upsilon^{-1}\| \|\delta\tilde{b}\| \tag{C.3}$$

multiplicando as desigualdades (C.2) e (C.3) temos:

$$\frac{\|\delta p\|}{\|p\|} \leq \|\Upsilon\| \|\Upsilon^{-1}\| \frac{\|\delta \tilde{b}\|}{\|\tilde{b}\|}$$

As quantidades

$$\frac{\|\delta p\|}{\|p\|} \text{ e } \frac{\|\delta \tilde{b}\|}{\|\tilde{b}\|}$$

medem a alteração relativa no vetor solução p e no vetor \tilde{b} , respectivamente. Logo, podemos verificar que a razão entre a alteração relativa na solução e a alteração relativa no vetor \tilde{b} é limitada pelo número $\|\Upsilon\| \|\Upsilon^{-1}\|$ também conhecido como o número de condição da matriz Υ , representado por $Cond(\Upsilon)$, que depende da norma usada.

Caso 2

O vetor \tilde{b} é mantido inalterado e a matriz $\Upsilon = (\tilde{H} + \alpha I)$ sofre uma perturbação $\Delta\Upsilon$ de tal modo que $(\Upsilon + \Delta\Upsilon)^{-1}$ exista, então

$$\begin{aligned} p &= \Upsilon^{-1}\tilde{b} \\ (p + \delta p) &= (\Upsilon + \Delta\Upsilon)^{-1}\tilde{b} \\ \delta p &= -\Upsilon^{-1}\Delta\Upsilon(p + \delta p) \end{aligned}$$

aplicando a norma:

$$\|\delta p\| \leq \|\Upsilon^{-1}\| \|\Delta\Upsilon\| \|p + \delta p\| \quad (C.4)$$

$$\frac{\|\delta p\|}{\|p + \delta p\|} \leq \|\Upsilon\| \|\Upsilon^{-1}\| \frac{\|\Delta\Upsilon\|}{\|\Upsilon\|} \quad (C.5)$$

Análogo ao caso 1 o número $\|\Upsilon\| \|\Upsilon^{-1}\|$ é chamado número de condição da matriz Υ .

Quando o número de condição é pequeno dizemos que o sistema é bem condicionado, isto é, uma pequena mudança na matriz Υ ou no vetor \tilde{b} irá causar uma pequena variação na solução.

Comentário 7. *Embora o número de condição, como definido anteriormente, seja dependente de uma norma, em geral, se uma matriz é bem ou mal condicionada com relação a uma norma, ela é também bem ou mal condicionada com relação a algumas outras normas.*