

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

DIEGO ECKHARD

**PROJETO DE CONTROLADORES
BASEADO EM DADOS:
CONVERGÊNCIA DOS MÉTODOS
ITERATIVOS**

Porto Alegre
2008

DIEGO ECKHARD

**PROJETO DE CONTROLADORES
BASEADO EM DADOS:
CONVERGÊNCIA DOS MÉTODOS
ITERATIVOS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Controle e Automação

ORIENTADOR: Prof. Dr. Alexandre Sanfelice Bazanella

Porto Alegre
2008

DIEGO ECKHARD

**PROJETO DE CONTROLADORES
BASEADO EM DADOS:
CONVERGÊNCIA DOS MÉTODOS
ITERATIVOS**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Alexandre Sanfelice Bazanella, UFRGS

Doutor pela UFSC – Florianópolis, Brasil

Banca Examinadora:

Prof. Dr. Luis Antonio Aguirre, UFMG
Doutor pela University of Sheffield – Sheffield, Inglaterra

Prof. Dr. João Manoel Gomes da Silva Jr., UFRGS
Doutor pela Université Paul Sabatier – Toulouse, França

Prof. Dr. Walter Fetter Lages, UFRGS
Doutor pelo Instituto Tecnológico de Aeronáutica – São José dos Campos,
Brasil

Coordenador do PPGEE: _____

Prof. Dr. Arturo Suman Bretas

Porto Alegre, Dezembro de 2008.

AGRADECIMENTOS

Gostaria de agradecer à Priscila, por todo amor, apoio e compreensão.

Aos meus pais pela correção do texto e por todo suporte nestes anos de estudo.

À Luciola e Jeferson pelas muitas discussões, horas de trabalho e pela amizade.

Ao Bender, Marcela, Ivan e Maurício, pelas risadas e descontração na hora do café.

Ao bolsista pelas horas gastas na plantinha.

Ao professor João Manoel pelas inúmeras explicações.

Ao professor Bazanella, pela excelente orientação durante todo período do mestrado.

Ao PPGEE e ao CNPq pelo ensino de qualidade e concessão da bolsa.

RESUMO

O projeto de controladores baseado em dados consiste no ajuste dos parâmetros do controlador diretamente das bateladas de dados do processo, sem a necessidade de um modelo. O ajuste é feito resolvendo um problema de otimização, onde procura-se o argumento que minimize uma determinada função custo. Para resolver o problema de otimização são utilizados nesses métodos o algoritmo do gradiente, o algoritmo de Newton e variações destes. O algoritmo do gradiente apenas necessita informação do gradiente da função custo enquanto que os outros utilizam mais informações como a hessiana. Para obter estas últimas informações são utilizados experimentos mais longos e mais complexos, o que torna a aplicação mais complicada. Nesta linha o algoritmo do gradiente se apresenta como a melhor alternativa, por este motivo foi escolhido como foco deste trabalho. A convergência do algoritmo do gradiente para o mínimo global da função custo, no contexto de projeto de controladores, não é encontrada na bibliografia, decidiu-se portanto estudá-la. Essa convergência depende das condições iniciais do algoritmo e do tamanho do passo de iteração utilizado. É mostrado que as condições iniciais precisam estar dentro de uma certa região de atração. Formas de aumentar esta região de atração são tratadas na metodologia chamada Shaping da Função Custo. A principal contribuição deste trabalho é apresentar um método eficiente para a escolha do tamanho do passo de iteração que garante a convergência para o mínimo global da função custo. Algumas informações do processo são necessárias para o cálculo do tamanho do passo de iteração, também são apresentadas maneiras de obter estimativas para estas informações. Simulações e experimentos demonstram o funcionamento dos métodos.

Palavras-chave: Projeto de Controladores baseado em Dados, Critério H_2 , Algoritmo do Gradiente, Mínimo Global.

ABSTRACT

Data-based control design methods consist of adjusting the parameters of the controller directly from batches of input-output data of the process; no process model is used. The adjustment is done by solving an optimization problem, which searches the argument that minimizes a specific cost function. Iterative algorithms based on the gradient are applied to solve the optimization problem, like the steepest descent algorithm, Newton algorithm and some variations. The only information utilized for the steepest descent algorithm is the gradient of the cost function, while the others need more information like the hessian. Longer and more complex experiments are used to obtain more informations, that turns the application more complicated. For this reason, the steepest descent method was chosen to be studied in this work. The convergence of the steepest descent algorithm to the global minimum is not fully studied in the literature. This convergence depends on the initial conditions of the algorithm and on the step size. The initial conditions must be inside a specific domain of attraction, and how to enlarge this domain is treated by the methodology Cost Function Shaping. The main contribution of this work is a method to compute efficiently the step size, to ensure convergence to the global minimum. Some informations about the process are utilized, and this work presents how to estimate these informations. Simulations and experiments demonstrate how the methods work.

Keywords: Data-based tuning, H_2 criterium, steepest descent algorithm, global minimum.

LISTA DE ILUSTRAÇÕES

Figura 1:	Diagrama de Blocos	17
Figura 2:	Exemplo 1: Resposta ao salto em malha-fechada controlador inicial. . .	43
Figura 3:	Exemplo 1: Resposta ao salto em malha-fechada controlador final. . .	43
Figura 4:	Exemplo 1: Parâmetros do controlador a cada iteração - $r_1(t)$ - Condição Inicial $\rho_1 = [2 \ 2]^T$	44
Figura 5:	Exemplo 1: Custo a cada iteração - $r_1(t)$	44
Figura 6:	Exemplo 1: Fase das funções de sensibilidade.	45
Figura 7:	Exemplo 1: Parâmetros do controlador a cada iteração - $r_2(t)$ - Condição Inicial $\rho_1 = [2 \ 2]^T$	46
Figura 8:	Exemplo 1: Custo a cada iteração - $r_2(t)$	46
Figura 9:	Exemplo 2: Resposta ao salto em malha-fechada - método proposto. . .	51
Figura 10:	Exemplo 2: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [2 \ 2]^T$	52
Figura 11:	Exemplo 2: Custo a cada iteração - método proposto.	53
Figura 12:	Exemplo 2: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [2 \ 2]^T$	54
Figura 13:	Exemplo 2: Resposta ao salto em malha-fechada - método clássico. . .	54
Figura 14:	Exemplo 3: Resposta em malha-fechada - método proposto.	58
Figura 15:	Exemplo 3: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [0, 4 \ 0, 4]^T$	58
Figura 16:	Exemplo 3: Custo a cada iteração - método proposto.	59
Figura 17:	Exemplo 3: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [0, 4 \ 0, 4]^T$	60
Figura 18:	Exemplo 3: Custo a cada iteração - método clássico.	60
Figura 19:	Exemplo 4: Resposta ao salto em malha-fechada - método proposto. . .	65
Figura 20:	Exemplo 4: Sinal de controle da resposta ao salto em malha-fechada - método proposto.	65
Figura 21:	Exemplo 4: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [1 \ 1]^T$	66
Figura 22:	Exemplo 4: Custo a cada iteração - método proposto.	67
Figura 23:	Exemplo 4: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [1 \ 1]^T$	67
Figura 24:	Exemplo 4: Resposta ao salto em malha-fechada - método clássico. . .	69
Figura 25:	Exemplo 4: Sinal de controle da resposta ao salto em malha-fechada - método clássico.	69
Figura 26:	Exemplo 4: Custo a cada iteração - método clássico.	70

Figura 27:	Exemplo 5: Resposta ao salto com controlador inicial.	73
Figura 28:	Exemplo 5: Simulação em malha-fechada.	73
Figura 29:	Exemplo 5: Resposta ao salto com controlador $\hat{C}_d(z)$	74
Figura 30:	Exemplo 5: Resposta ao salto com controlador final.	75
Figura 31:	Exemplo 5: Custo a cada iteração.	76
Figura 32:	Exemplo 6: Intervalo de confiança do modelo identificado.	78
Figura 33:	Exemplo 6: Tamanho do passo para limites do intervalo de confiança.	78
Figura 34:	Foto do processo.	80
Figura 35:	Experimento de identificação em malha-aberta.	82
Figura 36:	Resposta ao salto com controlador $C_1(z, \rho)$	83
Figura 37:	Resposta ao salto com controlador $C_2(z, \rho)$	84
Figura 38:	Resposta ao salto com controlador $C_3(z, \rho)$	85
Figura 39:	Resposta ao salto com controlador $C_4(z, \rho)$	86
Figura 40:	Resposta ao salto com controlador $C_5(z, \rho)$	86
Figura 41:	Resposta ao salto com controlador $C_6(z, \rho)$	88

LISTA DE TABELAS

Tabela 1:	Exemplo 2: Evolução dos controladores - método proposto.	53
Tabela 2:	Exemplo 3: Evolução dos controladores - método proposto.	59
Tabela 3:	Exemplo 4: Evolução dos controladores - método proposto.	68
Tabela 4:	Exemplo 5: Evolução dos controladores.	75
Tabela 5:	Exemplo 6: Estimativa de γ	77
Tabela 6:	Experimento 1: Evolução dos controladores - $r_1(t)$	83
Tabela 7:	Experimento 1: Evolução dos controladores - $r_2(t)$	84
Tabela 8:	Experimento 2: Evolução dos controladores - $r_3(t)$	87

LISTA DE ABREVIATURAS

CbT	Correlation based Tuning
FDT	Frequency Domain Tuning
IFT	Iterative Feedback Tuning
PE p	Persistentemente excitante de ordem p
PID	Proporcional-Integral-Derivativo
SPR	Estritamente positiva-real
VRFT	Virtual Reference Feedback Tuning

LISTA DE SÍMBOLOS

$eig(\cdot)$	Autovalor
$(\cdot)^*$	Conjugado Transposto
$\Phi(e^{j\omega})$	Espectro
∇	Gradiente
$min(\cdot)$	Mínimo
$\ \cdot \ $	Norma 2
z	Operador de avanço
ρ	Parâmetros do controlador
$\Re\{\cdot\}$	Parte real de um número complexo
$\Im\{\cdot\}$	Parte imaginária de um número complexo
γ	Tamanho do passo de iteração
$(\cdot)^T$	Transposto
$E(\cdot)$	Valor Esperado
$square\left(\frac{2\pi t}{N}\right)$	Onda quadrada de período N

SUMÁRIO

1	INTRODUÇÃO	13
2	CONTROLE BASEADO EM DADOS	16
2.1	Introdução	16
2.2	Definições	16
2.3	Problema de otimização	18
2.4	Algoritmo do gradiente	20
2.5	Métodos de projeto de controladores baseados em dados	21
2.5.1	VRFT - <i>Virtual Reference Feedback Tuning</i>	22
2.5.2	IFT - <i>Iterative Feedback Tuning</i>	24
2.5.3	FDT - <i>Frequency Domain Tuning</i>	26
2.5.4	CbT - <i>Correlation based Tuning</i>	27
2.6	Considerações finais	28
3	CONVERGÊNCIA DO ALGORITMO DO GRADIENTE	29
3.1	Introdução	29
3.2	Domínio de atração	29
3.3	Seguimento de referência	31
3.3.1	Modelo de referência atingível	31
3.3.2	Modelo de referência inatingível	33
3.4	Atenuação da sensibilidade ao ruído	34
3.4.1	Modelo de referência atingível	35
3.4.2	Modelo de referência inatingível	36
3.5	Minimização do esforço de controle	37
3.6	Combinando os três objetivos	39
3.7	Shaping da Função Custo	39
3.7.1	Controle cauteloso	40
3.7.2	Manipulação do espectro da referência	40
3.7.3	Escolha da janela de dados	45
3.8	Considerações finais	48
4	PROJETO DO PASSO DE ITERAÇÃO	49
4.1	Introdução	49
4.2	Seguimento de referência	49
4.2.1	Modelo de referência atingível	49
4.2.2	Modelo de referência inatingível	55
4.3	Atenuação da sensibilidade ao ruído	56

4.3.1	Modelo de referência atingível	56
4.3.2	Modelo de referência inatingível	60
4.4	Minimização do esforço de controle	62
4.5	Combinando os três objetivos	63
4.6	Considerações finais	70
5	ESTIMADOR DE $M(\rho)$	71
5.1	Introdução	71
5.2	Estimativa baseada em modelo	71
5.3	Aumentando a robustez	74
5.4	Considerações finais	77
6	EXPERIMENTOS	80
6.1	O processo	80
6.2	Shaping da Função Custo	82
6.3	Utilização do passo de iteração proposto	85
6.4	Considerações finais	87
7	CONCLUSÕES	89
	REFERÊNCIAS	90

1 INTRODUÇÃO

Muitos objetivos de controle podem ser expressos em termos de funções custo. Para resolver os problemas de otimização com frequência é necessário o conhecimento do processo e das perturbações que agem sobre ele. Na prática, raramente se conhece o processo e seus distúrbios, portanto, estas informações precisam ser obtidas. A partir de dados de experimentos pode-se identificar modelos para o processo e para os distúrbios. Estes modelos podem ser utilizados para o projeto de controladores ditos baseados em modelo.

A identificação de modelos tipicamente é feita ajustando parâmetros de uma estrutura fixa previamente escolhida, de modo que o modelo venha a representar da melhor forma possível o processo considerando aquele conjunto de dados (AGUIRRE, 2004) (LJUNG, 1987). A utilidade do modelo está no fato que ele pode ser utilizado para prever o comportamento do processo para outros pontos de operação, diferentes daqueles do conjunto de dados, no caso em que o processo é linear. Ao mesmo tempo que esta abordagem é poderosa, também possui suas deficiências. Um modelo incorreto pode gerar comportamentos indesejáveis e possivelmente catastróficos. O problema de como garantir que um modelo identificado seja suficientemente bom para o projeto de controladores é um assunto de grande interesse, tratado na literatura como identificação para controle (GEVERS, 1993) (HOF; SCHRAMA, 1995) (HJALMARSSON, 2005) (GEVERS, 2005). Outro fator limitante dos métodos baseados em modelo é que os controladores obtidos comumente possuem ordem próxima a do processo. Em aplicações industriais é bastante desejável utilizar controladores de estrutura fixa, como por exemplo, os controladores proporcional-integral-derivativo (PID).

No contexto de controladores com estrutura fixa, uma abordagem existente é a de redução do modelo do processo, para que a dimensão do controlador projetado também seja diminuída. De forma alternativa, controladores de ordem reduzida podem ser obtidos a partir de um controlador de ordem completa seguido de um passo de redução do controlador (ANDERSON; LIU, 1989).

Ainda no contexto de controladores baseados em modelos, existem os esquemas de identificação iterativa e projeto de controladores que também tratam do problema de controladores com estrutura fixa (RAYMOND A. DE CALLAFON, 1997) (LEE et al., 1993) (SCHRAMA; HOF, 1993). Estes esquemas iterativos identificam um modelo para o processo, atualizam o controlador por métodos baseados em modelo e utilizam este controlador no processo real. Por trás destes esquemas está o conceito de que experimentos em malha-fechada, com o controlador atual, devem gerar dados que são informativos para a identificação de um modelo, e que o controlador baseado neste modelo, que está a cada iteração mais preciso, deve gerar um comportamento em malha-fechada cada vez melhor. Maiores informações consultar (HOF; SCHRAMA, 1995) (GEVERS, 1993)

(HJALMARSSON, 2005).

Contudo, poucos trabalhos apresentam resultados que dão suporte para esses esquemas, exceto para o caso não realista de modelos e controladores de ordem plena. Em (HJALMARSSON; GEVERS; BRUYNE, 1996) foi mostrado que um controlador específico na malha pode gerar um controlador estimado que obtém o melhor desempenho possível para o sistema real. Além disso, um esquema iterativo é proposto para encontrar estas condições ideais. No caso de controladores de ordem reduzida, são encontrados alguns exemplos de sucesso, incluindo casos experimentais e industriais, para o acima mencionado esquema de identificação iterativa e projeto de controladores (PARTANEN; BITMEAD, 1995). Porém existem exemplos que estes esquemas não convergem para o controlador que minimiza a função custo, exceto para os casos de modelos e controladores de ordem completa (HJALMARSSON; GUNNARSSON; GEVERS, 1995) (LJUNG; SODERSTROM, 1983).

Os autores de (HJALMARSSON; GUNNARSSON; GEVERS, 1995), estudando as propriedades de convergência/divergência dos esquemas de identificação iterativa e projeto de controladores, tiveram a idéia de reformular o problema de otimização em outro direto nos parâmetros do controlador, abandonando o passo intermediário de identificação. Esta abordagem é análoga ao controle adaptativo direto, sendo as principais diferenças: a possibilidade de utilizar controladores de ordem reduzida e que a atualização dos parâmetros do controlador não é feita a cada amostragem, mas sim a cada batelada de dados. Neste tipo de abordagem, os parâmetros do controlador são obtidos na resolução de um problema de otimização, diretamente a partir de dados de experimentos realizados no processo. Daqui surge o nome *Controle baseado em Dados*.

Em (HJALMARSSON; GUNNARSSON; GEVERS, 1994) foi proposto um método de projeto de controladores baseado em dados chamado *Iterative Feedback Tuning* (IFT). Utilizando um critério de projeto H_2 , o problema de otimização é resolvido utilizando algoritmos iterativos baseados no gradiente da função custo. Um estimador não polarizado para o gradiente da função custo é proposto, de forma que os problemas de seguimento de referência, atenuação de ruído e minimização de esforço de controle são levados em conta. Em (KAMMER; BITMEAD; BARTLETT, 2000) um procedimento chamado *Frequency Domain Tuning* (FDT) foi proposto para minimizar o critério de desempenho H_2 para um sistema com referência nula. Nele uma abordagem parecida com a apresentada pelo IFT é proposta, diferenciando-se pelo fato de toda a sua análise ser feita no domínio da frequência. Novamente algoritmos iterativos são utilizados na resolução do problema de otimização. Já o método *Correlation based Tuning* (CbT) (KARIMI; MISKOVIC; BONVIN, 2004) utiliza o conceito de variáveis instrumentais que reduzem efeitos indesejados do ruído para alcançar o objetivo de seguimento de referência. A resolução do problema de otimização encontra os parâmetros do controlador que solucionam um sistema de equações de correlação.

Outro método não iterativo, mas também baseado em dados, é o *Virtual Reference Feedback Tuning* (VRFT) (CAMPI; LECCHINI; SAVARESI, 2002). Através de uma manipulação de variáveis o critério de desempenho H_2 é transformado em um critério quadrático, o qual pode ser minimizado diretamente sem a necessidade de um algoritmo iterativo. O método pode ser visto como um esquema de identificação dos parâmetros do controlador desejado.

Um fator limitante nas aplicações dos métodos iterativos é a possível falta de convergência para o mínimo global do critério. Comumente são utilizadas modificações no algoritmo de otimização para remediar esta falta de convergência. Estas modificações

muitas vezes são formas empíricas baseadas na experiência do projetista. Uma abordagem diferente foi dada ao problema de falta de convergência em (BAZANELLA et al., 2008): o foco é dado nas propriedades da função custo, o que permite analisar a existência de convergência dos algoritmos. Além disso é demonstrado que se a função custo for suficientemente “bem comportada” então é possível garantir convergência para o mínimo global do critério utilizando qualquer algoritmo correto. No mesmo trabalho são apresentadas formas de escolher parâmetros de projeto de forma a tornar a função custo mais “bem comportada”. Esta metodologia foi batizada de Shaping da Função Custo.

A escolha natural para o algoritmo que resolve o problema de otimização é o algoritmo do gradiente. Este algoritmo requer apenas uma estimativa para o gradiente da função custo para ser utilizado. Outros algoritmos como o de Newton ou quase-Newton, necessitam de mais informações sobre o processo, como o conhecimento da função Hessiana. Algoritmos mais complexos exigem um maior número de dados, e muitas vezes de experimentos mais complexos.

Neste trabalho os resultados obtidos em (BAZANELLA et al., 2008) foram estendidos de forma a garantir a convergência do algoritmo do gradiente para o mínimo global do critério. É mostrado que a convergência depende das condições iniciais do algoritmo bem como do tamanho do passo de iteração. Este trabalho tem por objetivo demonstrar o método de cálculo do tamanho do passo “ótimo” para o algoritmo do gradiente, ou seja, o maior tamanho de passo para o qual a convergência para o mínimo global da função custo é garantida. Escolhendo o tamanho do passo dessa maneira, o algoritmo converge tão rápido quanto possível.

Esta dissertação apresenta no Capítulo 2 definições utilizadas no trabalho, o algoritmo do gradiente e detalhes sobre alguns métodos de projeto de controladores baseado em dados. O Capítulo 3 apresenta os resultados de (BAZANELLA et al., 2008) que apresenta propriedades de convergência do algoritmo do gradiente. O método de escolha do tamanho do passo de iteração, que é o resultado principal deste trabalho, é apresentado no Capítulo 4. O Capítulo 5 mostra como obter os dados necessários para efetivamente computar o tamanho do passo de iteração com o procedimento proposto. O capítulo 6 apresenta resultados experimentais dos tópicos tratados neste trabalho. No Capítulo 7 estão as conclusões e perspectivas de trabalhos futuros.

2 CONTROLE BASEADO EM DADOS

2.1 Introdução

Muitos métodos de projeto de controladores são baseados em problemas de otimização onde deseja-se encontrar o argumento que minimiza uma dada função custo. Esta função custo representa os requisitos de desempenho do sistema em malha-fechada, de forma que quanto menor o custo, mais próximo está o sistema em malha-fechada do comportamento que é desejado.

Para resolver estes problemas de otimização, é necessário em boa parte dos métodos, o conhecimento completo do processo e das características das perturbações. Frequentemente ainda é necessário que a estrutura do controlador possa ser livremente escolhida. Estes requisitos comumente não são observados na prática, o que motivou o desenvolvimento de métodos de projeto de controladores de estrutura fixa para processos dos quais os modelos são parcialmente ou completamente desconhecidos. Nestes métodos o problema de otimização é realizado diretamente a partir de dados coletados de um experimento, o que motiva a denominação de *controle baseado em dados*.

Nos últimos anos vários métodos de projeto de controladores baseado em dados têm surgido na literatura. Cada método apresenta uma abordagem diferente para o problema e utiliza uma função custo específica. Esta função custo pode expressar um ou mais objetivos fundamentais de controle: seguimento de referência, atenuação de ruído e minimização do esforço de controle.

Neste capítulo são apresentadas os conceitos que fundamentam o controle baseado em dados de um ponto de vista comum para todos os métodos. Na seqüência será apresentado o algoritmo do gradiente, que é um algoritmo iterativo utilizado para minimização da função custo. Por fim serão descritas características específicas de cada um dos métodos de projeto de controladores baseado em dados.

2.2 Definições

Considera-se um processo monovariável linear discreto e invariante no tempo, dado por

$$y(t) = G(z)u(t) + \nu(t), \quad (1)$$

onde z é o operador de avanço no tempo, $G(z)$ é a função de transferência discreta do processo, $u(t)$ é a entrada de controle e $\nu(t)$ é o ruído de processo. O ruído é um processo quase-estacionário descrito por $\nu(t) = H(z)e(t)$, onde $e(t)$ é ruído branco com variância σ_e^2 . Um sinal é dito quase-estacionário se a média e autocorrelação convergem para um valor finito quando o tamanho da amostra cresce. As funções de transferência $G(z)$ e

$H(z)$ são funções racionais, próprias e desconhecidas. Assume-se que $H(\infty) = 1$, ou seja, que a resposta impulsiva $h(t)$ do filtro $H(z)$ satisfaz $h(0) = 1$.

Esse processo é controlado por um controlador com um grau de liberdade dado por:

$$u(t) = C(z, \rho)(r(t) - y(t)), \quad (2)$$

onde $r(t)$ é o sinal de referência e ρ é o vetor de parâmetros do controlador. Na Figura 1 pode-se visualizar um diagrama com os sinais e blocos do sistema.

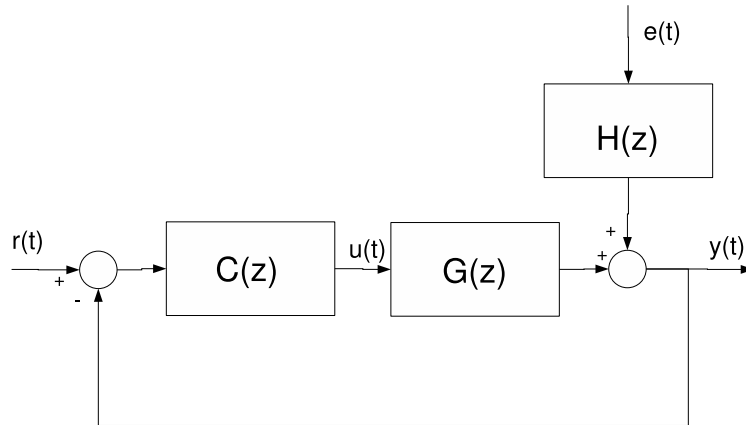


Figura 1: Diagrama de Blocos

Definição 1 (LJUNG, 1987) Um sinal $s(t)$ é dito quasi-estacionário se:

- $E[s(t)] = m_s(t), |m_s(t)| \leq C, \forall t$
- $E[s(t)s(r)] = R_s(t, r), |R_s(t, r)| \leq C, \forall t$
- $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_s(t, \tau) = R_s(\tau), \forall t$

Assume-se que $r(t)$ é quasi-estacionário e não correlacionado com o ruído, então

$$\bar{E}[r(t)e(\varsigma)] = 0 \quad \forall t, \varsigma,$$

onde é definido para um sinal quasi-estacionário, como em (LJUNG, 1987), que

$$\bar{E}[f(t)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E[f(t)]$$

onde $E[f(t)]$ significa o valor esperado.

Assume-se que o controlador é linear e invariante no tempo com estrutura fixa e pertence a uma classe \mathcal{C} de funções de transferência lineares que é escolhida pelo projetista. Essa classe é tal que $C(z)G(z)$ tenha grau relativo positivo para todo $C(z) \in \mathcal{C}$. Assume-se que a estrutura é tal que a parametrização seja linear.

Hipótese 1 *Parametrização Linear*

$$C(z, \rho) = \rho^T \bar{C}(z),$$

onde $\rho = [\rho_1 \ \rho_2 \ \dots \ \rho_p]^T \in \mathbb{R}^p$ é um vetor que representa os parâmetros do controlador que podem ser ajustados. O vetor $\bar{C}(z) = [\bar{C}_1(z) \ \bar{C}_2(z) \ \dots \ \bar{C}_p(z)]^T$ é composto de funções de transferência lineares conhecidas.

Os controladores industriais comumente são parametrizados de maneira linear. O controlador PID com pólo derivativo fixo, por exemplo, pode ser escrito como

$$C(z, \rho) = [k_p \ k_i \ k_d] \begin{bmatrix} 1 & \frac{z}{z-1} & \frac{z-1}{z} \end{bmatrix}^T.$$

O sistema (1)-(2) em malha-fechada é dado por,

$$y(t, \rho) = T(z, \rho)r(t) + S(z, \rho)v(t),$$

onde

$$T(z, \rho) = \frac{C(z, \rho)G(z)}{1 + C(z, \rho)G(z)} = C(z, \rho)G(z)S(z, \rho).$$

É necessário incluir ainda algumas outras definições. Um sinal $x(t)$ quase-estacionário é persistentemente excitante de ordem k - PE k - se o seu espectro $\Phi_x(e^{j\omega})$ tem pelo menos k componentes diferentes de zero (AGUIRRE, 2004). Um campo vetorial

$$V(e^{j\omega}) = [v_1(e^{j\omega}) \ v_2(e^{j\omega}) \ \dots \ v_k(e^{j\omega})]^T$$

, onde cada $v_i(e^{j\omega})$ é uma função da variável ω , tem posto completo se as funções $v_i(e^{j\omega})$ formam um conjunto linearmente independente (LI) sobre os reais, ou seja, $\nexists \eta \in \mathbb{R}^k, \eta \neq 0 : \eta^T V(e^{j\omega}) = 0 \ \forall \omega \in \mathbb{R}$.

2.3 Problema de otimização

Uma maneira clássica de especificar o desempenho desejado para o sistema em malha-fechada é por meio de um modelo de referência. Este modelo de referência nada mais é do que a função de transferência desejada para o sistema em malha-fechada:

$$y_d(t) = T_d(z)r(t)$$

Uma maneira de formular o problema de otimização, com o objetivo encontrar o controlador que torna o sistema em malha-fechada o mais próximo possível do modelo de referência, é por meio de uma função custo que dependa da diferença entre a saída do sistema em malha-fechada e do modelo de referência.

Uma maneira clássica de escrever esta diferença é por meio do erro quadrático (ANDERSON; MOORE, 1971). A função custo pode ser dada por

$$J(\rho) = \bar{E} [(y(t, \rho) - y_d(t))^2].$$

Se for desejado que o problema de otimização minimize também o esforço de controle é possível acrescentar mais um termo nesta função custo

$$J(\rho) \triangleq \bar{E} [\lambda \tilde{y}(t, \rho)^2 + (1 - \lambda)u(t, \rho)^2] \quad (3)$$

$$\tilde{y}(t, \rho) = y(t, \rho) - y_d(t), \quad (4)$$

onde $\lambda \in [0 \ 1]$ é uma constante pré-definida pelo projetista que determina um peso entre o seguimento do modelo de referência e a minimização do esforço de controle.

Este tipo de função custo é utilizado por muitos métodos de projeto de controladores baseados em modelo, tais como LQR/LQG (ANDERSON; MOORE, 1971) e controle preditivo generalizado (BITMEAD; GEVERS; WERTZ, 1990), nos quais a solução de (3)

pode ser obtida utilizando ferramentas como equações de Riccati, Inequações Matriciais Lineares (LMI), Inequações Matriciais Bilineares (BMI) (BOYD et al., 1994).

Considerando que o sinal de referência e o ruído são descorrelacionados, a função custo (3) pode ser dividida em três partes (GOODWIN; SIN, 1984):

$$J(\rho) = \lambda [J_y(\rho) + J_e(\rho)] + (1 - \lambda)J_u(\rho), \quad (5)$$

onde são definidos

$$\begin{aligned} J_y(\rho) &= \bar{E} [((T(z, \rho) - T_d(z))r(t))^2], \\ J_e(\rho) &= \bar{E} [(S(z, \rho)\nu(t))^2] e \\ J_u(\rho) &= \bar{E} [(u(t, \rho))^2] \end{aligned}$$

Cada uma das parcelas em (5) representa o desempenho do sistema em malha-fechada com respeito a cada um dos objetivos fundamentais de um sistema de controle: seguimento de referência - representado por $J_y(\rho)$, minimização da sensibilidade ao ruído - $J_e(\rho)$ - e minimização de esforço de controle - $J_u(\rho)$. Pode-se estudar em separado cada um desses objetivos de controle.

O problema de seguimento de referência surge da necessidade que a resposta do sistema em malha-fechada seja a mais próxima possível de uma determinada referência não nula. Se apenas o problema de seguimento de referência for considerado, o controlador ideal $C_d(z)$ com o qual o sistema em malha-fechada atinge exatamente o modelo de referência $T_d(z)$ é dado por

$$C_d(z) = \frac{T_d(z)}{G(z)(1 - T_d(z))}. \quad (6)$$

Considerando Γ o conjunto de todos os parâmetros dos controladores que tornam o sistema em malha-fechada BIBO-estável pode-se reescrever a parcela do seguimento de referência utilizando o teorema de Parseval. Se $\rho \in \Gamma$:

$$J_y(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |T(e^{j\omega}, \rho) - T_d(e^{j\omega})|^2 \Phi_r(e^{j\omega}) d\omega. \quad (7)$$

onde $\Phi_r(e^{j\omega})$ é o espectro de frequência do sinal de referência e $T(e^{j\omega}, \rho)$ é a DTFT do sistema em malha-fechada $T(z)$.

O projeto de controladores pela minimização do custo associado ao ruído - $J_e(\rho)$ - é tópico clássico na literatura conhecido por controle de mínima variância. Uma versão adaptativa do problema é apresentada em (ASTROM; WITTENMARK, 1973). Dado um sistema com referência nula, deseja-se que o ruído $\nu(t)$ presente no sistema seja atenuado. A solução para este problema, quando $G(z)$ é de fase mínima, é dado pelo seguinte controlador (ASTROM, 1970):

$$C_e(z) = \frac{H(z) - 1}{G(z)}. \quad (8)$$

Novamente podemos utilizar o teorema de Parseval para escrever a parcela da minimização de ruído, para $\rho \in \Gamma$, como:

$$J_e(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S(e^{j\omega}, \rho)|^2 \Phi_\nu(e^{j\omega}) d\omega.$$

onde $\Phi_\nu(e^{j\omega})$ é o espectro de freqüência do sinal do ruído.

O problema de minimização de esforço de controle muitas vezes aparece combinado com os outros problemas citados anteriormente. A solução ótima para este problema isolado é $\rho = 0$, o que significa que nenhuma ação de controle é utilizada. Esta solução pode ser encontrada para qualquer controlador linearmente parametrizado. Utilizando o teorema de Parseval esta parcela pode ser escrita como para $\rho \in \Gamma$:

$$J_u(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |C(e^{j\omega}, \rho)|^2 |S(e^{j\omega}, \rho)|^2 [\Phi_r(e^{j\omega}) + \Phi_\nu(e^{j\omega})] d\omega. \quad (9)$$

O problema de otimização pode ser definido por encontrar os parâmetros do controlador que minimizam a função custo (3). O projeto do controlador se resume em resolver o seguinte problema de otimização:

$$\begin{aligned} \min_{\rho} J(\rho) \\ J(\rho) \triangleq \bar{E} [\lambda \tilde{y}(t, \rho)^2 + (1 - \lambda)u(t, \rho)^2]. \end{aligned} \quad (10)$$

Diversos algoritmos foram propostos para resolver este problema de otimização. Como a função custo raras vezes é convexa, esses algoritmos podem convergir para mínimos locais da função custo. Algoritmos cada vez mais complexos têm sido propostos para evitar que a convergência para esses mínimos locais ocorra (KAMMER; BITMEAD; BARTLETT, 2000) (KARIMI; MISKOVIC; BONVIN, 2004).

2.4 Algoritmo do gradiente

Vários algoritmos existem para resolver o problema de otimização apresentado na seção anterior. Deseja-se, utilizando esses algoritmos, encontrar o mínimo global da função custo, no menor número possível de iterações. Muitos algoritmos são baseados no gradiente da função custo para alcançar um mínimo da função, como o algoritmo do gradiente, algoritmo de Newton e variações destes algoritmos.

O algoritmo do gradiente apenas utiliza informação do gradiente da função custo enquanto que os outros algoritmos necessitam de informações adicionais como a hessiana. Para obter estas informações são utilizados experimentos mais longos e mais complexos, o que vai contra a principal vantagem dos métodos baseados em dados, que é a simplicidade. Nesta linha, o algoritmo do gradiente se sobressai e foi escolhido como foco neste trabalho.

A convergência do algoritmo do gradiente para o mínimo global da função custo não existe na bibliografia, decidiu-se por isso mesmo estudá-la. Propõe-se estudar com profundidade a função custo e estabelecer métodos que garantam a convergência para o mínimo global da função custo, utilizando o algoritmo do gradiente.

O algoritmo do gradiente foi proposto inicialmente por Cauchy (CAUCHY, 1847) para encontrar mínimos de funções. Neste algoritmo as iterações são dadas por:

$$\rho_{i+1} = \rho_i - \gamma_i \nabla J(\rho_i), \quad (11)$$

onde $\nabla J(\rho) = \frac{\partial J(\rho)}{\partial \rho}$ e $\gamma_i > 0$ é o tamanho do passo dado a cada iteração. O raciocínio por trás deste algoritmo é simples: sucessivas atualizações são feitas no sentido do gradiente mas em sua direção oposta, com isso, pelo menos para γ_i pequeno o suficiente, um valor menor do custo é alcançado a cada iteração.

Nota-se que não é necessário o conhecimento da função custo para utilizar este algoritmo. A única informação necessária é o valor do gradiente da função custo no ponto ρ_i . A função custo (3) e seu gradiente dependem do modelo do processo, o qual é desconhecido. Na prática é utilizado um estimador para o gradiente da função custo dado por $\widehat{\nabla J}(\rho)$. Os vários métodos Iterative Feedback Tuning, Correlation based Tuning e Frequency Domain Tuning propõem metodologias diferentes para o cálculo deste estimador do gradiente da função custo.

A convergência do algoritmo do gradiente para o mínimo global da função custo depende tanto das condições iniciais do algoritmo (ρ_0) quanto do tamanho dos passos de iteração γ_i .

Foram encontrados na bibliografia algumas formas de escolher o tamanho do passo de iteração, no entanto, nenhuma delas se preocupa em tornar a convergência mais rápida, apenas tratam de garantir a convergência. Em (HJALMARSSON et al., 1998) é proposto um método baseado em (ROBBINS; MONRO, 1951) que garante a convergência para um mínimo local da função custo. Este método garante a convergência estocástica do algoritmo, se for utilizada uma estimativa não polarizada para o gradiente da função custo. Nesse método o tamanho do passo deve respeitar duas condições:

$$\sum_{i=1}^{\infty} \gamma_i = \infty \text{ e } \sum_{i=1}^{\infty} \gamma_i^2 < \infty$$

Foi provado que respeitando estas condições o tamanho do passo decresce suficientemente rápido para garantir a convergência. Encontrar uma seqüência de tamanhos de passo que satisfazem as duas condições não é uma tarefa simples. A série harmônica satisfaz estas condições, e foi a única encontrada como exemplo. O tamanho do passo é escolhido de forma que:

$$\gamma_i = \frac{\gamma_0}{i}; i > 0.$$

O artigo (HILDEBRAND et al., 2005) apresenta uma análise da taxa de convergência do algoritmo, para o caso de atenuação de ruído, utilizando como tamanho do passo de iteração a expressão acima. Como pode ser observado, esse método gera passos muito pequenos após algumas iterações, o que torna a taxa de convergência muito lenta.

No contexto de minimização de funções convexas e quase-convexas são utilizados os tradicionais algoritmos de busca linear (CURRY, 1944), (GOLDSTEIN, 1962), (ARMJO, 1966), (POLAK, 1973) e (KIWIEL; MURTY, 1996), com os quais dada uma direção do gradiente, testa-se inúmeros valores do passo e escolhe-se aquele que encontrar o primeiro mínimo local na direção do gradiente. A problemática desta metodologia está na necessidade de muitos testes e, portanto, de muitos experimentos o que torna o método muito custoso e lento quando utilizado para sintonia de controladores.

Um estudo detalhado da convergência para o mínimo global da função custo é descrito no Capítulo 3 enquanto que no Capítulo 4 será proposto um método para a escolha do tamanho do passo de iteração com o objetivo de garantir a convergência para o mínimo global da função custo no menor número possível de iterações.

2.5 Métodos de projeto de controladores baseados em dados

Nesta seção serão apresentados os métodos mais conhecidos de projeto de controladores baseado em dados. O primeiro a ser apresentado, o VRFT, é um método direto e

portanto não utiliza algoritmos iterativos. Os outros três métodos são iterativos e utilizam algoritmos do gradiente e de Newton para encontrar os mínimos da função custo.

Uma vantagem do método VRFT é que por não ser iterativo, são necessários poucos experimentos para o projeto do controlador. Com apenas uma batelada de dados determina-se os novos parâmetros do controlador, com o intuito de alcançar o modelo de referência.

Essa vantagem vista pelo lado do custo de experimentação representa uma desvantagem considerando o lado da segurança do processo. Com os métodos iterativos pequenas modificações são feitas nos parâmetros do controlador a cada passo. Dessa forma espera-se que o comportamento do sistema em malha-fechada mude apenas um pouco a cada iteração. Essa operação possui vantagens do ponto de vista de segurança do processo pois mudanças bruscas nos parâmetros do controlador podem ser vistas como perturbações no processo, que podem causar comportamentos bastante indesejáveis. Como podem existir dinâmicas não observadas e não-linearidades do processo, pequenas variações consecutivas são mais seguras que procedimentos baseados em uma única batelada de dados.

2.5.1 VRFT - *Virtual Reference Feedback Tuning*

O VRFT é um método direto de projeto de controladores baseado em dados, ou seja, não é um método iterativo. Apenas um ensaio é realizado para obter os dados necessários para o projeto do controlador (CAMPI; LECCHINI; SAVARESI, 2002). Uma aplicação em um sistema de suspensão ativa pode ser visto em (CAMPI; LECCHINI; SAVARESI, 2003). A extensão para processos não lineares é apresentada em (CAMPI; SAVARESI, 2006). O objetivo do método é o seguimento de referência, portanto, apenas a parcela $J_y(\rho)$ do custo é utilizada.

Utilizam-se dados coletados da planta, contendo medidas de entrada $u(t)$ e as correspondentes medidas na saída da planta $y(t)$ no projeto do controlador. Dada a medida $y(t)$, computa-se uma referência $\bar{r}(t)$ tal que $T_d(z)\bar{r}(t) = y(t)$, onde $T_d(z)$ é o modelo de referência desejado para o sistema em malha fechada. Essa referência é chamada de virtual pois a mesma não é usada para gerar $y(t)$. Nota-se que $y(t)$ é a saída desejada do sistema em malha fechada quando a referência utilizada nesse sistema é $\bar{r}(t)$.

Então, calcula-se o erro de seguimento de referência $e(t) = \bar{r}(t) - y(t)$. Mesmo que a planta $G(z)$ seja desconhecida, sabe-se que quando $G(z)$ é alimentada com $u(t)$ (sinal de entrada medido), a mesma gera $y(t)$ como sinal de saída. Assim, um bom controlador é aquele que gera $u(t)$ quando alimentado por $e(t)$. A idéia então é procurar tal controlador. Uma vez que ambos sinais $u(t)$ e $e(t)$ são conhecidos, o projeto resume-se ao problema de identificação que descreve uma relação dinâmica entre $e(t)$ e $u(t)$.

Dessa forma, o método VRFT pode ser descrito em dois passos:

1. Dadas as medidas de entrada e saída da planta $\{u(t), y(t)\}_{t=1, \dots, N}$ calcular:
 - uma referência virtual $\bar{r}(t)$ tal que $y(t) = T_d(z)\bar{r}(t)$, e
 - o erro correspondente $e(t) = \bar{r}(t) - y(t)$, assumindo $T_d(z) \neq 1$;
2. Selecionar o vetor de parâmetros do controlador $\hat{\rho}_N$ que minimiza o seguinte critério

$$J_{VR}^N(\rho) = \frac{1}{N} \sum_{t=1}^N (u(t) - C(z, \rho)e(t))^2$$

Note que quando $C(z, \rho) = \rho^T \bar{C}(z)$, o critério de minimização pode ser descrito por

$$J_{VR}^N(\rho) = \frac{1}{N} \sum_{t=1}^N (u_L(t) - \rho^T \varphi^T(t))^2,$$

onde $\varphi(t) = \bar{C}(z)e_L(t)$ e o vetor de parâmetros $\hat{\rho}_N$ é dado por

$$\hat{\rho}_N = \left[\sum_{t=1}^N \varphi(t)\varphi(t)^T \right]^{-1} \sum_{t=1}^N \varphi(t)u_L(t)$$

e os sinais $e_L(t)$ e $u_L(t)$ são os sinais $e(t)$ e $u(t)$ filtrados por um apropriado filtro $L(z)$. Ver detalhes em (CAMPI; LECCHINI; SAVARESI, 2002).

2.5.1.1 Uso de dados com ruído

Quando saída do processo é afetada por ruído, o vetor de parâmetros obtido pelo método acima é polarizado, resultando em uma deterioração significativa do desempenho do sistema. Isto pode ser verificado através da análise da expressão do critério $J_{VR}(\rho)$ no domínio da frequência, quando utilizados dados com e sem ruído:

- Critério assintótico usando dados sem ruído $\{u(t), y(t)\}_{t=1, \dots, N}$:

$$J_{VR}(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G|^2 |C(\rho) - C_d|^2 |1 - T_d|^2 \frac{|L|^2}{|T_d|^2} \Phi_u d\omega. \quad (12)$$

- Critério assintótico usando dados com ruído $\{u(t), y(t)\}_{t=1, \dots, N}$:

$$J_{VR}(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|G|^2 |C(\rho) - C_d|^2 |1 - T_d|^2 \frac{|L|^2}{|T_d|^2} \Phi_u + \frac{|C(\rho)|^2}{|G|^2 |C_d|^2} |L|^2 \Phi_\nu \right] d\omega. \quad (13)$$

onde Φ_ν é a densidade espectral do ruído.

Aparentemente, os mínimos de (12) e (13) são diferentes pelo fato de o termo relacionado ao ruído em (13) depender de ρ . Para neutralizar o efeito do ruído, propõe-se a utilização de uma variável instrumental (LJUNG, 1987).

Seja $\zeta(t)$ a variável instrumental e

$$\varphi_L(t) = \bar{C}(z)L(z) (T_d^{-1}(z) - 1) y(t),$$

o vetor dos regressores filtrado, o vetor de parâmetros é estimado utilizando as variáveis instrumentais de acordo com a equação

$$\hat{\rho}_N^{IV} = \left[\sum_{t=1}^N \zeta(t)\varphi_L(t)^T \right]^{-1} \left[\sum_{t=1}^N \zeta(t)u_L(t) \right].$$

Em (CAMPI; LECCHINI; SAVARESI, 2002) duas maneiras diferentes de se escolher as variáveis instrumentais são propostas. A primeira garante que $\hat{\rho}^{IV} = \hat{\rho}$ assintoticamente. Porém, é necessário um experimento adicional na planta. A segunda não garante que $\hat{\rho}^{IV} = \hat{\rho}$ rigorosamente, mas presume-se que o erro residual seja pequeno. Nesse caso, não é necessário ensaio adicional na planta. As duas maneiras propostas podem ser descritas por:

- *Experimento adicional.* Realiza-se um segundo experimento na planta, utilizando o mesmo sinal de entrada $\{u(t)\}_{t=1,\dots,N}$ e coleta-se a seqüência de dados da saída correspondente $\{y'(t)\}_{t=1,\dots,N}$. Com estes dados, obtém-se as variáveis instrumentais através de:

$$\zeta(t) = \bar{C}(z)L(z) (T_d^{-1}(z) - 1) y'(t) \quad (14)$$

Nota-se que $\{y'(t)\}_{t=1,\dots,N}$ é diferente de $\{y(t)\}_{t=1,\dots,N}$, uma vez que as duas seqüências são afetadas por realizações diferentes de ruído. Porém, assume-se que os ruídos nos dois experimentos são não-correlacionados e, portanto, (14) proporciona $\hat{\rho}$ assintoticamente, como no caso sem ruído.

- *Identificação da planta.* Identifica-se um modelo $\hat{G}(z)$ da planta por meio dos dados $\{u(t), y(t)\}_{t=1,\dots,N}$ e simula-se a saída através de $\hat{y}(t) = \hat{G}(z)u(t)$. Então, constrói-se a variável instrumental como:

$$\zeta(t) = \bar{C}(z)L(z) (T_d^{-1}(z) - 1) \hat{y}(t)$$

Devido à possibilidade de erros na estimativa de $\hat{G}(z)$, este método não garante que a estimativa tenda assintoticamente para $\hat{\rho}$.

2.5.2 IFT - Iterative Feedback Tuning

O método de projeto de controladores *Iterative Feedback Tuning* foi proposto inicialmente em (HJALMARSSON; GUNNARSSON; GEVERS, 1994). Uma discussão profunda é encontrada em (HJALMARSSON et al., 1998) e (HJALMARSSON, 2002), e a extensão que engloba características de robustez em (PROCHAZKA et al., 2005). O método otimiza uma função custo geral H_2 (3). Um estimador não polarizado do gradiente da função custo é obtido diretamente de dados de experimentos realizados em malha-fechada. Esta estimativa do gradiente é utilizada no algoritmo do gradiente para encontrar um mínimo da função custo.

O método IFT descreve um estimador do gradiente para a função custo (3). O gradiente desta função pode ser escrito como

$$\nabla J(\rho) = 2\lambda \bar{E}[\tilde{y}(t, \rho) \nabla \tilde{y}(t, \rho)] + 2(1 - \lambda) E[u(t, \rho) \nabla u(t, \rho)]. \quad (15)$$

onde ∇ é o operador de derivada parcial com relação a vetor de parâmetros do controlador $\nabla = \frac{\partial}{\partial \rho}$.

Na equação (15) pode-se notar que $\nabla J(\rho)$ depende de $\tilde{y}(\rho)$ que é a diferença entre a saída do sistema em malha-fechada e a saída desejada, $u(\rho)$ que é o sinal de controle além dos respectivos gradientes destes sinais $\nabla \tilde{y}(\rho)$ e $\nabla u(\rho)$.

A partir de um simples experimento realizado em malha-fechada é possível coletar dados e obter os sinais $\tilde{y}(\rho)$ e $u(\rho)$. Os sinais dos gradientes $\nabla \tilde{y}(\rho)$ e $\nabla u(\rho)$ podem ser computados utilizando também dados coletados de experimentos realizados em malha-fechada, como será descrito na seqüência.

Em relação a $\nabla \tilde{y}(\rho)$, tem-se a seguinte expressão

$$\nabla \tilde{y}(\rho) = \frac{1}{C(z, \rho)} \nabla C(z, \rho) [T(z, \rho)r(t) - T(z, \rho)y(t, \rho)] \quad (16)$$

Nessa expressão, as quantidades $C(z, \rho)$, $\nabla C(z, \rho)$ e $r(t)$ são conhecidas *a priori*, enquanto que $T(z, \rho)$ e $y(t, \rho)$ dependem de informações do processo que é desconhecido. Assim, para o cálculo de $\nabla \tilde{y}(\rho)$ é necessário obter informações sobre os termos $T(z, \rho)y(t, \rho)$ e $T(z, \rho)r(t)$ presentes em (16).

O termo $T(z, \rho)y(t, \rho)$ representa o sinal de saída filtrado pelo sistema em malha-fechada. Já o termo $T(z, \rho)r(t)$ representa o sinal referência filtrado pelo sistema em malha-fechada. Destas expressões decorre a idéia de realizar experimentos no sistema em malha-fechada para obter informações destes termos desconhecidos.

Para computar $T(z, \rho)y(t, \rho)$ são utilizados dois experimentos. O primeiro experimento é realizado em malha-fechada do qual se obtém o sinal de saída

$$y_1(t, \rho) = T(z, \rho)r(t) + S(z, \rho)\nu_1(t)$$

onde o índice 1 indica que o sinal foi obtido no primeiro experimento. Este sinal de saída do primeiro experimento é utilizado como sinal de referência no segundo experimento, de forma a obter

$$y_2(t, \rho) = T(z, \rho)y_1(t, \rho) + S(z, \rho)\nu_2(t).$$

Como pode-se ver, o sinal $y_2(t, \rho)$ é composto pelo termo $T(z, \rho)y_1(t, \rho)$, o qual procura-se conhecer, adicionado do ruído filtrado $S(z, \rho)\nu_2(t)$.

Já para computar o termo $T(z, \rho)r(t)$ de (16) é realizado um terceiro experimento em malha-fechada. Neste experimento o sinal de referência é o mesmo utilizado no primeiro experimento. Como sinal de saída tem-se

$$y_3(t, \rho) = T(z, \rho)r(t) + S(z, \rho)\nu_3(t).$$

Este sinal $y_3(t, \rho)$ é composto pelo termo $T(z, \rho)r(t)$ adicionado do ruído filtrado $S(z, \rho)\nu_3(t)$.

Os sinais $y_2(t, \rho)$ e $y_3(t, \rho)$ são então utilizados para gerar a seguinte estimativa:

$$\widehat{\nabla \tilde{y}}(t, \rho) = \frac{\nabla C(z, \rho)}{C(z, \rho)}(y_3(t, \rho) - y_2(t, \rho))$$

que é uma “versão perturbada” de $\nabla \tilde{y}(y, \rho)$ por $\nu_2(t)$ e $\nu_3(t)$ uma vez que

$$\widehat{\nabla \tilde{y}}(t, \rho) = \nabla \tilde{y}_i(t, \rho) + \frac{S(z, \rho)\nabla C(z, \rho)}{C(z, \rho)}(\nu_3(t) - \nu_2(t)). \quad (17)$$

Na equação (17) pode-se ver que apenas os ruídos do segundo e terceiro experimentos introduzem erros na estimativa do gradiente de $\tilde{y}(t, \rho)$. O sinal de saída do primeiro experimento, com o ruído, é exatamente o sinal necessário para ser utilizado como referência no segundo experimento e, portanto, o ruído não é prejudicial.

O procedimento de cálculo de $\widehat{\nabla u}(\rho)$ é similar ao de $\widehat{\nabla \tilde{y}}(t, \rho)$.

A partir dos sinais definidos anteriormente, uma estimativa do gradiente $\widehat{\nabla J}(\rho)$ pode ser obtida por

$$\widehat{\nabla J}(\rho) = \frac{2}{N} \sum_{t=1}^N \left(\lambda \tilde{y}(t, \rho) \widehat{\nabla \tilde{y}}(t, \rho) + (1 - \lambda) u(t) \widehat{\nabla u}(t, \rho) \right).$$

Além disso, como $\tilde{y}(t, \rho)$ é descorrelacionado de $\widehat{\nabla \tilde{y}}(t, \rho)$ e $u(t)$ é descorrelacionado de $\widehat{\nabla u}(t, \rho)$ tem-se que

$$E[\widehat{\nabla J}(\rho)] = \nabla J(\rho).$$

Esta condição é necessária para que a aproximação estocástica do gradiente possa ser utilizada no algoritmo. Esta é uma condição chave nas provas de convergência para mínimos locais, utilizando aproximações do gradiente, como apresentado em (ROBBINS; MONRO, 1951) (HJALMARSSON et al., 1998).

2.5.3 FDT - *Frequency Domain Tuning*

O método de projeto de controladores *Frequency domain tuning* é utilizado para melhorar o desempenho de sistemas em malha-fechada com referência nula, ou seja, não trata do problema de seguimento de referência (KAMMER; BITMEAD; BARTLETT, 2000). A partir de uma análise espectral do problema é proposto um estimador não polarizado para o gradiente da função custo.

O método FDT descreve um estimador para o gradiente da seguinte função custo:

$$J(\rho) = \lambda J_e(\rho) + (1 - \lambda) J_u(\rho).$$

A partir do Teorema de Parseval esta função custo pode ser descrita por

$$J(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \lambda \Phi_y(e^{j\omega}, \rho) + (1 - \lambda) \Phi_u(e^{j\omega}, \rho) d\omega.$$

Como a referência tem valor nulo os espectros podem ser descritos por

$$\begin{aligned} \Phi_y(e^{j\omega}, \rho) &= \left| \frac{H(e^{j\omega})}{1 + G(e^{j\omega})C(e^{j\omega}, \rho)} \right|^2 \sigma_e^2, \\ \Phi_u(e^{j\omega}, \rho) &= \left| \frac{H(e^{j\omega})C(e^{j\omega}, \rho)}{1 + G(e^{j\omega})C(e^{j\omega}, \rho)} \right|^2 \sigma_e^2. \end{aligned}$$

A função custo pode agora se representada por:

$$J(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{j\omega}, \rho) h(e^{j\omega}, \rho) d\omega.$$

onde

$$\begin{aligned} g(e^{j\omega}, \rho) &= \lambda + (1 - \lambda) |C(e^{j\omega}, \rho)|^2, \\ h(e^{j\omega}, \rho) &= \frac{|H(e^{j\omega})|^2}{|1 + G(e^{j\omega})C(e^{j\omega}, \rho)|^2} \sigma_e^2 = \Phi_y(e^{j\omega}, \rho). \end{aligned}$$

O gradiente da função custo é calculado como:

$$\nabla J(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial g(e^{j\omega}, \rho)}{\partial \rho} h(e^{j\omega}, \rho) + g(e^{j\omega}, \rho) \frac{\partial h(e^{j\omega}, \rho)}{\partial \rho} \right] d\omega. \quad (18)$$

onde

$$\begin{aligned} \frac{\partial g(e^{j\omega}, \rho)}{\partial \rho} &= 2(1 - \lambda) \Re \left\{ C(e^{j\omega}, \rho)^* \frac{\partial C(e^{j\omega}, \rho)}{\partial \rho} \right\}, \\ \frac{\partial h(e^{j\omega}, \rho)}{\partial \rho} &= -2\Phi_y(e^{j\omega}, \rho) \Re \left\{ \frac{G(e^{j\omega})}{1 + C(e^{j\omega}, \rho)G(e^{j\omega})} \frac{\partial C(e^{j\omega}, \rho)}{\partial \rho} \right\}. \end{aligned}$$

Observando as expressões acima é fácil notar que a única parcela desconhecida é dada por.

$$\frac{G(e^{j\omega})}{1 + C(e^{j\omega}, \rho)G(e^{j\omega})}.$$

Realizando um experimento em malha-fechada com uma referência estacionária pode-se obter o espectro cruzado entre $r(t)$ e $y(t)$:

$$\Phi_{yr}(e^{j\omega}, \rho) = \frac{C(e^{j\omega}, \rho)G(e^{j\omega})}{1 + C(e^{j\omega}, \rho)G(e^{j\omega})} \Phi_r(e^{j\omega}) \quad (19)$$

A partir de (19) pode-se obter a parcela desconhecida citada anteriormente

$$\frac{G(e^{j\omega})}{1 + C(e^{j\omega}, \rho)G(e^{j\omega})} = \frac{\Phi_{yr}(e^{j\omega}, \rho)}{C(e^{j\omega}, \rho)\Phi_r(e^{j\omega})}.$$

Como todas as parcelas são conhecidas pode-se calcular o estimador para o gradiente da função custo utilizando (18). Esta estimativa pode ser utilizada no algoritmo do gradiente para procurar um mínimo da função custo.

Em (KAMMER; BITMEAD; BARTLETT, 2000) também é apresentada uma forma de calcular uma estimativa para a Hessiana da função custo de forma a utilizar o algoritmo de Newton em vez do algoritmo do gradiente.

2.5.4 CbT - Correlation based Tuning

O método de projeto de controladores *Correlation based Tuning* também é baseado em otimização utilizando o algoritmo do gradiente (KARIMI; MISKOVIC; BONVIN, 2004). Uma aplicação em um sistema de suspensão magnética é apresentado em (KARIMI; MISKOVIC; BONVIN, 2003). A prova da convergência local do método é apresentada em (KARIMI; MISKOVIC; BONVIN, 2002). A idéia principal do método é minimizar a correlação cruzada entre a referência e o sinal $\tilde{y}(t, \rho)$ (diferença entre a saída desejada e a saída efetiva do sistema), em vez de procurar o critério que minimiza o custo H_2 . Desse modo, o sistema obtido em malha-fechada irá capturar as dinâmicas do sistema desejado, compensando as dinâmicas do processo não modelado, de modo que não sobre nenhuma informação do sinal de referência no sinal de erro em malha-fechada $\tilde{y}(t, \rho)$.

O problema de otimização é baseado em encontrar os parâmetros do controlador que solucionam a seguinte equação de correlação

$$\sum_{i=1}^N \xi(t) \tilde{y}(t, \rho) = 0,$$

onde N é o tamanho dos dados do experimento e $\xi(t)$ é uma variável instrumental que deve ser correlacionada com a referência e descorrelacionada com o ruído.

Este sistema de equações não é linear e portanto é utilizado o algoritmo do gradiente para encontrar o argumento que soluciona o sistema de equações. O gradiente do algoritmo é calculado como:

$$\nabla J(\rho) = \sum_{i=1}^N \xi(t, \rho) \tilde{y}(t, \rho).$$

Várias são as possíveis escolhas para a variável instrumental, mas a escolha que tem gerado melhores resultados é uma estimativa do gradiente da saída com relação aos parâmetros do controlador. Neste caso é utilizado o seguinte gradiente para a função custo

$$\nabla J(\rho) = \sum_{i=1}^N \tilde{y}(t, \rho) \nabla \tilde{y}(t, \rho)$$

Apesar desse método partir de uma idéia diferente dos métodos anteriores, ele é muito parecido com o método IFT. Fica claro que $\nabla J(\rho)$ é o mesmo para o método CbT e IFT. A diferença está no modo que é computado o vetor $\nabla \tilde{y}(t, \rho)$. Enquanto que no método IFT esse vetor é calculado a partir de dados de três experimentos, no método CbT ele é computado a partir de um modelo reduzido identificado do processo.

Esta estimativa para $\nabla J(\rho)$ é utilizada no algoritmo do gradiente para procurar um mínimo da função custo.

Como no método FDT, o método CbT também apresenta uma forma de estimar a Hessiana da função custo, de forma que o algoritmo de Newton pode ser utilizado para resolver o problema de otimização.

2.6 Considerações finais

Este capítulo apresentou os quatro métodos de projeto de controladores baseado em dados. Inicialmente foram definidos conceitos gerais utilizados por todos os métodos. Desta forma é possível analisar semelhanças e diferenças entre os diversos métodos.

O problema de otimização foi apresentado na forma de um critério H_2 , que foi dividido em três termos representativos dos principais objetivos de controle: seguimento de referência, atenuação de ruído e minimização do esforço de controle.

O algoritmo do gradiente foi apresentado como ferramenta utilizada para resolver o problema de otimização. Por sua simplicidade ele foi escolhido para ser estudado neste trabalho. É utilizado por todos os métodos iterativos de projeto de controladores baseado em dados. A convergência deste algoritmo para o mínimo global da função custo para o problema de controle baseado em dados não foi encontrada na literatura, portanto optou-se em estudá-la.

Na seqüência foram apresentados quatro métodos de projeto de controladores baseado em dados. O método VRFT é um método direto, enquanto que o IFT, FDT e CbT são métodos iterativos. Todos os métodos iterativos utilizam o algoritmo do gradiente bem como variações baseadas no algoritmo de Newton para solucionar o problema de otimização. Como a convergência para o mínimo global da função custo é o objetivo fundamental destes métodos, e ainda não está compreendida decidiu-se estudá-la com profundidade.

Nos capítulos seguintes será apresentado um estudo sobre a convergência destes métodos iterativos baseado em dados utilizando o algoritmo do gradiente. Neste trabalho optou-se por não procurar algoritmos mais complexos para resolver o problema de otimização, mas sim em estudar com profundidade a função custo de forma a garantir convergência com o algoritmo mais simples, no caso o algoritmo do gradiente.

3 CONVERGÊNCIA DO ALGORITMO DO GRADIENTE

3.1 Introdução

Neste capítulo serão apresentadas análises sobre a convergência do algoritmo do gradiente para o mínimo global da função custo. Será demonstrado que a convergência depende de propriedades da função custo, das condições iniciais do algoritmo e do tamanho do passo de iteração utilizado. Toda esta análise foi apresentada primeiramente em (BAZANELLA et al., 2007) e (BAZANELLA et al., 2008), que são as bibliografias que introduziram muitos conceitos e idéias utilizados neste trabalho. O capítulo seguinte deste trabalho, o qual é a principal contribuição, pode ser considerado um extensão dos assuntos abordados nesses dois artigos.

Será demonstrado, neste capítulo, que para que ocorra a convergência para o mínimo global da função custo é necessário que as condições iniciais do algoritmo estejam dentro de um certo domínio de atração. Esse domínio de atração depende tanto da função custo como do tamanho do passo de iteração utilizado. Serão apresentadas ainda propriedades desse domínio de atração. Na seção 3.7 será mostrado como manipular algumas variáveis de projeto de modo a aumentar esse domínio de atração. No capítulo seguinte será apresentado um modo de cálculo do tamanho do passo de iteração para garantir a convergência para o mínimo global da função custo.

3.2 Domínio de atração

Definição 2 *Seja ρ_* o mínimo global de uma função $J(\rho) : \mathbb{R}^n \rightarrow \mathbb{R}^+$. Um conjunto $\Omega \subset \mathbb{R}^n$ é um domínio de atração de um algoritmo $\rho_{i+1} = f(\rho_i)$ para a função $J(\rho)$ se $\lim_{i \rightarrow \infty} \rho_i = \rho_*, \forall \rho_0 \in \Omega$.*

Durante todo este trabalho serão estudados domínios de atração em forma de bola. Estes conjuntos são definidos por:

$$\mathcal{B}_\alpha(\rho_*) = \{\rho : (\rho - \rho_*)^T(\rho - \rho_*) < \alpha\}.$$

A convergência do algoritmo do gradiente é analisada utilizando uma função de Lyapunov quadrática do tipo $V(\rho) = (\rho - \rho_*)^T(\rho - \rho_*)$ (KHALIL, 1996). Nesta análise pretende-se provar que para qualquer condição inicial dentro de um conjunto $\mathcal{B}_\alpha(\rho_*)$ o algoritmo do gradiente converge para o mínimo global da função custo.

Pela teoria de Lyapunov, o mínimo global da função custo ρ_* é um ponto de equilíbrio assintoticamente estável, e o conjunto $\mathcal{B}_\alpha(\rho_*)$ é um domínio de atração para este ponto de equilíbrio, se o gradiente com relação ao parâmetro ρ da função candidata de

Lyapunov for definido negativo para todos os elementos do conjunto diferentes do ponto de equilíbrio, portanto no caso discreto,

$$V(\rho_{i+1}) - V(\rho_i) < 0; \forall \rho \in \mathcal{B}_\alpha(\rho_*), \rho \neq \rho_*. \quad (20)$$

Substituindo (11) em (20) tem-se

$$\begin{aligned} V(\rho_{i+1}) - V(\rho_i) &= (\rho_i - \gamma_i \nabla J(\rho_i) - \rho_*)^T (\rho_i - \gamma_i \nabla J(\rho_i) - \rho_*) - (\rho_i - \rho_*)^T (\rho_i - \rho_*) \\ &= -2\gamma_i (\rho_i - \rho_*)^T \nabla J(\rho_i) + \gamma_i^2 \nabla J^T(\rho_i) \nabla J(\rho_i) < 0. \end{aligned} \quad (21)$$

Como γ_i é definido positivo a condição (21) somente é satisfeita se

$$\gamma_i < 2 \frac{(\rho_i - \rho_*)^T \nabla J(\rho_i)}{\nabla J^T(\rho_i) \nabla J(\rho_i)}. \quad (22)$$

Algumas observações podem ser feitas neste ponto. Como o denominador de (22) é semi-definido positivo por construção visto que é composto por um vetor multiplicado pelo mesmo transposto, a condição (22) é satisfeita se e somente se

$$(\rho_i - \rho_*)^T \nabla J(\rho_i) > 0 \quad \forall \rho_i \in \mathcal{B}_\alpha(\rho_*), \rho_i \neq \rho_*. \quad (23)$$

A condição (23) revela algumas características da função custo no domínio $\mathcal{B}_\alpha(\rho_*)$. Esta condição implica que:

- o ângulo entre o gradiente e o vetor $(\rho - \rho_*)$ está sempre no intervalo $(-\pi/2, \pi/2)$;
- o gradiente nunca é zero - não existem pontos de máximo, pontos de mínimo ou pontos de sela que não são o mínimo global ρ_* .

Pode-se enunciar agora o seguinte teorema:

Teorema 1 (BAZANELLA *et al.*, 2008)

Considere uma função duplamente diferenciável $J(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^+$. Considere que esta função tem um mínimo global isolado ρ_ . Se a condição (23) for respeitada então existe uma seqüência $\gamma_i, i = 1, \dots, \infty$ tal que $\mathcal{B}_\alpha(\rho_*)$ é um domínio de atração do algoritmo (11) para $J(\rho)$.*

A condição (23) diz respeito às características da função custo, mas não se refere ao algoritmo utilizado. Se o controlador inicial for tal que $\rho_0 \in \mathcal{B}_\alpha(\rho_*)$, então existe um algoritmo do gradiente que garante a convergência para o mínimo global da função custo. Este algoritmo é descrito pela seqüência passos γ_i utilizados em cada iteração. Quando a condição (23) é satisfeita em um dado conjunto, diz-se que este conjunto é candidato a domínio de atração, pois existe uma seqüência de valores de γ_i , tal que este conjunto é efetivamente um domínio de atração.

Definição 3 *Seja ρ_* o mínimo global de uma função $J(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^+$. O conjunto Ω é candidato a domínio de atração se $\rho_* \in \Omega$ e (23) é satisfeito para todo $\rho_i \in \Omega$.*

Nas próximas seções será analisada mais atentamente a condição (23) para cada um dos casos particulares da função custo: seguimento de referência, rejeição a ruído e minimização do esforço de controle.

3.3 Seguimento de referência

Nesta seção será analisada a parcela da função custo do seguimento de referência. Considera-se que o sistema não está sujeito a perturbações. Deseja-se, resolvendo o problema de otimização, encontrar os parâmetros do controlador que tornam o desempenho do sistema em malha-fechada o mais próximo possível do modelo de referência $T_d(z)$.

O sistema em malha-fechada somente poderá se comportar exatamente como o modelo de referência se a estrutura do controlador for tal que o controlador ideal $C_d(z)$ pode ser representado por esta estrutura. Esta afirmação significa que $C_d(z)$ deve pertencer à classe de controladores \mathcal{C} considerada no projeto (BAZANELLA et al., 2008), ou seja, que a seguinte hipótese é satisfeita:

Hipótese 2 $C_d \in \mathcal{C}$ ou, de maneira equivalente,

$$\exists \rho_d : C(z, \rho_d) = C_d(z) = \rho_d^T \bar{C}(z).$$

Quando a Hipótese 2 é satisfeita diz-se que o modelo de referência é atingível.

3.3.1 Modelo de referência atingível

Nesta subseção será analisada a condição (23) para o caso de seguimento de referência quando a Hipótese 2 é satisfeita.

Para analisar mais a fundo a condição (23) precisa-se de uma expressão para o gradiente de $J_y(\rho)$. A partir de (7) tem-se que

$$J_y(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |T(e^{j\omega}, \rho) - T_d(e^{j\omega})|^2 \Phi_r(e^{j\omega}) d\omega$$

Na seqüência será obtida uma expressão para o gradiente de $J_y(\rho)$. Foram omitidas a dependência das variáveis z e ω para poupar espaço. A partir de (7) tem-se que o gradiente de $J_y(\rho)$ pode ser descrito por:

$$\nabla J_y(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ [T(\rho) - T_d]^* \frac{\partial}{\partial \rho} [T(\rho) - T_d] + [T(\rho) - T_d] \frac{\partial}{\partial \rho} [T^*(\rho) - T_d^*] \right\} \Phi_r d\omega \quad (24)$$

As seguintes expressões serão utilizadas para simplificar as expressão de $\nabla J_y(\rho)$:

$$T(\rho) - T_d = \frac{C(\rho)G}{1 + C(\rho)G} - \frac{C_d G}{1 + C_d G} = (C(\rho) - C_d) G S_d S(\rho) \quad (25)$$

$$(T(\rho) - T_d)^* = (C(\rho) - C_d)^* G^* S_d^* S^*(\rho) \quad (26)$$

$$\frac{\partial T(\rho)}{\partial \rho} = \frac{\partial}{\partial \rho} \frac{C(\rho)G}{1 + C(\rho)G} = G S^2(\rho) \frac{\partial C(\rho)}{\partial \rho} \quad (27)$$

$$\frac{\partial T^*(\rho)}{\partial \rho} = G^* S^{*2}(\rho) \frac{\partial C^*(\rho)}{\partial \rho} \quad (28)$$

Inserindo (25), (26), (27) e (28) em (24):

$$\begin{aligned} \nabla J_y(\rho) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ (C(\rho) - C_d)^* G^* S_d^* S^*(\rho) G S^2(\rho) \frac{\partial C(\rho)}{\partial \rho} \right. \\ &\quad \left. + (C(\rho) - C_d) G S_d S(\rho) G^* S^{*2}(\rho) \frac{\partial C^*(\rho)}{\partial \rho} \right\} \Phi_r d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_r |G S(\rho)|^2 2\Re \left\{ (C(\rho) - C_d) G S_d S(\rho) S^{*2}(\rho) \frac{\partial C^*(\rho)}{\partial \rho} \right\} d\omega \end{aligned}$$

onde $\Re \{ \cdot \}$ indica a parte real de um número complexo.

Utilizando a Hipótese 1 então:

$$\begin{aligned}\nabla J_y(\rho) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 2 \{ (C(\rho) - C_d)^* S_d^* S(\rho) \bar{C} + (C(\rho) - C_d) S_d S^*(\rho) \bar{C}^{*T} \} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 2 \Re \{ (C(\rho) - C_d)^* S_d^* S(\rho) \bar{C} \} d\omega\end{aligned}\quad (29)$$

Finalmente se a Hipótese 2 também é satisfeita então:

$$\nabla J_y(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} 2\Phi_r |GS(\rho)|^2 \Re \{ S_d^* S(\rho) \bar{C} \bar{C}^* \} d\omega (\rho - \rho_d)$$

O gradiente de $J_y(\rho)$ pode então ser descrito por:

$$\nabla J_y(\rho) = M_y(\rho) (\rho - \rho_d) \quad (30)$$

$$M_y(\rho) \triangleq \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |G|^2 |S(\rho)|^2 \Re \{ S_d^* S(\rho) \bar{C} \bar{C}^* \} d\omega \quad (31)$$

Agora que é conhecida uma expressão para $\nabla J_y(\rho)$, pode-se utilizá-la para estudar mais profundamente a condição (23). A expressão do gradiente pode ser introduzida em (23) resultando em uma nova condição mais específica:

$$(\rho_i - \rho_d)^T \nabla J_y(\rho_i) = (\rho_i - \rho_d)^T M_y(\rho) (\rho - \rho_d) > 0 \quad \forall \rho_i \in \mathcal{B}_\alpha(\rho_d), \rho_i \neq \rho_d. \quad (32)$$

Nota-se que as propriedades da condição (32) dependem das propriedades da função matricial $M_y(\rho)$. Portanto serão analisadas algumas propriedades desta função matricial. É simples mostrar que

$$M_y(\rho) \triangleq M_{ys}(\rho) + M_{ya}(\rho)$$

onde

$$\begin{aligned}M_{ys}(\rho) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |G|^2 |S(\rho)|^2 \Re \{ S_d^* S(\rho) \} \Re \{ \bar{C} \bar{C}^* \} d\omega \\ M_{ya}(\rho) &= -\frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |G|^2 |S(\rho)|^2 \Im \{ S_d^* S(\rho) \} \Im \{ \bar{C} \bar{C}^* \} d\omega\end{aligned}$$

A matriz $M_{ys}(\rho)$ é simétrica enquanto que $M_{ya}(\rho)$ é anti-simétrica.

Sabe-se que dada uma matriz $Q \in \mathbb{R}^{n \times n}$ anti-simétrica, uma propriedade importante é que $x^T Q x = 0$ para qualquer $x \in \mathbb{R}^n$. Esta propriedade será utilizada para simplificar várias expressões durante todo o trabalho.

Utilizando as propriedades do termo anti-simétrico de $M_y(\rho)$ pode-se simplificar a condição (32):

$$(\rho_i - \rho_d)^T M_y(\rho) (\rho_i - \rho_d) = (\rho_i - \rho_d)^T M_{ys}(\rho) (\rho - \rho_d) > 0 \quad \forall \rho_i \in \mathcal{B}_\alpha(\rho_d), \rho_i \neq \rho_d \quad (33)$$

que agora deixa de depender da parte anti-simétrica.

A condição (33) também pode ser simplificada visto que a matriz $M_{ys}(\rho)$ precisa ser definida positiva para que a condição seja satisfeita. Desta maneira uma condição muito mais simples pode ser encontrada:

$$M_{ys}(\rho) > 0; \quad \forall \rho \in \mathcal{B}_\alpha(\rho_d), \rho \neq \rho_d. \quad (34)$$

Neste ponto podem ser feitas algumas observações. Se a condição (34) é satisfeita então $\mathcal{B}_\alpha(\rho_d)$ é um domínio de atração candidato para o mínimo da função custo $J_y(\rho)$. Esta condição será utilizada na Seção 3.7 para determinar parâmetros de projeto com a finalidade de aumentar o tamanho do domínio de atração.

Esta condição (34) também pode ser analisada mais detalhadamente. Se o campo vetorial $\bar{C}(e^{j\omega})$ tiver posto completo, a soma de $\bar{C}\bar{C}^*$ por um número suficiente grande de frequências é positiva definida. Portanto se o termo $\Re\{S_d^*S(\rho)\}$ for positivo definido, também será $M_{ys}(\rho)$. Pode-se então enunciar o seguinte teorema:

Teorema 2 (BAZANELLA *et al.*, 2008)

Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEP. Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo tal que $\rho_d \in \Gamma$, e para todo $\rho \in \Gamma$

$$\Re\{S_d(e^{j\omega})^*S(e^{j\omega}, \rho)\} > 0 \quad \forall \omega \quad (35)$$

Então

$$(\rho - \rho_d)^T \nabla J_y(\rho) > 0 \quad \forall \rho \in \Upsilon, \rho \neq \rho_d.$$

Este teorema apresenta condições suficientes relativamente simples para garantir que (23) é verificada. A condição (35) também pode ser expressa por:

$$\max |\angle S(e^{j\omega}, \rho) - \angle S_d(e^{j\omega})| < \pi/2 \quad \forall \rho \in \Upsilon \quad \forall \omega \quad (36)$$

Esta máxima diferença de fase entre as duas funções de transferência é uma medida que pode ser usada para medir a distância entre um dado ρ e o mínimo global ρ_d .

O Teorema 2 permite algumas interpretações:

- o sinal de entrada deve ser persistentemente excitante de ordem pelo menos igual à dimensão do vetor de parâmetros, para que a soma de $\bar{C}\bar{C}^*$ tenha posto completo;
- a parametrização do controlador não pode ser redundante, ou seja, ela deve representar a classe de controladores considerada com um número mínimo de parâmetros, para que $\bar{C}(e^{j\omega})$ tenha posto completo;
- os parâmetros em Υ devem ser suficientemente próximos do parâmetro ideal ρ_d , para que as funções de sensibilidade não sejam tão diferentes, no sentido que as fases devem ser próximas, mas sem restrição alguma com respeito às amplitudes.

Esse resultado do Teorema 2 será utilizado na Seção 3.7 para determinar parâmetros de projeto com a finalidade de aumentar o tamanho do domínio de atração.

3.3.2 Modelo de referência inatingível

Vamos analisar agora o caso em que o modelo de referência não pode ser atingido, ou seja, o caso que o controlador ideal não pertence à classe de controladores \mathcal{C} e portanto não pode ser alcançado. Nesse caso, a Hipótese 2 precisa ser substituída por outra mais abrangente, para analisar as propriedades de $J_y(\rho)$.

Hipótese 3 A função custo $J_y(\rho)$ apresenta um mínimo global ρ_* , isto é,

$$\exists \rho_* : J_y(\rho) > J_y(\rho_*) \quad \forall \rho \neq \rho_*, \quad \nabla J(\rho_*) = 0, \quad \left. \frac{\partial^2 J_y(\rho)}{\partial \rho^2} \right|_{\rho_*} > 0$$

Quando a Hipótese 2 é satisfeita, ρ_d é o mínimo global ($\rho_* = \rho_d$) e $C(z, \rho) = C_d(z)$, portanto o controlador ideal pode ser encontrado. Quando a Hipótese 2 não é satisfeita, o melhor controlador que pode ser encontrado é $C(z, \rho_*) \neq C_d(z)$. A diferença entre o melhor controlador que pode ser encontrado e o controlador ideal é definida pela seguinte função de transferência

$$K(z) \triangleq C_d(z) - C(z, \rho_*) \quad (37)$$

Agora substituindo (37) em (29) tem-se que

$$\begin{aligned} \nabla J_y(\rho) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 \Re \{ (C(\rho) - C(\rho_*) - K)^* S_d^* S(\rho) \bar{C} \} d\omega \\ &= M_y(\rho)(\rho - \rho_*) - m_y(\rho) \end{aligned} \quad (38)$$

onde $M_y(\rho)$ é como definido anteriormente e define-se

$$m_y(\rho) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 \Re \{ S_d^* S(\rho) \bar{C} K^* \} d\omega.$$

A equação (38) é similar a (30) mas perturbada pela função vetorial $m_y(\rho)$. Esta perturbação é desconhecida, contínua e satisfaz $m_y(\rho_*) = 0$. Além disso, $m_y(\rho)$ é limitada para todo $\rho \in \Gamma$. Portanto, $|m_y(\rho)|$ pode ser linearmente limitada, ou seja, para qualquer conjunto $\Upsilon \subseteq \Gamma$ contendo o mínimo global ρ_* :

$$\exists \kappa_y \in \mathbb{R}^+ : |m_y(\rho)| < \kappa_y |\rho - \rho_*| \quad \forall \rho \in \Upsilon \quad (39)$$

Teorema 3 (BAZANELLA et al., 2008)

Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEP e κ_y como definido em (39). Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo tal que $\rho_* \in \Upsilon$, e para todo $\rho \in \Upsilon$

$$\Re \{ S_d(e^{j\omega})^* S(e^{j\omega}, \rho) \} > 0 \quad \forall \omega$$

Se, além disso, o termo da perturbação $m_y(\rho)$ é tal que o limitante κ_y em (39) satisfaz $M(\rho) > \kappa_y I \quad \forall \rho \in \Upsilon$, então

$$(\rho - \rho_d)^T \nabla J(\rho) > 0 \quad \forall \rho \in \Upsilon, \rho \neq \rho_d.$$

Este resultado é provado pela simples substituição de (38) em (23) utilizando as hipóteses do teorema.

3.4 Atenuação da sensibilidade ao ruído

Nesta seção será analisada a parcela da função custo de atenuação de ruído. Portanto considera-se que o sistema possui referência nula. Deseja-se resolvendo o problema de otimização encontrar os parâmetros do controlador que minimizam o efeito do ruído. O sistema em malha-fechada somente tem menor custo de ruído possível se o controlador ideal $C_e(z)$ pertencer à classe de controladores \mathcal{C} considerada no projeto (BAZANELLA et al., 2008), ou seja, se a seguinte hipótese for satisfeita:

Hipótese 4 $C_e \in \mathcal{C}$ ou, de maneira equivalente,

$$\exists \rho_e : C(z, \rho_e) = C_e(z) = \rho_e^T \bar{C}(z).$$

Quando a Hipótese 4 é satisfeita diz-se que o modelo de referência é atingível.

3.4.1 Modelo de referência atingível

Nesta subsecção será analisada a condição (23) para o caso de atenuação da sensibilidade ao ruído quando a Hipótese 2 é satisfeita.

Será desenvolvida na seqüência uma expressão para o gradiente da função custo. Como $e(t)$ é ruído branco com variância σ_e^2 :

$$J_e(\rho) = \sigma_e^2 \int_{-\pi}^{\pi} |S(\rho)H|^2 d\omega = \sigma_e^2 \int_{-\pi}^{\pi} |1 + HS(\rho) - 1|^2 d\omega$$

$$J_e(\rho) = \sigma_e^2 \left[1 + \int_{-\pi}^{\pi} |H|^2 + |S(\rho) - S_e|^2 d\omega \right]$$

onde foi utilizado o fato que $HS(\rho) - 1$ é uma função de transferência estritamente própria, o que implica que a integral da sua parte real é zero. Então

$$\nabla J_e(\rho) = \sigma_e^2 \int_{-\pi}^{\pi} |H|^2 \frac{\partial}{\partial \rho} |S(\rho) - S_e(\rho)|^2 d\omega \quad (40)$$

Sobre o gradiente da função de sensibilidade tem-se que

$$\frac{\partial}{\partial \rho} S(\rho) = \frac{\partial}{\partial \rho} \frac{1}{1 + C(\rho)G} = -GS(\rho)^2 \frac{\partial C(\rho)}{\partial \rho}$$

e portanto:

$$\frac{\partial}{\partial \rho} |S(\rho) - S_e|^2 = \left[S^*(\rho)S_e(C(\rho) - C_e) \frac{\partial C^*(\rho)}{\partial \rho} + S(\rho)S_e^*(C(\rho) - C_e)^* \frac{\partial C(\rho)}{\partial \rho} \right] |S(\rho)|^2 |G|^2$$

Se as Hipóteses 1 e 4 forem satisfeita:

$$\frac{\partial}{\partial \rho} |S(\rho) - S_e|^2 = |S(\rho)|^2 |G|^2 [S^*(\rho)S_e \bar{C} \bar{C}^* + S(\rho)S_e^* (\bar{C} \bar{C}^*)^T] (\rho - \rho_e) \quad (41)$$

Substituindo (41) em (40) tem-se uma expressão para o gradiente de $J_e(\rho)$:

$$\begin{aligned} \nabla J_e(\rho) &= M_e(\rho)(\rho - \rho_e) \\ M_e(\rho) &\triangleq \frac{\sigma_e^2}{\pi} \int_{-\pi}^{\pi} |H|^2 |G|^2 |S(\rho)|^2 \Re \{ S_e^* S(\rho) \bar{C} \bar{C}^* \} d\omega \end{aligned} \quad (42)$$

No caso de existir apenas a parcela de rejeição de ruído na função custo, a equação (23) pode ser reescrita

$$(\rho_i - \rho_e)^T \nabla J(\rho_i) = (\rho_i - \rho_e)^T M_e(\rho)(\rho - \rho_e) > 0 \quad \forall \rho_i \in \mathcal{B}_\alpha(\rho_e), \rho_i \neq \rho_e. \quad (43)$$

As propriedades de convergência do algoritmo dependem então de propriedades da matriz $M_e(\rho)$. Esta matriz também pode ser dividida em dois termos, como feito no caso de seguimento de referência:

$$\begin{aligned} M_e(\rho) &\triangleq M_{es}(\rho) + M_{ea}(\rho) \\ M_{es}(\rho) &= \frac{\sigma_e^2}{\pi} \int_{-\pi}^{\pi} |H|^2 |G|^2 |S(\rho)|^2 \Re \{ S_e^* S(\rho) \} \Re \{ \bar{C} \bar{C}^* \} d\omega \\ M_{ea}(\rho) &= -\frac{\sigma_e^2}{\pi} \int_{-\pi}^{\pi} |H|^2 |G|^2 |S(\rho)|^2 \Im \{ S_e^* S(\rho) \} \Im \{ \bar{C} \bar{C}^* \} d\omega \end{aligned}$$

É simples verificar que $M_{es}(\rho)$ é simétrica enquanto que $M_{ea}(\rho)$ é anti-simétrica.

Utilizando a propriedade das matrizes anti-simétricas de $M_{ea}(\rho)$, tem-se de (43):

$$(\rho_i - \rho_e)^T M_{es}(\rho)(\rho_i - \rho_e) > 0.$$

Portanto para que $\mathcal{B}_\alpha(\rho_e)$ seja um domínio de atração para o algoritmo (11), é suficiente que

$$M_{es}(\rho) > 0; \forall \rho \in \mathcal{B}_\alpha(\rho_e), \rho \neq \rho_e. \quad (44)$$

Neste ponto podem ser feitas algumas observações. Se a condição (44) é satisfeita então $\mathcal{B}_\alpha(\rho_e)$ é um domínio de atração candidato para o mínimo da função custo $J_e(\rho)$. Esta condição (44) também pode ser analisada mais detalhadamente. Como $\bar{C}\bar{C}^*$ é positivo definido por construção, a sua soma por um número suficiente grande de frequências é positiva definida, se o campo vetorial $\bar{C}(e^{j\omega})$ tiver posto completo. Portanto se o termo $\Re\{S_e^* S(\rho)\}$ for definido positivo, também será $M_{es}(\rho)$ e pode-se então enunciar o seguinte teorema:

Teorema 4 (BAZANELLA *et al.*, 2008)

Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEP. Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo tal que $\rho_e \in \Gamma$, e para todo $\rho \in \Gamma$

$$\Re\{S_e(e^{j\omega})^* S(e^{j\omega}, \rho)\} > 0 \quad \forall \omega \quad (45)$$

Então

$$(\rho - \rho_e)^T \nabla J_e(\rho) > 0 \quad \forall \rho \in \Gamma, \rho \neq \rho_e.$$

Este teorema apresenta condições suficientes relativamente simples para garantir que (23) é verificada. A condição (45) também pode ser expressa por:

$$\max |\angle S(e^{j\omega}, \rho) - \angle S_e(e^{j\omega})| < \pi/2 \quad \forall \rho \in \Upsilon \quad \forall \omega \quad (46)$$

Esta máxima diferença de fase entre as duas funções de transferência também é uma medida que pode ser usada para medir a distância entre um dado ρ e o mínimo global ρ_e . Observações equivalentes àquelas apresentadas para o caso de seguimento de referência podem ser feitas também neste caso.

3.4.2 Modelo de referência inatingível

Será analisado agora o caso em que o modelo de referência não pode ser atingido, ou seja, o caso que o controlador ideal não pertence à classe de controladores \mathcal{C} e portanto não pode ser alcançado. Para analisar o comportamento de $J_e(\rho)$ a Hipótese 4 precisa ser substituída por outra mais fraca:

Hipótese 5 A função custo $J_e(\rho)$ apresenta um mínimo global ρ_* , isto é,

$$\exists \rho_* : J_e(\rho) > J_e(\rho_*) \quad \forall \rho \neq \rho_*, \quad \nabla J(\rho_*) = 0, \quad \left. \frac{\partial^2 J_e(\rho)}{\partial \rho^2} \right|_{\rho_*} > 0$$

Quando a Hipótese 4 é satisfeita, ρ_e é o mínimo global ($\rho_* = \rho_e$) e $C(z, \rho) = C_e(z)$, portanto o controlador ideal pode ser encontrado. Quando a Hipótese 4 não é satisfeita, o melhor controlador que pode ser encontrado é $C(z, \rho_*) \neq C_e(z)$. A diferença entre o

melhor controlador que pode ser encontrado e o controlador ideal é definida pela seguinte função de transferência

$$K(z) \triangleq C_e(z) - C(z, \rho_*)$$

Neste caso o gradiente da função custo é

$$\begin{aligned} \nabla J_e(\rho) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 \Re \{ (C(\rho) - C(\rho_*) - K)^* S_e^* S(\rho) \bar{C} \} d\omega \\ &= M_e(\rho)(\rho - \rho_*) - m_e(\rho) \end{aligned} \quad (47)$$

onde $M_e(\rho)$ é como definido anteriormente e define-se

$$m_e(\rho) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |GS(\rho)|^2 \Re \{ S_e^* S(\rho) \bar{C} K^* \} d\omega$$

A equação (47) é similar a (42) mas perturbada pela função vetorial $m_e(\rho)$. Esta perturbação é desconhecida contínua e satisfaz $m_e(\rho_*) = 0$. Além disso, $m_e(\rho)$ é limitada para todo $\rho \in \Gamma$. Portanto, $|m_e(\rho)|$ pode ser linearmente limitada, ou seja, para qualquer conjunto $\Upsilon \subseteq \Gamma$ contendo o mínimo global ρ_* :

$$\exists \kappa_e \in \mathbb{R}^+ : |m_e(\rho)| < \kappa_e |\rho - \rho_*| \quad \forall \rho \in \Upsilon \quad (48)$$

Teorema 5 (BAZANELLA *et al.*, 2008)

Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEP e κ_e como definido em (48). Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo tal que $\rho_* \in \Upsilon$, e para todo $\rho \in \Upsilon$

$$\Re \{ S_e(e^{j\omega})^* S(e^{j\omega}, \rho) \} > 0 \quad \forall \omega$$

Se, além disso, o termo da perturbação $m_e(\rho)$ é tal que o limitante κ_e em (48) satisfaz $M(\rho) > \kappa_e I \quad \forall \rho \in \Upsilon$, então

$$(\rho - \rho_e)^T \nabla J(\rho) > 0 \quad \forall \rho \in \Upsilon, \rho \neq \rho_e.$$

Este resultado é provado pela simples substituição de (47) em (23) utilizando as hipóteses do teorema.

3.5 Minimização do esforço de controle

Nesta seção vamos analisar a parte minimização do esforço de controle da função custo. Como considera-se que o controlador é parametrizado linearmente, ou seja, a Hipótese 1 é sempre respeitada, então o controlador que minimiza o esforço de controle $\rho_u = 0$ sempre pode ser alcançado. Então não existe o caso de modelo de referência não atingível.

Para analisar mais a fundo a condição (23) precisa-se de uma expressão para o gradiente de $J_u(\rho)$. A partir de (9) tem-se que

$$J_y(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |C(e^{j\omega}, \rho)|^2 |S(e^{j\omega}, \rho)|^2 (\Phi_r(e^{j\omega}) + \Phi_\nu(e^{j\omega})) d\omega$$

Na seqüência será obtida uma expressão para o gradiente de $J_u(\rho)$. Foram omitidas a dependência das variáveis z e ω para poupar espaço. A partir de (9) e (7) fica claro que

a parte dependente de ρ nas duas expressões são iguais. A partir do resultado para o caso do seguimento de referência pode-se obter:

$$\begin{aligned}\nabla J_u(\rho) &= \frac{1}{\pi} \int_{-\pi}^{\pi} (\Phi_r + \Phi_\nu) |S(\rho)|^2 \Re \left\{ C^*(\rho) S(\rho) \frac{\partial C^*(\rho)}{\partial \rho} \right\} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\Phi_r + \Phi_\nu) \left\{ C^*(\rho) S(\rho) \frac{\partial C(\rho)}{\partial \rho} + C(\rho) S^*(\rho) \frac{\partial C^*(\rho)}{\partial \rho} \right\} d\omega.\end{aligned}$$

Utilizando a Hipótese 1 então:

$$\nabla J_u(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\Phi_r + \Phi_\nu) |S(\rho)|^2 \Re \{ S(\rho) \bar{C} \bar{C}^* \} d\omega \rho$$

Pode-se calcular o gradiente de $J_u(\rho)$ como

$$\begin{aligned}\nabla J_u(\rho) &= M_u(\rho)(\rho) \tag{49} \\ M_u(\rho) &\triangleq \frac{1}{\pi} \int_{-\pi}^{\pi} [\Phi_r + \Phi_\nu] |S(\rho)|^2 \Re \{ S(\rho) \bar{C} \bar{C}^* \} d\omega\end{aligned}$$

Estudando apenas a parcela de minimização do esforço de controle, a equação (23) pode ser reescrita

$$\rho_i^T \nabla J_u(\rho_i) = \rho_i^T M_u(\rho) \rho_i > 0 \quad \forall \rho_i \in \mathcal{B}_\alpha(\rho_u), \rho_i \neq \rho_u. \tag{50}$$

As propriedades de convergência do algoritmo dependem novamente das propriedades da matriz $M_u(\rho)$. Esta matriz também pode ser dividida em sua parte simétrica e anti-simétrica:

$$\begin{aligned}M_u(\rho) &\triangleq M_{us}(\rho) + M_{ua}(\rho) \\ M_{us}(\rho) &= \frac{1}{\pi} \int_{-\pi}^{\pi} [\Phi_r + \Phi_\nu] |S(\rho)|^2 \Re \{ S(\rho) \} \Re \{ \bar{C} \bar{C}^* \} d\omega \\ M_{ua}(\rho) &= -\frac{1}{\pi} \int_{-\pi}^{\pi} [\Phi_r + \Phi_\nu] |S(\rho)|^2 \Im \{ S(\rho) \} \Im \{ \bar{C} \bar{C}^* \} d\omega\end{aligned}$$

De (50) utilizando a propriedade das matrizes anti-simétricas tem-se:

$$\rho_i^T M_{us}(\rho) \rho_i > 0.$$

Portanto para que $\mathcal{B}_\alpha(\rho_u)$ seja um domínio de atração para o algoritmo (11), é suficiente que

$$M_{us}(\rho) > 0; \quad \forall \rho \in \Omega, \rho \neq \rho_* \tag{51}$$

Neste ponto podem ser feitas algumas observações. Se a condição (51) é satisfeita então $\mathcal{B}_\alpha(0)$ é um domínio de atração candidato para o mínimo da função custo $J_u(\rho)$. Esta condição (51) também pode ser analisada mais detalhadamente. Se o campo vetorial $\bar{C}(e^{j\omega})$ tiver posto completo, a soma de $\bar{C} \bar{C}^*$ por um número suficiente grande de frequências é positiva definida. Portanto se o termo $\Re \{ S(\rho) \}$ for positivo definido, também será $M_{us}(\rho)$. Pode-se então enunciar o seguinte teorema:

Teorema 6 (BAZANELLA *et al.*, 2008) *Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEp. Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo tal que $0 \in \Gamma$, e para todo $\rho \in \Gamma$*

$$\Re \{ S(e^{j\omega}, \rho) \} > 0 \quad \forall \omega$$

Então

$$\rho^T \nabla J_u(\rho) > 0 \quad \forall \rho \in \Upsilon, \rho \neq 0.$$

Este teorema apresenta condições suficientes relativamente simples para garantir que (23) é verificada.

3.6 Combinando os três objetivos

Diferentes métodos de projeto baseado em dados minimizam diferentes combinações dessas três parcelas, representando diferentes casos particulares de $J(\rho)$. Em um caso geral a função custo é uma combinação das três parcelas que representam diferentes objetivos de controle. Neste caso, o gradiente da função custo também será uma combinação dos gradientes de cada uma das parcelas:

$$\nabla J(\rho) = \lambda [\nabla J_y(\rho) + \nabla J_e(\rho)] + (1 - \lambda) \nabla J_u(\rho)$$

então

$$\begin{aligned} \nabla J(\rho) &= [\lambda M_y(\rho) + \lambda M_e(\rho) + (1 - \lambda) M_u(\rho)] (\rho - \rho_*) \\ \nabla J(\rho) &= M(\rho) (\rho - \rho_*) \end{aligned}$$

Novamente pode-se dividir $M(\rho)$ em uma parte simétrica e outra anti-simétrica.

$$\begin{aligned} M(\rho) &= M_s(\rho) + M_a(\rho) \\ M_s(\rho) &= \lambda M_{ys}(\rho) + \lambda M_{es}(\rho) + (1 - \lambda) M_{us}(\rho) \\ M_a(\rho) &= \lambda M_{ya}(\rho) + \lambda M_{ea}(\rho) + (1 - \lambda) M_{ua}(\rho) \end{aligned}$$

Uma condição suficiente para que $M(\rho)$ seja não singular é que cada uma das matrizes $M_y(\rho)$, $M_e(\rho)$ e $M_u(\rho)$ tenha sua parte simétrica definida positiva. Sobre esta condição o custo total $J(\rho)$ tem propriedades similares e pode-se então apresentar um resultado similar àqueles apresentados anteriormente para cada termo isolado.

Teorema 7 (BAZANELLA *et al.*, 2008)

Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEp. Seja $\Upsilon \subseteq \Gamma$ um conjunto conexo contendo ρ_d , ρ_e , 0 e ρ_* , e para todo $\rho \in \Upsilon$

$$\begin{aligned} \Re \{S_d(e^{j\omega})^* S(e^{j\omega}, \rho)\} &> 0 \quad \forall \omega \\ \Re \{S_e(e^{j\omega})^* S(e^{j\omega}, \rho)\} &> 0 \quad \forall \omega \\ \Re \{S(e^{j\omega}, \rho)\} &> 0 \quad \forall \omega \end{aligned}$$

Então

$$(\rho - \rho_*)^T \nabla J(\rho) > 0 \quad \forall \rho \in \Upsilon, \rho \neq \rho_*$$

3.7 Shaping da Função Custo

Como visto anteriormente, uma condição suficiente para garantir que um conjunto Ω seja um domínio de atração para o algoritmo do gradiente é $M_s(\rho) > 0, \forall \rho \in \Omega$. Esta função matricial $M_s(\rho)$ depende de muitas variáveis, sendo que algumas podem ser escolhidas pelo projetista enquanto que outras não. Vamos analisar cada uma destas variáveis e verificar quais estão à disposição do projetista e como podem ser manipuladas.

As variáveis do processo $G(z)$, $H(z)$ e σ_e^2 são desconhecidas e com certeza não podem ser manipuladas. A classe de controladores \mathcal{C} é conhecida, mas na maioria dos casos práticos não pode ser manipulada pois o hardware disponível impõe restrições. Por outro lado, a parametrização do controlador pode ser livremente escolhida, tomando o devido cuidado de que esta parametrização seja mínima, ou seja, que a classe de controladores \mathcal{C} seja representada pelo número mínimo possível de parâmetros.

Os parâmetros do controlador inicial ρ_0 podem ser escolhidos livremente. Contudo garantir estabilidade em malha-fechada é difícil pelo não conhecimento do processo. Neste caso, métodos diretos como o VRFT podem ser de grande valia, pois estes não necessitam de condições iniciais e muitas vezes geram controladores bastante razoáveis. Em situações práticas, na maioria das vezes os métodos baseados em dados são utilizados para melhorar o desempenho de sistemas que já estão funcionando em malha-fechada. Sendo assim, apesar de ser possível escolher ρ_0 , utilizam-se os valores atuais do sistema em malha-fechada.

Ainda sobraram as variáveis correspondentes ao critério de desempenho: $T_d(z)$, λ e Φ_r . Quando o critério de desempenho é escolhido, estas variáveis são fixas, e modificar estas variáveis implica minimizar um outro critério diferente do desejado. Todavia, se o critério desejado é muito difícil de ser otimizado iniciando com o controlador ρ_0 , pode-se considerar minimizar um critério intermediário mais fácil. Na verdade vários critérios intermediários podem ser utilizados, e na medida que um critério é alcançado passa-se para outro mais próximo do critério desejado. Esta é a idéia central do Shaping da Função Custo: manipular as variáveis $T_d(z)$, λ e Φ_r , pouco a pouco de forma a utilizar critérios de desempenho intermediários para aumentar o domínio de atração do mínimo global do critério desejado.

3.7.1 Controle cauteloso

Começando de um controlador inicial que atinge um dado desempenho $T_0(z)$ o qual é considerado ruim, propõe-se a utilização de um modelo de referência intermediário $T_d^1(z)$. Este modelo de referência é escolhido de forma *cautelosa* de modo que escolha-se $T_d^1(z)$ muito mais perto de $T_0(z)$ do que de $T_d(z)$.

Fazendo esta escolha deseja-se que as fases de $S_0(z)$ e $S_d^1(z)$ sejam bastante próximas de modo a garantir que a condição (36) seja satisfeita e portanto que ρ_0 esteja no domínio de atração de ρ_d^1 .

Uma vez que o mínimo global deste novo critério (ρ_*^1) é alcançado, escolhe-se um segundo modelo de referência mais ambicioso e portanto mais perto de $T_d(z)$. Inicia-se então o problema de otimização a partir de ρ_*^1 . Este processo pode ser repetido sucessivamente com vários modelos de referência até que o modelo de referência desejado $T_d(\rho)$ seja alcançado.

Este conceito de controle cauteloso é familiar em projeto de experimentos baseados em dados (KAMMER; BITMEAD; BARTLETT, 2000) (KAMMER, 2005), identificação iterativa e projeto de controladores (ZANG; BITMEAD; GEVERS, 1995) e também é um conceito fundamental da abordagem *windsurf* de controle adaptativo (LEE et al., 1995) (LEE et al., 1993). Em (ANDERSON; GEVERS, 1998), o problema de objetivo de controle impraticável é abordado e é sugerida a utilização de sucessivas iterações do critério aproximando-se do desejado.

3.7.2 Manipulação do espectro da referência

Foi provado nas seções anteriores que para um ponto extremo seja único em um determinado conjunto é suficiente que a diferença entre as fases de determinadas funções de sensibilidade seja limitada em $\pi/2$. Contudo, esta não é uma condição necessária e pode ser contornada por uma escolha adequada do espectro da referência.

Para demonstrar este fato primeiramente serão apresentadas propriedades da função de sensibilidade que faz parte daquela condição (36).

Lema 1 *Seja $S(z, \rho) = (1 + \rho^T \bar{C}(z)G(z))^{-1}$ e seja Γ o conjunto de todos os parâmetros que tornam o sistema em malha-fechada estável. Para todo $\rho_1, \rho_2 \in \Gamma$:*

$$\begin{aligned}\angle S(1, \rho_1) &= \angle S(1, \rho_2) \\ \angle S(-1, \rho_1) &= \angle S(-1, \rho_2)\end{aligned}$$

Prova:

$$\angle S(e^{j\omega}, \rho) = \sum_{i=1}^n \angle(e^{j\omega} - b_i) - \sum_{i=1}^n \angle(e^{j\omega} - a_i(\rho))$$

onde b_i são os pólos da função de laço $\bar{C}(e^{j\omega})G(e^{j\omega})$ e $a_i(\rho)$ são os pólos de malha-fechada. Para $\omega = 0$ tem-se

$$\angle S(1, \rho) = \sum_{i=1}^n \angle(1 - b_i) - \sum_{i=1}^n \angle(1 - a_i(\rho))$$

Mas $\sum_{i=1}^n \angle(1 - b_i)$ não depende do parâmetro ρ e $\sum_{i=1}^n \angle(1 - a_i(\rho)) = 0 \forall i, \rho \in \Gamma$ porque $a_i(\rho)$ pertence ao círculo unitário para todo $\rho \in \Gamma$. O mesmo argumento é válido para $\omega = \pi$. \square

Desta propriedade sobre a função de sensibilidade $S(e^{j\omega}, \rho)$ sabendo que a função é contínua pode-se concluir que a diferença de fase entre duas funções de sensibilidade é pequena para frequências próximas de $\omega = 0$ e $\omega = \pi$.

Lema 2

$$\begin{aligned}|\angle S(e^{j\omega}, \rho) - \angle S_d(e^{j\omega})| &< \frac{\pi}{2} \forall \omega \leq \omega_l \\ |\angle S(e^{j\omega}, \rho) - \angle S_d(e^{j\omega})| &< \frac{\pi}{2} \forall \omega \geq \omega_h\end{aligned}$$

De acordo com este resultado, se o espectro da referência tiver componentes apenas nas frequências que a diferença de fase $|\angle S(e^{j\omega}, \rho) - \angle S_d(e^{j\omega})|$ for pequena, então a integral (31) permanece positiva, mesmo se a diferença de fase exceder o limite $\pi/2$ em algumas faixas de frequência.

Portanto se o espectro de referência estiver concentrado nas bordas do espectro de frequência, onde a diferença de fase é pequena, então $M_s(\rho)$ será definida positiva e o mínimo é único no conjunto de estabilidade Γ . Este fato é formalizado no seguinte teorema.

Teorema 8 *Tenha $\bar{C}(e^{j\omega})$ posto completo e seja $r(t)$ PEP. Então $\exists \omega_l, \omega_h$ tal que $\Phi_r(\omega) = 0 \forall \omega \in (\omega_l, \omega_h)$ implica que ρ_d é o único extremo para $J_y(\rho)$ em Γ .*

O Teorema 8 fala que é sempre possível encontrar uma referência para o qual o domínio de atração candidato para o mínimo ρ_d é todo o conjunto Γ . Portanto, não importando qual referência tem-se, pode-se sempre escolher uma outra referência para o qual o problema de otimização pode ser executado de qualquer condição inicial e gerará o mesmo resultado.

Em outras palavras, se a função custo é difícil de minimizar, troca-se a função para outra mais fácil mas que possui o mesmo mínimo. Esta também é a idéia central do método VRFT (CAMPI; LECCHINI; SAVARESI, 2002) e alguns particulares métodos de projeto MRAC (ASTROM; WITTENMARK, 1994). Note, contudo, que este argumento não é válido para o custo combinado $J_y(\rho) + J_e(\rho)$, que é o custo minimizado na presença

de ruído. Apesar disso, pode-se esperar que esta abordagem seja eficiente se a parcela $J_e(\rho)$ for muito menor que $J_y(\rho)$ para parâmetros distantes de ρ_d .

Resultados similares podem ser obtidos para $J_e(\rho)$ e $J_y(\rho)$ apesar de não ser possível utilizá-los para moldar a função custo, dado que não é possível manipular $\Phi_\nu(\omega)$.

Proposição 1 *Tenha $\bar{C}(e^{j\omega})$ posto completo. Então:*

- $\exists \omega_l^v, \omega_h^v$ tal que $\Phi_\nu \approx 0 \forall \omega \in (\omega_l^v, \omega_h^v)$ implica ρ_e ser o único extremo de $J_e(\rho)$ em Γ ;
- $\exists \omega_l^u, \omega_h^u$ tal que $\Phi_u \approx 0 \forall \omega \in (\omega_l^u, \omega_h^u)$ implica ρ_u ser o único extremo de $J_u(\rho)$ em Γ .

Será apresentado na sequência um exemplo que demonstra que a escolha do sinal de referência têm grande influência na convergência para o mínimo global. Se o espectro da referência for trocado por outro mais concentrado nas frequências baixas ou altas, o domínio de atração do mínimo global pode ser aumentado, e uma dada condição inicial que não pertencia ao domínio de atração pode passar a pertencer.

Exemplo 1 *Considere o seguinte sistema sem ruído*

$$y(t) = \frac{0,15}{z - 0,5} u(t).$$

Este sistema é controlado por um controlador do tipo PI

$$C(z, \rho) = \begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix} \begin{bmatrix} \frac{z}{z-1} & \frac{1}{z-1} \end{bmatrix}^T.$$

Como requisito de projeto considera-se apenas o seguimento de referência:

$$y_d(t) = \frac{0,1}{z - 0,9} r(t).$$

Este modelo de referência pode ser alcançado se o seguinte controlador é utilizado

$$C_d(z, \rho) = \begin{bmatrix} 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} \frac{z}{z-1} & \frac{1}{z-1} \end{bmatrix}^T.$$

Como pode-se facilmente notar a Hipótese 2 é respeitada.

Considere que inicialmente o seguinte controlador é utilizado no sistema em malha fechada

$$C(z, \rho) = \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} \frac{z}{z-1} & \frac{1}{z-1} \end{bmatrix}^T.$$

Utiliza-se então o método de ajuste de controladores IFT para melhorar o desempenho do sistema em malha fechada, fazendo atualizações dos parâmetros do controlador a cada iteração. São utilizadas 100 iterações com o seguinte sinal de referência nos ensaios do IFT

$$r_1(t) = 1 + \text{sen} \left(\frac{2\pi t}{10} \right).$$

Na Figura 2 são apresentadas as respostas do sistema em malha fechada a uma entrada do tipo salto unitário. Nota-se na figura que o controlador inicial possui uma resposta bastante diferente da resposta desejada dada pelo modelo de referência. Na Figura 3 vê-se que o controlador final gera uma resposta bastante insatisfatória. Na Figura 4

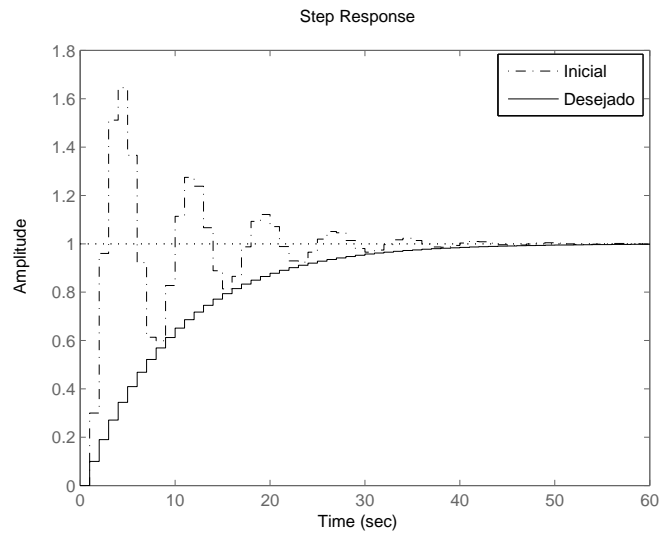


Figura 2: Exemplo 1: Resposta ao salto em malha-fechada controlador inicial.

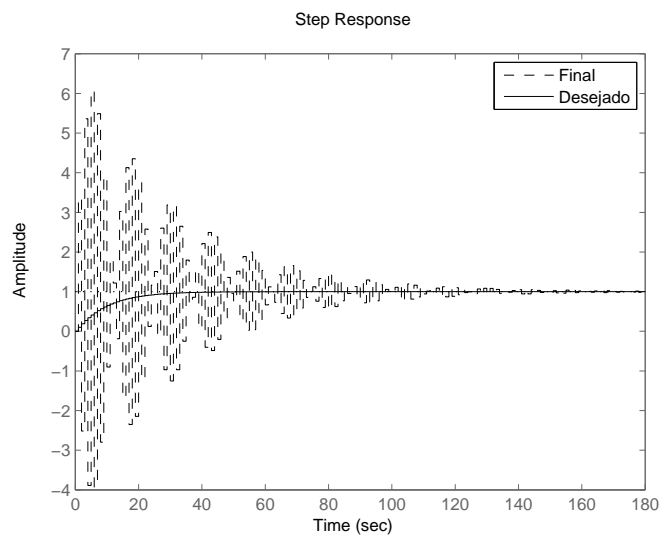


Figura 3: Exemplo 1: Resposta ao salto em malha-fechada controlador final.

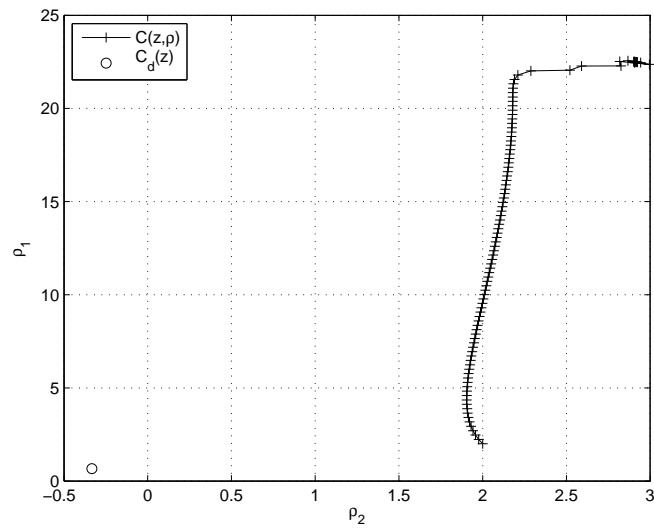


Figura 4: Exemplo 1: Parâmetros do controlador a cada iteração - $r_1(t)$ - Condição Inicial $\rho_1 = [2 \ 2]^T$.

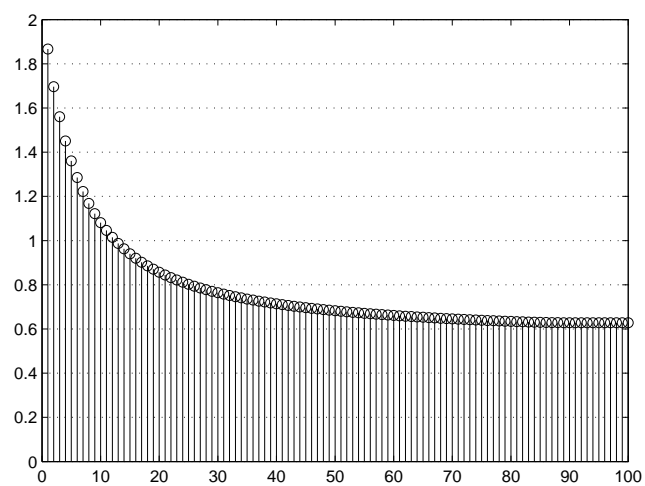


Figura 5: Exemplo 1: Custo a cada iteração - $r_1(t)$.

pode-se ver que os parâmetros do controlador estão se afastando do mínimo global com o passar das iterações. Na Figura 5 pode-se ver que com o passar das iterações o custo não diminui, o que sugere que os parâmetros do controlador acomodaram em extremo da função custo, no caso um mínimo local.

Claramente a condição inicial do algoritmo não pertence ao domínio de atração do mínimo global da função custo. Pode-se, seguindo as idéias de shaping da função custo, escolher outro sinal de referência para os ensaios realizados no processo. A Figura 6 apresenta a fase da função de sensibilidade $S(z, \rho_0)$, da função de sensibilidade desejada $S_d(z)$, e a diferença de fase destas duas funções. Nota-se que a diferença entre as fases entre $S(z, \rho_0)$ e $S_d(z)$ é maior que $\pi/2$ para a frequência $2\pi/10$, que é a frequência do sinal de referência utilizado no experimento. O Teorema 8 sugere que seja utilizado um sinal de referência com componentes do espectro de potência apenas em frequências mais altas ou baixas. Escolheu-se então um sinal de referência que não viole a restrição da diferença entre as fases. Utilizou-se o seguinte sinal de referência:

$$r_2(t) = 1 + \text{sen} \left(\frac{2\pi t}{5} \right).$$

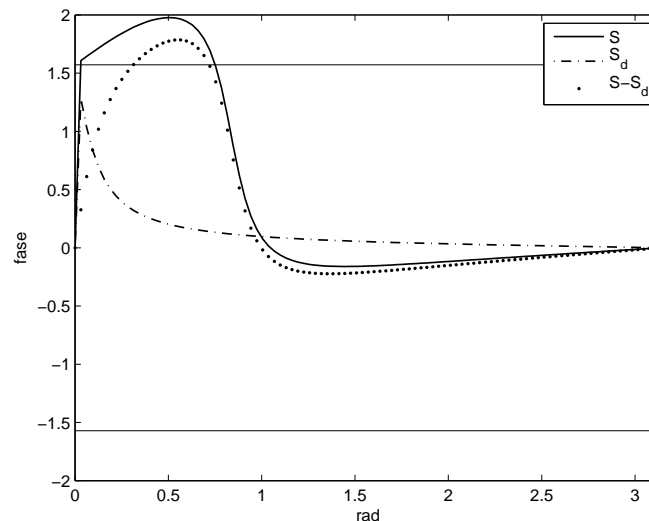


Figura 6: Exemplo 1: Fase das funções de sensibilidade.

Pode-se ver nas Figuras 7 e 8 que a condição inicial do algoritmo pertence ao domínio de atração, para a função custo que utiliza $r_2(t)$, dado que com este novo sinal de referência o algoritmo converge para o mínimo global da função custo.

Pode-se notar com este exemplo que o sinal de referência tem um papel fundamental na convergência para o mínimo global da função custo. Apesar do mínimo global ser o mesmo, para duas funções custo geradas de sinais de referência diferentes, o domínio de atração deste mínimo global é diferente. Funções custo geradas por sinais de referência com espectro concentrado nas baixas e altas frequências são mais “comportadas”, apresentando menos pontos extremos.

3.7.3 Escolha da janela de dados

Foi mostrado que a convergência pode ser alcançada aplicando no sistema em malha-fechada um sinal de referência escolhido adequadamente. Este procedimento requer que

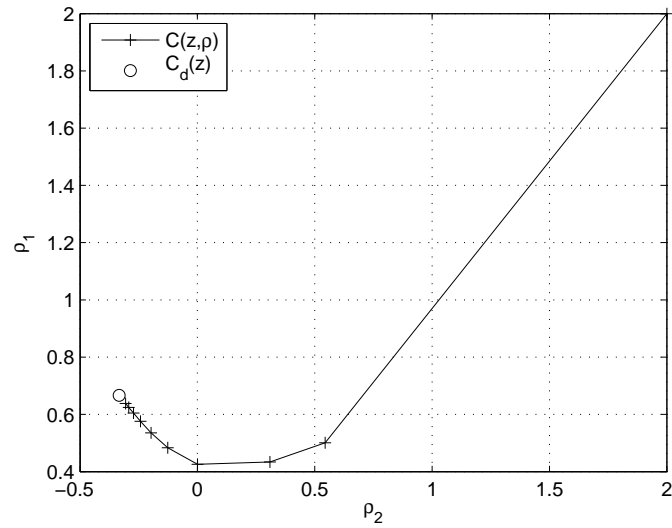


Figura 7: Exemplo 1: Parâmetros do controlador a cada iteração - $r_2(t)$ - Condição Inicial $\rho_1 = [2 \ 2]^T$.

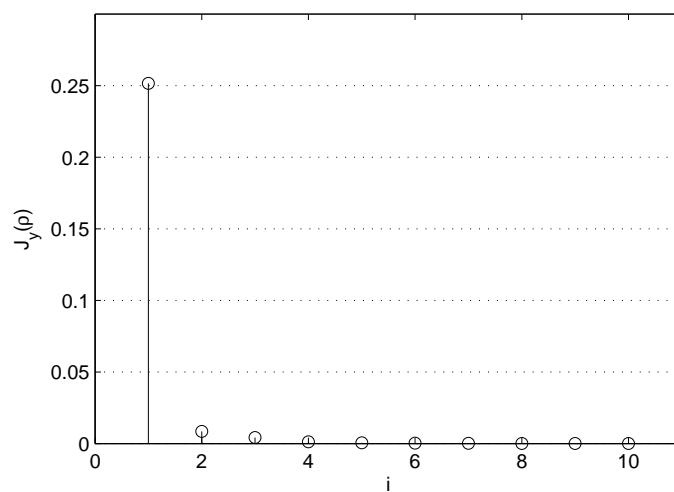


Figura 8: Exemplo 1: Custo a cada iteração - $r_2(t)$.

seja aplicado ao processo um sinal de referência diferente daquele que supostamente o processo deveria seguir. Muitas vezes esta troca do sinal de referência não pode ser implementada na prática. Contudo, a função custo pode ser manipulada de maneira que ela “veja” um espectro do sinal de referência diferente, mesmo que este sinal não seja modificado. Para fazer tal “truque” pode-se utilizar janelas aos dados coletados.

Inicialmente é importante notar que a função custo $J_y(\rho)$ como definida originalmente não é computada na prática. O que pode ser computado é

$$\hat{J}_y(\rho, N) = \frac{1}{N} \sum_{t=1}^N (y(t, \rho) - y_d(t))^2.$$

Se a razão sinal-ruído é grande, então

$$\hat{J}_y(\rho, N) \approx \frac{1}{N} \sum_{t=1}^N [(T(z, \rho) - T_d(z)) r(t)]^2.$$

Sob a premissa que todos os sinais são quase-estacionários, é um resultado padrão que a soma acima converge para $J_y(\rho)$ quando N cresce, ou seja,

$$\lim_{N \rightarrow \infty} \hat{J}_y(\rho, N) = J_y(\rho)$$

Também é claro que $\hat{J}_y(\rho, N)$ é uma função quadrática de $\rho - \rho_d$, portanto ρ_d é um mínimo global isolado de $\hat{J}_y(\rho, N)$, dado que $N > p$ (lembrando que p é a dimensão do vetor de parâmetros).

É desejável que o espectro do sinal de referência seja concentrado nas frequências muito baixas ou muito altas. O espectro de um sinal muda quando medido através de janelas, portanto esta técnica pode ser utilizada para alocar os componentes do espectro do sinal de referência.

Na seqüência alguns fatos apresentados serão formalizados. Seja $e(t) = (T(z, \rho) - T_d(z)) r(t)$ e o sinal $e^N(t)$ obtido como a repetição periódica do sinal $e(t)$ truncado em $t = N$, ou seja,

$$e^N(t + kN) = e(t) \quad t = 1, \dots, N$$

para qualquer k . Calcula-se o critério de desempenho para esse sinal:

$$\begin{aligned} J^N(\rho) &= E [e^N(t)]^2 = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{t=1}^m [e^N(t)]^2 = \lim_{k \rightarrow \infty} \frac{1}{kN} k \sum_{t=1}^N [e^N(t)]^2 \\ &= \frac{1}{N} \sum_{t=1}^N [e^N(t)]^2 = \frac{1}{N} \sum_{t=1}^N [e(t)]^2 = \hat{J}_y(\rho, N) \end{aligned}$$

O custo aproximado $\hat{J}_y(\rho, N)$ é exatamente o custo que seria obtido se a referência $r^N(t) = (T(z, \rho) - T_d(z))^{-1} e^N(t)$ fosse aplicada ao sistema. Portanto utilizar valores diferentes para N nos cálculos tem efeito semelhante com o de utilizar outro sinal de referência. Por exemplo, se um sinal tipo salto é utilizado então pegar mais dados após o transitório tem o mesmo efeito que utilizar um sinal com espectro concentrado em frequências mais baixas. A utilização desta ferramenta ainda não está completamente estudada e ainda é assunto para investigação futura. Para estas referências do tipo salto, as quais são as mais comuns na prática, o raciocínio acima é apropriado.

3.8 Considerações finais

Neste capítulo foram apresentados resultados sobre a convergência para o mínimo global da função custo, do algoritmo do gradiente, no contexto de controle baseado em dados. Utilizando a teoria de Lyapunov e uma região de atração candidata no formato de bola, estabeleceu-se propriedades importantes da convergência do algoritmo.

A convergência depende fundamentalmente das condições iniciais que precisam estar dentro do domínio de atração, e do tamanho do passo de iteração utilizado no algoritmo. Utilizando propriedades da função custo foi desenvolvida uma metodologia chamada Shaping da Função Custo. Nesta metodologia são apresentadas formas de escolher e manipular o espectro do sinal de referência, bem como de escolher modelos de referência intermediários para aumentar o domínio de atração.

No capítulo seguinte será apresentada uma maneira de cálculo do tamanho do passo de iteração que garante a convergência para o mínimo global no menor número possível de passos. O método foi desenvolvido como extensão dos assuntos tratados nesse capítulo.

4 PROJETO DO PASSO DE ITERAÇÃO

4.1 Introdução

Neste capítulo será apresentado um método para o projeto do tamanho do passo de iteração γ , o qual é o resultado principal desse trabalho. O método surge de manipulações algébricas e majorações a partir de (23), de modo a garantir que o algoritmo do gradiente convirja para o mínimo global da função custo, desde que o controlador inicial pertença a um domínio de atração candidato do mínimo global.

Novamente os resultados serão apresentados para os casos: seguimento de referência, atenuação de ruído e minimização do esforço de controle. Da mesma forma que no capítulo anterior, os resultados serão obtidos em separado para os casos de modelo de referência atingível e inatingível.

4.2 Seguimento de referência

4.2.1 Modelo de referência atingível

Nesta seção será analisada a parte da função custo do seguimento de referência quando a Hipótese 2 é satisfeita, portanto o caso em que o modelo de referência é atingível.

Substituindo (30) em (22) tem-se:

$$\gamma_i < 2 \frac{(\rho_i - \rho_d)^T M_y(\rho)(\rho_i - \rho_d)}{(\rho_i - \rho_d)^T M_y^T(\rho) M_y(\rho)(\rho_i - \rho_d)}$$

Utilizando a propriedade de anti-simétrica de $M_{ya}(\rho)$, sabe-se que $(\rho_i - \rho_d)^T M_{ya}(\rho)(\rho_i - \rho_d)$ e pode-se reescrever a equação anterior na forma

$$\gamma_i < 2 \frac{(\rho_i - \rho_d)^T M_{ys}(\rho)(\rho_i - \rho_d)}{(\rho_i - \rho_d)^T M_y(\rho)^T M_y(\rho)(\rho_i - \rho_d)}$$

Se o controlador inicial dado por ρ_0 pertencer a um domínio de atração candidato $\mathcal{B}_\alpha(\rho_d)$, então utilizando (34) sabe-se que $M_{ys}(\rho) > 0$ e portanto, utilizando propriedades de normas de matrizes positivas definidas simétricas, tem-se que

$$\frac{(\rho_i - \rho_d)^T M_{ys}(\rho)(\rho_i - \rho_d)}{(\rho_i - \rho_d)^T M_y(\rho)^T M_y(\rho)(\rho_i - \rho_d)} > \frac{\|(\rho_i - \rho_d)\|^2 \min(\text{eig}(M_{ys}(\rho)))}{\|(\rho_i - \rho_*)^T M_y(\rho)^T M_y(\rho)(\rho_i - \rho_d)\|}$$

e novamente utilizando propriedades de normas matriciais

$$\frac{\|(\rho_i - \rho_d)\|^2 \min(\text{eig}(M_{ys}(\rho)))}{(\rho_i - \rho_d)^T M_y(\rho)^T M_y(\rho)(\rho_i - \rho_d)} > \frac{\|(\rho_i - \rho_d)\|^2 \min(\text{eig}(M_{ys}(\rho)))}{\|(\rho_i - \rho_d)\|^2 \|M_y(\rho)^T M_y(\rho)\|} = \frac{\min(\text{eig}(M_{ys}(\rho)))}{\|M_y(\rho)^T M_y(\rho)\|}$$

onde $\text{eig}(M_{ys}(\rho))$ denota os autovalores da matriz $M_{ys}(\rho)$.

O desenvolvimento acima prova o seguinte teorema:

Teorema 9 *Seja $\mathcal{B}_\alpha(\rho_d) \subset \Gamma$, se as seguintes restrições forem respeitadas*

$$M_{ys}(\rho) > 0 \quad \forall \rho \in \mathcal{B}_\alpha(\rho_d); \quad \rho \neq \rho_d$$

e

$$\gamma_i < 2 \frac{\min(\text{eig}(M_{ys}(\rho_i)))}{\|M_y(\rho_i)^T M_y(\rho_i)\|}, \quad (52)$$

então $\mathcal{B}_\alpha(\rho_d)$ é um domínio de atração de ρ_d .

O Teorema 9 diz que se o controlador inicial ρ_0 estiver dentro de um domínio de atração candidato, então utilizando a regra (52) para a escolha do tamanho do passo de iteração, garante-se que o algoritmo do gradiente converge para o mínimo global da função custo ρ_d .

A expressão (52) é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição (52) é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

Para o cálculo de γ_i utilizando (52) é necessário conhecer $M_y(\rho)$. Essa matriz é função do processo o qual é desconhecido. Para utilizar (52) é preciso de um estimador para $M_y(\rho)$. Um estimador para esta matriz é proposto neste trabalho e apresentado com detalhes no Capítulo 5.

Observação 1 *Garante-se com o Teorema 9 que a cada iteração os parâmetros do controlador estarão em um conjunto $\mathcal{B}_\alpha(\rho_d)$ com raio menor ($\alpha_i < \alpha_{i-1}$) e portanto os parâmetros do controlador estarão mais próximos do mínimo global ρ_d .*

Esse fato não significa que o valor da função custo será menor a cada iteração, pois podem existir controladores mais próximos do mínimo global cujo custo é maior do que de controladores mais distantes.

Observação 2 *Quando a condição inicial do algoritmo pertence a estimativa do domínio de atração, ou seja, $\rho_0 \in \mathcal{B}_\alpha(\rho_d)$, garante-se que os controladores de todas as iterações tornarão o sistema em malha-fechada estável, e que a cada iteração os parâmetros do controlador estarão mais próximos do mínimo global.*

Em casos práticos não é possível saber a priori se $\rho_0 \in \mathcal{B}_\alpha(\rho_d)$. Se a condição inicial estiver fora de $\mathcal{B}_\alpha(\rho_d)$, tanto pode ocorrer convergência para um mínimo local, como o controlador obtido pode tornar o sistema em malha-fechada instável.

Os trabalhos (KAMMER; BITMEAD; BARTLETT, 2000) (KAMMER, 2005) utilizam a norma ν - gap (VINNICOMBE, 1993) como medida de estabilidade do sistema, de modo que dado um controlador que estabiliza o sistema em malha-fechada, pode-se saber se um outro controlador também estabiliza esse sistema. Esta medida ν -gap pode ser estimada a partir de dados de experimentos realizados em malha-fechada. Um tópico de trabalho futuro é estudar a utilização dessa norma, para restringir o tamanho do passo de iteração, para que mesmo se os parâmetros do controlador estiverem fora de $\mathcal{B}_\alpha(\rho_d)$, o sistema em malha-fechada não se torne instável.

Na sequência será apresentado um exemplo que demonstra o funcionamento do ajuste dos parâmetros do controlador, utilizando o algoritmo do gradiente e o tamanho do passo de iteração proposto para o caso de seguimento de referência. Paralelamente será utilizado o método “clássico” de escolha do tamanho do passo, para fins de comparação.

Exemplo 2 Considere o seguinte sistema sem ruído

$$y(t) = \frac{0,15}{z - 0,5} u(t).$$

Este sistema é controlado por um controlador do tipo PI

$$C(z, \rho) = [k_p \ k_i] \left[1 \quad \frac{z}{z-1} \right]^T$$

que pode ser reescrito como

$$C(z, \rho) = [\rho_1 \ \rho_2] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T.$$

Como requisito de projeto considera-se apenas o seguimento de referência:

$$y_d(t) = \frac{0,1}{z - 0,9} r(t).$$

Este modelo de referência pode ser alcançado se o seguinte controlador é utilizado

$$C_d(z, \rho) = [2/3 \ -1/3] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T.$$

Como pode-se facilmente notar a Hipótese 2 é respeitada.

Considere que inicialmente o seguinte controlador está sendo utilizado no sistema em malha fechada

$$C(z, \rho) = [2 \ 2] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T.$$

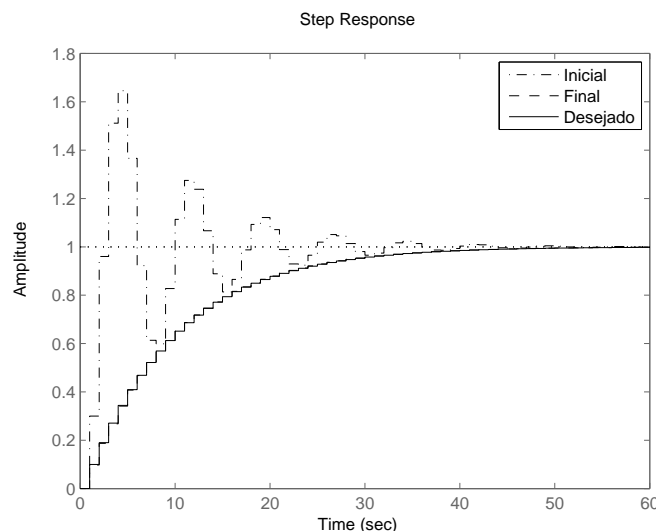


Figura 9: Exemplo 2: Resposta ao salto em malha-fechada - método proposto.

Na Figura 9 pode-se ver as respostas do sistema em malha fechada a uma entrada do tipo salto unitário. Pode-se notar na figura que o controlador inicial gera uma resposta bastante diferente da resposta desejada dada pelo modelo de referência.

Utiliza-se então o método de ajuste de controladores IFT para melhorar o desempenho do sistema em malha fechada, fazendo atualizações dos parâmetros do controlador a cada iteração. É utilizado o seguinte sinal de referência nos ensaios do IFT

$$r(t) = 1 + \text{sen} \left(\frac{2\pi t}{5} \right).$$

Neste momento considera-se que existe um estimador perfeito para a função matricial $M_y(\rho)$ de tal forma que o tamanho do passo do algoritmo γ_i possa ser calculado utilizando a equação 52.

São utilizadas 10 iterações com o algoritmo IFT para melhorar o desempenho em malha fechada do sistema. Na Figura 10 são mostrados os parâmetros do controlador a cada iteração. Nota-se que o algoritmo está convergindo para o controlador ótimo $C_d(z)$. Na Figura 11 é apresentado o custo $J(\rho)$ a cada iteração. Nota-se que o custo cai drasticamente nas primeiras 10 iterações tornando-se muito próximo de 0, o que indica que a resposta do sistema está muito próxima da resposta desejada.

Olhando mais atentamente para a Figura 9 pode-se notar que também está plotada a resposta do controlador após as 10 iterações. Esta resposta está tão próxima da resposta ideal que as duas resposta tornam-se indistinguíveis graficamente.

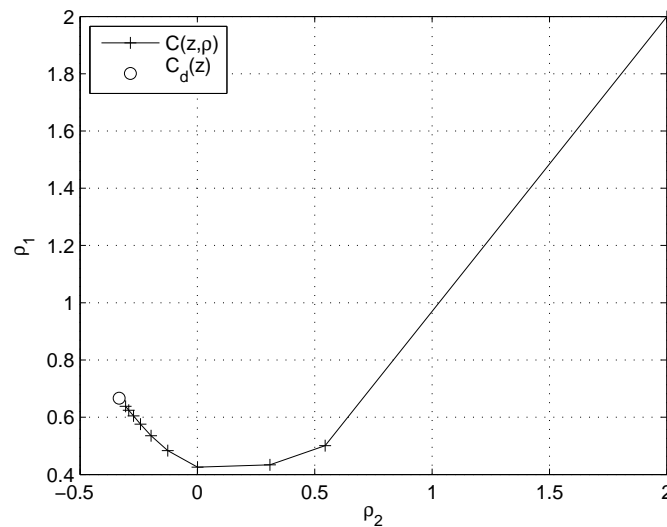


Figura 10: Exemplo 2: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [2 \ 2]^T$.

Para fins de comparação é apresentado na Figura 12 a evolução dos parâmetros do controlador quando o método “clássico” é utilizado para o cálculo do tamanho do passo de iteração:

$$\gamma_i = \frac{\gamma_0}{i}; \quad i > 0.$$

onde i é o número da iteração.

O método clássico permite calcular o tamanho do passo para todas iterações exceto a primeira. O tamanho do primeiro passo deve ser escolhido pelo projetista, sendo que

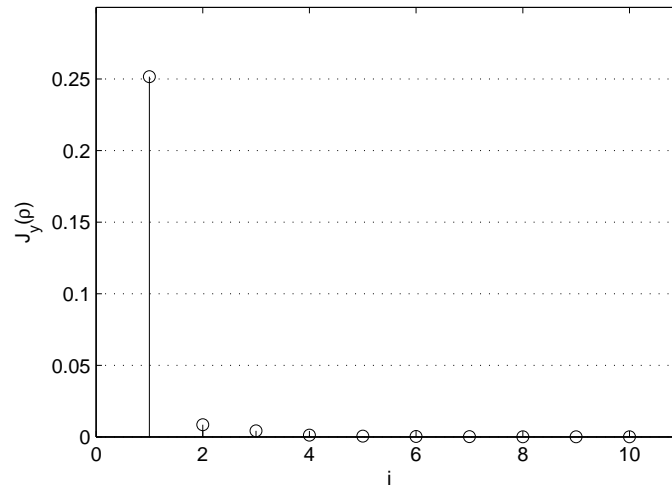


Figura 11: Exemplo 2: Custo a cada iteração - método proposto.

Tabela 1: Exemplo 2: Evolução dos controladores - método proposto.

i	$J(\rho)$	ρ_1	ρ_2	γ
1	0,25166	2,00000	2,00000	6.7828
2	0,008466	0,50077	0,54432	21.9079
3	0,0042011	0,43355	0,30892	22.6770
4	0,0012646	0,42595	0,00060	22.1835
5	0,00055415	0,48352	-0,12615	22.0459
6	0,00025667	0,53557	-0,19775	22.1847
7	0,00011908	0,57561	-0,24279	22.3689
8	$5,4756e-05$	0,60453	-0,27241	22.5236
9	$2,4959e-05$	0,62469	-0,29226	22.6365
10	$1,1302e-05$	0,63847	-0,30566	22.7143

não é encontrada na literatura maneiras de fazer esta escolha. Neste exemplo é utilizado o mesmo tamanho de passo da primeira iteração do método proposto, para que a comparação seja mais clara.

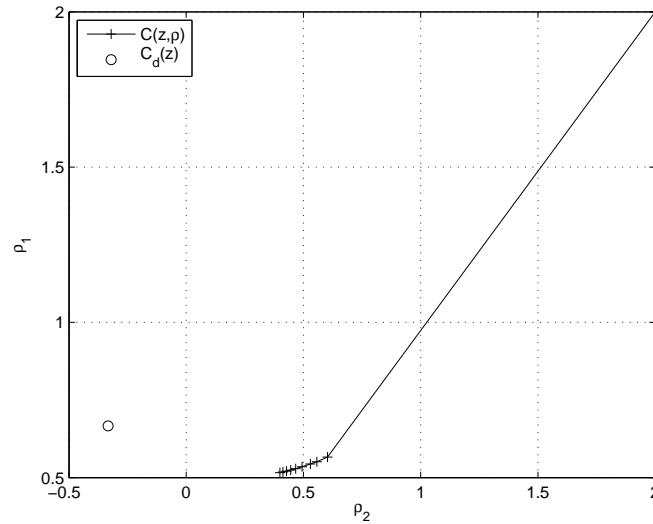


Figura 12: Exemplo 2: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [2 \ 2]^T$.

Também são utilizadas 10 iterações com o método IFT. Como visto na Figura 12, o método clássico gera passos muito pequenos, o que torna a convergência muito lenta. O primeiro passo é grande, pois foi escolhido pelo método proposto, já os outros passos determinados pelo método clássico são muito pequenos. Na Figura 13 é mostrada a resposta ao salto do sistema em malha-fechada, para o controlador inicial, controlador alcançado após as 10 iterações e para o modelo de referência. Após as 10 iterações a resposta ainda é bastante insatisfatória, enquanto que com o método proposto a resposta é praticamente aquela desejada.

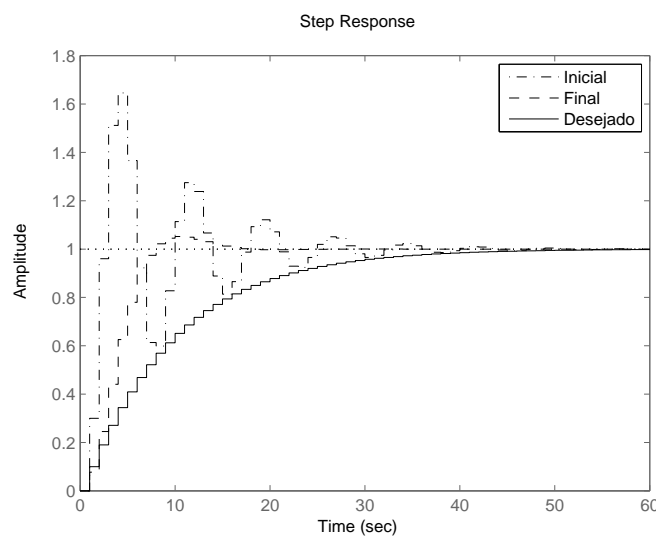


Figura 13: Exemplo 2: Resposta ao salto em malha-fechada - método clássico.

4.2.2 Modelo de referência inatingível

Nesta seção será analisada a parte da função custo do seguimento de referência quando a Hipótese 2 não é satisfeita, portanto o caso em que o modelo de referência não é atingível.

Substituindo (38) em (22) tem-se:

$$\gamma < 2 \frac{(\rho - \rho_*)^T \nabla J_y(\rho)}{\nabla J_y^T(\rho) \nabla J_y(\rho)} = \frac{(\rho - \rho_*)^T [M_y(\rho)(\rho - \rho_*) - m_y(\rho)]}{\nabla J_y^T(\rho) \nabla J_y(\rho)}.$$

Utilizando a propriedade sobre a matriz anti-simétrica $M_{ya}(\rho)$ obtém-se:

$$\frac{(\rho - \rho_*)^T [M_y(\rho)(\rho - \rho_*) - m(\rho)]}{\nabla J_y^T(\rho) \nabla J_y(\rho)} = \frac{(\rho - \rho_*)^T M_{ys}(\rho)(\rho - \rho_*) - (\rho - \rho_*)^T m(\rho)}{\nabla J_y^T(\rho) \nabla J_y(\rho)}$$

Utilizando a condição (39):

$$\frac{(\rho - \rho_*)^T M_{ys}(\rho)(\rho - \rho_*) - (\rho - \rho_*)^T m_y(\rho)}{\nabla J_y^T(\rho) \nabla J_y(\rho)} > \frac{(\rho - \rho_*)^T (M_{ys}(\rho) - \kappa_y I)(\rho - \rho_*)}{\nabla J_y^T(\rho) \nabla J_y(\rho)}$$

Se $M_{ys}(\rho) > \kappa_y I$, utilizando propriedades de normas matriciais :

$$\frac{(\rho - \rho_*)^T (M_{ys}(\rho) - \kappa_y I)(\rho - \rho_*)}{\nabla J_y^T(\rho) \nabla J_y(\rho)} > \frac{\min(\text{eig}(M_{ys}(\rho) - \kappa_y I)) \|(\rho - \rho_*)\|^2}{\nabla J_y^T(\rho) \nabla J_y(\rho)}$$

Sobre o denominador dessa inequação pode-se escrever:

$$\nabla J_y^T(\rho) \nabla J_y(\rho) = [M(\rho)(\rho - \rho_*) - m_y(\rho)]^T [M_y(\rho)(\rho - \rho_*) - m_y(\rho)]$$

e portanto:

$$\nabla J_y^T(\rho) \nabla J_y(\rho) = (\rho - \rho_*)^T M_y^T(\rho) M_y(\rho)(\rho - \rho_*) + m_y^T(\rho) m_y(\rho) - 2(\rho - \rho_*)^T M_y(\rho) m_y(\rho)(\rho - \rho_*)$$

E, sobre a norma deste denominador pode-se estabelecer que:

$$\begin{aligned} \|\nabla J_y^T(\rho) \nabla J_y(\rho)\| &< \|(\rho - \rho_*)^T M_y^T(\rho) M_y(\rho)(\rho - \rho_*)\| + \|m_y^T(\rho) m_y(\rho)\| \\ &\quad + \|2(\rho - \rho_*)^T M_y(\rho) m_y(\rho)(\rho - \rho_*)\| \end{aligned}$$

e novamente utilizando a condição (39):

$$\|\nabla J_y^T(\rho) \nabla J_y(\rho)\| < \|(\rho - \rho_*)\|^2 [\|M_y^T(\rho) M_y(\rho)\| + \kappa_y^2 + 2\kappa_y \|M_y(\rho)\|]$$

Logo, pode-se estabelecer a seguinte condição sobre o tamanho do passo de iteração:

$$\gamma < 2 \frac{\min(\text{eig}(M_{ys}(\rho) - \kappa_y I))}{\|M_y^T(\rho) M_y(\rho)\| + \kappa_y^2 + 2\kappa_y \|M_y(\rho)\|}$$

onde $\text{eig}(M_{ys}(\rho))$ denota os autovalores da matriz $M_{ys}(\rho)$.

O desenvolvimento acima prova o seguinte teorema:

Teorema 10 *Seja $\mathcal{B}_\alpha(\rho_d) \subset \Gamma$, se as seguintes restrições forem respeitadas*

$$I. \exists \kappa_y \in \mathbb{R}^+ : |m_y(\rho)| < \kappa_y |\rho - \rho_*| \quad \forall \rho \in \mathcal{B}_\alpha(\rho_*)$$

2. $M_{ys}(\rho) > \kappa_y I \forall \rho \in \mathcal{B}_\alpha(\rho_*); \rho \neq \rho_*$

3.

$$\gamma_i < 2 \frac{\min(\text{eig}(M_{ys}(\rho) - \kappa_y I))}{\|M_y^T(\rho)M_y(\rho)\| + \kappa_y^2 + 2\kappa_y \|M_y(\rho)\|} \quad (53)$$

então $\mathcal{B}_\alpha(\rho_*)$ é um domínio de atração de ρ_d .

O Teorema 10 afirma que se o controlador inicial ρ_0 estiver dentro do domínio de atração candidato $\mathcal{B}_\alpha(\rho_*)$, então utilizando a regra (53) para a escolha do tamanho do passo de iteração, garante-se que o algoritmo do gradiente converge para o mínimo global da função custo ρ_* .

A expressão (53) é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição (53) é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

4.3 Atenuação da sensibilidade ao ruído

4.3.1 Modelo de referência atingível

Nesta seção será analisada a parcela da função custo de atenuação da sensibilidade ao ruído quando a Hipótese 4 é satisfeita, portanto o caso em que o modelo de referência é atingível.

Substituindo (42) em (22) tem-se:

$$\gamma_i < 2 \frac{(\rho_i - \rho_e)^T M_e(\rho)(\rho_i - \rho_e)}{(\rho_i - \rho_e)^T M_e(\rho)^T M_e(\rho)(\rho_i - \rho_e)}$$

Utilizando a propriedade de anti-simétrica de $M_{ea}(\rho)$, sabe-se que $(\rho_i - \rho_e)^T M_{ea}(\rho)(\rho_i - \rho_e)$ e pode-se reescrever a equação anterior na forma

$$\gamma_i < 2 \frac{(\rho_i - \rho_e)^T M_{es}(\rho)(\rho_i - \rho_e)}{(\rho_i - \rho_e)^T M_e(\rho)^T M_e(\rho)(\rho_i - \rho_e)}$$

Se o controlador inicial dado por ρ_0 pertencer a um domínio de atração candidato $\mathcal{B}_\alpha(\rho_e)$, então utilizando (44) sabe-se que $M_{es}(\rho) > 0$ e portanto utilizando propriedades de normas de matrizes positivas definidas simétricas tem-se que

$$\frac{(\rho_i - \rho_e)^T M_{es}(\rho)(\rho_i - \rho_e)}{(\rho_i - \rho_*)^T M_e(\rho)^T M_e(\rho)(\rho_i - \rho_e)} > \frac{\|(\rho_i - \rho_e)\|^2 \min(\text{eig}(M_{es}(\rho)))}{(\rho_i - \rho_*)^T M_e(\rho)^T M_e(\rho)(\rho_i - \rho_*)}$$

e novamente utilizando propriedades de normas matriciais

$$\frac{\|(\rho_i - \rho_e)\|^2 \min(\text{eig}(M_{es}(\rho)))}{(\rho_i - \rho_e)^T M_e(\rho)^T M_e(\rho)(\rho_i - \rho_e)} > \frac{\|(\rho_i - \rho_e)\|^2 \min(\text{eig}(M_{es}(\rho)))}{\|(\rho_i - \rho_e)\|^2 \|M_e(\rho)^T M_e(\rho)\|} = \frac{\min(\text{eig}(M_{es}(\rho)))}{\|M_e(\rho)^T M_e(\rho)\|}$$

onde $\text{eig}(M_{es}(\rho))$ denota os autovalores da matriz $M_{es}(\rho)$.

O desenvolvimento descrito acima prova o seguinte teorema:

Teorema 11 *Se as seguintes restrições forem respeitadas*

$$M_{es}(\rho) > 0 \quad \forall \rho \in \mathcal{B}_\alpha(\rho_e); \quad \rho \neq \rho_e$$

e

$$\gamma_i < 2 \frac{\min(\text{eig}(M_{es}(\rho_i)))}{\|M_e(\rho_i)^T M_e(\rho_i)\|}, \quad (54)$$

garante-se que para qualquer $\rho_0 \in \mathcal{B}_\alpha(\rho_e)$ o algoritmo (11) converge para ρ_e .

A expressão (54) é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição (54) é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

Na seqüência será apresentado um exemplo que demonstra o funcionamento do ajuste dos parâmetros do controlador, utilizando o algoritmo do gradiente e o tamanho do passo de iteração proposto para o caso de atenuação da sensibilidade ao ruído. Paralelamente será utilizado o método “clássico” de escolha do tamanho do passo, para fins de comparação.

Exemplo 3 *Considere o seguinte sistema com ruído*

$$y(t) = \frac{1}{z - 0,7} u(t) + \frac{z - 0,6}{z - 0,8} e(t).$$

onde o ruído gaussiano $e(t)$ tem variância $\sigma_e^2 = 1$

Este sistema é controlado por um controlador do avanço/atraso com pólo fixo:

$$C(z, \rho) = \frac{\rho_1 z + \rho_2}{z - 0,8}$$

Como requisito de projeto considera-se apenas a rejeição de ruído. O mínimo ruído na saída pode ser alcançado utilizando o controlador calculado utilizando 8:

$$C_e(z, \rho) = [0, 2 \quad -0, 14] \left[\frac{z}{z-0,8} \quad \frac{1}{z-0,8} \right]^T.$$

Como pode-se facilmente notar a Hipótese 4 é respeitada.

Considere que inicialmente o seguinte controlador é utilizado no sistema em malha fechada

$$C(z, \rho) = [0, 4 \quad 0, 4] \left[\frac{z}{z-0,8} \quad \frac{1}{z-0,8} \right]^T.$$

Utiliza-se então o método de ajuste de controladores IFT, para melhorar o desempenho do sistema em malha fechada, fazendo atualizações dos parâmetros do controlador a cada iteração. O sinal de referência é nulo, portanto trata-se apenas do problema de regulação.

$$r(t) = 0.$$

Neste momento considera-se que existe um estimador perfeito para a função matricial $M_e(\rho)$ de tal forma que o tamanho do passo do algoritmo γ_i possa ser calculado utilizando a equação (54).

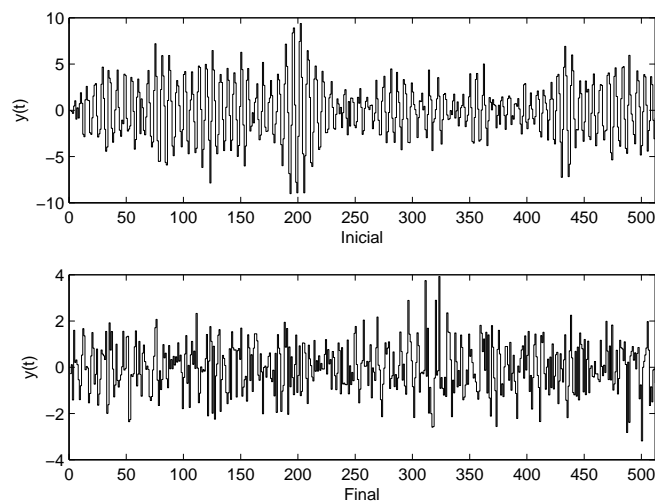


Figura 14: Exemplo 3: Resposta em malha-fechada - método proposto.

São utilizadas 20 iterações com o algoritmo IFT para melhorar o desempenho em malha fechada do sistema. Na Figura 15 são mostrados os parâmetros do controlador a cada iteração. Nota-se que o algoritmo está convergindo para o controlador ótimo $C_e(z)$. Na Figura 16 é apresentado o custo $J(\rho)$ a cada iteração. Nota-se que o custo cai com o aumento das iterações tornando-se muito próximo de 1, como era esperado.

Na Figura 14 pode-se ver a saída do sistema em malha fechada. Pode-se notar na figura que o controlador inicial gera uma resposta pior que a resposta alcançada, o que é consistente com o valores dos custos inicial e final.

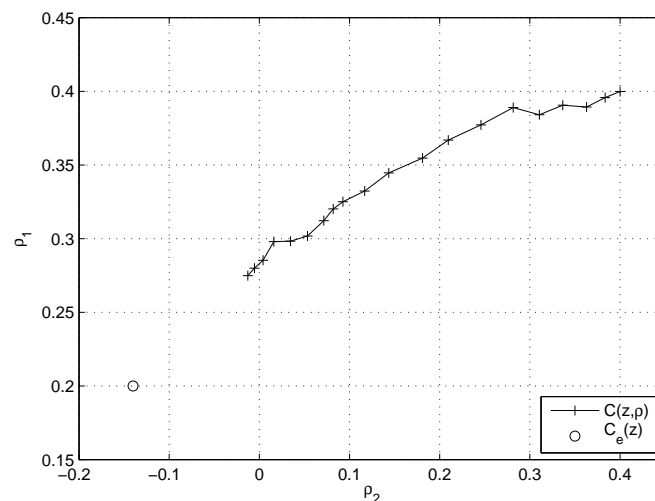


Figura 15: Exemplo 3: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [0, 4, 0, 4]^T$.

Para fins de comparação é apresentado na Figura 17 a evolução dos parâmetros do controlador quando o método “clássico” é utilizado para o cálculo do tamanho do passo de iteração. Como para o caso de seguimento de referência, o tamanho do passo da primeira iteração é o mesmo tamanho de passo do método proposto.

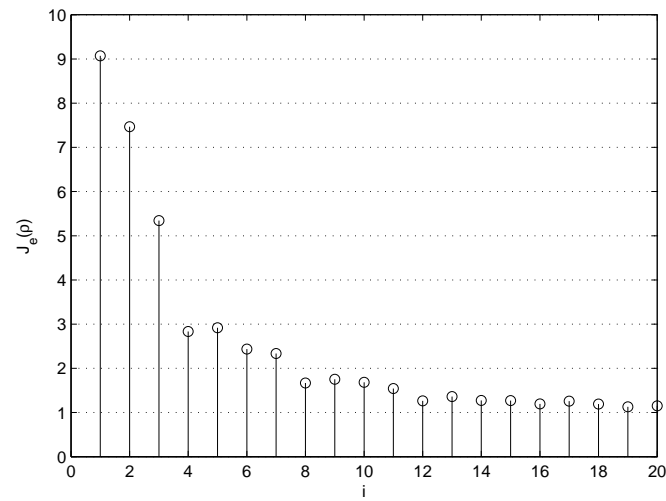


Figura 16: Exemplo 3: Custo a cada iteração - método proposto.

Tabela 2: Exemplo 3: Evolução dos controladores - método proposto.

i	J	ρ_1	ρ_2	γ
1	9,0711	0,40000	0,40000	$7.2970e - 05$
2	7,4674	0,39584	0,38352	$3.4113e - 04$
3	5,3427	0,38939	0,36290	$3.6875e - 04$
4	2,8327	0,39066	0,33643	$4.2315e - 04$
5	2,9169	0,38420	0,31050	$4.5024e - 04$
6	2,4353	0,38901	0,28145	$4.9651e - 04$
7	2,3369	0,37731	0,24569	$4.4921e - 04$
8	1,6680	0,36697	0,20956	$4.9380e - 04$
9	1,7515	0,35479	0,18100	$4.7433e - 04$
10	1,6836	0,34471	0,14357	$5.2538e - 04$
11	1,5433	0,33238	0,11684	$5.2396e - 04$
12	1,2646	0,32514	0,09271	$5.1083e - 04$
13	1,3626	0,32021	0,08191	$5.4788e - 04$
14	1,2735	0,31231	0,07144	$5.7572e - 04$
15	1,2711	0,30179	0,05331	$5.4187e - 04$
16	1,1948	0,29839	0,03457	$5.5809e - 04$
17	1,2583	0,29807	0,01600	$6.3935e - 04$
18	1,1926	0,28536	0,00411	$5.6515e - 04$
19	1,1280	0,28006	-0,00545	$6.3162e - 04$
20	1,1500	0,27498	-0,01272	$6.5181e - 04$

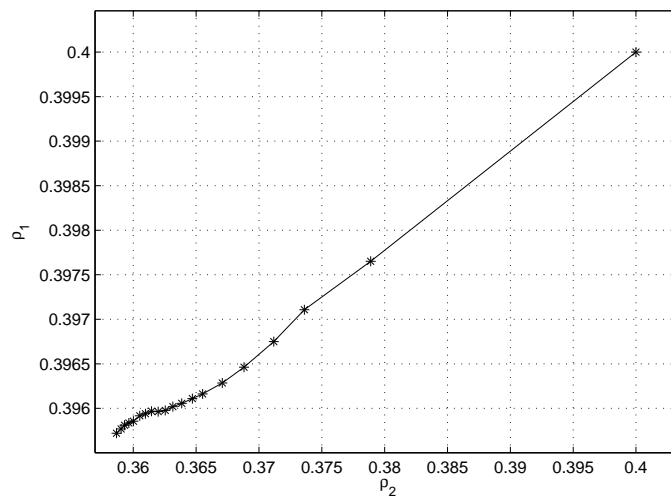


Figura 17: Exemplo 3: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [0, 4, 0, 4]^T$.

Também são utilizadas 20 iterações com o método IFT. Como visto na Figura 17, o método clássico gera passos muito pequenos, o que torna a convergência muito lenta. O primeiro passo é razoável pois foi escolhido pelo método proposto, já os outros passos, determinados pelo método clássico, são muito pequenos. Na Figura 18 é mostrado o custo $J_e(\rho)$ a cada iteração. Após as 20 iterações o custo ainda é maior do que 4, enquanto que com o método proposto o custo alcançado é muito próximo do custo ótimo, que é 1.

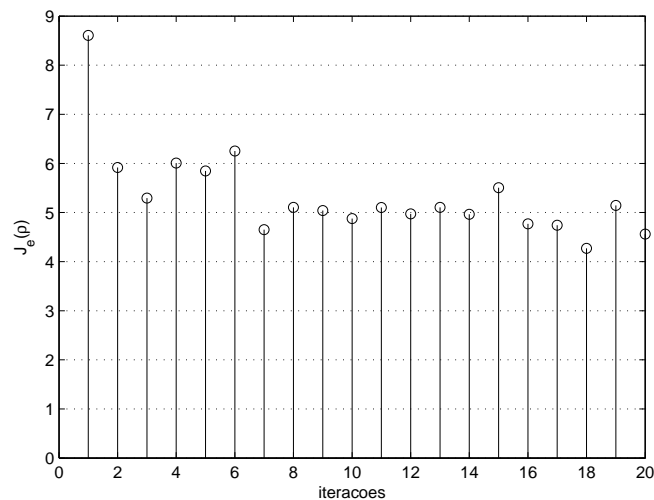


Figura 18: Exemplo 3: Custo a cada iteração - método clássico.

4.3.2 Modelo de referência inatingível

Nesta seção será analisada a parte da função custo de atenuação da sensibilidade ao ruído quando a Hipótese 4 não é satisfeita, portanto o caso de modelo de referência inatingível.

Substituindo (47) em (22) tem-se:

$$\gamma < 2 \frac{(\rho - \rho_*)^T \nabla J_e(\rho)}{\nabla J_e^T(\rho) \nabla J_e(\rho)} = \frac{(\rho - \rho_*)^T [M_e(\rho)(\rho - \rho_*) - m_e(\rho)]}{\nabla J_e^T(\rho) \nabla J_e(\rho)}$$

Utilizando a propriedade sobre a matriz anti-simétrica $M_e a(\rho)$ obtém-se:

$$\frac{(\rho - \rho_*)^T [M_e(\rho)(\rho - \rho_*) - m_e(\rho)]}{\nabla J_e^T(\rho) \nabla J_e(\rho)} = \frac{(\rho - \rho_*)^T M_{es}(\rho)(\rho - \rho_*) - (\rho - \rho_*)^T m_e(\rho)}{\nabla J_e^T(\rho) \nabla J_e(\rho)}$$

Utilizando a condição (48):

$$\frac{(\rho - \rho_*)^T M_{es}(\rho)(\rho - \rho_*) - (\rho - \rho_*)^T m_e(\rho)}{\nabla J_e^T(\rho) \nabla J_e(\rho)} > \frac{(\rho - \rho_*)^T (M_{es}(\rho) - \kappa_e I)(\rho - \rho_*)}{\nabla J_e^T(\rho) \nabla J_e(\rho)}$$

Se $M_{es}(\rho) > \kappa_e I$, utilizando propriedades de normas matriciais :

$$\frac{(\rho - \rho_*)^T (M_{es}(\rho) - \kappa_e I)(\rho - \rho_*)}{\nabla J_e^T(\rho) \nabla J_e(\rho)} > \frac{\min(\text{eig}(M_{es}(\rho) - \kappa_e I)) \|(\rho - \rho_*)\|^2}{\nabla J_e^T(\rho) \nabla J_e(\rho)}$$

Sobre o denominador dessa inequação pode-se escrever:

$$\nabla J_e^T(\rho) \nabla J_e(\rho) = [M_e(\rho)(\rho - \rho_*) - m_e(\rho)]^T [M_e(\rho)(\rho - \rho_*) - m_e(\rho)]$$

e portanto:

$$\nabla J_e^T(\rho) \nabla J_e(\rho) = (\rho - \rho_*)^T M_e^T M_e(\rho)(\rho - \rho_*) + m_e^T(\rho) m_e(\rho) - 2(\rho - \rho_*)^T M_e(\rho) m_e(\rho)(\rho - \rho_*)$$

Agora, sobre a norma deste denominador pode-se estabelecer que:

$$\begin{aligned} \|\nabla J_e^T(\rho) \nabla J_e(\rho)\| &< \|(\rho - \rho_*)^T M_e^T(\rho) M_e(\rho)(\rho - \rho_*)\| + \|m_e^T(\rho) m_e(\rho)\| \\ &\quad + \|2(\rho - \rho_*)^T M_e(\rho) m_e(\rho)(\rho - \rho_*)\| \end{aligned}$$

e novamente utilizando a condição (48):

$$\|\nabla J_e^T(\rho) \nabla J_e(\rho)\| < \|(\rho - \rho_*)\|^2 [\|M_e^T(\rho) M_e(\rho)\| + \kappa_e^2 + 2\kappa_e \|M_e(\rho)\|]$$

Portanto pode-se estabelecer a seguinte condição sobre o tamanho do passo de iteração:

$$\gamma < 2 \frac{\min(\text{eig}(M_{es}(\rho) - \alpha I))}{\|M_e^T(\rho) M_e(\rho)\| + \kappa_e^2 + 2\kappa_e \|M_e(\rho)\|}$$

onde $\text{eig}(M_{ys}(\rho))$ denota os autovalores da matriz $M_{es}(\rho)$.

O desenvolvimento descrito acima prova o seguinte teorema:

Teorema 12 *Seja $\mathcal{B}_\alpha(\rho_e) \subset \Gamma$, se as seguintes restrições forem respeitadas*

1. $\exists \kappa_e \in \mathbb{R}^+ : |m_e(\rho)| < \kappa_e |\rho - \rho_*| \forall \rho \in \mathcal{B}_\alpha(\rho_*)$

2. $M_{es}(\rho) > \kappa_e I \forall \rho \in \mathcal{B}_\alpha(\rho_*); \rho \neq \rho_*$

- 3.

$$\gamma_i < 2 \frac{\min(\text{eig}(M_{es}(\rho) - \kappa_e I))}{\|M_e^T(\rho) M_e(\rho)\| + \kappa_e^2 + 2\kappa_e \|M_e(\rho)\|}$$

então $\mathcal{B}_\alpha(\rho_*)$ é um domínio de atração de ρ_e .

O Teorema 12 afirma que se o controlador inicial ρ_0 estiver dentro do domínio de atração candidato $\mathcal{B}_\alpha(\rho_*)$, então utilizando a regra (3) para a escolha do tamanho do passo de iteração, garante-se que o algoritmo do gradiente converge para o mínimo global da função custo ρ_* .

A expressão 3 é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição 3 é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

4.4 Minimização do esforço de controle

Nesta seção será analisada a parte da função custo do esforço de controle. Como estamos considerando que o controlador é parametrizado linearmente, ou seja, que a Hipótese 1 é sempre respeitada, então o controlador que minimiza o esforço de controle $\rho_u = 0$ sempre pode ser alcançado. Então não existe o caso de modelo de referência inatingível.

Substituindo (49) em (22) tem-se:

$$\gamma_i < 2 \frac{\rho_i^T M_u(\rho) \rho_i}{\rho_i^T M_u(\rho_i)^T M(\rho_i) \rho_i}$$

Utilizando novamente a propriedade de anti-simétrica de $M_{ua}(\rho_i)$ podemos reescrever a equação anterior na forma

$$\gamma_i < 2 \frac{\rho_i^T M_{us}(\rho)(\rho_i)}{\rho_i^T M_u(\rho_i)^T M_u(\rho_i) \rho_i}$$

Se o controlador inicial dado por ρ_0 pertencer a um domínio de atração candidato $\mathcal{B}_\alpha(\rho_u)$, então utilizando (51) sabe-se que $M_{us}(\rho) > 0$ e portanto utilizando propriedades de normas de matrizes positivas definidas simétricas tem-se que

$$\frac{\rho_i^T M_{us}(\rho_i) \rho_i}{\rho_i^T M_u^T(\rho_i) M_u(\rho_i) \rho_i} > \frac{\|\rho_i\|^2 \min(\text{eig}(M_{us}(\rho_i)))}{\rho_i^T M_u(\rho_i)^T M_u(\rho_i) \rho_i}$$

e

$$\frac{\|\rho_i\|^2 \min(\text{eig}(M_{us}(\rho_i)))}{\rho_i^T M_u^T(\rho_i) M_u(\rho_i) \rho_i} > \frac{\|\rho_i\|^2 \min(\text{eig}(M_{us}(\rho_i)))}{\|\rho_i\|^2 \|M^T(\rho_i) M(\rho_i)\|} = \frac{\min(\text{eig}(M_{us}(\rho_i)))}{\|M_u(\rho_i)^T M_u(\rho_i)\|}$$

onde $\text{eig}(M_{us}(\rho))$ denota os autovalores da matriz $M_{us}(\rho)$.

O desenvolvimento descrito acima prova o seguinte teorema:

Teorema 13 *Seja $\mathcal{B}_\alpha(\rho_u) \subset \Gamma$, se as seguintes restrições forem respeitadas*

$$M_{us}(\rho) > 0 \quad \forall \rho \in \mathcal{B}_\alpha(0); \quad \rho \neq 0$$

e

$$\gamma_i < 2 \frac{\min(\text{eig}(M_{us}(\rho)))}{\|M_u(\rho)^T M_u(\rho)\|}, \quad (55)$$

garante-se que $\mathcal{B}_\alpha(0)$ é um domínio de atração.

A expressão (55) é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição (55) é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

4.5 Combinando os três objetivos

Diferentes métodos de projeto de controladores baseado em dados minimizam diferentes combinações dessas três parcelas, representando diferentes casos particulares de $J(\rho)$.

Em um caso geral, a função custo é uma combinação das três parcelas que representam diferentes objetivos de controle. Neste caso o gradiente da função custo também será uma combinação dos gradientes de cada uma das parcelas:

Vamos agora projetar γ_i para garantir a convergência do algoritmo quando é utilizada uma combinação de objetivos de controle. Substituindo (52) em (22) tem-se:

$$\gamma_i < 2 \frac{(\rho_i - \rho_*)^T M(\rho)(\rho_i - \rho_*)}{(\rho_i - \rho_*)^T M(\rho)^T M(\rho)(\rho_i - \rho_*)}$$

Utilizando novamente a propriedade de anti-simétrica de $M_a(\rho)$ podemos reescrever a equação anterior na forma

$$\gamma_i < 2 \frac{(\rho_i - \rho_*)^T M_s(\rho)(\rho_i - \rho_*)}{(\rho_i - \rho_*)^T M(\rho)^T M(\rho)(\rho_i - \rho_*)}$$

Utilizando propriedades de normas matriciais, e a restrição do domínio de atração (23) tem-se que

$$\frac{(\rho_i - \rho_*)^T M_s(\rho)(\rho_i - \rho_*)}{(\rho_i - \rho_*)^T M(\rho)^T M(\rho)(\rho_i - \rho_*)} > \frac{\|(\rho_i - \rho_*)\|^2 \min(\text{eig}(M_s(\rho)))}{(\rho_i - \rho_*)^T M(\rho)^T M(\rho)(\rho_i - \rho_*)}$$

e

$$\frac{\|(\rho_i - \rho_*)\|^2 \min(\text{eig}(M_s(\rho)))}{(\rho_i - \rho_*)^T M(\rho)^T M(\rho)(\rho_i - \rho_*)} > \frac{\|(\rho_i - \rho_*)\|^2 \min(\text{eig}(M_s(\rho)))}{\|(\rho_i - \rho_*)\|^2 \|M(\rho)^T M(\rho)\|} = \frac{\min(\text{eig}(M_s(\rho)))}{\|M(\rho)^T M(\rho)\|}$$

onde $\text{eig}(M_s(\rho))$ denota os autovalores da matriz $M_s(\rho)$.

O desenvolvimento descrito acima prova o seguinte teorema:

Teorema 14 *Seja $\mathcal{B}_\alpha(\rho_*) \subset \Gamma$, se as seguintes restrições forem respeitadas*

$$M_s(\rho) > 0 \quad \forall \rho \in \mathcal{B}_\alpha(\rho_*); \quad \rho \neq \rho_*$$

e

$$\gamma_i < 2 \frac{\min(\text{eig}(M_s(\rho_i)))}{\|M(\rho_i)^T M(\rho_i)\|}, \quad (56)$$

então garante-se $\mathcal{B}_\alpha(\rho_*)$ é um domínio de atração de ρ_* .

A expressão (56) é uma inequação, e para garantir a convergência para o mínimo global deve-se respeitar esta condição. Para que a convergência dê-se o mais rápido possível deve-se escolher o maior tamanho do passo que respeita esta condição. Em casos práticos pode-se, na maioria dos casos, extrapolar este limite e transformar esta condição em uma igualdade com o fim de calcular o tamanho do passo. Como a condição (56) é na maioria dos casos bastante conservadora, utilizar a expressão em forma de igualdade não prejudica a convergência do algoritmo.

Na seqüência será apresentado um exemplo que demonstra o funcionamento do ajuste dos parâmetros do controlador, utilizando o algoritmo do gradiente e o tamanho do passo de iteração proposto para o caso seguimento de referência e minimização do esforço de controle. Paralelamente será utilizado o método “clássico” de escolha do tamanho do passo, para fins de comparação.

Exemplo 4 *Considere o seguinte sistema sem ruído*

$$y(t) = \frac{0,15}{z - 0,5} u(t). \quad (57)$$

Este sistema é controlado por um controlador do tipo PI

$$C(z, \rho) = [\rho_1 \ \rho_2] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T. \quad (58)$$

Com este controlador a Hipótese 2 é respeitada, o que é uma condição necessária para utilizar o Teorema 14.

Como requisito de projeto considera-se a minimização do esforço de controle, com $\lambda = 0,98$, e o seguimento de referência com

$$T_d(t) = \frac{0,1}{z - 0,9}. \quad (59)$$

O melhor controlador que pode ser alcançado com esta estrutura é dado por

$$C(z, \rho) = [0,50041490 \quad -0,27615490] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T. \quad (60)$$

que proporciona um custo de $J(\rho) = 0,22618$. Este controlador foi obtido utilizando métodos numéricos, por uma procura extensiva.

Considere que inicialmente o seguinte controlador está sendo utilizado no sistema em malha fechada

$$C(z, \rho) = [1 \ 1] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T. \quad (61)$$

Na Figura 19 pode-se visualizar as respostas do sistema em malha fechada à uma entrada do tipo salto unitário. Pode-se notar na figura que o controlador inicial possui uma resposta bastante diferente da melhor resposta que pode ser obtida. Na Figura 20 pode-se ver o sinal de controle das respostas do sistema em malha fechada à uma entrada do tipo salto unitário. Pode-se notar que o esforço de controle é bastante reduzido com o método. Na Tabela 3 pode-se ver o custo do seguimento de referência J_y , e o custo do esforço de controle J_u .

Utiliza-se então o método de ajuste de controladores IFT para melhorar o desempenho do sistema em malha fechada, fazendo atualizações dos parâmetros do controlador a cada iteração. É utilizado o seguinte sinal de referência nos ensaios do IFT

$$r(t) = 1 + \text{sen} \left(\frac{2\pi t}{10} \right).$$

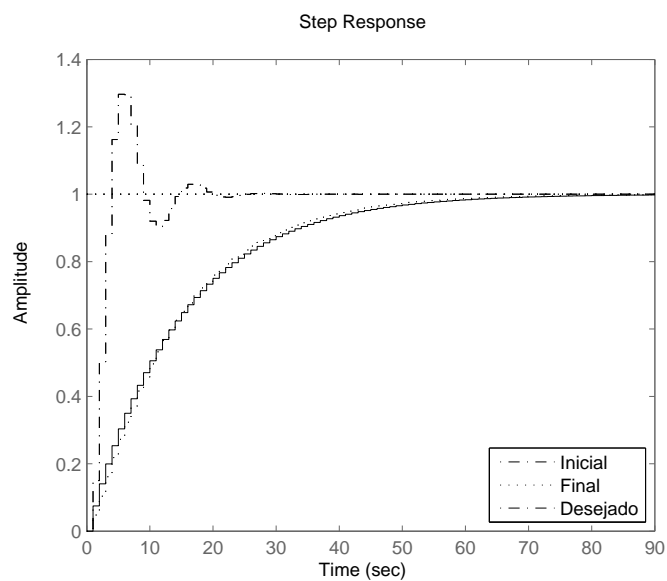


Figura 19: Exemplo 4: Resposta ao salto em malha-fechada - método proposto.

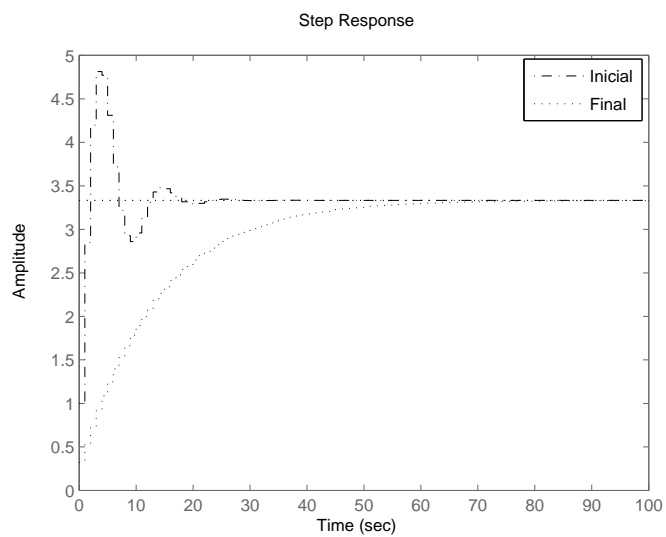


Figura 20: Exemplo 4: Sinal de controle da resposta ao salto em malha-fechada - método proposto.

São utilizadas 50 iterações com o algoritmo do gradiente e estimativa do gradiente obtida com o IFT, para melhorar o desempenho em malha fechada do sistema. Na Figura 21 são apresentados os parâmetros do controlador a cada iteração. Nota-se que o algoritmo está convergindo para o controlador ótimo. Na Figura 22 é apresentado o custo $J(\rho)$ a cada iteração. Nota-se que apesar dos parâmetros do controlador estarem bastante diferentes do controlador ótimo, o custo alcançado após 50 iterações é quase o custo ótimo. Na Figura 19 pode-se observar a resposta ao salto na referência do controlador alcançado com o algoritmo, do controlador inicial e do controlador ótimo. A resposta alcançada é muito próxima da resposta desejada.

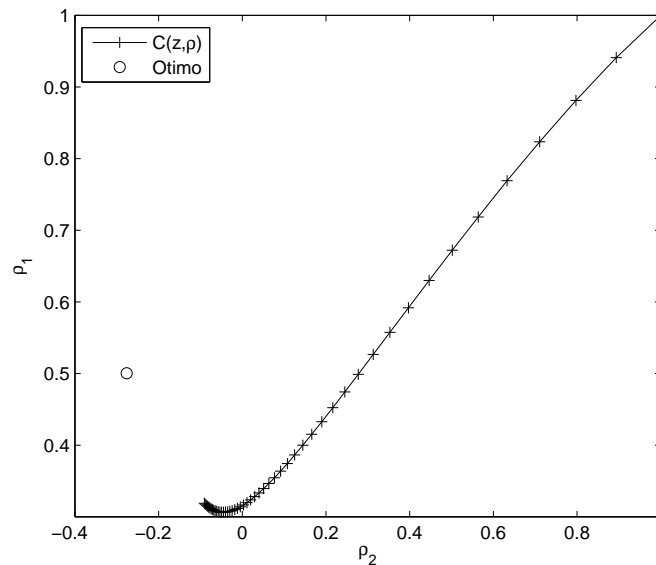


Figura 21: Exemplo 4: Parâmetros do controlador a cada iteração - método proposto - Condição Inicial $\rho_1 = [1 \ 1]^T$.

Para fins de comparação é apresentado na Figura 12 a evolução dos parâmetros do controlador quando o método “clássico” é utilizado para o cálculo do tamanho do passo de iteração. Neste exemplo novamente é utilizado o mesmo na primeira iteração o tamanho de passo do método proposto.

Também são utilizadas 50 iterações com o método IFT. Como visto na Figura 23, o método clássico gera passos muito pequenos, o que torna a convergência muito lenta. O primeiro passo é grande pois foi escolhido pelo método proposto, já os outros passos determinados pelo método clássico são muito pequenos.

Na Figura 24 é mostrada a resposta ao salto do sistema em malha-fechada, para o controlador inicial, controlador alcançado após as 50 iterações e para o modelo de referência. Após as 50 iterações a resposta ainda é bastante insatisfatória, enquanto que com o método proposto a resposta é praticamente aquela desejada. Na Figura 25 é mostrado o sinal de controle da resposta ao salto do sistema em malha-fechada. O esforço do controle com o método clássico foi reduzido, mas a redução com o método proposto foi muito maior. Na Figura 26 é mostrado o custo $J(\rho)$ a cada iteração. Após as 50 iterações o custo ainda é maior do que 0,6, enquanto que com o método proposto o custo alcançado é 0,227, muito próximo do custo ótimo que é 0,226.

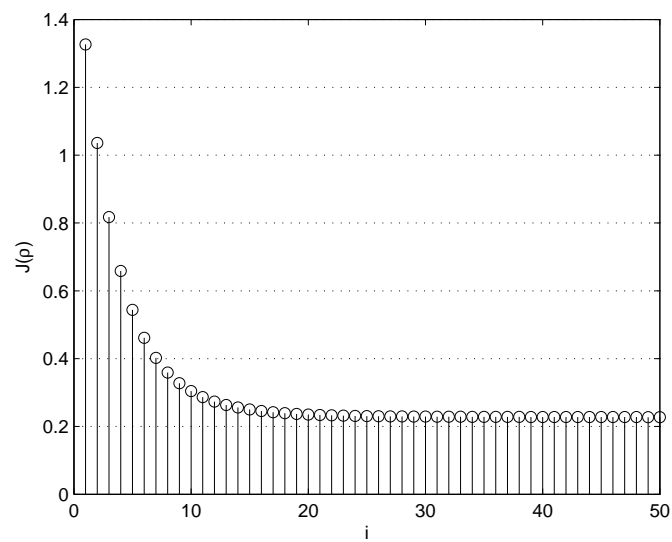


Figura 22: Exemplo 4: Custo a cada iteração - método proposto.

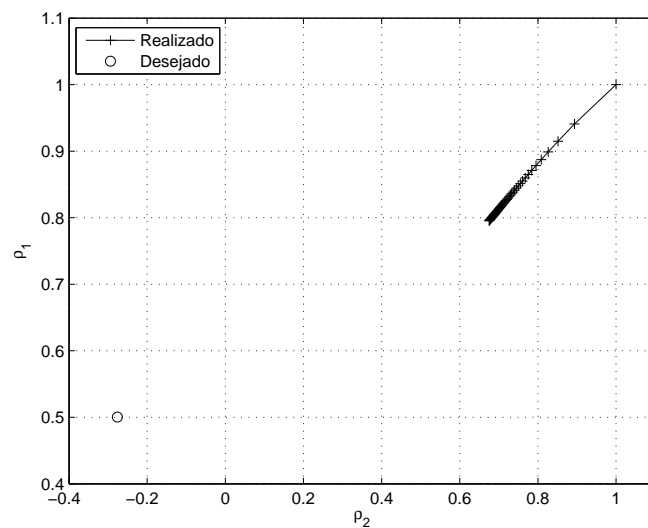


Figura 23: Exemplo 4: Parâmetros do controlador a cada iteração - método clássico - Condição Inicial $\rho_1 = [1 \ 1]^T$.

Tabela 3: Exemplo 4: Evolução dos controladores - método proposto.

i	J	λJ_y	$(1 - \lambda)J_u$	ρ_1	ρ_2	γ
1	1,32670	0,72539	0,60127	1,00000	1,00000	0.0922
2	1,03590	0,52802	0,50792	0,94101	0,89373	0.1057
3	0,81755	0,38104	0,43651	0,88136	0,79727	0.1123
4	0,65837	0,27489	0,38348	0,82360	0,71052	0.1166
5	0,54385	0,19925	0,34460	0,76912	0,63291	0.1198
6	0,46160	0,14549	0,31611	0,71858	0,56360	0.1224
7	0,40228	0,10716	0,29512	0,67222	0,50172	0.1245
8	0,35917	0,07965	0,27952	0,63005	0,44644	0.1263
9	0,32756	0,05976	0,26780	0,59188	0,39699	0.1279
10	0,30417	0,04527	0,25890	0,55750	0,35268	0.1293
11	0,28671	0,03464	0,25207	0,52662	0,31294	0.1305
12	0,27357	0,02679	0,24677	0,49897	0,27722	0.1316
13	0,26360	0,02097	0,24262	0,47427	0,24507	0.1327
14	0,25597	0,01662	0,23935	0,45226	0,21608	0.1336
15	0,25010	0,01337	0,23673	0,43268	0,18990	0.1345
16	0,24556	0,01092	0,23463	0,41530	0,16623	0.1353
17	0,24202	0,00908	0,23293	0,39990	0,14478	0.1361
18	0,23925	0,00770	0,23155	0,38630	0,12531	0.1368
19	0,23706	0,00665	0,23041	0,37430	0,10762	0.1375
20	0,23534	0,00587	0,22947	0,36376	0,09150	0.1381
21	0,23397	0,00528	0,22868	0,35451	0,07681	0.1388
22	0,23287	0,00483	0,22803	0,34643	0,06338	0.1393
23	0,23199	0,00450	0,22748	0,33940	0,05108	0.1399
24	0,23128	0,00426	0,22702	0,33330	0,03981	0.1404
25	0,23070	0,00407	0,22662	0,32806	0,02945	0.1409
26	0,23023	0,00394	0,22629	0,32356	0,01990	0.1414
27	0,22985	0,00384	0,22600	0,31975	0,01110	0.1419
28	0,22953	0,00377	0,22576	0,31654	0,00297	0.1423
29	0,22927	0,00372	0,22555	0,31387	-0,004575	0.1427
30	0,22905	0,00368	0,22537	0,31169	-0,011565	0.1432
35	0,22836	0,00360	0,22476	0,30647	-0,040114	0.1450
40	0,22800	0,00354	0,22446	0,30762	-0,061112	0.1466
45	0,22777	0,00346	0,22430	0,31233	-0,077428	0.1480
50	0,22759	0,00335	0,22423	0,31895	-0,090763	0.1493

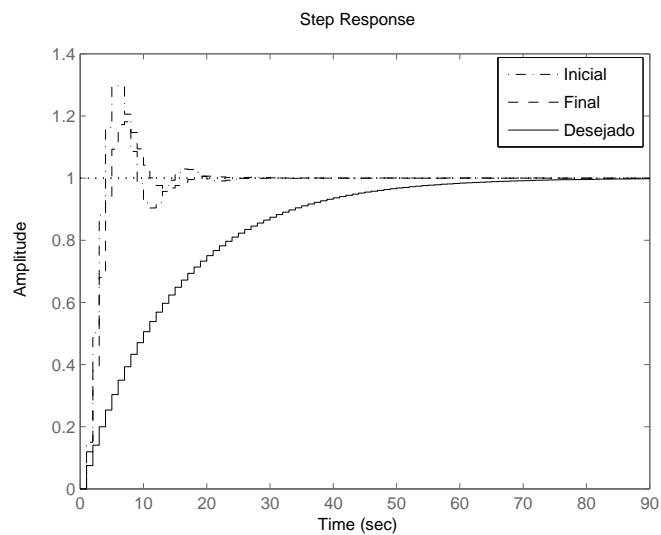


Figura 24: Exemplo 4: Resposta ao salto em malha-fechada - método clássico.

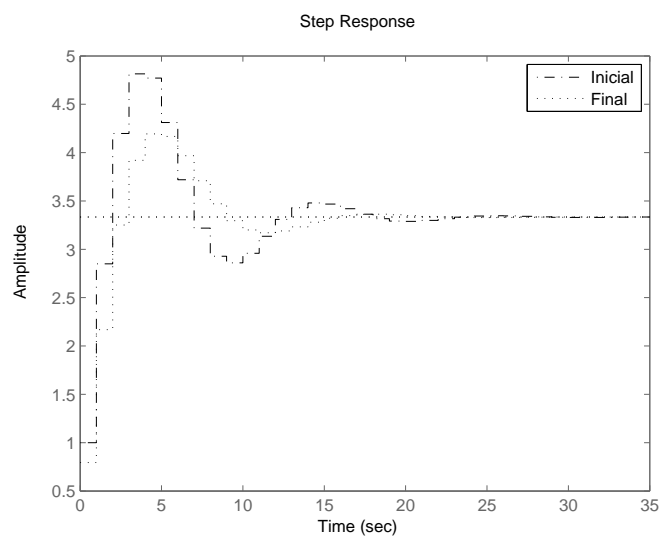


Figura 25: Exemplo 4: Sinal de controle da resposta ao salto em malha-fechada - método clássico.

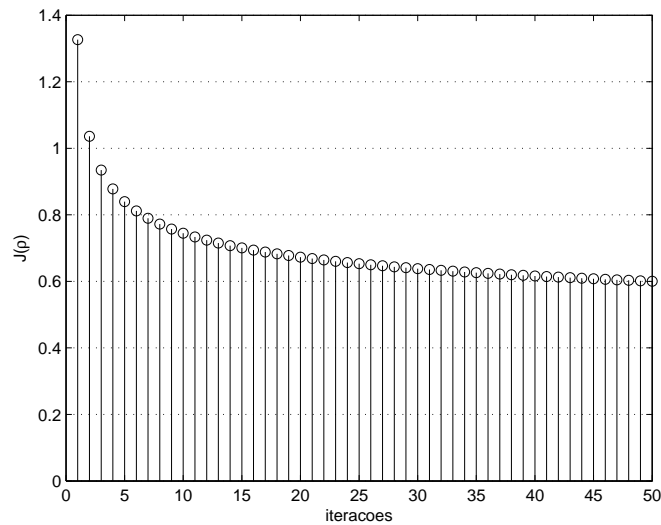


Figura 26: Exemplo 4: Custo a cada iteração - método clássico.

4.6 Considerações finais

Neste capítulo foram apresentados métodos de cálculo do tamanho do passo de iteração do algoritmo do gradiente, que garantem a convergência para o mínimo global da função custo. A prova de convergência é baseada na teoria de Lyapunov, utilizando funções candidatas quadráticas e estimativas do domínio de atração em formato de bola. Quando as condições iniciais do algoritmo estão dentro do domínio de atração, o algoritmo converge para o mínimo global da função custo, que é o único ponto de equilíbrio do algoritmo dentro da região.

Os métodos de escolha do tamanho do passo foram propostos inicialmente em separado para cada uma das parcelas da função custo: seguimento de referência, atenuação da sensibilidade ao ruído e seguimento de referência. Por fim, estas três parcelas foram combinadas, resultando em um método que garante a convergência para a função custo completa.

Para cada um dos casos propostos foram apresentados exemplos que demonstram a eficácia do método. Comparações mostraram que o método proposto converge mais rapidamente que o método clássico. Enquanto que o método clássico preocupa-se apenas em garantir a convergência, o método proposto também se preocupa com a taxa de convergência, de forma que o algoritmo converge da maneira mais rápida que foi possível provar.

São necessárias informações do processo para o cálculo do tamanho do passo pelos métodos propostos. Estas informações são obtidas por meio da matriz $M(\rho)$ e pelo limitante κ . Em princípio estas informações são desconhecidas e, portanto, precisam ser estimadas. No capítulo seguinte é apresentada uma maneira de estimar a matriz $M(\rho)$ de forma simples e eficaz. Não foi desenvolvida nesta dissertação uma maneira de estimar o limitante κ . Como este valor é necessário ao cálculo do tamanho do passo nos casos em que o modelo de referência não pode ser atingido, faz-se necessário estudar com mais profundidade maneiras de obter uma estimativa para este valor.

5 ESTIMADOR DE $M(\rho)$

5.1 Introdução

No capítulo anterior foram apresentadas formas de calcular o tamanho do passo de iteração do algoritmo do gradiente, para garantir a convergência para o mínimo global de uma dada função custo, que representa os objetivos de controle. As propostas apresentadas para o projeto do tamanho do passo de iteração são baseadas no conhecimento da matriz $M(\rho)$. Esta matriz é função do processo que é desconhecido. Desta forma, esta matriz também é desconhecida e faz-se necessário utilizar um estimador para esta função, para que os métodos propostos possam ser utilizados. Nas seções seguintes será apresentada uma maneira de estimar $M(\rho)$. Esta estimativa será obtida a partir dos mesmos dados utilizados para cálculo do gradiente da função custo, de maneira que não são necessários experimentos adicionais.

5.2 Estimativa baseada em modelo

Nesta seção será apresentada uma maneira de computar um estimador para a matriz $M(\rho)$. Será apresentado em detalhes um estimador para $M_y(\rho)$, que é o caso específico quando apenas o problema de seguimento de referência é tratado. Os demais casos e generalizações podem ser obtidas de forma análoga à apresentada na seqüência.

A matriz $M_y(\rho)$ pode ser definida por

$$M_y(\rho) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_r |G|^2 |S(\rho)|^2 \Re \{ S_d^* S(\rho) \bar{C} \bar{C}^* \} d\omega.$$

Nessa expressão é fácil notar que a função depende da referência, do controlador e do modelo de referência bem como do modelo do processo utilizado. A referência, o controlador e o modelo de referência são dados conhecidos pelo projetista, restando apenas o conhecimento do modelo do processo para o cálculo da matriz $M_y(\rho)$.

Surge então a idéia de obter uma estimativa para o modelo do processo e utilizá-la para computar o valor de $M_y(\rho)$. Este modelo pode ser obtido por um processo de identificação que utiliza os mesmos dados de entrada e saída do processo que são necessários para computar o gradiente da função custo. Desta maneira não é necessário realizar experimentos adicionais para obter os dados para computar o modelo.

Os dados utilizados para o cálculo do gradiente da função custo muitas vezes são obtidos de experimentos realizados em malha-fechada. Estes experimentos são afetados por ruído e têm seu espectro concentrado em algumas faixas de frequência. Além disso, em casos práticos estes são experimentos curtos e portanto fornecem pouca informação dinâmica sobre o processo.

Todos esses fatores apresentados contribuem para que a qualidade do modelo que será obtido com os dados desses experimentos seja bastante duvidosa. Em muitos casos esse modelo não é preciso o suficiente para que possa ser utilizado para o projeto de controladores baseado em modelo. Os controladores obtidos a partir deste modelo tornam a resposta em malha-fechada bastante diferente daquela desejada, descrita pelo modelo de referência.

Por outro lado, apesar dos vários aspectos negativos do modelo obtido com os dados utilizados para o cálculo do gradiente, esse modelo é satisfatório para gerar uma estimativa da matriz $M_y(\rho)$ que será utilizada para o cálculo do tamanho do passo de iteração. Essa estimativa pode ser utilizada para o cálculo do tamanho do passo de iteração porque são utilizadas apenas algumas informações sobre esta matriz para o cálculo do passo.

No capítulo anterior foi mostrado que são utilizados no cálculo do tamanho do passo: o menor autovalor da parte simétrica de $M_y(\rho)$ e a norma de $M_y^T(\rho)M_y(\rho)$ (que é exatamente igual ao maior autovalor). Nota-se que apenas algumas medidas da matriz $M(\rho)$ são utilizadas para computar o tamanho do passo, não a matriz completa. Pode-se demonstrar que o cálculo baseado nessas medidas é bastante conservativo na maioria dos casos, ou seja, existem tamanhos de passo maior que ainda garantem a convergência (mas não existem formas conhecidas de calculá-los). Desta forma, pequenas variações nesses valores da matriz não afetam a convergência para o mínimo global, dado que o tamanho dos passos é suficientemente pequeno para garantir a convergência.

No exemplo a seguir será demonstrado que se a classe de modelos utilizada na identificação for diferente da classe do processo, o modelo identificado não é preciso o suficiente para o projeto direto do controlador, mas pode perfeitamente ser utilizado para calcular o tamanho do passo de iteração. O projeto de controladores baseado em dados se mostrará muito mais eficiente para alcançar o modelo de referência.

Exemplo 5 *Considere o seguinte processo*

$$y(t) = \frac{1}{(z - 0,7)(z - 0,8)}u(t).$$

e o seguinte modelo de referência

$$y_d(t) = \frac{0,25}{(z - 0,5)(z - 0,5)}r(t)$$

Este modelo pode ser alcançado pelo sistema em malha fechada se for utilizado o seguinte controlador PID

$$C_d(z) = \frac{0,25(z - 0,7)(z - 0,8)}{z(z - 1)}$$

Considere que o processo é controlado inicialmente pelo seguinte controlador PID

$$C(z, \rho) = \frac{0,25z^2 - 0,35z + 0,14}{z(z - 1)}. \quad (62)$$

Este controlador gera uma resposta bastante insatisfatória para o sistema em malha-fechada. Esta resposta pode ser visualizada na Figura 27.

Com dados obtidos de um experimento realizado em malha-fechada com o controlador (62), foi identificado um modelo para o processo com a seguinte estrutura:

$$G(z) = \frac{\theta_1 z}{z^2 + \theta_2 z + \theta_3}.$$

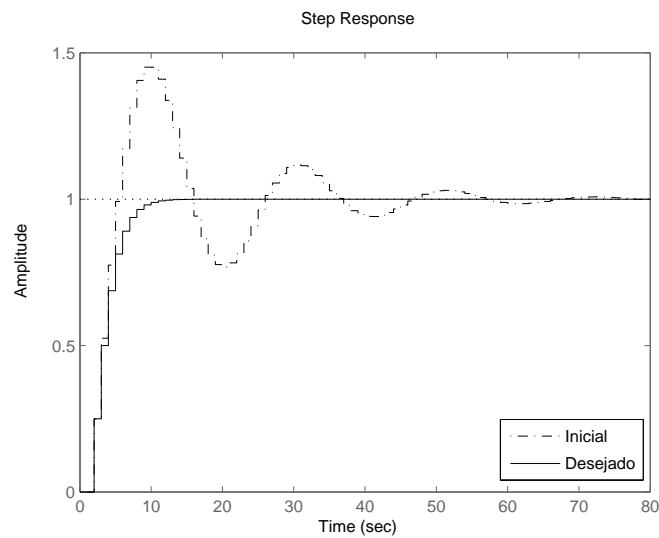


Figura 27: Exemplo 5: Resposta ao salto com controlador inicial.

Com os dados apresentados na Figura 28, foi identificado o seguinte modelo

$$\hat{G}(z) = \frac{0,49885z}{z^2 - 1,722z + 0,7553}.$$

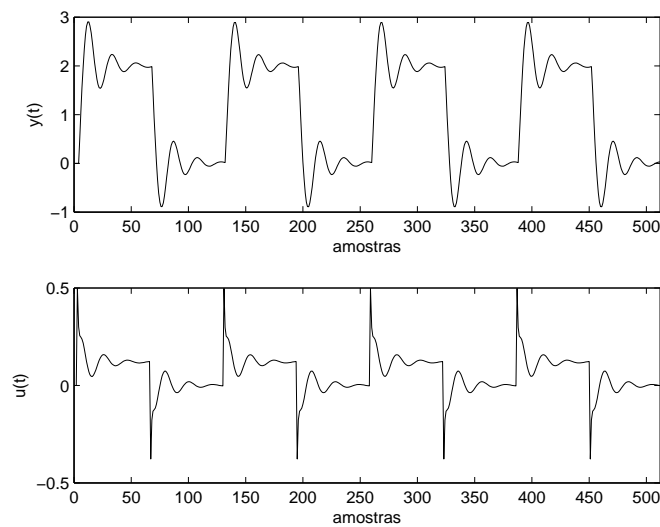


Figura 28: Exemplo 5: Simulação em malha-fechada.

Quando o modelo $\hat{G}(z)$ é utilizado para projetar um controlador que obtenha o modelo de referência $T_d(z)$, tem-se como resultado o seguinte controlador:

$$\hat{C}_d(z) = \frac{0,50115(z^2 - 1,722z + 0,7553)}{z^2(z - 1)}.$$

Espera-se que esse controlador proporcione uma resposta no sistema real em malha-fechada bastante parecida com o modelo de referência. Na verdade não é o que ocorre.

Apesar da estrutura do modelo identificado não ser muito diferente do processo real, o controlador obtido com o modelo identificado gera uma resposta em malha-fechada bastante insatisfatória. Pode-se ver esta resposta na Figura 29.

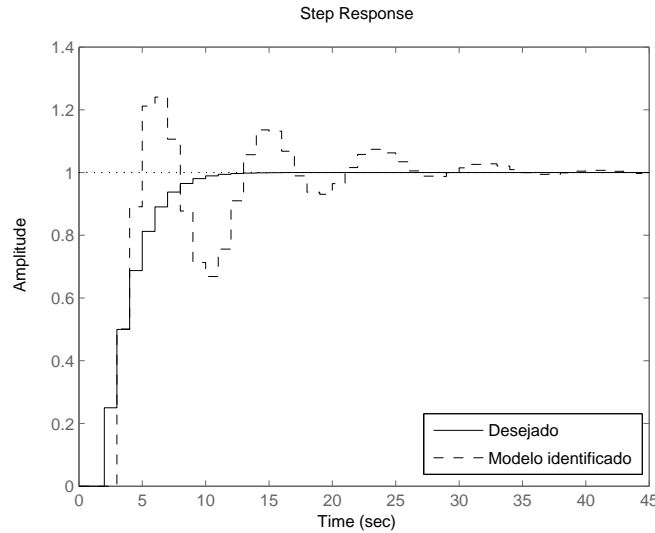


Figura 29: Exemplo 5: Resposta ao salto com controlador $\hat{C}_d(z)$.

Utilizando o algoritmo do gradiente para ajustar os parâmetros de um controlador PID pode-se obter uma resposta em malha-fechada muito melhor que aquela gerada com $\hat{C}_d(z)$. Utilizou-se então o método IFT para estimar o gradiente da função custo, e o método proposto no capítulo anterior para projetar o tamanho do passo de iteração. Para tanto foi utilizado o modelo obtido pelo processo de identificação para estimar $M_y(\rho)$ e calcular o tamanho do passo.

Foi utilizada uma onda quadrada como sinal de referência nos experimentos:

$$r(t) = 1 + \text{square} \left(\frac{2\pi t}{16} \right).$$

onde $\text{square} \left(\frac{2\pi t}{N} \right)$ é um sinal tipo onda quadrada de período N .

Após 10 iterações o comportamento do sistema em malha-fechada é bastante próximo ao comportamento desejado, pode-se ver esse fato na Figura 30. A Tabela 4 apresenta a evolução dos controladores e a Figura 31 mostra como o custo é reduzido a cada iteração.

5.3 Aumentando a robustez

Se for possível estimar a polarização e a covariância da estimativa do modelo, estes valores podem também ser utilizados para gerar uma estimativa mais robusta para o tamanho do passo de iteração.

A partir dos dados da polarização e covariância pode-se estabelecer o intervalo de confiança com probabilidade p no qual confia-se que o modelo correto está incluído. Nesse caso existe uma região à qual os parâmetros do modelo pertencem, e não mais apenas um vetor de parâmetros isolados. Uma forma conservativa, mas mais robusta de escolher o tamanho do passo de iteração, é selecionar dos parâmetros dentro do intervalo de confiança, aquele que gera o menor tamanho do passo de iteração. Desta forma, será obtido um valor para γ , que garantirá com no mínimo probabilidade p a convergência do algoritmo.

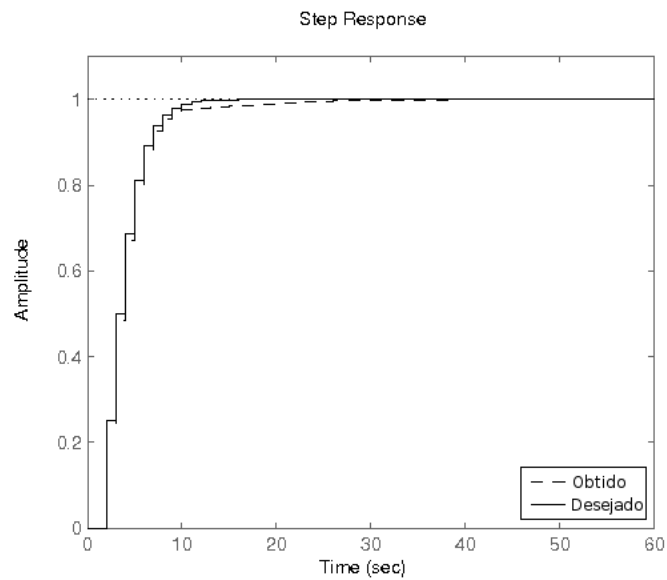


Figura 30: Exemplo 5: Resposta ao salto com controlador final.

Tabela 4: Exemplo 5: Evolução dos controladores.

i	$J_y(\rho)$	ρ_1	ρ_2	ρ_3
1	0,62298	0,25000	-0,35000	0,14000
2	0,21881	0,24736	-0,35268	0,13770
3	0,07228	0,24519	-0,35493	0,13569
4	0,02360	0,24367	-0,35656	0,13422
5	0,007953	0,24270	-0,35760	0,13326
6	0,002797	0,24212	-0,35824	0,13266
7	0,001022	0,24177	-0,35863	0,13229
8	0,000392	0,24156	-0,35887	0,13206
9	0,000163	0,24144	-0,35901	0,13192
10	$8,2812e - 05$	0,24137	-0,35910	0,13183

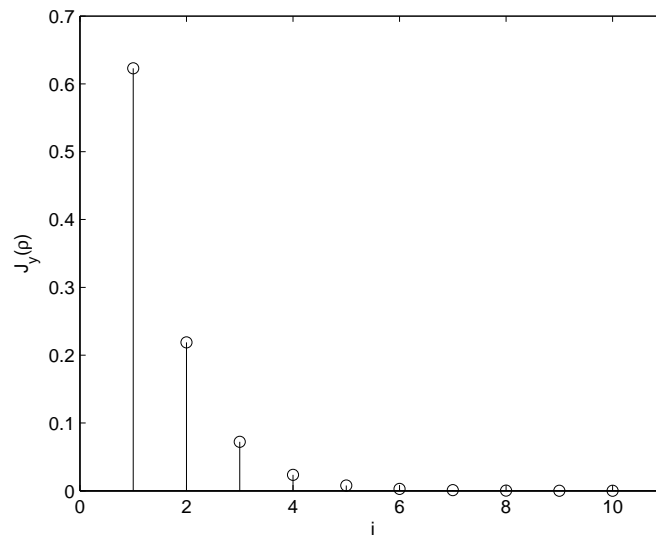


Figura 31: Exemplo 5: Custo a cada iteração.

Se o processo de identificação é realizado utilizando a minimização do erro de predição, e a estrutura escolhida para a identificação do modelo do processo for tal que a polarização do modelo é nula, então a região do intervalo de confiança do modelo é definida apenas pela variância do modelo. Nesse caso, a variância pode ser estimada utilizando as seguintes equações:

$$Cov(\theta) \approx \frac{1}{N} \hat{P} \quad (63)$$

$$\hat{P} = \hat{\lambda} \left[\frac{1}{N} \sum_{t=1}^N \psi(t) \psi^T(t) \right]^{-1} \quad (64)$$

$$\hat{\lambda} = \frac{1}{N} \sum_{t=1}^N \epsilon^2(t) \quad (65)$$

onde θ são os parâmetros do modelo, $\epsilon(t)$ é o resíduo e $\psi(t)$ é o vetor de regressores (LJUNG, 1987) (AGUIRRE, 2004). Com esta estimativa de $Cov(\theta)$, é possível traçar elipsóides que definem o intervalo de confiança do modelo.

Exemplo 6 Considere o seguinte sistema

$$y(t) = \frac{0,15}{z - 0,5} u(t) + \nu(t)$$

onde $\nu(t)$ é ruído branco com variância $\sigma_e^2 = 0,1$.

Este sistema é controlado por um controlador do tipo PI

$$C(z, \rho) = [2 \ 2] \left[\frac{z}{z-1} \quad \frac{1}{z-1} \right]^T.$$

Considere o seguinte modelo de referência

$$y_d(t) = \frac{0,1}{z - 0,9} r(t)$$

Tabela 5: Exemplo 6: Estimativa de γ .

Modelo	$G(z)$	γ
Modelo Nominal	$\frac{0,15}{z-0,5}$	2,7907
Modelo Identificado	$\frac{0,1484}{z-0,5191}$	3,0303
Modelo menor γ	$\frac{0,1484}{z-0,5018}$	2,7548

e que o seguinte sinal de referência é utilizado no sistema em malha-fechada.

$$r(t) = 1 + \text{sen} \left(\frac{2\pi t}{32} \right).$$

Para calcular o tamanho do passo de iteração é necessário uma estimativa da função $M(\rho)$ que por sua vez depende do modelo da planta $G(z)$. Um processo de identificação é feito com o sistema descrito acima com o qual deseja-se identificar os parâmetros da seguinte classe de modelos:

$$G(z, \theta) = \frac{\theta_1}{z - \theta_2}.$$

O processo de identificação resulta no seguinte modelo

$$G(z, \theta) = \frac{0,1484}{z - 0,5191},$$

que é bastante próximo do modelo nominal. Na Figura 32 vê-se o modelo nominal, o modelo identificado e uma elipse que é o intervalo de confiança de 95% centrado no modelo identificado. Este intervalo de confiança é traçado com a informação da matriz de covariância:

$$\text{Cov}(\theta) = \begin{bmatrix} 0,00010995 & -2,2666e - 05 \\ -2,2666e - 05 & 8,6918e - 06 \end{bmatrix}$$

Na Figura 33 vê-se o mesmo gráfico anterior, mas agora com mais uma eixo que representa o tamanho do passo de iteração γ calculado com cada um dos modelos representados. Nota-se que γ calculado com o modelo identificado é maior que o valor do modelo nominal, e portanto uma precaução deve ser tomada na utilização deste tamanho de passo.

Calcula-se então o valor de γ para inúmeros pontos na fronteira da região representada na figura. Escolhe-se o menor valor de γ encontrado para utilizar no algoritmo. Esta escolha é conservativa, porém mais robusta.

Na Tabela 5 temos os valores γ calculados com o modelo nominal, o modelo identificado e o modelo dentro do intervalo de confiança que gera o menor tamanho de γ . Nota-se que o tamanho do passo calculado com o modelo identificado é maior que o tamanho do passo calculado com o modelo nominal. Por outro lado, o menor valor de γ encontrado no intervalo de confiança de 95% é menor que o valor calculado com o modelo nominal, portanto é um valor adequado para garantir a convergência.

5.4 Considerações finais

Neste capítulo foi apresentada uma forma de calcular informações necessárias para o cálculo do tamanho do passo de iteração, utilizando a metodologia proposta neste trabalho. Foi mostrado que uma estimativa da matriz $M(\rho)$ pode ser obtida a partir de um

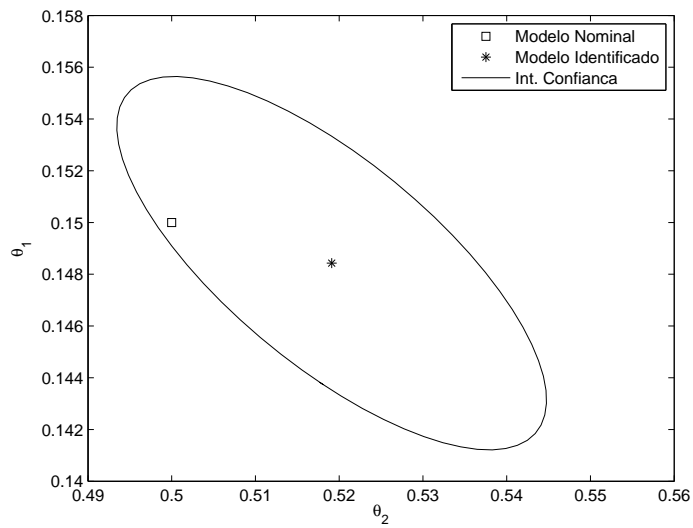


Figura 32: Exemplo 6: Intervalo de confiança do modelo identificado.

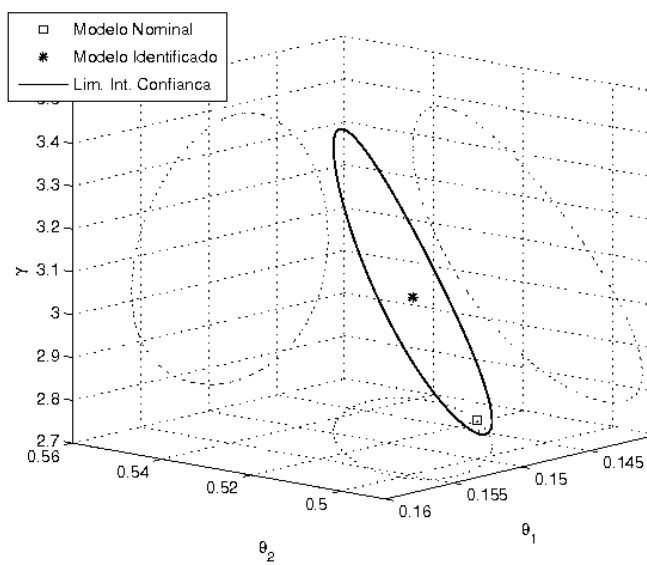


Figura 33: Exemplo 6: Tamanho do passo para limites do intervalo de confiança.

modelo identificado do processo. Esse modelo não precisa ser bastante preciso. Por meio de exemplos mostrou-se que mesmo um modelo com estrutura diferente da estrutura do processo pode gerar boas estimativas para o tamanho do passo de iteração. Por outro lado, mostrou-se que esse mesmo modelo não é preciso o suficiente para projetar um controlador por métodos baseados em modelo.

Na seqüência foi apresentada uma idéia de como aumentar a robustez do métodos escolhendo tamanhos de passos mais conservativos. Se forem obtidas estimativas de polarização e variância do modelo obtido, estes dados podem ser utilizados para a escolha do tamanho do passo. Com estas informações pode-se estabelecer intervalos de confiança, e a proposta apresentada é de escolher o menor tamanho do passo de iteração dentro deste intervalo de confiança.

6 EXPERIMENTOS

Este capítulo tem o propósito de apresentar resultados experimentais dos temas apresentados neste trabalho.

Os experimentos foram realizados em um processo de controle de temperatura. Este processo consiste em uma resistência, montada dentro de um cilindro metálico, cuja temperatura pode ser controlada. Para a medição da temperatura da resistência é utilizado um termopar tipo K. A alimentação da resistência é feita pelo chaveamento da tensão da rede elétrica através de um relé de potência. Um controlador PID roda em um PC como um algoritmo do Elipse E3. Uma foto do processo é apresentada na Figura 34.



Figura 34: Foto do processo.

Nas seções seguintes será apresentado o processo com mais detalhes bem como resultados utilizando o conceito Shaping da Função Custo, demonstrações da utilização do tamanho do passo de iteração proposto para o algoritmo do gradiente e utilização da estimativa da matriz $M(\rho)$.

6.1 O processo

Para que análises quantitativas possam ser feitas sobre os métodos de projetos de controladores é útil ter em mãos um modelo preciso do processo. Nesta seção será apresentado um modelo do processo obtido por um técnicas de identificação.

Naturalmente um processo de controle de temperatura é contínuo no tempo. Para utilizar os métodos apresentados neste trabalho é necessário utilizar um modelo discreto do processo. No processo de discretização, um parâmetro chave é a escolha do período de amostragem. Esse valor não pode ser muito grande, pois nesse caso não é possível obter características da dinâmica do processo, nem muito curto, pois causa problemas do ponto de vista numérico (AGUIRRE, 2004). No caso de haver atraso de transporte no processo, é bastante interessante que este atraso seja múltiplo do período de amostragem para que esse atraso possa ser representado fielmente pelo modelo discretizado.

Analisando a resposta ao degrau do processo, pode-se notar a existência de um atraso de transporte de aproximadamente 3 segundos e o tempo de acomodação em torno de 6 minutos. Com base nestes dados foi escolhido o período de amostragem com valor de 3 segundos.

Foi escolhida a seguinte estrutura de segunda ordem com grau relativo unitário para o processo:

$$G(z) = \frac{kz}{(z-a)(z-b)}$$

com $a, b \in (-1, 1)$. Esta classe de processo representa com bastante fidelidade o processo térmico pelas duas dinâmicas: a primeira representa a transferência de calor do atuador para o meio, enquanto que a segunda dinâmica representa a transferência de calor do meio para o sensor. Além disso, como o período de amostragem foi escolhido igual ao atraso de transporte a estrutura do modelo foi escolhida para ajustar-se a esta escolha.

Gostar-se-ia de controlar este processo com erro zero em regime permanente para referências e distúrbios do tipo salto, que são especificações industriais típicas.

Se o modelo de referência for tal que respeita a seguinte estrutura

$$T_d(z) = \frac{(1-c)}{(z-c)},$$

então a Hipótese 2 é respeitada utilizando um controlador PID e o controlador ideal é dado por

$$C_d(z) = \frac{(1-c)(z-a)(z-b)}{kz(z-1)}.$$

A partir de uma técnica de identificação foi obtido o seguinte modelo para o processo para a faixa de operação entre $70^\circ C$ e $90^\circ C$:

$$G(z) = \frac{0,067209z}{(z-0,9703)(z-0,5492)}.$$

O sinal de entrada utilizado na identificação foi uma soma de senóides com período 180 e 60 segundos. O experimento foi repetido duas vezes para poder ser observada a influência do ruído na saída do processo. O sinal de saída foi filtrado para retirar as componentes do espectro com frequências diferentes das componentes do espectro do sinal de entrada. Os parâmetros do modelo foram identificados utilizando mínimos quadrados, com erro menor que 3×10^{-3} , para um intervalo de confiança de 95%. O sinal de entrada e os sinais de saída podem ser vistos na Figura 35. Nota-se que o ruído tem um efeito bastante pequeno nas observações.

Para estimar a qualidade dos controladores que serão projetados será utilizada a seguinte estimativa para a função custo

$$\hat{J}(\rho) = \frac{1}{N} \sum_{t=1}^N (y(t, \rho) - y_d(t))^2$$

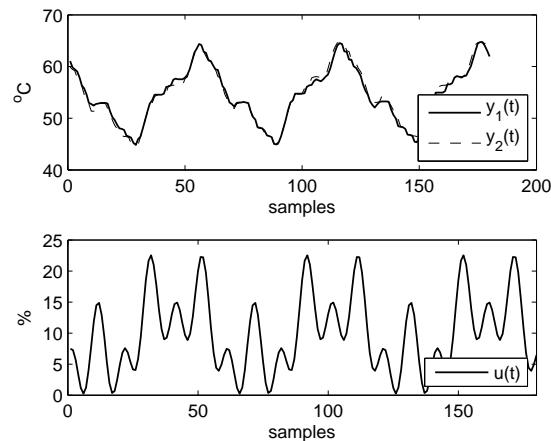


Figura 35: Experimento de identificação em malha-aberta.

Aqui N é o número de amostras utilizadas para o cálculo da estimativa. Para que a medida tenha um significado físico, escolheu-se utilizar o seguinte critério de qualidade, o qual possui a mesma unidade de medida da saída do processo.

$$\check{J}(\rho) = \sqrt{\hat{J}(\rho)}$$

O modelo de referência escolhido para o sistema em malha-fechada nos experimentos é o seguinte:

$$T_d(z) = \frac{0,1}{z - 0,9}$$

O modelo identificado foi utilizado para projetar um controlador PID que alcance esse modelo de referência escolhido, utilizando (6):

$$C_d(z) = [0,675 \ 0,02 \ 0,793] \left[1 \quad \frac{z}{z-1} \quad \frac{z-1}{z} \right]^T$$

A resposta ao salto de 20°C na referência do sistema em malha-fechada com o controlador $C_d(z)$ resulta em um custo $\check{J}(\rho_d) = 0,22^\circ\text{C}$. Este custo é a composição da parcela de seguimento de referência e da parcela de atenuação de ruído. O valor do custo é baixo, o que indica que o controlador é bom.

6.2 Shaping da Função Custo

Nesta seção será demonstrado que a escolha da referência utilizada nos experimentos tem fundamental importância para convergência dos métodos de projeto de controladores baseados em dados.

O controlador inicial rodando no sistema é dado por:

$$C_1(z, \rho) = [0,55 \ 0,20 \ 0,00] \left[1 \quad \frac{z}{z-1} \quad \frac{z-1}{z} \right]^T$$

A Figura 36 mostra a resposta ao salto com o controlador $C_1(z, \rho)$. O custo estimado para a resposta com este controlador é $\check{J}(\rho) = 6,34^\circ\text{C}$.

Para melhorar o comportamento do sistema em malha-fechada é utilizado o método IFT. Inicialmente será utilizado o seguinte sinal de referência nos experimentos realizados

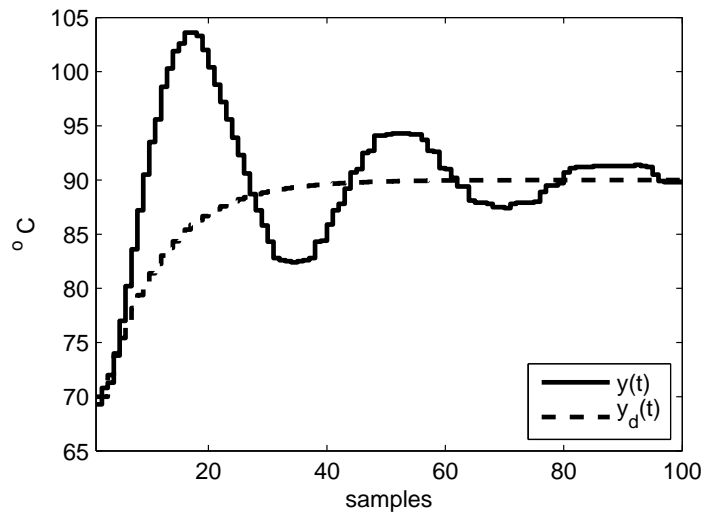


Figura 36: Resposta ao salto com controlador $C_1(z, \rho)$.

Tabela 6: Experimento 1: Evolução dos controladores - $r_1(t)$.

i	ρ	$\check{J}(\rho)$
1	[0, 5500 0, 20000 0, 00000]	6,34
2	[0, 6973 0, 03970 - 0, 04672]	1,76
3	[0, 6641 0, 01956 - 0, 05345]	0,67
4	[0, 7069 0, 02960 - 0, 04977]	1,23
5	[0, 7118 0, 02457 - 0, 05141]	0,88
6	[0, 7168 0, 01954 - 0, 05304]	0,49

para computar a estimativa do gradiente da função custo $\widehat{\nabla J}(\rho)$

$$r_1(t) = 80 + 10 \text{square} \left(\frac{2\pi t}{200} \right).$$

onde $\text{square} \left(\frac{2\pi t}{T} \right)$ é um sinal tipo onda quadrada com período T .

A evolução dos parâmetros do controlador é mostrada na Tabela 6 com o custo calculado a cada iteração. Após algumas iterações o custo foi reduzido para $\check{J}(\rho) = 0,495$ °C com o controlador

$$C_2(z) = [0,7168 \ 0,01954 \ -0,053] \left[1 \ \frac{z}{z-1} \ \frac{z-1}{z} \right]^T.$$

Os parâmetros do controlador estão convergindo para o mínimo global. É importante ressaltar que o mínimo global não é dado por ρ_d definido anteriormente, porque o critério minimizado pelo IFT envolve $J_y(\rho)$ e $J_e(\rho)$. Porém como o nível de ruído é baixo espera-se que o mínimo global seja bastante próximo do valor ρ_d . Pode-se observar que os parâmetros k_p e k_i estão bastante próximos dos valores correspondentes de ρ_d . O ganho derivativo k_d , por outro lado, é muito menor; isto é esperado visto que a ação derivativa tende a amplificar o ruído, e este efeito é penalizado pelo critério de desempenho $J(\rho)$. A resposta ao salto com $C_2(z, \rho)$ é mostrada na Figura 37, onde pode ser visto que o desempenho é bastante próximo do desejado.

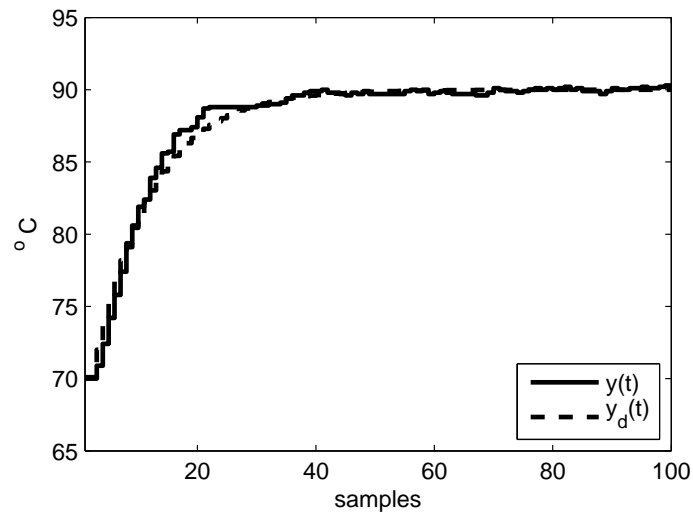


Figura 37: Resposta ao salto com controlador $C_2(z, \rho)$.

Tabela 7: Experimento 1: Evolução dos controladores - $r_2(t)$.

i	ρ	$\check{J}(\rho)$
1	[0, 5500 0, 2000 0, 0000]	3, 60
2	[0, 5327 0, 2343 0, 0029]	3, 51
3	[0, 4996 0, 2738 0, 0128]	3, 43
4	[0, 4730 0, 3116 0, 0197]	3, 39

Na seqüência será mostrado o que pode ocorrer se o sinal de referência utilizado nos experimentos for inadequado. Escolheu-se utilizar o seguinte sinal de referência nos experimentos utilizados para calcular a estimativa do gradiente da função custo

$$r_2(t) = 80 + 10 \text{sen} \left(\frac{2\pi t}{58} \right).$$

Quando esta referência é utilizada o controlador inicial $C_1(z, \rho)$ proporciona uma resposta ao sistema em malha-fechada com um valor de custo dado por $\check{J}(\rho) = 3, 60$ °C. Apesar do controlador ser o mesmo, o valor do custo é diferente pois como a referência é outra, a função custo também se modifica.

Pode-se ver na Tabela 7 a evolução dos controladores e o custo estimado. Na iteração 4 o controlador $C_3(z, \rho)$ é alcançado e o custo estimado reduziu de $\check{J}(\rho) = 3, 60$ °C para $\check{J}(\rho) = 3, 39$ °C.

$$C_3(z) = [0, 4730 \ 0, 3116 \ 0, 0197] \begin{bmatrix} 1 & \frac{z}{z-1} & \frac{z-1}{z} \end{bmatrix}^T$$

Apesar do custo ser reduzido, os parâmetros do controlador estão a cada iteração mais longe do mínimo global. Este fato pode facilmente ser visto da Figura 38, o qual mostra a resposta ao salto do controlador $C_4(z, \rho)$. Isto pode ser explicado pelo fato que o algoritmo está convergindo para um mínimo local da função custo.

Contudo, como mostrado nos capítulos anteriores, o critério de desempenho depende fortemente do espectro da referência, portanto reduzir o valor do custo para o sinal de referência $r_2(t)$ não implica necessariamente em reduzir o valor para a referência tipo salto.

Isto explica porque o controlador $C_3(z)$ apresenta um valor de custo $\check{J}(\rho) = 10,00 \text{ }^\circ\text{C}$, maior que os $\check{J}(\rho) = 6,34 \text{ }^\circ\text{C}$ obtidos para o controlador inicial $C_1(z, \rho)$ para a mesma referência tipo salto.

Sob a Hipótese 2, o mínimo global das funções custo é o mesmo para qualquer sinal de referência. A diferença fundamental existente entre as diversas escolhas do sinal de referência está que a região de atração do mínimo global é diferente para cada referência. O controlador inicial $C_1(z, \rho)$ pertence ao domínio de atração da função custo para a referência $r_1(t)$, mas não pertence ao domínio de atração da função custo para a referência $r_2(t)$.

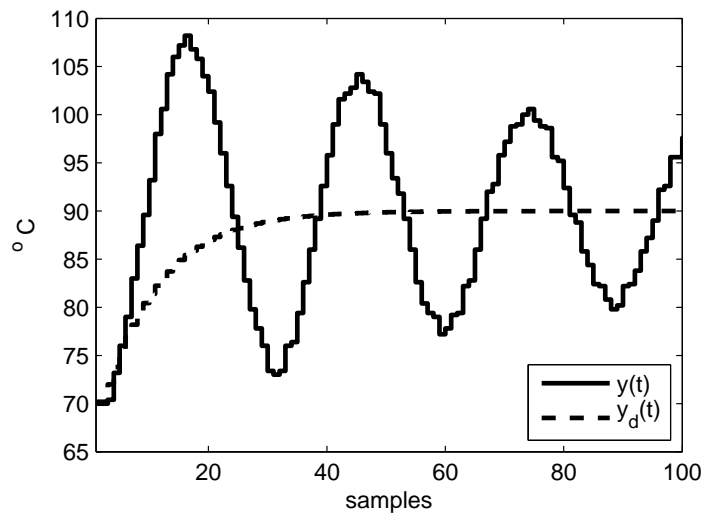


Figura 38: Resposta ao salto com controlador $C_3(z, \rho)$.

6.3 Utilização do passo de iteração proposto

Nesta seção serão apresentados experimentos que mostram a aplicabilidade do algoritmo do gradiente para o ajuste dos parâmetros do controlador, utilizando a metodologia apresentada nos capítulos anteriores.

Inicialmente o sistema em malha-fechada está rodando com o seguinte controlador:

$$C_4(z, \rho) = [1, 6 \quad -2, 4 \quad 0, 9] \left[\begin{array}{ccc} \frac{z^2}{z^2-z} & \frac{z}{z^2-z} & \frac{1}{z^2-z} \end{array} \right]^T.$$

Este controlador gera uma resposta em malha-fechada bastante insatisfatória cujo custo é $\check{J}(\rho) = 3,34 \text{ }^\circ\text{C}$. Esta resposta pode ser vista na Figura 39.

Com os dados de entrada e saída deste experimento foi identificado um modelo de ordem 1 para o processo com a seguinte estrutura:

$$G(z) = \frac{\theta_1}{z - \theta_2}$$

Esta estrutura parece ser bastante adequada para o modelo do processo dado que existe uma dinâmica dominante, dada pela transferência de calor do atuador para o meio.

O processo de identificação resultou no seguinte modelo:

$$\hat{G}(z) = \frac{0,1166}{z - 0,9855}.$$

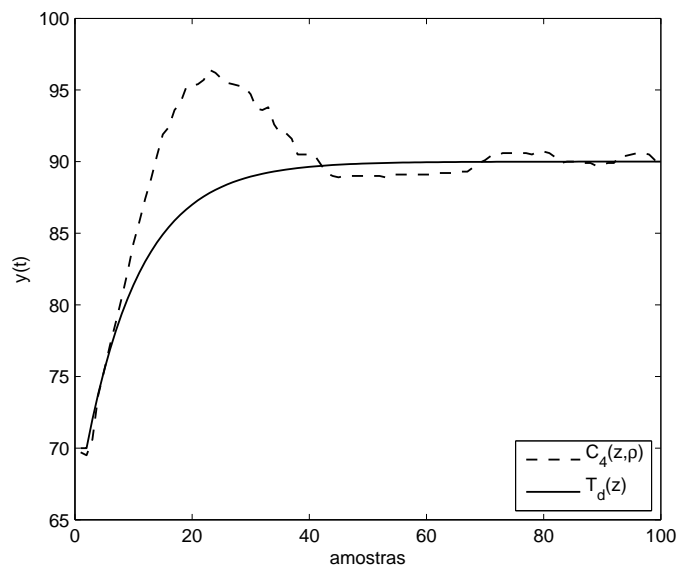


Figura 39: Resposta ao salto com controlador $C_4(z, \rho)$.

Esse modelo identificado foi utilizado para o projeto de um controlador baseado em modelo, que minimiza o critério de seguimento de referência. O controlador projetado é dado por:

$$C_5(z, \rho) = \frac{0,857632 - 0,84519z}{z - 1}.$$

Este controlador gera uma resposta no sistema em malha-fechada com custo $\check{J}(\rho) = 1,41$ °C. Este custo é menor que a metade do custo alcançado pelo sistema com o controlador $C_4(z, \rho)$. A resposta ao salto pode ser vista na Figura 40. Apesar do custo ter diminuindo vê-se claramente que a resposta alcançada está mais lenta que a resposta desejada.

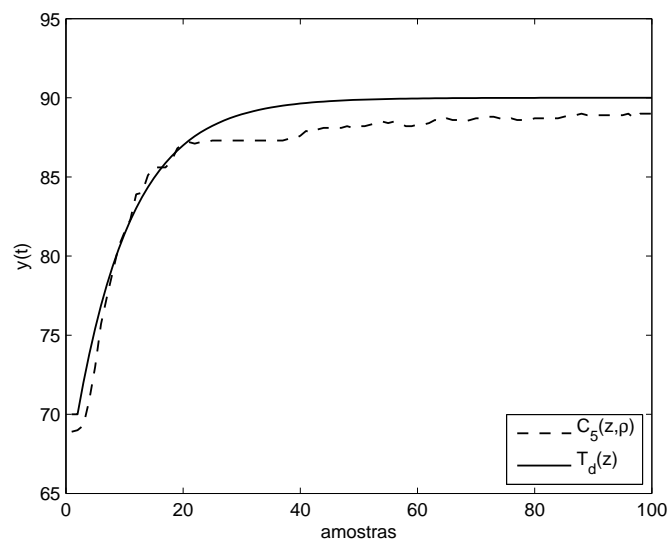


Figura 40: Resposta ao salto com controlador $C_5(z, \rho)$.

Tabela 8: Experimento 2: Evolução dos controladores - $r_3(t)$.

i	ρ	$\check{J}(\rho)$
1	[1, 6000 2, 4000 0, 9000]	7, 84
2	[1, 5883 2, 4132 0, 8852]	4, 99
3	[1, 5748 2, 4278 0, 8695]	0, 86
4	[1, 5749 2, 4277 0, 8695]	0, 79

Utilizando o algoritmo do gradiente, com o gradiente da função custo estimado pelo método do IFT, e o tamanho do passo de iteração escolhido conforme apresentado nos capítulos anteriores, a resposta que pode ser alcançada em malha-fechada é muito melhor que a alcançada pelo método baseado em modelo.

Como visto no Capítulo 5, é necessário obter uma estimativa para a matriz $M(\rho)$ para calcular o tamanho do passo de iteração. Neste exemplo esta estimativa será calculada utilizando o modelo identificado $\hat{G}(z)$ em todas as iterações.

O sinal de referência utilizado nos experimentos para o cálculo da estimativa do gradiente da função custo é o seguinte:

$$r_3(t) = 80 + 10square\left(\frac{2\pi t}{64}\right).$$

Esta referência possui um período mais curto para que seja dada ênfase no transitório da resposta do sistema.

O controlador $C_4(z, \rho)$ gera com a referência $r_3(t)$ uma resposta cujo custo é $\check{J}(\rho) = 7,84$ °C. Após 4 iterações o custo foi reduzido para $\check{J}(\rho) = 0,79$ °C, praticamente 10 vezes menor que o custo inicial. A evolução dos parâmetros do controlador pode ser vista na Tabela 8.

O controlador alcançado na quarta iteração é dado por

$$C_6(z) = [1, 5749 \quad -2, 4277 \quad 0,8695] \left[\begin{array}{ccc} \frac{z^2}{z^2-z} & \frac{z}{z^2-z} & \frac{1}{z^2-z} \end{array} \right]^T$$

A resposta ao salto do sistema em malha-fechada com $C_6(z, \rho)$ pode ser vista na Figura 41. O custo calculado com a referência sendo um salto de 20 °C é de $\check{J}(\rho) = 0,38$ °C, que representa menos que um oitavo do custo do controlador inicial, e menos que um terço do custo calculado com o controlador projetado pelo método baseado em modelo.

6.4 Considerações finais

Neste capítulo foram apresentados alguns resultados experimentais dos assuntos tratados neste trabalho.

Foi mostrado que a metodologia Shaping da Função Custo é válida e pode ser facilmente utilizada na prática. Escolhendo o espectro do sinal de referência utilizado nos experimentos pode-se ampliar ou diminuir o tamanho do domínio de atração do mínimo global.

Como apresentado no Capítulo 3, referências cujos componentes do espectro estão concentrados nas baixas ou altas frequências proporcionam um maior domínio de atração. No exemplo foi mostrado que utilizando um sinal de referência no meio do espectro pode-se convergir para mínimos locais da função custo.

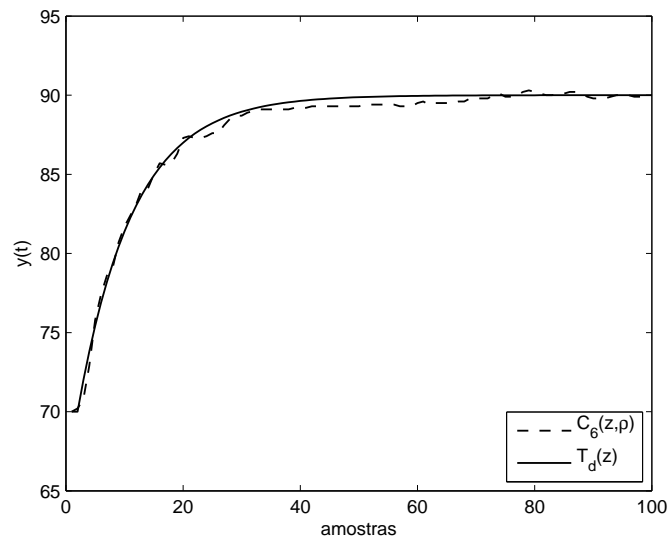


Figura 41: Resposta ao salto com controlador $C_6(z, \rho)$.

A metodologia de escolha do tamanho do passo de iteração também foi utilizada nos exemplos. O cálculo do tamanho do passo foi realizado com base na estimativa da matriz $M(\rho)$ calculada com um modelo estimado do processo. Mostrou-se que este modelo pode ser uma aproximação grosseira do comportamento do processo.

O algoritmo do gradiente com o tamanho do passo proposto mostrou-se bastante eficiente para ajustar os parâmetros do controlador. Apenas com 4 iterações o valor do custo alcançado foi menor que um oitavo do custo inicial e menor que um terço do custo alcançado pelo controlador projetado com o modelo grosseiro identificado.

7 CONCLUSÕES

Este trabalho apresentou várias propriedades de convergência para o mínimo global da função custo, no problema de projeto de controladores baseado em dados.

Foi mostrado que se o critério minimizado for “bem comportado” então pode-se garantir a convergência utilizando algoritmos de otimização bastante simples. Utilizando o algoritmo do gradiente, foram apresentadas características da função custo e sua dependência dos parâmetros de projeto. A metodologia Shaping da Função Custo diz como realizar os experimentos de forma a aumentar o tamanho da região de atração candidata do mínimo global. Dessa maneira para condições iniciais mais distantes do mínimo global pode-se ainda garantir a convergência.

O principal resultado apresentado neste trabalho foi um método para a escolha do tamanho do passo de iteração para o algoritmo do gradiente, de forma a garantir que, se as condições iniciais estiverem dentro de um certo domínio de atração, o mínimo global é alcançado. Esse método é baseado em informações contidas na matriz $M(\rho)$. Esta matriz a princípio é desconhecida e portanto precisa ser estimada. Formas de estimar essa matriz foram apresentados e algumas características de robustez desta estimativa.

Como trabalhos futuros pretende-se obter formas mais precisas de obter as informações necessárias para o projeto do tamanho do passo de iteração. Além disso, em todo o trabalho foram consideradas regiões de atração candidatas com formato de bola. Pretende-se expandir essas regiões para formatos elipsoidais para aumentar as possibilidades de condições iniciais que convergem para o mínimo global da função custo. Também pretende-se estudar formas de garantir a estabilidade do sistema em malha-fechada quando a condição inicial não pertence ao domínio de atração.

REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas**. Belo Horizonte, Brasil: UFMG, 2004.
- DOROTHEE, N.-C. (Ed.). **Perspectives in control: theory and applications**. London: Springer, 1998.
- ANDERSON, B. D. O.; LIU, Y. Controller reduction: concepts and approaches. **IEEE transactions on automatic control**, Brookline, USA, v.34, n.8, p.802–812, 1989.
- ANDERSON, B. D. O.; MOORE, J. B. **Linear optimal control**. Englewood Cliffs, New Jersey: Prentice Hall, 1971.
- ARMIJO, L. Minimization of functions having continuous partial derivatives. **Pacific j. math.**, [S.l.], v.16, p.1–3, 1966.
- ASTROM, K. J. **Introduction to stochastic control theory**. New York, NY - USA: Academic Press, 1970.
- ASTROM, K. J.; WITTENMARK, B. On self tuning regulators. **Automatica**, Tarrytown, NY, USA, v.9, n.2, p.185 – 199, 1973.
- ASTROM, K. J.; WITTENMARK, B. **Adaptive control**. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1994.
- BAZANELLA, A. S.; GEVERS, M.; MISKOVIC, L.; ANDERSON, B. On the shape of H_2 performance criteria. In: IFAC SYMPOSIUM ON SYSTEM, STRUCTURE AND CONTROL, 2007, Foz do Iguassu, Brazil. **Proceedings...** Porto Alegre: IFAC, 2007.
- BAZANELLA, A. S.; GEVERS, M.; MISKOVIC, L.; ANDERSON, B. Iterative minimization of H_2 control performance criteria. **Automatica**, Tarrytown, NY, USA, v.44, n.10, p.2549–2559, 2008.
- BITMEAD, R. R.; GEVERS, M.; WERTZ, V. **Adaptive optimal control: the thinking man's gpc**. New York: Prentice Hall International, 1990.
- BOYD, S.; GHAOUI, L. E.; FERON, E.; BALAKRISHNAN, V. **Linear matrix inequalities in system & control theory**. Philadelphia: Society for Industrial and Applied Mathematic, 1994.
- CAMPI, M. C.; LECCHINI, A.; SAVARESI, S. An application of the virtual reference feedback tuning method to a benchmark problem. **European journal of control**, Milano, Italy, v.9, p.66–76, 2003.

CAMPI, M. C.; LECCHINI, A.; SAVARESI, S. M. Virtual reference feedback tuning: a direct method for the design of feedback controllers. **Automatica**, Kidlington, v.38, n.8, p.1337–1346, 2002.

CAMPI, M.; SAVARESI, S. Direct nonlinear control tuning design: the virtual reference feedback tuning approach. **IEEE transactions on automatic control**, Brookline, USA, v.51, p.14–26, 2006.

CAUCHY, A. L. Méthode générale pour la résolution des systèmes d'équations simultanées. **Comptes rendus de l'academie des sciences**, Paris, v.25, p.536–538, 1847.

CURRY, H. The method of steepest descent for nonlinear minimization problems. **Quart appl. math.**, [S.l.], v.2, p.258–261, 1944.

TRENTELMAN, H. L.; WILLEMS, J. C. (Ed.). **Essays on control: perspectives in the theory and its applications**. Boston: Birkhauser, 1993.

GEVERS, M. Identification for Control: from the early achievements to the revival of experiment design. In: IEEE CONFERENCE ON DECISION AND CONTROL AND EUROPEAN CONTROL CONFERENCE 2005. CDC-ECC 05., 44., 2005, Seville, Spain. **Proceedings...** Seville: IEEE, 2005. p.12–12.

GOLDSTEIN, A. A. Cauchy's method of minimization. **Numer. math.**, [S.l.], v.4, p.146–150, 1962.

GOODWIN, G. C.; SIN, K. S. **Adaptive filtering, prediction and control**. Englewood Cliffs, NJ, EUA: Prentice Hall, 1984.

HILDEBRAND, R.; LECCHINI, A.; SOLARI, G.; GEVERS, M. Asymptotic Accuracy of Iterative Feedback Tuning. **IEEE transactions on automatic control**, Brookline, USA, v.50, n.8, p.1182–1185, Aug 2005.

HJALMARSSON, H. Iterative feedback tuning - an overview. **International journal of adaptive control and signal processing**, Glasgow, UK, v.16, n.5, p.373–395, 2002.

HJALMARSSON, H. From experiment design to closed-loop control. **Automatica**, Tarrytown, NY, USA, v.41, n.3, p.393 – 438, 2005.

HJALMARSSON, H.; GEVERS, M.; BRUYNE, F. D. For model-based control design, closed loop identification gives better performance. **Automatica**, Tarrytown, NY, USA, v.32, n.12, p.1659–1673, 1996.

HJALMARSSON, H.; GEVERS, M.; GUNNARSSON, S.; LEQUIN, O. Iterative feedback tuning: theory and applications. **IEEE control systems magazine**, Piscataway, USA, v.18, n.4, p.26–41, 1998.

HJALMARSSON, H.; GUNNARSSON, S.; GEVERS, M. A convergent iterative restricted complexity control design scheme. In: CONFERENCE ON DECISION AND CONTROL, 33., 1994, Lake Buena Vista, FL, USA. **Proceedings...** New York: IEEE, 1994. p.1735–1740.

HJALMARSSON, H.; GUNNARSSON, S.; GEVERS, M. Optimality and sub-optimality of iterative identification and control design schemes. In: AMERICAN CONTROL CONFERENCE, 1995, Seattle, USA. **Proceedings...** [S.l.: s.n.], 1995. v.4, p.2559–2563.

HOF, P. M. J. van den; SCHRAMA, R. J. P. Identification and control: closed-loop issues. **Automatica**, Tarrytown, NY, USA, v.31, n.12, p.1751–1770, 1995.

KAMMER, L. C. Stability assessment for cautious iterative controller tuning. **Automatica**, Tarrytown, NY, USA, v.41, n.10, p.1829 – 1834, 2005.

KAMMER, L. C.; BITMEAD, R. R.; BARTLETT, P. L. Direct iterative tuning via spectral analysis. **Automatica**, Tarrytown, NY, USA, v.36, n.9, p.1301–1307, 2000.

KARIMI, A.; MISKOVIC, L.; BONVIN, D. Convergence analysis of an iterative correlation-based controller tuning method. In: IFAC WORLD CONGRESS, 15., 2002, Barcelona, Spain (July 2002). **Proceedings...** New York: IFAC, 2002. p.1546.

KARIMI, A.; MISKOVIC, L.; BONVIN, D. Iterative Correlation-Based Controller Tuning: Application to a Magnetic Suspension System. **Control engineering practice**, London, UK, v.11, n.9, p.1069–1078, 2003.

KARIMI, A.; MISKOVIC, L.; BONVIN, D. Iterative correlation-based controller tuning. **International journal of adaptive control and signal processing**, Glasgow, UK, v.18, n.8, p.645–664, 2004.

KHALIL, H. K. **Nonlinear systems**. Upper Saddle River, NJ, USA: Prentice Hall, 1996.

KIWIEL, K. C.; MURTY, L. Convergence of the steepest descent method for minimizing quasiconvex functions. **Journal of optimization theory and applications**, Houston, TX, USA, v.89, p.221–226, 1996.

LEE, W.; ANDERSON, B.; KOSUT, R.; MAREELS, I. On robust performance improvement through the windsurfer approach to adaptive robust control. In: CONFERENCE ON DECISION AND CONTROL, 1993, San Antonio, USA. **Proceedings...** New York: IEEE, 1993. v.3, p.2821–2827.

LEE, W. S.; ANDERSON, B. D. O.; MAREELS, I. M. Y.; KOSUT, R. L. On some key issues in the Windsurfer approach to adaptive robust control. **Automatica**, Tarrytown, NY, USA, v.31, n.11, p.1619–1636, 1995.

LEE, W. S.; ANDERSSON, B. D. O.; KOSUT, R. L.; MAREELS, I. M. Y. A new approach to adaptive robust control. **International journal of adaptive control and signal processing**, Glasgow, UK, v.7, p.183–211, 1993.

LJUNG, L. **System identification: theory for the user**. Englewood Cliffs, NJ, USA: Prentice Hall, 1987.

LJUNG, L.; SODERSTROM, T. **Theory and practice of recursive identification**. Cambridge, Massachusetts: MIT Press, 1983.

PARTANEN, A. G.; BITMEAD, R. R. The application of an iterative identification and controller design to a sugar cane crushing mill. **Automatica**, Tarrytown, NY, USA, v.31, n.11, p.1547 – 1563, 1995.

POLAK, E. An historical survey of computational methods in optimal control. **SIAM review**, [S.l.], v.15, n.2, p.553–584, 1973.

PROCHAZKA, H.; GEVERS, M.; ANDERSON, B. D. O.; FERREIRA, C. Iterative feedback tuning for robust controller design and optimization. In: IEEE CONFERENCE ON DECISION AND CONTROL - EUROPEAN CONTROL CONFERENCE, 2005, Seville, Spain. **Proceedings...** New York: IEEE, 2005. p.3602–3607.

RAYMOND A. DE CALLAFON, P. M. J. V. d. H. Suboptimal Feedback Control by a Scheme of Iterative Identification and Control Design. **Mathematical modelling of systems**, [S.l.], v.3, n.1, p.77 – 101, 1997.

ROBBINS, H.; MONRO, S. A Stochastic Approximation Method. **The annals of mathematical statistics**, [S.l.], v.22, n.3, p.400–407, 1951.

SCHRAMA, R. J. P.; HOF, P. M. J. V. den. Iterative identification and control design: a three step procedure with robustness analysis. In: EUROPEAN CONTROL CONFERENCE, 1993, Groningen, Netherlands. **Proceedings...** [S.l.: s.n.], 1993. p.237–241.

VINNICOMBE, G. Frequency domain uncertainty and the graph topology. **IEEE transactions on automatic control**, Brookline, USA, v.38, n.9, p.1371–1383, Sep 1993.

ZANG, Z.; BITMEAD, R. R.; GEVERS, M. Iterative weighted least-squares identification and weighted LQG control design. **Automatica**, Tarrytown, NY, USA, v.31, n.11, p.1577–1594, 1995.