

Universidade Federal de Minas Gerais  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia Elétrica

# CONTRIBUIÇÕES AO PROBLEMA DE ESTIMAÇÃO DE PARÂMETROS EM SISTEMAS LPV

Luiz Cláudio Andrade Souza

Tese de Doutorado apresentada ao  
Programa de Pós-Graduação em  
Engenharia Elétrica da Universidade  
Federal de Minas Gerais, como  
requisito parcial para obtenção do  
título de Doutor em Engenharia  
Elétrica.

**Orientador:** Prof. Reinaldo Martinez Palhares, Dr. (UFMG)

Belo Horizonte, Março de 2010

"Contribuições ao Problema de Estimação  
de Parâmetros de Sistemas LPV"

**Luiz Cláudio Andrade Souza**

Tese de Doutorado submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de *Doutor em Engenharia Elétrica*.

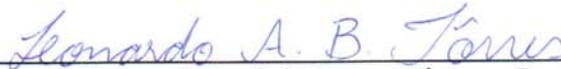
Aprovada em 05 de março de 2010.

Por:



---

Reinaldo Martinez Palhares - Dr.  
Prof. DELT/UFMG - Orientador



---

Leonardo Antônio Borges Tôres - Dr.  
Prof. DELT/UFMG



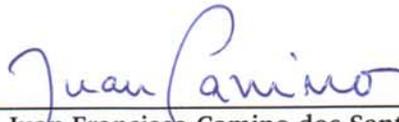
---

Bruno Otávio Soares Teixeira - Dr.  
Prof. DELT/UFMG



---

Eduardo Nunes Gonçalves - Dr.  
Prof. CEFET/MG



---

Juan Francisco Camino dos Santos - Dr.  
Prof. UNICAMP

## Resumo

Esta tese investiga o problema da estimação de parâmetros variantes no tempo em sistemas dinâmicos lineares denominados LPV's, do inglês *linear parameter-varying* (LPV) *systems*. Como modelos matemáticos, tais sistemas descrevem diretamente vários problemas relevantes em engenharia, e podem funcionar como aproximações razoáveis para uma gama ainda mais ampla de problemas não-lineares. Embora possa parecer tentador tratar a variação paramétrica como incerteza e aplicar a esses sistemas técnicas de análise e controle robusto, tal estratégia teria como limitações a introdução de conservadorismo, a validade limitada à variação lenta dos parâmetros e, mesmo nesse último caso, a possibilidade de falha em determinados casos. Alguns trabalhos na literatura procuram contornar esses problemas introduzindo a taxa máxima da variação paramétrica, suposta limitada, nos cálculos. Com isso, introduzem a necessidade de se conhecer ou de se estimar tal informação *a priori*. A outra alternativa, utilização de métodos variantes no tempo como *gain scheduling* e controle adaptativo, é aqui enfocada. Nessa linha o conhecimento do valor instantâneo dos parâmetros possibilitaria utilizar controladores baseados nos mesmos, sendo essa a principal motivação deste trabalho. Outras razões para se desejar conhecer o valor dos parâmetros, como por exemplo forma de medição de alguma grandeza de interesse, são aqui preocupações secundárias.

Em termos de apresentação, o texto começa com uma breve revisão dos problemas de identificação de sistemas e de estimação de estados, incluindo uma possível interpretação da estimação de parâmetros como estimação de estados em sistemas não-lineares. Situado o problema em termos de linearização do modelo em relação aos parâmetros, percebe-se que uma avaliação da largura de banda do sistema LPV seria fundamental não só para a implementação de técnicas que dependam de sua correta amostragem, mas também para abordagens no domínio do tempo contínuo. Os resultados alcançados nesse estudo preliminar, em termos de avaliação do raio espectral de matrizes LPV, poderão ser também relevantes para outras aplicações. Determinada a questão da amostragem, o problema de estimação de parâmetros é estudado primeiro no domínio do tempo discreto, como um problema de otimização. Percebem-se aí algumas limitações quanto ao que é possível de se estimar ou não, dadas características da função objetivo como o seu gradiente. Outra abordagem estudada é a transformação da estimação de parâmetros individuais em problemas escalares, obtendo uma nova interpretação das limitações antes encontradas. O trabalho se encerra com as conclusões e discussão de

algumas perspectivas para trabalhos futuros. Dentre essas últimas, as bases para uma reinterpretação dos resultados alcançados em termos de geometria diferencial são propostas em um dos apêndices.

## Abstract

This thesis is an investigation on time-varying parameter estimation in linear dynamic systems. Such systems, henceforth referred to as linear parameter-varying (LPV) systems, model many engineering problems, and are workable approximations to even more nonlinear problems. Treating parameter variation as uncertainty and applying robust analysis and control techniques to LPV systems, tempting as it may seem, is not an ideal approach: conservative, valid only for slowly varying parameters, and even then it may fail. Although some work in the literature try to sidestep those issues by introducing bounds on maximum parameter variation rates, it just becomes another piece of information that must be known or estimated *a priori*. This work is more in concert with the alternative, time-varying methods such as gain scheduling or adaptive control. System parameter values are a natural input to variable control structures, and when their actual values are not available, their estimates may be just as valuable. Parameter measurement for other purposes is a secondary concern here.

The text begins with a brief review of system identification and the state estimation problem, for parameter estimation is an identification problem that may be posed as a nonlinear state estimation. From the perspective of linearizing the problem with respect to the parameters, evaluating LPV system bandwidth may be perceived as fundamental not only to discrete-time study of the sampled system, but also to continuous-time analysis. The results obtained in that area, regarding the LPV matrix's spectral radius, may be useful for other applications as well. Given a discrete or properly sampled continuous LPV system, parameter estimation is then posed as an optimization problem: some limits on what can and cannot be estimated are detected via the objective function's gradient. After that, a second approach is proposed: turning individual parameters' estimation into scalar problems. Another set of limitations is obtained. Finally, conclusions and suggestions of possible further work are listed and discussed. In particular, some pointers for a reinterpretation of the results in the differential geometry framework are presented in one of the appendices.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Sistemas Lineares com Parâmetros	
	Variantes no Tempo . . . . .	1
1.1.1	Formulação das Matrizes LPV . . . . .	2
1.2	Análise e Projeto de Controladores para Sistemas LPV . . . . .	4
1.2.1	Exemplo . . . . .	6
1.3	Objetivos e Metodologia . . . . .	7
1.3.1	Estrutura do Texto . . . . .	8
<b>2</b>	<b>Identificação de Sistemas, Estimação de Estados e de Parâmetros</b>	<b>11</b>
2.1	Identificação de Sistemas . . . . .	12
2.2	Estimação de Estados . . . . .	15
2.2.1	Filtro de Kalman . . . . .	15
2.2.2	Extensões Não-Lineares do Filtro de Kalman . . . . .	18
2.2.3	Outros Filtros . . . . .	18
2.3	Estimação de Parâmetros . . . . .	18
2.3.1	Limitações . . . . .	20
<b>3</b>	<b>Estimativa do Raio Espectral e Amostragem de Sistemas LPV</b>	<b>22</b>
3.1	Estimação do Raio Espectral . . . . .	23
3.1.1	Raio Espectral Mínimo . . . . .	24
3.1.2	Raio Espectral Máximo . . . . .	26
3.1.3	Amostragem de Sistemas com Parâmetros Estáticos . . . . .	27
3.2	Sistemas LPV de Segunda Ordem . . . . .	28
3.2.1	Estudo dos Autovalores . . . . .	28
3.2.2	Comportamento em Coordenadas Polares . . . . .	31
3.3	Taxa de Amostragem . . . . .	33
3.4	Exemplo . . . . .	35
3.4.1	Análise Preliminar . . . . .	35
3.4.2	Autovalores . . . . .	36

3.4.3	Coordenadas Polares . . . . .	37
3.4.4	Raios Espectrais . . . . .	40
3.4.5	Amostragem . . . . .	43
<b>4</b>	<b>Estimação de Parâmetros via Otimização</b>	<b>46</b>
4.1	Função Objetivo . . . . .	47
4.2	Implementação . . . . .	48
4.2.1	Exemplo . . . . .	48
<b>5</b>	<b>Estimação de Parâmetros por Métodos Escalares</b>	<b>54</b>
5.1	Premissas e Definições . . . . .	54
5.1.1	Função Escalar do Estado . . . . .	55
5.2	Desenvolvimento . . . . .	56
5.2.1	Método em Malha Aberta . . . . .	56
5.2.2	Métodos em Malha Fechada . . . . .	58
5.3	Extensão para Múltiplos Parâmetros . . . . .	62
5.4	Exemplos . . . . .	63
5.4.1	Primeiro Exemplo . . . . .	63
5.4.2	Segundo Exemplo . . . . .	68
5.4.3	Terceiro Exemplo (Conclusão de 1.2.1) . . . . .	71
<b>6</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>73</b>
6.1	Conclusões . . . . .	73
6.2	Perspectivas de Trabalhos Futuros . . . . .	74
<b>A</b>	<b>Teoria de Magnus</b>	<b>75</b>
A.1	Solução de Sistemas Lineares Variantes no Tempo . . . . .	75
A.1.1	Solução Padrão . . . . .	75
A.1.2	Solução via Série de Magnus . . . . .	79
A.2	Sistemas LPV e a Série de Magnus . . . . .	81
<b>B</b>	<b>Transformações de Equivalência e de Lyapunov</b>	<b>83</b>

# Lista de Acrônimos, Notação Matemática e Terminologia

## Acrônimos

BLUE	<i>Best Linear Unbiased Estimate.</i>
EKF	<i>Extended Kalman Filter.</i>
LI	Linearmente Independente(s).
LMI	<i>Linear Matrix Inequality.</i>
LPV	<i>Linear Parameter Varying.</i>
LS	<i>Least Squares.</i>
LTI	<i>Linear Time-Invariant.</i>
LTV	<i>Linear Time-Varying.</i>
MLE	<i>Maximum Likelihood Estimate.</i>
NP-hard	<i>(Nondeterministic Polynomial time)-hard</i> , classe de complexidade.
PRBS	<i>PseudoRandom Binary Signal.</i>
SISO	<i>Single-Input, Single-Output.</i>
U-D	Decomposição U-D, $P = UDU^T$ , $U$ é uma matriz unitária triangular, e $D$ uma matriz diagonal.
UKF	<i>Unscented Kalman Filter.</i>
WLS	<i>Weighted Least Squares.</i>

## Notação Matemática

$\dot{x}(t)$	Derivada de $x(t)$ em relação ao tempo, $\dot{x}(t) = \frac{d}{dt}x(t)$ ou $\dot{x}(t) = \frac{\partial}{\partial t}x(t)$ conforme apropriado.
$ a $	Valor absoluto de $a$ .
$ \mathcal{A} $	Cardinalidade (número de elementos) do conjunto $\mathcal{A}$ .
$A^*$	Adjunta da matriz $A$ .
$\arctan(x)$	Arco tangente de $x$ .
$\mathcal{B}$	“Bola” (no sentido topológico).
$B$	Largura de banda (Capítulo 3, <i>cf.</i> $B$ em (1.13)).
$\text{Co}$	Casco convexo.
$a_{i,j}$	Componente da matriz $A$ , linha $i$ , coluna $j$ , $A = [a_{i,j}(\Theta)]$ ; $a_{i,j,k}$ = componente da matriz $A_k$ .
$\emptyset$	Conjunto vazio.
$\cos(x)$	Cosseno de $x$ .
$\Delta, \Delta(A)$	Determinante, determinante da matriz $A$ .
$\Delta A$	Incerteza associada à matriz $A$ ; também $\Delta\Theta$ = incerteza associada ao vetor $\Theta$ , etc. Exceções: $\Delta F$ definido em (3.24), e $\Delta F_1(t)$ em (5.10).
$\delta_{i,j}$	Função delta de Kronecker.
$\delta t$	Duração infinitesimal de um intervalo de tempo, $\delta t \rightarrow 0^+$ .
$\Delta t$	Duração finita de um intervalo de tempo.
$\det(A)$	Determinante da matriz $A$ .
$x \sim (\mu, \sigma^2)$	Distribuição da variável aleatória (ou vetor aleatório) $x$ com média $\mu$ e variância (ou matriz de covariância) $\sigma^2$ .
$x \sim \mathcal{N}(\mu, \sigma^2)$	Distribuição da variável aleatória $x$ normal (Gaussiana), média $\mu$ e variância $\sigma^2$ .
$e$	Número de Euler.

$\exp(x)$	Função exponencial, $\exp(x) = e^x$ .
$e_{m,n}$	$m$ -ésimo vetor elementar de $\mathbb{R}^n$ , e.g. $e_{2,3} = [0 \ 1 \ 0]^T \in \mathbb{R}^3$ .
$E[x]$	Valor esperado da variável aleatória $x$ .
$\mathfrak{g}$	Álgebra de Lie, associada ao grupo de Lie $G$ .
$I, I_p$	Matriz identidade, matriz identidade em $\mathbb{R}^p$ .
$\Im\{z\}$	Parte imaginária da variável complexa $z$ , cf. parte real $\Re\{z\}$ . Se $z = x + jy$ , $\Re\{z\} = x$ e $\Im\{z\} = y$ .
$(a, b)$	Intervalo $x \in \mathbb{R} \mid a < x < b$ .
$[a, b]$	Intervalo $x \in \mathbb{R} \mid a \leq x \leq b$ .
$(a, b]$ ou $[a, b)$	Intervalos $x \in \mathbb{R} \mid a < x \leq b$ ou $x \in \mathbb{R} \mid a \leq x < b$ .
$j$	Unidade imaginária, $j = \sqrt{-1}$ .
$\lambda(A)$	Autovalores da matriz $A$ .
$[\cdot, \cdot]$	Lie <i>bracket</i> , $[A, B] \triangleq AB - BA$ .
$\log(x)$	Logaritmo natural (base $e$ ) de $x$ .
$\ \cdot\ $	Norma.
$\mathcal{N}(A)$	Espaço nulo da matriz $A$ .
$\forall$	Para todo (valor da variável que se segue).
$\preceq$	Precede (menor, componente a componente).
$\otimes$	Produto de Kronecker.
$\mathbb{R}$	Conjunto dos números reais.
$\text{range}(A)$	<i>Range</i> (espaço das colunas ou imagem) da matriz $A$ .
$\rho(A)$	Raio espectral da matriz $A$ .
$\mathbb{R}^+$	Conjunto dos números reais não-negativos.
$\sin(x)$	Seno de $x$ .

$\text{sgn}(a)$	Função sinal de $a$ , $-1$ se $a < 0$ , $0$ se $a = 0$ , $+1$ se $a > 0$ .
t.o.s.	Termos de ordem superior.
$\text{tr}(A)$	Traço da matriz $A$ .
$A^T$	Transposição: $A^T =$ matriz $A$ transposta.
$\mathcal{L}$	Transformada de Laplace. Transformada inversa: $\mathcal{L}^{-1}$ .
$\mathcal{Z}$	Transformada Z. Transformada inversa: $\mathcal{Z}^{-1}$ .

## Terminologia

Espaço	No sentido topológico, é usado em referência a dois espaços de Hilbert (vide [1]) específicos: o espaço Euclidiano dos estados, $\mathbb{R}^n$ , e o espaço dos parâmetros $\mathbb{R}^p$ (ou $\mathbb{R}^{p+1}$ , em alguns casos). Sem qualificações, palavras como <i>espaço</i> , <i>trajetória</i> e <i>estimação</i> se referem a parâmetros. O espaço $\mathbb{R}^{n+p}$ , onde estados e parâmetros seriam estimados em conjunto, é relevante no contexto da Seção 2.3 .
Estático	Referindo-se a parâmetros, o adjetivo é usado no sentido de invariante no tempo. O sistema LPV continua sendo dinâmico em seus estados, tornando-se LTI. Relevante principalmente no Capítulo 3 .
Estimação	Neste trabalho a palavra se refere a estimação de parâmetros em sistemas LPV, no contexto do problema de filtragem mencionado na entrada para <i>filtro</i> . No Capítulo 2 a palavra subentende estimação de estados, <i>cf.</i> identificação de sistemas, pois lá se sugere como o problema LPV pode ser reinterpretado como tal .
Filtro	A palavra é utilizada em duas acepções relacionadas mas distintas neste trabalho. A primeira advém da teoria de processos estocásticos, especificamente o <i>problema de filtragem</i> , obtenção da “melhor estimativa” do valor real de alguma grandeza em um sistema, dada uma série de observações potencialmente ruidosas do mesmo. Nesse sentido têm-se os filtros de Wiener, Kalman, etc. O segundo sentido, mais prosaico, é aquele de processamento de sinais: remoção ou atenuação de componentes indesejadas

de um sinal, *e.g.* filtros passa-baixas, filtros Butterworth, etc. Filtros nos sentidos de teoria dos conjuntos, ordenação ou topologia não são discutidos neste trabalho .

Identificação No Capítulo 2, “identificação” seria o processo de selecionar e ajustar os parâmetros de uma estrutura matemática para se obter um modelo de um sistema físico. Embora o conceito seja aplicável a sistemas variantes no tempo e portanto a parâmetros em sistemas LPV, neste trabalho tal sentido é referido como estimação de parâmetros .

Trajetória O entendimento mais comum quando a palavra é mencionada individualmente seria o de “trajetória do sistema”, evolução do estado  $x(t)$  ao longo do tempo de acordo com uma solução específica de  $\dot{x}(t) = f(x(t), \dots)$ . Neste trabalho, porém, a palavra quase sempre se refere à trajetória dos parâmetros  $\Theta(t) \in \mathbb{R}^p$  que se deseja estimar .

# Lista de Figuras

1.1	Respostas do sistema LPV (1.5) a ser controlado (matrizes (1.15)) e da referência (matrizes (1.16)). . . . .	7
3.1	Resposta do sistema autônomo, $\omega = 1$ , à condição inicial $x(0) = [1, 0]^T$ . . . . .	37
3.2	Resposta do sistema autônomo, $\omega = 1$ , à condição inicial $x(0) = [1, 0]^T$ , no espaço de estados. . . . .	38
3.3	Magnitude $r(t)$ , com crescimento exponencial, em (a), e o respectivo logaritmo $\ell(t)$ , linear, em (b). . . . .	39
3.4	Fase $\phi(t) = -\omega t$ , linear, mapeada entre $-\pi$ e $\pi$ ( $\omega = 1$ ). . . . .	39
3.5	Subníveis $\rho(A(\Theta)) \leq \frac{7}{4}$ (a), e superníveis $\rho(A(\Theta)) \geq 1$ (b). . . . .	44
4.1	Resposta do sistema autônomo, exemplo da Seção 3.4, com parâmetros $\Theta(t) = \Theta^*(\omega t) = [\cos^2(\omega t), \frac{1}{2} + \cos(\omega t)\sin(\omega t), \sin^2(\omega t)]^T$ e $\omega = 1 + \frac{1}{\sqrt{2}}$ , para a condição inicial $x(0) = [1, 0]^T$ , em coordenadas cartesianas (a) e polares (b). Os parâmetros são mostrados em (c) e (d). . . . .	49
4.2	Estimativa dos parâmetros, otimização sem restrições, partindo de $\Theta_0 = [0.5, 0.5, 0.5]^T$ a cada instante. . . . .	50
4.3	Estimativa dos parâmetros, otimização sem restrições, partindo de $\Theta_0 = [0, 0, 0]^T$ a cada instante. . . . .	51
4.4	Componentes do gradiente $\nabla f(\Theta)$ ao longo de $\Theta^*(t)$ , em coordenadas cartesianas (a) e no espaço dos parâmetros (b) (informação de (a) em cores sobreposta à trajetória $\Theta^*(t)$ ). . . . .	52
4.5	Mapa de contorno $\frac{\partial f}{\partial \theta_1}$ , em (a) (valor absoluto proporcional à intensidade da cor azul), e sobreposição de $\frac{\partial f}{\partial \theta_1} = 0$ (uma das famílias de curvas de nível em (a)) com o valor exato de $\theta_1(t)$ e a estimativa desse parâmetro, $\hat{\theta}_1(t)$ , em (b). . . . .	52
4.6	Estimativa dos parâmetros, restritos a $0 \leq \hat{\theta}_k(t) \leq 1$ , $k = 1, 2, 3$ . . . . .	53
5.1	Cálculo do erro, método em malha aberta, um único parâmetro. . . . .	59

5.2	Cálculo da estimativa, método em malha aberta, um único parâmetro (incluindo o artifício numérico, ocasionalmente usado, de calcular a divisão por meio da diferença de logaritmos). . . . .	59
5.3	Cálculo do erro, método em malha fechada, um único parâmetro. . . . .	61
5.4	Primeiro exemplo, estimação de $\theta_1(t)$ : entrada e saída do sistema (a), método em malha aberta (b), método em malha fechada (c) e método em malha fechada simplificado (d). . . . .	64
5.5	Primeiro exemplo, estimação de $\theta_1(t)$ via o método em malha fechada simplificado, com ruído. . . . .	65
5.6	Primeiro exemplo, $\theta_2(t)$ : entrada e saída (a), malha aberta (b), malha fechada (c) e malha fechada simplificado (d). . . . .	66
5.7	Primeiro exemplo, $\theta_3(t)$ : entrada e saída (a), malha aberta (b), malha fechada (c) e malha fechada simplificado (d). . . . .	67
5.8	Segundo exemplo, estimação de um único parâmetro: métodos em malha aberta (a), malha fechada (b), e malha fechada simplificado com (c) e sem (d) ruído $\sim \mathcal{N}(0, 0.01)$ . . . . .	69
5.9	Segundo exemplo, estimação simultânea de $\theta_1(t)$ (a) e de $\theta_2(t)$ (b). . . . .	70
5.10	Segundo exemplo, estimação simultânea de $\theta_2(t)$ (a) e de $\theta_3(t)$ (b). . . . .	70
5.11	Segundo exemplo, estimação simultânea e bem-sucedida de $\theta_1(t)$ (a) e de $\theta_2(t)$ (b). . . . .	71
5.12	Terceiro exemplo, resposta ao degrau, controlador de [2] com o valor real do parâmetro (a), e respectivo erro (c). Em (b) o mesmo experimento é repetido com o controlador usando a estimativa mostrada na Figura 5.13, obtida simultaneamente; (d) é o erro correspondente a (b). . . . .	72
5.13	Terceiro exemplo, estimativa de $\theta_1(t)$ (a) e erro (b), para sistema e controlador descritos em [2], com a diferença que o valor estimado é que parametriza o controlador. . . . .	72

# Lista de Tabelas

3.1	Combinações factíveis de sinais dos elementos da matriz. . . . .	40
3.2	Limites máximos e mínimos dos raios espectrais, para cada combinação factível dos sinais dos elementos da matriz. . . . .	43

# Capítulo 1

## Introdução

### 1.1 Sistemas Lineares com Parâmetros Variantes no Tempo

Seja

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t), w(t), t) , \\ y(t) &= h(x(t), u(t), v(t), t)\end{aligned}\tag{1.1}$$

o modelo matemático de um sistema dinâmico não-linear.  $x(t)$  é denominado *estado*,  $u(t)$  *entrada* exógena,  $w(t)$  *ruído de processo*,  $t$  *tempo*,  $y(t)$  *saída*, e  $v(t)$  *ruído de medição*. Se as funções  $f(\cdot)$  e  $h(\cdot)$  são lineares, então o sistema é dito linear variante no tempo (*linear time-varying*, LTV), e pode ser escrito como:

$$\begin{aligned}\dot{x}(t) &= A(t)x(t) + B(t)u(t) + w(t) , \\ y(t) &= C(t)x(t) + D(t)u(t) + v(t) .\end{aligned}\tag{1.2}$$

Supondo que as dimensões dos vetores sejam  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  e  $y(t) \in \mathbb{R}^q$ , as matrizes terão dimensões  $A(t) \in \mathbb{R}^{n \times n}$ ,  $B(t) \in \mathbb{R}^{n \times m}$ ,  $C(t) \in \mathbb{R}^{q \times n}$  e  $D(t) \in \mathbb{R}^{q \times m}$ . Os vetores  $w(t) \in \mathbb{R}^n$  e  $v(t) \in \mathbb{R}^q$  são processos aleatórios, conhecidos apenas por suas estatísticas – neste trabalho, ambos são considerados ruídos brancos, de média zero e matrizes de covariância  $Q(t)$  e  $R(t)$ , independentes entre si e dos demais valores do modelo:  $w(t) \sim (0, Q(t))$  e  $v(t) \sim (0, R(t))$ , respectivamente. Ambos serão frequentemente omitidos a fim de simplificar a notação, mas mencionados quando relevantes para a discussão.

O objeto de estudo deste trabalho, sistemas lineares com parâmetros variantes no tempo (*linear parameter-varying*, LPV), são modelos lineares vari-

antes no tempo em que as matrizes de (1.2) são combinações afins<sup>1</sup> de funções escalares do tempo com coeficientes matriciais constantes:

$$\begin{aligned} \dot{x}(t) &= \left( A_0 + \sum_{k=1}^p A_k \theta_k(t) \right) x(t) + \left( B_0 + \sum_{k=1}^p B_k \theta_k(t) \right) u(t) , \\ y(t) &= \left( C_0 + \sum_{k=1}^p C_k \theta_k(t) \right) x(t) + \left( D_0 + \sum_{k=1}^p D_k \theta_k(t) \right) u(t) . \end{aligned} \quad (1.3)$$

As funções escalares do tempo  $\theta_k(t) : \mathbb{R}^+ \mapsto \mathbb{R}$ ,  $k = 1, \dots, p$  em (1.3) são denominadas *parâmetros* do sistema. Elas podem ser agrupadas em um vetor

$$\Theta(t) \triangleq [\theta_1(t) \ \dots \ \theta_p(t)]^T \in \mathbb{R}^p , \quad (1.4)$$

produzindo a notação mais concisa:

$$\begin{aligned} \dot{x}(t) &= A(\Theta(t))x(t) + B(\Theta(t))u(t) , \\ y(t) &= C(\Theta(t))x(t) + D(\Theta(t))u(t) . \end{aligned} \quad (1.5)$$

A dependência afim dos parâmetros ficará implícita nas notações  $A(\Theta(t))$ ,  $B(\Theta(t))$ ,  $C(\Theta(t))$  e  $D(\Theta(t))$  acima. Detalhes dessas formulações são discutidos a seguir, embora não sejam essenciais para a compreensão do restante do texto.

### 1.1.1 Formulação das Matrizes LPV

Sem perda de generalidade, apenas a matriz LPV

$$A(\Theta(t)) = \left( A_0 + \sum_{k=1}^p A_k \theta_k(t) \right) \in \mathbb{R}^{n \times n} \quad (1.6)$$

será considerada. Definam-se ainda

$$e_{m,n} \triangleq [\delta_{m,1} \ \dots \ \delta_{m,n}]^T \in \mathbb{R}^n , \quad (1.7)$$

onde

$$\delta_{i,j} = \begin{cases} 1 , & i = j \\ 0 , & i \neq j \end{cases}$$

---

<sup>1</sup>Devido ao termo  $A_0$ . É possível obter também uma formulação linear, como se discute na Seção 1.1.1.

é a função delta de Kronecker, e

$$A_S \triangleq \begin{bmatrix} A_1 \\ \vdots \\ A_p \end{bmatrix} \in \mathbb{R}^{np \times n}, \quad (1.8)$$

uma matriz em blocos que armazena as constantes  $A_1$  a  $A_p$ . Note-se que  $e_{m,n}$ ,  $m = 1, \dots, n$  são apenas os vetores elementares em  $\mathbb{R}^n$ .

Com as duas definições acima, mais o produto de Kronecker  $\otimes$  e  $I_p$  como matriz identidade em  $\mathbb{R}^p$ , é possível escrever:

$$A(\Theta(t)) = A_0 + \sum_{i=1}^p ((e_{i,p} \otimes I_p)^T (\Theta(t))^T \otimes A_S) (e_{i,p} \otimes I_p). \quad (1.9)$$

A formulação (1.9) pode ser feita linear, e ainda mais compacta, se for introduzido um parâmetro formal  $\theta_0(t)$  multiplicando  $A_0$ : com as definições alternativas

$$\Theta^+(t) \triangleq [\theta_0(t) \quad \dots \quad \theta_p(t)]^T \in \mathbb{R}^{p+1} \quad (1.10)$$

e

$$A_S^+ \triangleq \begin{bmatrix} A_0 \\ \vdots \\ A_p \end{bmatrix} \in \mathbb{R}^{n(p+1) \times n}, \quad (1.11)$$

$A(\Theta(t))$  torna-se

$$A(\Theta(t)) = \sum_{i=1}^{p+1} ((e_{i,p+1} \otimes I_{p+1})^T (\Theta^+(t))^T \otimes A_S^+) (e_{i,p+1} \otimes I_{p+1}). \quad (1.12)$$

A formulação acima é conveniente para manipulações algébricas, pois a condição  $\theta_0(t) = 1$  pode ser facilmente reintroduzida quando se fizer necessário. Cabe ainda comentar que, em (1.12),  $A(\Theta(t))$  corresponderia grosso modo a um “traço” da matriz  $\Theta^+(t)^T \otimes A_S^+$ , considerada em blocos  $n \times n$ . Embora não se faça aqui nenhuma alegação de rigor matemático, tal idéia é compatível com a noção de  $A(\Theta(t))$  ser uma função convexa de  $\Theta(t)$ .

## 1.2 Análise e Projeto de Controladores para Sistemas LPV

Se as matrizes forem constantes, (1.2) se reduz a um sistema linear invariante no tempo (*linear time-invariant*, LTI):

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) , \\ y(t) &= Cx(t) + Du(t) . \end{aligned} \quad (1.13)$$

Há um vasto corpo de conhecimentos sobre sistemas da forma acima: são de particular interesse para este trabalho os problemas de identificação de sistemas e de estimação de estados a partir de relações entre entradas e saídas, como também o problema de controle, tanto no seu aspecto de regulação<sup>2</sup>, quanto no de rastreamento<sup>3</sup>.

Modelos LTI onde há algum grau de incerteza sobre os coeficientes podem ser escritos como

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A)x(t) + (B + \Delta B)u(t) \\ y(t) &= (C + \Delta C)x(t) + (D + \Delta D)u(t) . \end{aligned} \quad (1.14)$$

Tais incertezas poderiam, à primeira vista, modelar a variação paramétrica de um sistema LPV: isso seria interessante pois a teoria de controle robusto oferece diversas ferramentas para análise e projeto de controladores, com as mesmas funcionalidades daqueles para sistemas LTI, satisfazendo ainda critérios de desempenho como  $\mathcal{H}_2$  ou  $\mathcal{H}_\infty$ .

Para utilizar (1.14) como aproximação para (1.3), é necessário assumir primeiro que a faixa de variação dos parâmetros seja limitada. Seja  $\mathcal{B} \subset \mathbb{R}^p$  uma “bola”<sup>4</sup> tal que  $\Theta(t) \in \mathcal{B}$ ,  $\forall t$ ; se os parâmetros forem normalizados tais que  $0 \leq \theta_i(t) \leq 1$ ,  $i = 1, \dots, p$ , por exemplo,  $\mathcal{B} = [0, 1]^p$ , um hipercubo unitário. Se então

$$A(\Theta(t)) \subset A + \Delta A \quad \forall \Theta \in \mathcal{B} ,$$

e repetindo o procedimento para as demais matrizes, (1.5) pode ser reinterpretada como um sistema incerto na forma (1.14) que contém todas as

---

<sup>2</sup>Controladores cujo objetivo é gerar uma entrada  $u(t)$  que leve o sistema a um ponto de equilíbrio, *e.g.*  $x(t) \rightarrow 0$ .

<sup>3</sup>O objetivo do controlador é fazer com que a saída do sistema acompanhe uma trajetória desejada, *e.g.*  $y(t) \rightarrow \bar{y}(t)$ , onde  $\bar{y}(t)$  é essa trajetória desejada.

<sup>4</sup>No sentido topológico, *i.e.* não necessariamente uma esfera.

configurações possíveis do sistema LPV original. De fato, incluem-se aí configurações que o sistema LPV jamais assumiria, o que torna essa aproximação conservadora em dois aspectos distintos. O primeiro está relacionado ao desacoplamento entre as variações das diferentes matrizes: por exemplo, se  $A(\Theta(t_1)) = A_{t_1} \neq A(\Theta(t_2)) = A_{t_2}$  e  $B(\Theta(t_1)) = B_{t_1} \neq B(\Theta(t_2)) = B_{t_2}$ , o sistema LPV nunca assumiria as configurações  $(A_1, B_2)$  ou  $(A_2, B_1)$ ; já o sistema incerto inclui ambas as possibilidades. O segundo aspecto tem a ver com o conjunto de todas as trajetórias possíveis de  $\Theta(t)$ , supondo  $\Theta(t) \in \Delta\Theta$ ,  $\forall t$ . Se  $\Delta\Theta$  for uma região não convexa, e  $\mathcal{B} = \text{Co}(\Delta\Theta)$  for o casco convexo dessa incerteza, haverá pontos de  $\mathcal{B}$  que nunca são atingidos. O volume de  $\mathcal{B} - \Delta\Theta$  representa em parte esse conservadorismo, que pode ainda ser amplificado se as incertezas matriciais, funções de  $\mathcal{B}$ , (*e.g.*  $\Delta A = \Delta A(\mathcal{B})$ ) introduzirem ganhos.

De qualquer forma, a rigor a aproximação robusta descrita acima só seria válida se parâmetros não variassem no tempo, *i.e.* se fossem estáticos. Na prática, a aproximação poderia continuar valendo enquanto a variação paramétrica fosse “lenta” em relação à dinâmica do sistema incerto. Trabalhos como [3, 4, 5, 6] analisam mais a fundo estabilidade quadrática e controle LPV nessas condições, valendo destacar métodos que incluem nessa análise também a taxa de variação máxima  $\Delta\dot{\Theta}$  dos parâmetros, a fim de compensar eventuais efeitos dessa variação, como apresentado em [7, 8]. No entanto, o que pode limitar definitivamente o valor de tais soluções é a existência de sistemas LPV em que instabilidade pode ser induzida não pela variação rápida dos parâmetros, mas sim por trajetórias específicas desses parâmetros [9].

Uma alternativa para o controle de sistemas LPV, que evita os problemas descritos, é o uso de técnicas de *gain scheduling* [10], válidas para sistemas variantes no tempo de forma geral. A idéia básica é alterar a estrutura de controle, seja pelo chaveamento de controladores fixos, seja pelo uso de controladores parametrizados de acordo com alguma grandeza do sistema: estado, entradas, saídas ou, no caso LPV, o valor parâmetros propriamente ditos [11, 12, 2]. Trabalhos na literatura, como [2], constroem suas análises assumindo que os parâmetros sejam conhecidos *a priori* ou ao menos que estejam disponíveis *on-line*; aqui se investiga o que fazer caso os parâmetros não estejam disponíveis, especificamente a viabilidade de estimá-los e de usar essa estimativa no controle. Antes disso, a próxima seção introduz o exemplo de [2], a título de motivação.

### 1.2.1 Exemplo

Deseja-se, em [2], que o sistema LPV (1.5), com as matrizes<sup>5</sup>

$$\begin{aligned} A(\Theta(t)) &= \begin{bmatrix} \theta_1(t) & 1 \\ 1 & 0 \end{bmatrix}, & B(\Theta(t)) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ C(\Theta(t)) &= [0 \ 1], \end{aligned} \quad (1.15)$$

siga o modelo de referência

$$\begin{aligned} \bar{A}(\Theta(t)) &= \begin{bmatrix} 2.8 & 2 \\ 1 & 0 \end{bmatrix}, & \bar{B}(\Theta(t)) &= \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \\ \bar{C}(\Theta(t)) &= [0 \ 1]. \end{aligned} \quad (1.16)$$

O parâmetro é dado como  $\theta_1(t) = \cos(y(t)) \in [-1, 1]$ .

Os autores propõem o controlador

$$u(t) = K(\Theta) (x(t) - G\bar{x}(t)) + M(\Theta)\bar{x}(t) + Q\bar{u}(t),$$

onde  $u(t)$  é a entrada do sistema LPV,  $\bar{x}(t)$  o estado do modelo de referência, e  $\bar{u} = 1(t)$ , degrau unitário, a entrada dessa referência. As matrizes de projeto  $G$ ,  $Q$ ,  $M(\Theta)$  e  $K(\Theta)$  devem ser tais que

$$\begin{aligned} \bar{C} &= CG, & G\bar{A}(\Theta) &= BM(\Theta) + A(\Theta)G, \\ G\bar{B} &= BQ, & K(\Theta) &= L(\Theta)P^{-1}. \end{aligned}$$

Ainda, as matrizes  $L(\Theta)$  e  $P = P^T > 0$  acima devem satisfazer a LMI

$$H(\Theta) = (A(\Theta)P + BL(\Theta)) + (A(\Theta)P + BL(\Theta))^T < 0.$$

Para se obter uma solução, escolhe-se  $L(\Theta)$  com a estrutura igual à de  $A(\Theta)$ ,

$$L(\Theta) = L_0 + \sum_{k=1}^p L_k \theta_k. \quad (1.17)$$

Com isso, no artigo calculam-se

$$\begin{aligned} L_0 &= \begin{bmatrix} -73.7192 & 0 \end{bmatrix}, & L_1 &= \begin{bmatrix} -81.9103 & 32.7641 \end{bmatrix}, \\ P &= \begin{bmatrix} 81.9103 & -32.7641 \\ -32.7641 & 81.9103 \end{bmatrix}, & K(\Theta) &= \begin{bmatrix} -1.0714 - \theta_1(t) & -0.4286 \end{bmatrix}, \\ G &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & M(\Theta) &= \begin{bmatrix} -2.8 - \theta_1(t) & -1 \end{bmatrix} \text{ e} \\ Q &= 2, \end{aligned}$$

---

<sup>5</sup>N.B. a notação dependente de parâmetros é usada no artigo mesmo para matrizes constantes.

que satisfazem as condições de projeto.

Sem o controlador, a resposta ao degrau do sistema LP com as matrizes (1.15) seria como mostra a Figura 1.1. Embora a figura não mostre, com o controlador a resposta do sistema se torna indistinguível da referência. Ao final desta tese pretende-se mostrar que, se o parâmetro não estiver disponível, é possível estimá-lo e obter o mesmo resultado.

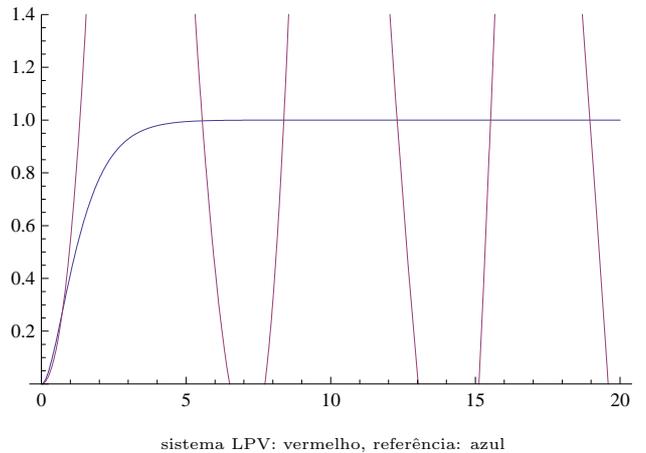


Figura 1.1: Respostas do sistema LPV (1.5) a ser controlado (matrizes (1.15)) e da referência (matrizes (1.16)).

### 1.3 Objetivos e Metodologia

Introduzidos os temas principais pela discussão apresentada até aqui, pode-se dizer que objetivo primário deste trabalho seria estudar o problema de estimação de parâmetros em sistemas LPV tendo em vista utilizar a informação obtida para aprimorar o controle dos mesmos. É importante destacar essa finalidade para desde já eliminar uma possível fonte de mal-entendidos: não se pretende aqui medir os parâmetros, no sentido de se obter um valor preciso para os mesmos a todo tempo. A influência paramétrica sobre o comportamento observável do sistema é variável, e nem sempre há informação suficiente para se obter uma medida. Informação limitada, porém, pode ainda ser útil. Nesse sentido, não foi feita uma análise quantitativa dos erros de estimação encontrados, mas apenas uma análise subjetiva. Análise quantitativa, como erro quadrático médio, só seria aplicável ou no contexto de medição, não estudado aqui, ou no contexto de aplicação, onde poderia ter sido usada para comparar as medidas de erro das Figuras 5.12 (c) e (d).

Em termos de metodologia, o trabalho se iniciou com a construção dos métodos de estimação a partir de matemática básica. Durante esse processo percebeu-se a necessidade de justificar a amostragem do sistema LPV ou de estabelecer o tamanho de uma vizinhança, em torno de um dado instante de tempo, onde determinadas aproximações poderiam ser feitas. Isso levou ao estudo preliminar sobre largura de banda e amostragem de sistemas LPV, que é apresentado primeiro. No tocante à revisão da literatura pertinente a esses esforços, o que se percebeu foi uma falta de informação sobre o assunto: falta não no sentido de não ser possível encontrar artigos citando as palavras-chaves necessárias, mas sim no fato de a grande maioria deles só tratar de forma tangencial os pontos de interesse desta tese.

Uma questão interessante, “por quê não se usou filtragem de Kalman como método para estimar parâmetros?”, merece ser respondida à parte. Do ponto de vista da literatura, há bastante material sobre o assunto, sendo que a Seção 2.3 chega a apresentar o estimador adaptado. Tal idéia não foi levada adiante, e apresentada junto aos outros métodos, em parte por falta de tempo. Além disso, a contribuição extra teria sido possivelmente pequena. Para encerrar a discussão, vale apresentar uma justificativa filosófica para a abordagem pouco convencional adotada nesta tese, parafraseando Kuhn [13]. Segundo ele, o estabelecimento científico tem uma forte tendência à ortodoxia, no sentido de construir um corpo de conhecimentos consistente e trabalhar para mantê-lo assim. Até certo ponto obtém-se um desenvolvimento linear, mas eventualmente tal paradigma pode se tornar ultrapassado. Quando isso acontece os proponentes da necessidade de se inovar são marginalizados num primeiro momento, até que a substituição do paradigma antigo por um novo se torne inevitável. Têm-se assim as revoluções científicas, como heliocentrismo *vs.* geocentrismo, ou física quântica e relatividade *vs.* física clássica. Este trabalho não pretende ser revolucionário, mas não deixa de tentar perturbar o *status quo* pelo menos um pouco.

### 1.3.1 Estrutura do Texto

O Capítulo 2 faz uma breve revisão sobre os problemas de identificação de sistemas e de estimação de estados. A idéia não é explorar a fundo muito menos esgotar os assuntos, mas sim contextualizar o problema de estimação de parâmetros que, a rigor, seria um problema de identificação, mas que do ponto de vista prático, e de ferramental matemático, talvez se aproxime mais da estimação de estados. A adaptação do filtro de Kalman descrita na Seção 2.3 ilustra esse segundo ponto de vista. Este capítulo pode ser omitido

sem prejudicar a leitura do restante da tese.

Em seguida, o Capítulo 3 investiga largura de banda e amostragem de sistemas LPV. Embora lineares em sua estrutura externa, a variação temporal dos parâmetros faz com que a questão da amostragem de entradas e saídas não seja óbvia, repercutindo nos problemas de identificação e estimação em tempo discreto. Mesmo no domínio do tempo contínuo é necessário determinar o tamanho da vizinhança, dada a linearização do sistema num determinado instante de tempo, em que essa linearização permanece válida. Em contraste com trabalhos já mencionados onde a taxa máxima de variação paramétrica é suposta conhecida, o estudo nesta tese aponta na direção de não ser necessário conhecer ou limitar essa taxa, visto que a parte linear do modelo LPV funciona como limitante natural para seus efeitos, um filtro “passa-baixas”. Cabe ainda mencionar que o cálculo do raio espectral, usado como estimativa conservadora para largura de banda de sistemas incertos, tem aplicações também em sistemas amostrados e matemática discreta, aplicações essas que poderão ser exploradas em trabalhos futuros.

Uma primeira visão de estimação de parâmetros é apresentada no Capítulo 4. Nele o problema, considerado em termos de amostras do sistema LPV, é escrito como minimização de uma função de erro. As características dessa função objetivo permitem afirmar que o problema nem sempre possui solução exata, no sentido que em determinados momentos pode não ser possível estimar um ou mais parâmetros. Isso ocorre, conforme dito antes, pois nem sempre o comportamento observável do sistema é sensível a todos os parâmetros.

A segunda interpretação da estimação de parâmetros, na forma de problemas escalares, é apresentada no Capítulo 5. Os métodos propostos evoluem de uma abordagem em “malha aberta” (estimador em malha aberta, não significa que o sistema LPV seja uma malha aberta) a uma abordagem em “malha fechada” (mesmo comentário), onde o parâmetro é obtido pela demodulação de sua influência na resposta do sistema. As condições em que essa influência desaparece, e suas conseqüências para a estimação do correspondente parâmetro, bem como possível a interferência entre dois ou mais parâmetros, são vistas em um novo ângulo. O capítulo se encerra retomando o exemplo da Seção 1.2.1, que serve para demonstrar a viabilidade de se estimar um parâmetro e aplicar o valor obtido no controle LPV.

Finalmente, o Capítulo 6 lista as conclusões e discute alguns possíveis trabalhos futuros. Cabe também citar os apêndices A e B: o segundo é

uma revisão sobre transformações em matrizes variantes no tempo, usadas no Capítulo 5, enquanto o primeiro explora a solução de equações diferenciais de um ponto de vista alternativo, baseado em geometria diferencial. Originalmente essa teoria deveria servir como base para um terceiro método de estimação de parâmetros, mas por razões de tempo tal desenvolvimento ficou relegado a futuras investigações.

## Capítulo 2

# Identificação de Sistemas, Estimação de Estados e de Parâmetros

O problema de estimação dos parâmetros variantes no tempo em (1.3) está entre as áreas de identificação de sistemas e de estimação de estados, que compartilham grande parte do mesmo ferramental matemático. Os parâmetros são, a rigor, parte do modelo do sistema LPV; a outra parte, os coeficientes matriciais, são considerados conhecidos neste trabalho. Assumindo que os estados sejam mensuráveis, a estimação dos parâmetros ficaria assim caracterizada como um problema de identificação. De fato, se os parâmetros  $\theta_k(t)$  forem funções diferenciáveis do tempo<sup>1</sup>, para cada instante  $t_i$  haverá um valor  $\delta t > 0$  tal que, na vizinhança  $|t - t_i| \leq \delta t$ , os termos  $A_k \theta_k(t) x(t)$  de (1.3) podem ser aproximados como  $A_{k,i} x(t)$ ,  $A_{k,i} \triangleq A_k \bar{\theta}_{k,i}$ ,  $\bar{\theta}_{k,i}$  o valor médio do parâmetro nessa vizinhança. Por outro lado, os mesmos termos podem ser escritos também como  $A_k (\theta_k(t) x(t))$ : agrupando os parâmetros como os estados, e reescrevendo o sistema apropriadamente, o problema se torna estimação de um estado aumentado  $x^a(t)$ .

Embora este trabalho se concentre na primeira estratégia, uma abordagem da segunda alternativa é introduzida ao final deste capítulo; antes, porém, são apresentados os principais conceitos de identificação de sistemas e de estimação de estados, numa breve revisão. Em termos de referências, a visão geral sobre estimação de estados aqui apresentada vem de [14]. (Por isso, relativamente poucas outras referências são citadas ao longo do texto.) Quanto à identificação de sistemas, o material básico aqui apresentado vem de [15, 16]; para uma visão mais aprofundada, citam-se [17, 18].

---

<sup>1</sup>Ou mesmo sob condições um pouco menos estritas.

## 2.1 Identificação de Sistemas

Embora a identificação de sistemas passe primeiro por um estágio passivo de observação, determinação de princípios de funcionamento e levantamento de dados relevantes, o primeiro passo ativo normalmente é a obtenção de informações de entrada e de saída. O processo de excitação do sistema para aquisição de dados pode ser feito de diversas maneiras, como por exemplo:

- **Resposta transiente:** resposta a impulsos ou a degraus. Este tipo de procedimento é relativamente simples de implementar, e frequentemente representativo dos sinais de entrada que o sistema pode encontrar em operação normal. Por outro lado, a amplitude de excitação necessária para uma boa relação sinal-ruído pode ser elevada, e os resultados dos testes podem não ser diretamente usáveis para identificação ou controle;
- **Resposta em frequência:** amplitude e fase da resposta à excitação por senóides de diferentes frequências. A coleta de dados pode ser mais trabalhosa do que na alternativa anterior, particularmente se as constantes de tempo do sistema são altas. Amplitude da excitação também é fator na relação sinal-ruído, mas os dados obtidos podem ser usados diretamente em alguns métodos de projeto de controle;
- **Informação estocástica em estado estacionário:** utiliza a resposta ao ruído sobreposto a entradas e saídas do sistema, durante sua operação normal. Na prática o desempenho é limitado por dois fatores: a amplitude da perturbação pode ser insuficiente, ou a excitação dinâmica do sistema pode ser incompleta;
- **Excitação por sinal binário pseudo-aleatório** (*Pseudorandom binary signal*, PRBS): excitação do sistema por um sinal binário de amplitude  $+A$  ou  $-A$ , conforme saída de um registrador de deslocamento de  $n$  bits. Além da amplitude do sinal e da largura do registrador, a função de realimentação do mesmo é também parâmetro de projeto, produzindo uma seqüência pseudo-aleatória de período  $2^n - 1$  suficientemente longo.

Obtidos os dados, o passo seguinte é determinar um modelo matemático capaz de explicar o comportamento observado e prever respostas do sistema a futuras entradas. Tais modelos podem ser paramétricos – comple-

tamente caracterizados por um conjunto bem definido de fatores – ou não-paramétricos – descrições funcionais cuja caracterização precisa pode não ser trivial, como por exemplo respostas em frequência. Neste trabalho a discussão está restrita aos primeiros, cujas estruturas são conhecidas *a priori*. De particular interesse são os modelos de sistemas dinâmicos que podem ser expressos como funções matemáticas explícitas, como (1.1): tal equação será reescrita aqui ignorando os ruídos e destacando sua natureza paramétrica, via a dependência de um vetor de parâmetros  $\Theta(t)$ :

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), \Theta(t), t) , \\ y(t) &= h(x(t), u(t), \Theta(t), t) . \end{aligned}$$

Para fins de identificação, freqüentemente o sistema é amostrado. Discretizando a equação acima, tem-se:

$$\begin{aligned} x_{k+1} &= f_d(x_k, u_k, \Theta_k, k) , \\ y_k &= h_d(x_k, u_k, \Theta_k, k) . \end{aligned}$$

Para simplificar esta discussão inicial o sistema será considerado invariante no tempo, isto é, os parâmetros  $\Theta_k = \Theta$  serão considerados constantes. Outra simplificação será reescrever o sistema na forma de uma equação de diferenças,

$$y_k = g(y_{k-1}, \dots, y_{k-n}, u_k, \dots, u_{k-n}, \Theta) . \quad (2.1)$$

A idéia de um modelo de ordem finita  $n$ , dimensão do estado que armazena a história do sistema, está implícita acima. Uma outra observação é que todas as grandezas podem ser escalares ou vetores, a distinção é imaterial aqui.

Sejam  $\tilde{u}$  e  $\tilde{y}$  os valores mensurados (*i.e.*, incluindo ruídos) das entradas e saídas do sistema. Introduzindo esses valores na equação (2.1), num dado instante  $k$  o valor da saída seria previsto como

$$\hat{y}_k = g(\tilde{y}_{k-1}, \dots, \tilde{y}_{k-n}, \tilde{u}_k, \dots, \tilde{u}_{k-n}, \Theta) .$$

O único fator desconhecido na equação acima seriam os parâmetros  $\Theta$ . A função  $g$  a cada instante  $k$  pode ser reescrita como  $m_k(\Theta)$ , variante no tempo pois depende das entradas  $\tilde{u}_k, \dots, \tilde{u}_{k-n}$  e das saídas  $\tilde{y}_{k-1}, \dots, \tilde{y}_{k-n}$ , medidas no instante atual ( $\tilde{u}_k$ ) e antes. Define-se então o erro de predição no instante  $k$ , função dos parâmetros, como:

$$e_k \triangleq y_k - \hat{y}_k = m_k(\Theta) . \quad (2.2)$$

Conforme posto acima, o problema de identificação paramétrica pode ser colocado como uma otimização: minimizar um custo, *e.g.* o quadrado de (2.2) acima, como função dos parâmetros. Tal idéia, mínimos quadrados (*least squares*, LS), é fundamental para uma série de desenvolvimentos posteriores. Assumindo que a função  $m_k(\Theta)$  seja linear nos parâmetros, ela pode ser escrita como o produto de um vetor  $m_k$  pelo vetor de parâmetros  $\Theta$ . Tem-se assim

$$e_k = y_k - h_k^T \Theta .$$

Tomando-se um número  $N$  de medidas, é possível escrever:

$$\begin{bmatrix} e_k \\ e_{k+1} \\ \vdots \\ e_{k+N} \end{bmatrix} = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+N} \end{bmatrix} - \begin{bmatrix} m_{k,1} & m_{k,2} & \dots & m_{k,p} \\ m_{k+1,1} & m_{k+1,2} & \dots & m_{k+1,p} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k+N,1} & m_{k+N,2} & \dots & m_{k+N,p} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$

ou

$$e = y - M\Theta . \quad (2.3)$$

Se  $J$  é um custo quadrático, escalar, a ser minimizado:

$$\begin{aligned} J &= e^T e = y^T y - y^T M\Theta - \Theta^T M^T y + \Theta^T M^T M\Theta \\ \Rightarrow \frac{\partial}{\partial \Theta} e^T e &= -2M^T y + 2M^T M\Theta = 0 \\ \Rightarrow \Theta_{\text{LS}} &= (M^T M)^{-1} M^T y , \end{aligned} \quad (2.4)$$

com a matriz  $M$  definida implicitamente em (2.3). Por meio de alguns resultados de álgebra linear a técnica acima pode ser adaptada para permitir o recálculo da estimativa com adição de novas medidas, sem necessidade de inverter matrizes  $M$  cujas dimensões crescem a cada instante: tal extensão é denominada estimador recursivo por mínimos quadrados.

Outra variação importante de (2.4) é o estimador ponderado  $\Theta_{\text{WLS}} = (M^T W M)^{-1} M^T W y$  (*weighted least squares*, WLS), também com uma versão recursiva. A idéia é atribuir pesos diferentes, via a matriz  $W$ , para os diferentes pontos de dados: por exemplo, atribuir mais confiabilidade a medidas mais recentes. Isso pode ser reinterpretado do ponto de vista estatístico: de fato, se a variância do ruído for  $R = \sigma^2 I$ , então a estimativa WLS com

$W = R^{-1}$  é o chamada *best linear unbiased estimate* (BLUE).

Para finalizar a discussão sobre modelos paramétricos e o ponto de vista estatístico, cabe citar também os métodos de *maximum likelihood estimate* (MLE). Neste tipo de identificação as funções de densidade de probabilidade para as variáveis aleatórias relevantes são escritas em termos dos parâmetros, e ajustadas de acordo com as observações. Embora mais flexível, tal estratégia pode se tornar computacionalmente intensiva e só viável como procedimento numérico.

## 2.2 Estimação de Estados

No problema de estimação de estados, parte-se do pressuposto que o modelo em sistemas dinâmico seja conhecido. Além disso, conhecem-se também os valores de entradas e saídas, possivelmente combinados a ruídos. O processo de estimação é normalmente feito *on-line*, produzindo de forma recursiva o valor esperado do estado, condicionado à história anterior do sistema (conhecimento *a priori*), ao que o modelo prediz e à real saída mensurada (conhecimento *a posteriori*).

### 2.2.1 Filtro de Kalman

O filtro de Kalman é deduzido a partir da análise de como o valor esperado da estimativa de estado e sua covariância se propagam ao longo do tempo, na presença de ruídos com estatísticas também conhecidas, de acordo com as equações dinâmicas do sistema. Seja o modelo discreto

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} , \\ y_k &= H_k x_k + v_k , \end{aligned} \tag{2.5}$$

com ruídos de processo e de medição brancos, média zero, não correlacionados, cujas matrizes de covariâncias são dadas por

$$\begin{aligned} w_k &\sim (0, Q_k) , & v_k &\sim (0, R_k) , \\ E [w_k w_j^T] &= Q_k \delta(k - j) , & E [v_k v_j^T] &= R_k \delta(k - j) , \text{ e} \\ E [v_k w_j^T] &= 0 . \end{aligned}$$

O filtro de Kalman, estimador do estado e de sua covariância, é dado pelas seguintes equações:

$$\begin{aligned}
 \hat{x}_0^+ &= \mathbb{E}[x_0] , \\
 P_0^+ &= \mathbb{E}[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] , \\
 P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} , \\
 K_k &= P_k^-H_k^T(H_kP_k^-H_k^T + R_k)^{-1} \\
 &= P_k^+H_k^TR_k^{-1} , \\
 \hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} , \\
 \hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k\hat{x}_k^-) , \\
 P_k^+ &= (I - K_kH_k)P_k^-(I - K_kH_k)^T + K_kR_kK_k^T \\
 &= ((P_k^-)^{-1} + H_k^TR_k^{-1}H_k)^{-1} \\
 &= (I - K_kH_k)P_k^- .
 \end{aligned} \tag{2.6}$$

Acima,  $\hat{x}_k$  é a estimativa do estado no instante  $k$ , e  $P_k$  sua respectiva covariância.  $K_k$  é o ganho do filtro, e os sobrescritos  $-$  e  $+$  denotam conhecimento *a priori* e *a posteriori*, *i.e.* antes ou depois de as medidas  $y_k$  terem sido processadas. Os valores esperados do estado inicial e de sua covariância são usados para inicializar o algoritmo.

Para sistemas contínuos, o filtro de Kalman pode ser obtido através da discretização de

$$\begin{aligned}
 \dot{x}(t) &= Ax(t) + Bu(t) + w(t) , \\
 y(t) &= Cx(t) + v(t) , \\
 w(t) &\sim (0, Q_c) , \\
 v(t) &\sim (0, R_c) ,
 \end{aligned} \tag{2.7}$$

com intervalo de amostragem  $T$ . Para pequenos valores de  $T$ , toma-se a expansão em série  $e^{AT} = I + AT + \text{t.o.s.}$  e desprezam-se os termos de ordem superior, obtendo

$$\begin{aligned}
 F &= e^{AT} \approx (I + AT) , & G &= (e^{AT} - I)A^{-1}B \approx BT , \\
 H &= C , & Q &= Q_cT , \\
 R &= \frac{R_c}{T} .
 \end{aligned} \tag{2.8}$$

As matrizes acima são substituídas em (2.5), e como o sistema original é invariante no tempo,  $F_k = F$  e assim por diante. Quanto aos ruídos de

processo e de medição, supõe-se que as versões contínuas sejam tais que

$$\begin{aligned} w(t) &\sim (0, Q_c) \Leftrightarrow \text{E} [w(t)w(\tau)^T] = Q_c \delta(t - \tau) \text{ e} \\ v(t) &\sim (0, R_c) \Leftrightarrow \text{E} [v(t)v(\tau)^T] = R_c \delta(t - \tau) , \end{aligned}$$

de forma que as relações entre covariâncias contínuas e discretas sejam como em (2.8).

Após substituir as matrizes (2.8) em (2.6), o passo seguinte é retornar ao domínio contínuo. Primeiramente, toma-se a expressão para o ganho do filtro e calcula-se o limite quando  $T \rightarrow 0$ :

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= P_k^- C^T (C P_k^- C^T + R_c/T)^{-1} \\ \Rightarrow \frac{K_k}{T} &= P_k^- C^T (C P_k^- C^T T + R_c)^{-1} \\ \Rightarrow \lim_{T \rightarrow 0} \frac{K_k}{T} &= P_k^- C^T R_c^{-1} , \text{ ou } K(t) = P(t) C^T R_c^{-1} . \end{aligned}$$

As demais equações podem então ser obtidas diretamente (no caso das condições iniciais), ou utilizando o resultado acima e calculando uma vez mais limites (nos casos da covariância e da estimativa). O resultado final é

$$\begin{aligned} \hat{x}(0) &= \text{E} [x(0)] , \\ P(0) &= \text{E} [(x(0) - \hat{x}(0))(x(0) - \hat{x}(0))^T] , \\ K(t) &= P(t) C^T R_c^{-1} , \\ \dot{\hat{x}}(t) &= \lim_{T \rightarrow 0} \frac{\hat{x}_k^+ - \hat{x}_{k-1}^+}{T} \\ &= A\hat{x}(t) + Bu(t) + K(t)(y(t) - C\hat{x}(t)) , \\ \dot{P}(t) &= \lim_{T \rightarrow 0} \frac{P_{k+1}^- - P_k^-}{T} \\ &= -P(t) C^T R_c^{-1} C P(t) + AP(t) + P(t) A^T + Q_c . \end{aligned} \tag{2.9}$$

A solução da equação de Riccati,  $\dot{P}(t) = \dots$  acima, pode ter um alto custo computacional. Além disso, podem surgir problemas numéricos que resultam na perda da propriedade positiva definida de  $P(t)$ . Diferentes fatorações dessa matriz dão origem a métodos alternativos, *e.g.* raiz quadrada ou U-D, capazes de reduzir embora não eliminar esses problemas.

### 2.2.2 Extensões Não-Lineares do Filtro de Kalman

O filtro de Kalman como apresentado acima se aplica a sistemas dinâmicos lineares. Para sistemas não-lineares, uma adaptação básica seria linearizar as equações do modelo e estimar os estados a partir delas: tal idéia, o filtro de Kalman estendido (*extended Kalman filter*, EKF), é de implementação simples mas seu desempenho pode ser prejudicado dependendo de como a linearização distorça a propagação de médias e de covariâncias. O *unscented Kalman filter* (UKF) é uma alternativa geralmente superior em aplicações não-lineares, como [19, 20, 21]. Nesse filtro, ao invés de se usar o modelo linearizado para propagar as estatísticas, tomam-se uma série de amostras em torno da média, os chamados pontos sigma, que são transformados pela equação não-linear. Esses pontos são, por sua vez, usados para calcular médias e covariâncias mais precisas.

### 2.2.3 Outros Filtros

Para finalizar esta discussão, cabe mencionar outros filtros de importância histórica (Wiener) ou prática ( $\mathcal{H}_\infty$ , partículas). Os filtros de Wiener são calculados assumindo-se que sinais e ruídos sejam processos aleatórios estacionários, com características espectrais, autocorrelações e/ou correlações cruzadas conhecidas. Os resultados de projeto são sistemas lineares invariantes no tempo, causais ou não-causais, capazes de extrair a informação desejada. Já os filtros  $\mathcal{H}_\infty$  enfocam robustez: embora os filtros de Kalman também possam ser modificados para lidar com modelos imperfeitos, os filtros  $\mathcal{H}_\infty$  são projetados para minimizar o ganho máximo entre distúrbios e o erro de estimação. Informação estatística sobre ruídos não é requerida, embora possa ser usada. Finalmente, o filtro de partículas é uma alternativa não-linear para o EKF ou o UKF. Traçando um paralelo com os métodos MLE citados na seção 2.1, tanto o filtro de partículas quanto esses últimos utilizam densidades de probabilidade condicionais e podem ser computacionalmente caros, embora capazes de fornecer bons resultados.

## 2.3 Estimação de Parâmetros

Embora sua utilização primária seja estimar estados, os filtros discutidos na seção anterior podem ser adaptados para estimação de parâmetros. Dada a

equação (1.5), incluindo ruídos,

$$\begin{aligned}\dot{x}(t) &= A(\Theta(t))x(t) + B(\Theta(t))u(t) + w(t) , \\ y(t) &= C(\Theta(t))x(t) + D(\Theta(t))u(t) + v(t) ,\end{aligned}$$

se o estado for redefinido como

$$x^a(t) \triangleq \begin{bmatrix} x(t) \\ \Theta(t) \end{bmatrix} ,$$

o sistema LPV pode ser reinterpretado como um sistema não-linear. Supondo que o vetor de parâmetros  $\Theta(t)$  assumia um valor nominal constante, é possível somar a ele um ruído artificial  $w^\theta(t)$ , para que um filtro de Kalman possa atualizar sua estimativa. Definindo também um ruído de processo estendido,

$$w^a(t) \triangleq \begin{bmatrix} w(t) \\ w^\theta(t) \end{bmatrix} ,$$

tem-se:

$$\begin{aligned}\dot{x}^a(t) &= \begin{bmatrix} A(\Theta(t))x(t) + B(\Theta(t))u(t) + w(t) \\ w^\theta(t) \end{bmatrix} \\ &= f(x^a(t), u(t), w^a(t)) , \\ y(t) &= C(\Theta(t))x(t) + D(\Theta(t))u(t) + v(t) \\ &= h(x^a(t), u(t), v(t)) .\end{aligned}\tag{2.10}$$

Visando utilizar o filtro de Kalman estendido, seja  $\hat{x}^a(t)$  a estimativa do estado aumentado. Sejam também  $w^a(t) \sim (0, Q)$ ,  $v(t) \sim (0, R)$ , e as matrizes

$$\begin{aligned}\tilde{A} &= \left. \frac{\partial f}{\partial x^a} \right|_{\hat{x}^a} , & L &= \left. \frac{\partial f}{\partial w^a} \right|_{\hat{x}^a} , \\ \tilde{C} &= \left. \frac{\partial h}{\partial x^a} \right|_{\hat{x}^a} , & M &= \left. \frac{\partial h}{\partial v} \right|_{\hat{x}^a} , \\ \tilde{Q} &= LQL^T , & \tilde{R} &= MRM^T .\end{aligned}$$

Com essas definições, o filtro de Kalman

$$\begin{aligned}\hat{x}^a(0) &= E[x^a(0)] , \\ P(0) &= E[(x^a(0) - \hat{x}^a(0))(x^a(0) - \hat{x}^a(0))^T] , \\ \dot{\hat{x}}^a &= f(x^a(t), u(t), 0) + K(t)(y(t) - h(x^a(t), u(t), 0)) , \\ K(t) &= P(t)\tilde{C}^T\tilde{R}^{-1} , \\ \dot{P}(t) &= \tilde{A}P(t) + P(t)\tilde{A}^T + \tilde{Q} - P(t)\tilde{C}^T\tilde{R}^{-1}\tilde{C}P(t)\end{aligned}\tag{2.11}$$

pode ser usado para estimar o estado aumentado, incluindo os parâmetros  $\hat{\Theta}$ .

Para o caso discreto, seja o modelo

$$\begin{aligned} x_k &= F_{k-1}(\Theta)x_{k-1} + G_k(\Theta)u_{k-1} + L_{k-1}(\Theta)w_{k-1} , \\ y_k &= H_k x_k + v_k . \end{aligned} \quad (2.12)$$

Embora a equação acima assuma que  $H_k$  seja independente dos parâmetros, uma eventual dependência de  $\Theta$  não complicaria muito o problema. A idéia é, novamente, reescrever o sistema com o estado aumentado

$$x_k^a \triangleq \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} .$$

Se os parâmetros forem constantes, a evolução de  $\Theta$  no tempo poderia ser dada por

$$\Theta_k = \Theta_{k-1} + w_{k-1}^\theta ,$$

onde  $w_k^\theta$  é, uma vez mais, um termo artificial de ruído que possibilita ao filtro de Kalman ajustar a estimativa paramétrica. O modelo aumentado<sup>2</sup> torna-se

$$\begin{aligned} x_k^a &= F_{k-1}(\Theta)x_{k-1} + G_{k-1}(\Theta)u_{k-1} + L_{k-1}(\Theta)w_{k-1} \\ &= f(x_{k-1}^a, u_{k-1}, w_{k-1}, w_{k-1}^\theta) \\ y_k &= \begin{bmatrix} H_k & 0 \end{bmatrix} \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} + v_k , \end{aligned} \quad (2.13)$$

e um filtro não-linear pode então ser usado para estimar os parâmetros.

### 2.3.1 Limitações

Embora os filtros de Kalman e suas variações sejam amplamente utilizados, é possível apontar algumas limitações que os tornariam menos interessantes em determinadas aplicações. Por exemplo:

- Necessidade de conhecimento da média e da correlação dos ruídos  $w_k$  e  $v_k$ , a cada instante  $k$ . Isso não é observado na formulação apresentada aqui por causa da hipótese de ruídos com média zero e independentes entre si, adotada desde capítulo 1);

---

<sup>2</sup>N.B., embora a segunda equação,  $y_k = \dots$ , seja linear em  $x_k^a$ , a primeira não necessariamente o é.

- Estatisticamente, uma propriedade do filtro de Kalman é minimizar a variância do erro de estimação. Isso o torna ótimo se o ruído é Gaussiano, e a melhor alternativa linear se não. Outros filtros podem ser mais interessantes para outros tipos de sistemas, ruídos ou funções de custo, como por exemplo  $\mathcal{H}_\infty$ ;
- Necessidade de conhecer o modelo do sistema, *e.g.* as matrizes  $F_k$  e  $H_k$ , em todos os instantes. Cabe destacar que a versão para sistemas contínuos apresentada neste capítulo, em (2.9), assume sistemas LTI a fim de manter a apresentação simples.

Embora os aspectos estatísticos sejam importantes, o último ponto é talvez o mais relevante no que diz respeito à estimação de parâmetros de sistemas LPV. As adaptações (2.11) e (2.13) resultam em sistemas não lineares variantes no tempo, e essa variação está atrelada aos parâmetros que se deseja estimar. Do ponto de vista do ruído artificial  $w^\theta$ , os parâmetros são considerados ora constantes, quando da linearização local do modelo, ora variáveis, enquanto parte do estado estendido. Por causa dessa dualidade, a correta discretização de (1.3) torna-se crítica: diretamente em (2.13), indiretamente em (2.11). Esta é uma das motivações para o estudo a ser apresentado no próximo capítulo.

## Capítulo 3

# Estimativa do Raio Espectral e Amostragem de Sistemas LPV

A amostragem de sistemas LTI é um problema bem entendido, e está intimamente relacionada aos problemas de identificação desses sistemas e de estimação de seus estados. Na tentativa de se estender os resultados e métodos LTI para sistemas LPV, uma possível interpretação desses últimos seria a de superestruturas lineares invariantes no tempo, encapsulando subsistemas não-lineares variantes no tempo. O capítulo começa com um estudo sobre o raio espectral de uma matriz afim-dependente de parâmetros com os mesmos estáticos, *i.e.* invariantes no tempo: o cálculo desse valor, bem como do vetor  $\Theta$  que o atinge, permite avaliar um limitante superior para a largura de banda de sistemas LTI incertos.

A segunda parte do capítulo trata de sistemas LPV de segunda ordem que, devido ao espaço de estados plano, são fáceis de visualizar, ilustrando três categorias básicas no que diz respeito aos autovalores: reais distintos, reais repetidos, ou par de valores complexos conjugados. Tais categorias são importantes pois sistemas de ordem superior podem ser decompostos em blocos com essas características, por transformações de similaridade ou de equivalência<sup>1</sup>. Em particular, o caso de autovalores conjugados complexos permite estudar os efeitos da variação temporal na largura de banda, comparada à de sistemas LTI: aqui se propõe o uso de uma aproximação prática para adicionar uma margem de segurança que permite amostrar corretamente o sistema variante no tempo. O capítulo se encerra combinando os resultados da primeira parte com os da segunda, determinando uma abordagem para se calcular a taxa mínima de amostragem para sistemas LPV. Um resumo do desenvolvimento e dos resultados deverá ser publicado em [22].

---

<sup>1</sup>Vide apêndice B.

Em termos de idéias fundamentais e referências, citam-se primeiramente [23, 24]: a partir do estudo de  $\dot{x}(t) = A(t)x(t)$  e de seu caso particular  $A(t) = A(\Theta(t))$  mostra-se que, dadas certas restrições quanto ao intervalo mínimo entre possíveis descontinuidades (“*dwell time*”), é possível caracterizar uma taxa de crescimento exponencial. O problema discreto

$$x(n+1) = \left( A_0 + \sum_{k=1}^p A_k \theta_k(n) \right) x(n) \quad (3.1)$$

e os conceitos de raio espectral conjunto e generalizado, relacionados a produtos de matrizes, como na matriz de transição de estados para o sistema acima, também são relevantes visto que após amostragem (1.3) toma essa forma. Do ponto de vista computacional, a verificação da estabilidade de (3.1) é NP-hard [25, 26] e o cálculo dos raios mencionados um problema não-decidível [27], mas em alguns casos o resultado pode ser aproximado [28] ou mesmo calculado explicitamente, como nos problemas de conectividade de grafos estudados em [29].

### 3.1 Estimação do Raio Espectral

Dada a matriz  $A(\Theta(t))$  de (1.3), seus elementos  $a_{i,j}(\Theta(t))$  são funções afins dos parâmetros. Considerando os parâmetros invariantes no tempo (estáticos), a matriz  $A(\Theta(t)) = A(\Theta)$  e seus elementos podem ser escritos como:

$$\begin{aligned} A(\Theta) &= [a_{i,j}(\Theta)] = A_0 + \sum_{k=1}^p A_k \theta_k \\ \Rightarrow a_{i,j}(\Theta) &= a_{i,j,0} + \sum_{k=1}^p a_{i,j,k} \theta_k \quad . \end{aligned} \quad (3.2)$$

Os autovalores  $\lambda(A(\Theta))$  da matriz acima são as raízes da equação característica  $\det(\lambda I - A(\Theta)) = 0$ , e os coeficientes do polinômio característico  $\det(\lambda I - A(\Theta))$  são produtos dos elementos  $a_{i,j}(\Theta)$ . Por causa disso, os autovalores são funções contínuas de  $\Theta$ . O raio espectral corresponde ao menor círculo centrado na origem que engloba esses autovalores, sendo determinado pelo autovalor de maior valor absoluto. Como esse último varia com os parâmetros, a função raio espectral,  $\rho(A(\Theta))$ , é também contínua em  $\Theta$ .

Não há, porém, garantias quanto a sua concavidade ou convexidade, e nem quanto a sua diferenciabilidade. Maximização ou minimização de  $\rho(A(\Theta))$  em relação a  $\Theta$  não são triviais, podendo ficar presas em pontos críticos locais.

Uma possível alternativa para se calcular o raio espectral seria a fórmula de Gelfand<sup>2</sup> [30],

$$\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{1/k} .$$

Tal formulação evita os problemas mencionados acima, mas introduz várias outras dificuldades computacionais. Isso também vale para o cálculo do raio espectral por meio de artifícios como  $\rho(A(\Theta)) = \sqrt{\rho(A^*(\Theta)A(\Theta))}$ , que introduz produtos dos parâmetros originais  $\theta_k$ . Por outro lado, o corolário 6.1.8 de [31], cujo resultado final é reproduzido abaixo, permite calcular um limitante para o raio espectral que freqüentemente<sup>3</sup> coincide com o valor correto. Além disso, a fórmula preserva a dependência afim dos parâmetros:

$$\rho(A) \leq \min_{p_1, \dots, p_n > 0} \max_{1 \leq i \leq n} \left\{ \frac{1}{p_i} \sum_{j=1}^n p_j |a_{i,j}| \right\} . \quad (3.3)$$

Notas: (1) a equação está escrita acima em termos das linhas de  $A$ , embora também pudesse ser expressa em termos das colunas da matriz. A fim de deixar isso claro, as desigualdades obtidas expandindo os termos  $\max\{\dots\}$  de (3.3) serão sempre referidas como desigualdades de linha/coluna. (2) Introdução das variáveis auxiliares  $p_i \in \mathbb{R}^+$ , cf.  $p$ , número de parâmetros.

### 3.1.1 Raio Espectral Mínimo

Algumas transformações são necessárias para se ter (3.3) na forma de um problema de minimização (linear) padrão. Primeiramente, define-se

$$a_{i,j} \triangleq a_{i,j}^+ - a_{i,j}^- , \quad (3.4)$$

com  $|a_{i,j}| = a_{i,j}^+ + a_{i,j}^-$  e  $a_{i,j}^+, a_{i,j}^- \geq 0$ , tal que valores absolutos sejam eliminados da expressão (artifício relativamente comum em otimização): (3.3) se torna

$$\rho(A) \leq \min_{p_1, \dots, p_n > 0} \max_{1 \leq i \leq n} \left\{ \frac{1}{p_i} \sum_{j=1}^n p_j (a_{i,j}^+ + a_{i,j}^-) \right\} . \quad (3.5)$$

<sup>2</sup>A fórmula de Gelfand vale para qualquer norma matricial.

<sup>3</sup>Leia-se, em todos os exemplos calculados na elaboração deste trabalho. A rigor o corolário fornece apenas um limitante (*bound*, em inglês), teoricamente pode haver casos onde esse valor superestime o raio espectral, daí o sinal  $\leq$ .

Embora não seja imediatamente óbvio neste ponto, na solução final as restrições implícitas  $a_{i,j}^+ \neq 0 \Rightarrow a_{i,j}^- = 0$  e  $a_{i,j}^- \neq 0 \Rightarrow a_{i,j}^+ = 0$  serão automaticamente satisfeitas (o problema final será resolvido como linear).

O problema minimax dado por (3.5) e equações (3.2) e (3.4) combinadas tem  $2n^2 + n + p$  variáveis:  $a_{i,j}^+$ ,  $a_{i,j}^-$ ,  $p_i$  e  $\theta_k$ , com  $i, j = 1, \dots, n$ ,  $k = 1, \dots, p$ . Introduzindo uma variável auxiliar  $z$  tal que  $\rho(A) \leq z$  e reescrevendo as expressões indexadas por  $i$  na equação (3.5) como desigualdades, o problema se torna

$$\begin{aligned} \min_{z, p_i, \theta_k, a_{i,j}^+, a_{i,j}^-} \quad & z & (3.6) \\ \text{s. a.} \quad & p_i z \geq \sum_{j=1}^n p_j (a_{i,j}^+ + a_{i,j}^-), & i = 1, \dots, n \\ & a_{i,j}^+ - a_{i,j}^- = a_{i,j,0} + \sum_{k=1}^p a_{i,j,k} \theta_k, & i, j = 1, \dots, n \\ & 0 \leq \theta_k \leq 1, & k = 1, \dots, p \end{aligned}$$

A formulação acima é não-linear, mas se os valores de  $p_i$  forem fixados ela se torna linear em  $z$ ,  $a_{i,j}^+$ ,  $a_{i,j}^-$  e  $\theta_k$ . Dessa maneira ela pode ser utilizada para calcular um limitante inferior para  $\rho(A(\Theta))$  em função de  $\Theta$  e, se as dimensões do problema forem relativamente pequenas, pode mesmo ser resolvida simbolicamente. Assumindo que os valores  $p_i$  sejam constantes positivas, é possível aplicar o método simplex, por exemplo, a (3.6): os valores de  $z$  e de  $\Theta$  assim obtidos são funções não lineares dos  $p_i$ 's. Numa segunda etapa, minimiza-se  $z$  em relação aos  $p_i$ 's e verifica-se se as respostas numéricas, tanto  $z^{\text{Opt}}$  quanto  $\Theta^{\text{Opt}}$ , satisfazem o problema original. Normalmente  $\rho(A(\Theta^{\text{Opt}}))$  é o valor correto, comparado a  $z^{\text{Opt}}$  que pode incluir alguma folga.

A computação aqui descrita deve ser feita com cuidado, pois muitos detalhes de implementação, *e.g.* seleção de *pivots* ou condições de factibilidade, dependem implicitamente das faixas de valores assumidos pelos  $p_i$ 's – tais detalhes serão explorados no exemplo que se segue. Alternativamente, pode-se adotar um conjunto de valores arbitrários para os  $p_i$ 's, por exemplo  $P_0$ , e resolver o problema para um valor ótimo  $\Theta^0$ ; toma-se então esse  $\Theta^0$ , e resolve-se o problema original, agora não linear, para um novo conjunto de  $p_i$ 's,  $P_1$ . Tal processo iterativo deverá convergir para  $P_q = P^{\text{Opt}}$  e  $\Theta_q = \Theta^{\text{Opt}}$ ,  $q = 0, \dots, Q$ , resolvendo o problema original. De qualquer forma, a complexidade depende principalmente da ordem da matriz,  $n$ , pois somente as desigualdades de linha/coluna contém produtos dos  $p_i$ 's com outras variáveis;

os  $p$  parâmetros  $\theta_k$  são relevantes na parte do problema que inclui seus limites e as igualdades envolvendo  $a_{i,j}^+$  e  $a_{i,j}^-$ . Essas últimas variáveis e restrições de igualdade podem ser eliminadas completamente, como se descreve a seguir.

Alguns dos  $n^2$  elementos  $a_{i,j}(\Theta)$  da matriz podem nunca mudar de sinal: seja por não dependerem de nenhum parâmetro, seja porque dados parâmetros limitados, *e.g.*  $0 \leq \theta_k \leq 1$ , alguns  $a_{i,j}(\Theta)$  possam ser sempre não-positivos ou não-negativos. De fato, muitas das  $2^{n^2}$  combinações de sinais de elementos da matriz podem nunca acontecer. Além disso, os problemas de factibilidade que detectam combinações impossíveis de sinais são lineares e numéricos, portanto de baixo custo computacional. Isso possibilita significativa simplificação do problema. A equação (3.6) pode ser reescrita como

$$\begin{aligned} \min_{z, p_i, \theta_k} \quad & z & (3.7) \\ \text{s. a.} \quad & p_i z \geq \sum_{a_{i,j} \geq 0} p_j \sum_{k=1}^p (a_{i,j,k} \theta_k) - \sum_{a_{i,j} \leq 0} p_j \sum_{k=1}^p (a_{i,j,k} \theta_k), \quad i = 1, \dots, n, \end{aligned}$$

e os valores dos parâmetros são tais que o correspondente subproblema de factibilidade seja satisfeito. Os valores absolutos são eliminados das equações de linha/coluna pois  $a_{i,j}(\Theta) = a_{i,j}^+(\Theta)$  para  $a_{i,j}(\Theta) \geq 0$ , e  $a_{i,j}(\Theta) = -a_{i,j}^-(\Theta)$  para  $a_{i,j}(\Theta) \leq 0$ . A redução de complexidade computacional ao se particionar o domínio  $\Theta$ , eliminando  $2n^2$  variáveis comparado a (3.6), é significativo.

### 3.1.2 Raio Espectral Máximo

Como visto na seção anterior, o corolário usado para calcular o limitante do raio espectral transforma-se naturalmente num problema de minimização quanto aos parâmetros  $\Theta$ , pois é também uma minimização quanto às variáveis  $p_i$ . É possível utilizar a mesma estrutura e restrições para maximização quanto a  $\Theta$ , mas a complexidade computacional pode aumentar significativamente. O problema original envolve minimizar o máximo de um conjunto de restrições afins (minimax), e visto que o epígrafo  $z$  dessas restrições é uma região convexa, encontrar o mínimo global é fácil. Por outro lado, maximizar o mínimo das restrições (maximin) também é simples, com um hipografo côncavo: se todas as restrições estiverem ativas, esse é o caso.

O problema se torna mais complexo se nem todas as  $n$  desigualdades de linha/coluna estiverem ativas na solução que maximiza  $\Theta$ . Seja  $\mathcal{L} \triangleq \{\ell_1, \dots, \ell_n\}$  o conjunto dessas restrições, das quais somente aquelas no sub-

conjunto  $\mathcal{L}_A$  estão ativas. Obviamente, o caso descrito no parágrafo anterior ocorre quando há igualdade, *i.e.*  $|\mathcal{L}_A| = |\mathcal{L}| = n$ ; no geral, porém,  $1 \leq |\mathcal{L}_A| \leq n$ , e se  $\mathcal{P}(\mathcal{L})$  for o conjunto potência<sup>4</sup> de  $\mathcal{L}$ , há  $|\mathcal{P}(\mathcal{L})| - 1 = 2^n - 1$  subconjuntos candidatos a  $\mathcal{L}_A$ , que devem ser testados. Uma abordagem heurística para esse problema seria testar  $\mathcal{L}_A = \mathcal{L}$  primeiro, e se nem todas as desigualdades estiverem ativas, procurar em  $\mathcal{P}(\mathcal{L})$  o subconjunto  $\mathcal{L}_A$  que satisfaça duas condições: 1)  $z$  máximo, e 2) para todo  $\ell_i \in \mathcal{L} - \mathcal{L}_A$ ,  $\ell_i(\Theta) < z$ . Essas últimas desigualdades garantem que a solução obtida para o problema maximin com as  $\mathcal{L} - \mathcal{L}_A$  desigualdades removidas coincida com a solução do problema minimax original. Embora a complexidade algorítmica cresça exponencialmente, o problema é passível de computação paralela; além disso, em muitas aplicações a ordem  $n$  é pequena o bastante para que o problema não se torne intratável.

### 3.1.3 Amostragem de Sistemas com Parâmetros Estáticos

A Seção 3.1.2 anterior mostrou como calcular o valor máximo para o raio espectral da matriz  $A(\Theta)$  com parâmetros estáticos. Se tal valor é atingido em  $\Theta = \Theta^{\text{Opt}}$ , propõe-se aqui que  $\rho(A(\Theta^{\text{Opt}}))$  seja utilizado como uma estimativa conservadora para a largura de banda do sistema incerto. Retomando a teoria para sistemas invariantes no tempo, o teorema de Nyquist–Shannon [32, 33] (vide também [34], para uma perspectiva mais recente) garante que uma taxa de amostragem  $f_S$  maior que  $f_N = 2B$ , onde  $B = \max \Im\{\lambda(A)\}$  é a largura de banda de um sistema LTI como (1.13), garante a amostragem correta de seu estado  $x(t)$ . Interpretando (1.3) com parâmetros estáticos como (1.14),  $B = \max \Im\{\lambda(A(\Theta))\} \leq \rho(A(\Theta^{\text{Opt}}))$ , portanto fica clara a validade dessa proposta. Poderia se argumentar que  $\max \Im\{\lambda(A(\Theta^{\text{Opt}}))\}$  fosse um limitante menos conservador, mas não há garantia que exista um valor  $\tilde{\Theta}^{\text{Opt}}$  tal que  $\max \Im\{\lambda(A(\tilde{\Theta}^{\text{Opt}}))\} > \max \Im\{\lambda(A(\Theta^{\text{Opt}}))\}$ , embora  $\rho(A(\tilde{\Theta}^{\text{Opt}})) \leq \rho(A(\Theta^{\text{Opt}}))$ . De qualquer forma, para estender os resultados para (1.3) falta discutir os efeitos da variação temporal dos parâmetros.

---

<sup>4</sup>*Powerset*, conjunto de todos os subconjuntos de um dado conjunto.

## 3.2 Sistemas LPV de Segunda Ordem

Para avaliar os efeitos da variação temporal dos parâmetros num sistema LPV, esta seção começa estudando o caso de segunda ordem. A matriz  $A(\Theta(t)) \in \mathbb{R}^{2 \times 2}$  contém quatro elementos  $a_{i,j}(\Theta(t))$ ,  $i, j \in \{1, 2\}$ , e com quatro parâmetros independentes pode-se capturar qualquer comportamento pois existem apenas quatro graus de liberdade. Para manter a notação compacta, inclui-se também o parâmetro formal  $\theta_0(t) = 1$ , constante, tal que

$$A(\Theta(t)) = \sum_{i=0}^4 A_i \theta_i(t) . \quad (3.8)$$

Tais hipóteses,  $A(\Theta(t)) \in \mathbb{R}^{2 \times 2}$  e  $\Theta(t) \in \mathbb{R}^5 \cap \{\Theta^+(t) \mid \theta_0(t) = 1\} = \mathbb{R}^4$ , ficarão implícitas a partir daqui<sup>5</sup>.

### 3.2.1 Estudo dos Autovalores

Conforme dito antes, autovalores e autovetores de  $A(\Theta(t))$  são funções contínuas de seus parâmetros. Para estudar a estrutura dessas autofunções, seja primeiramente a matriz estática  $A \in \mathbb{R}^{2 \times 2}$ ,

$$A \triangleq \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} .$$

Seus autovalores  $\lambda(A)$  são as soluções de

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} (\lambda - a_{1,1}) & -a_{1,2} \\ -a_{2,1} & (\lambda - a_{2,2}) \end{vmatrix} = 0 \\ \Rightarrow (\lambda - a_{1,1})(\lambda - a_{2,2}) - a_{1,2}a_{2,1} &= 0 \\ \Rightarrow \lambda^2 - (a_{1,1} + a_{2,2})\lambda + a_{1,1}a_{2,2} - a_{1,2}a_{2,1} &= 0 \\ \Rightarrow \lambda^2 - \text{tr}(A)\lambda + \det(A) &= 0 , \end{aligned}$$

portanto,

$$\lambda = \frac{\text{tr}(A) \pm \sqrt{[-\text{tr}(A)]^2 - 4\det(A)}}{2} = \frac{\text{tr}(A) \pm \sqrt{\Delta}}{2} ,$$

<sup>5</sup>A notação  $\mathbb{R}^5 \cap \{\Theta^+(t) \mid \theta_0(t) = 1\} = \mathbb{R}^4$  denota a interseção do espaço de parâmetros artificialmente aumentado,  $\mathbb{R}^{p+1} = \mathbb{R}^5$ , com o hiperplano  $\{\Theta^+(t) \mid \theta_0(t) = 1\}$  (notação (1.10)). Isso corresponde ao espaço de parâmetros original  $\mathbb{R}^p = \mathbb{R}^4$

onde

$$\Delta \triangleq [\text{tr}(A)]^2 - 4 \det(A) . \quad (3.9)$$

Dependendo do sinal de  $\Delta$ ,  $\text{sgn}(\Delta)$ , há três possibilidades:

- $\Delta > 0$  :  $A$  tem dois autovalores reais e distintos,  $\lambda_1 \neq \lambda_2 \in \mathbb{R}$ ;
- $\Delta = 0$  :  $A$  tem um autovalor real com multiplicidade dois,  $\lambda_1 = \lambda_2 \in \mathbb{R}$ ;
- $\Delta < 0$  :  $A$  tem um par de autovalores complexos conjugados,  $\lambda_{1,2} = \alpha \pm j\beta \in \mathbb{C}$ .

O comportamento de um sistema autônomo  $\dot{x}(t) = Ax(t)$ , ou de (1.13) no caso geral, assume características diferentes em cada um desses casos.

Considerando agora o sistema autônomo variante no tempo

$$\dot{x}(t) = A(\Theta(t))x(t) , \quad (3.10)$$

se o estado  $x_i = x(t_i)$  pertence à trajetória do sistema, então uma aproximação LTI local no instante  $t = t_i$  terá as características descritas acima<sup>6</sup>. A medição do estado  $x_i$  e o cálculo do valor instantâneo  $\text{sgn}(\Delta(t_i))$  podem ser úteis na interpretação do comportamento local, mas uma análise *a priori* pode fornecer também informação global, relevante para a estimação de  $\Theta(t)$ . Definam-se antes

$$O \triangleq \begin{bmatrix} e_{1,10}^T \\ e_{3,10}^T \\ e_{5,10}^T \\ e_{7,10}^T \\ e_{9,10}^T \end{bmatrix} \text{ e } E \triangleq \begin{bmatrix} e_{2,10}^T \\ e_{4,10}^T \\ e_{6,10}^T \\ e_{8,10}^T \\ e_{10,10}^T \end{bmatrix} ,$$

com  $O, E \in \mathbb{R}^{5 \times 10}$ . Tomem-se também o vetor de parâmetros estendido  $\Theta^+(t)$  de (1.10), os vetores elementares  $e_{i,10}$  (elementares em  $\mathbb{R}^{10}$ , conforme (1.7)), e a matriz  $A_S^+ \in \mathbb{R}^{10 \times 2}$  de (1.11).

O valor do determinante  $\Delta(t) = \det(A(\Theta(t)))$  pode ser escrito em função do vetor de parâmetros  $\Theta^+(t)$  como

$$\Delta(\Theta^+(t)) = \Theta^+(t)^T M_\Delta \Theta^+(t) , \quad (3.11)$$

<sup>6</sup>Tal aproximação LTI local pode ser justificada pelo teorema da linearização (Hartman-Grobman, [35, 36, 37], vide também [38]). O sistema LPV é linear nos estados, mas não-linear considerando-se estados e parâmetros conjuntamente.

onde

$$M_{\Delta} \triangleq (OA_S^+ e_{1,2} - EA_S^+ e_{2,2})(OA_S^+ e_{1,2} - EA_S^+ e_{2,2})^T + 4(OA_S^+ e_{2,2})(EA_S^+ e_{1,2})^T. \quad (3.12)$$

A matriz  $M_{\Delta}$  é estática (invariante no tempo). A dependência temporal na forma quadrática (3.11) se deve apenas a  $\Theta^+(t)$ .

O estudo do sinal  $\text{sgn}(\Delta(t))$  permite fazer afirmações sobre os autovalores de  $A(\Theta(t))$ . Caso  $M_{\Delta}$  seja positiva definida calcula-se  $\Theta^{\text{Opt}} = \min_{\Theta} \Delta(\Theta)$ , e se  $\Delta(\Theta^{\text{Opt}}) > 0$  o sistema terá sempre um par de autovalores reais distintos. Da mesma forma, se  $M_{\Delta}$  for negativa definida calcula-se  $\Theta^{\text{Opt}} = \max_{\Theta} \Delta(\Theta)$ , e se  $\Delta(\Theta^{\text{Opt}}) < 0$  o sistema terá sempre um par de autovalores conjugados complexos. Os casos mistos  $M_{\Delta} > 0$ ,  $\Delta(\Theta^{\text{Opt}}) < 0$  e  $M_{\Delta} < 0$ ,  $\Delta(\Theta^{\text{Opt}}) > 0$  correspondem a sistemas que mudam de um par de autovalores reais distintos para um par de conjugados complexos, ou vice-versa; na transição  $\{\Theta \mid \Delta(\Theta) = 0\}$  haverá um par de autovalores reais, iguais.

Se  $M_{\Delta}$  é positiva semidefinida ou negativa semidefinida, a análise para os casos definidos pode ainda ser usada, com alguns cuidados: é possível que  $\text{sgn}(\Delta(\Theta(t)))$  seja independente de um ou mais parâmetros  $\theta_i(t)$ , e tais variáveis devem ser removidas antes de prosseguir com os cálculos. Um exemplo é a condição  $\theta_0(t) = 1$ ,  $\forall t$ , pela qual  $\theta_0(t)$  deve ser eliminada. Para tanto, (3.11) deve ser particionada como

$$\begin{aligned} \Delta(\Theta(t)) &= \Theta^+(t)^T M_{\Delta} \Theta^+(t) \\ &= \begin{bmatrix} 1 & \Theta(t)^T \end{bmatrix} \begin{bmatrix} m_{1,1} & \bar{m}_{1,i}^T \\ \bar{m}_{i,1} & \bar{M}_{\Delta} \end{bmatrix} \begin{bmatrix} 1 \\ \Theta(t) \end{bmatrix} \\ &= m_{1,1} + \Theta(t)^T \bar{m}_{i,1} + \bar{m}_{1,i}^T \Theta(t) + \Theta(t)^T \bar{M}_{\Delta} \Theta(t), \end{aligned} \quad (3.13)$$

com

$$\begin{aligned} \bar{m}_{i,1} &= [m_{2,1} \quad m_{3,1} \quad m_{4,1} \quad m_{5,1}]^T \text{ e} \\ \bar{m}_{1,i} &= [m_{1,2} \quad m_{1,3} \quad m_{1,4} \quad m_{1,5}]^T. \end{aligned}$$

A mesma idéia pode ser usada para eliminar outros parâmetros. Por exemplo, se o coeficiente  $A_i = 0$ ,  $M_{\Delta}$  teria ao menos um autovalor nulo: eliminando-se o parâmetro  $\theta_i(t)$ , o problema é resolvido.

Para encerrar esta discussão, cabe lembrar que uma matriz  $M_{\Delta}$  genérica pode não ser positiva nem negativa, definida ou semidefinida: nesse caso, as

raízes da equação  $\Delta(\Theta) = 0$  não são facilmente caracterizadas. Há também um fator que pode facilitar a análise: se a região em que os parâmetros variam é limitada, *e.g.*  $\Theta(t) \in \mathcal{B}$ , a análise pode se ater somente a ela. De qualquer forma, embora conhecer as características dos autovalores do sistema *a priori* possa ser interessante, optou-se neste trabalho por desenvolver métodos de estimação de parâmetros que não dependam desse tipo de informação.

### 3.2.2 Comportamento em Coordenadas Polares

As trajetórias temporais de um sistema dinâmico de segunda ordem se restringem ao plano  $\mathbb{R}^2$ , e são facilmente visualizadas. Seja  $x(t) = [x_1(t) \ x_2(t)]^T$  o estado de um desses sistemas, e sejam as transformações-padrão

$$\begin{aligned} r(t) &\triangleq \sqrt{(x_1(t))^2 + (x_2(t))^2} \quad \text{e} \\ \phi(t) &\triangleq \arctan\left(\frac{x_2(t)}{x_1(t)}\right), \end{aligned} \quad (3.14)$$

que definem as coordenadas polares  $r(t) \geq 0$  e  $\phi(t) \in [0, 2\pi)$  (por exemplo). A transformação inversa é dada por

$$\begin{aligned} x_1(t) &= r(t) \cos(\phi(t)) = r(t)c_\phi \quad \text{e} \\ x_2(t) &= r(t) \sin(\phi(t)) = r(t)s_\phi, \end{aligned} \quad (3.15)$$

com  $s_\phi \triangleq \sin(\phi(t))$  e  $c_\phi \triangleq \cos(\phi(t))$  tornando a notação mais compacta.

Com as mudanças de variáveis de (3.15) acima, um sistema autônomo variante no tempo genérico  $\dot{x}(t) = A(t)x(t)$  pode ser reescrito como:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} &= A(t) \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \\ \Rightarrow \begin{bmatrix} \dot{r}(t)c_\phi - r(t)s_\phi\dot{\phi}(t) \\ \dot{r}(t)s_\phi + r(t)c_\phi\dot{\phi}(t) \end{bmatrix} &= r(t)A(t) \begin{bmatrix} c_\phi \\ s_\phi \end{bmatrix} \\ \Rightarrow \begin{bmatrix} c_\phi & -s_\phi \\ s_\phi & c_\phi \end{bmatrix} \begin{bmatrix} \dot{r}(t) \\ r(t)\dot{\phi}(t) \end{bmatrix} &= r(t)A(t) \begin{bmatrix} c_\phi \\ s_\phi \end{bmatrix}. \end{aligned} \quad (3.16)$$

Seja

$$\ell(t) \triangleq \log(r(t)) \quad \Rightarrow \dot{\ell}(t) = \frac{\dot{r}(t)}{r(t)} \quad \Rightarrow \dot{r}(t) = r(t)\dot{\ell}(t),$$

tal que (3.16) se torna

$$r(t) \begin{bmatrix} \dot{\ell}(t) \\ \dot{\phi}(t) \end{bmatrix} = r(t) \begin{bmatrix} c_\phi & s_\phi \\ -s_\phi & c_\phi \end{bmatrix} A(t) \begin{bmatrix} c_\phi \\ s_\phi \end{bmatrix}.$$

Definindo ainda

$$v_{\ell\phi}(t) \triangleq \begin{bmatrix} \ell(t) \\ \phi(t) \end{bmatrix}, \quad M_\phi(t) \triangleq \begin{bmatrix} c_\phi & s_\phi \\ -s_\phi & c_\phi \end{bmatrix}, \quad \text{e} \quad v_\phi(t) \triangleq \begin{bmatrix} c_\phi \\ s_\phi \end{bmatrix}, \quad (3.17)$$

e assumindo que  $r(t) \neq 0$  para que seu valor possa ser cancelado, (3.16) se torna, na forma polar,

$$\dot{v}_{\ell\phi}(t) = M_\phi(t)A(t)v_\phi(t). \quad (3.18)$$

Pode-se observar que a matriz  $M_\phi(t)$  acima é uma rotação, portanto  $M_\phi(t)^{-1} = M_\phi(t)^T$ . Outro ponto é que, para certos valores de  $\phi(t)$ , *e.g.*  $\phi(t) = k\pi$  ou  $\phi(t) = \pi/2 + k\pi$ , alguns dos elementos de  $M_\phi$  e de  $v_\phi$  momentaneamente se anulam: as conseqüências disso serão comentadas em breve.

Com a matriz LPV  $A(t) = A(\Theta(t))$ , a equação (3.18) pode ser escrita como

$$\begin{aligned} \dot{v}_{\ell\phi}(t) &= M_\phi(t)A(\Theta(t))v_\phi(t) \\ &= M_\phi(t) \left( A_0 + \sum_{k=1}^p A_k \theta_k(t) \right) v_\phi(t) \\ &= M_\phi(t)A_0 v_\phi(t) + \sum_{k=1}^p M_\phi(t)A_k v_\phi(t) \theta_k(t). \end{aligned}$$

ou, expandida,

$$\begin{aligned} \begin{bmatrix} \dot{\ell}(t) \\ \dot{\phi}(t) \end{bmatrix} &= \begin{bmatrix} M_\phi(t)A_0 v_\phi(t) & \dots & M_\phi(t)A_p v_\phi(t) \end{bmatrix} \begin{bmatrix} 1 \\ \Theta(t) \end{bmatrix} \\ &= A_\phi(t) \begin{bmatrix} 1 \\ \Theta(t) \end{bmatrix}, \end{aligned} \quad (3.19)$$

com  $A_\phi(t)$  definida implicitamente acima. Demonstra-se assim que a evolução temporal de  $v_{\ell\phi}$  é determinada por uma equação diferencial linear (afim) em  $\Theta(t)$ . Os coeficientes do sistema LPV original são embutidos na matriz  $A_\phi(t)$ , por sua vez função de  $\phi(t)$ . Tal dependência da fase, e não da magnitude  $r(t)$  do estado – ou de seu logaritmo  $\ell(t)$  – será explorada a seguir. Num comentário à parte, pode-se ver também que em determinados instantes, quando elementos de  $M_\phi$  ou  $v_\phi$  se anulam, parte da informação sobre pelo menos alguns dos parâmetros pode desaparecer momentaneamente.

### 3.3 Taxa de Amostragem

Com os resultados da Seção 3.1 é possível estimar o raio espectral de uma matriz LPV de dimensões arbitrárias, assumindo parâmetros estáticos. Por outro lado, a análise feita na Seção 3.2 e o argumento que se segue assumem que o sistema LPV seja de segunda ordem. A fim de combinar os dois pontos de vista, defina-se  $\bar{\lambda} \triangleq \sigma \pm j\omega$  como o par de autovalores conjugados complexos com maior parte imaginária  $\omega$ , quando considerados os autovalores de  $A(\Theta^{\text{Opt}})$  correspondentes ao raio espectral máximo  $\rho(A(\Theta^{\text{Opt}}))$  calculado na Seção 3.1.3. Submetendo  $A(\Theta^{\text{Opt}})$  a uma transformação de similaridade (apêndice B), esse par de autovalores pode ser isolado em um bloco planar  $\bar{A}(\Theta^{\text{Opt}}) \in \mathbb{R}^{2 \times 2}$  tal que, com a reintrodução da variação temporal dos parâmetros, torna-se:

$$\bar{A}(\Theta(t))|_{\Theta(t)=\Theta^{\text{Opt}}} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}.$$

Localmente, esse bloco pode ser analisado como em mostrado em 3.2.2.

Tomando a equação para  $\dot{\phi}(t)$  em (3.19), com  $A(\Theta(t)) = \bar{A}(\Theta(t))$ , têm-se

$$\begin{aligned} \dot{\phi}(t) &= (\bar{a}_{2,1,0})c_{\phi}^2 + (\bar{a}_{2,2,0} - \bar{a}_{1,1,0})c_{\phi}s_{\phi} - (\bar{a}_{1,2,0})s_{\phi}^2 \\ &\quad + \sum_{k=1}^p ((\bar{a}_{2,1,k})c_{\phi}^2 + (\bar{a}_{2,2,k} - \bar{a}_{1,1,k})c_{\phi}s_{\phi} - (\bar{a}_{1,2,k})s_{\phi}^2) \theta_k(t) \\ \Rightarrow 2\dot{\phi}(t) &= (\bar{a}_{2,1,0} + \bar{a}_{1,2,0})c_{2\phi} + (\bar{a}_{2,2,0} - \bar{a}_{1,1,0})s_{2\phi} \\ &\quad + (\bar{a}_{2,1,0} - \bar{a}_{1,2,0}) + \sum_{k=1}^p ((\bar{a}_{2,1,k} + \bar{a}_{1,2,k})c_{2\phi} \\ &\quad + (\bar{a}_{2,2,k} - \bar{a}_{1,1,k})s_{2\phi} + (\bar{a}_{2,1,k} - \bar{a}_{1,2,k})) \theta_k(t), \end{aligned} \quad (3.20)$$

onde  $c_{2\phi} \triangleq \cos(2\phi(t))$ ,  $s_{2\phi} \triangleq \sin(2\phi(t))$ , e  $\bar{a}_{i,j,k}$  é o  $(i,j)$ -ésimo elemento de  $\bar{A}_k$ . Fixando  $\Theta(t)$ , essa equação diferencial poderia ser resolvida para  $\phi(t)$  independentemente; a equação para  $\dot{\ell}(t)$  depende de  $\phi(t)$ . Observa-se que  $\dot{\phi}(t)$  é uma função afim de  $c_{2\phi}$  e de  $s_{2\phi}$ , com coeficientes e termo independente também afins em  $\Theta(t)$ : sejam esses termos  $\alpha(\Theta(t))$ ,  $\beta(\Theta(t))$  e  $\gamma(\Theta(t))$ , tais que

$$\dot{\phi}(t) = c_{2\phi}\alpha(\Theta(t)) + s_{2\phi}\beta(\Theta(t)) + \gamma(\Theta(t)). \quad (3.21)$$

Ao invés de resolver (3.21) explicitamente, sejam os problemas

$$\begin{aligned} \max_{\phi, \Theta} \quad & c_{2\phi}\alpha(\Theta) + s_{2\phi}\beta(\Theta) + \gamma(\Theta) \\ \text{s.a.} \quad & -1 \leq c_{2\phi}, s_{2\phi} \leq 1 \end{aligned} \quad (3.22)$$

e

$$\begin{aligned} \min_{\phi, \Theta} \quad & c_{2\phi}\alpha(\Theta) + s_{2\phi}\beta(\Theta) + \gamma(\Theta) \\ \text{s.a.} \quad & -1 \leq c_{2\phi}, s_{2\phi} \leq 1 . \end{aligned} \quad (3.23)$$

As soluções de ambos resultam num limite inferior e noutra superior para  $\dot{\phi}(t)$ , válidos para o modo correspondente a  $\bar{A}(\Theta(t))$  na resposta temporal do sistema com  $A(\Theta(t))$ . Tais limites podem ser usados para calcular uma folga extra a ser adicionada à largura de banda proposta na Seção 3.1.3, adicionando à já conservadora estimativa do raio espectral os efeitos da variação temporal dos parâmetros.

A idéia será explicada em termos do sistema LPV plano com  $\bar{A}(\Theta(t))$ , nas coordenadas polares  $(r(t), \phi(t))$ : o efeito de  $\Theta(t)$  na trajetória do estado pode ser entendido como um caso de modulação de ângulo, especificamente modulação de frequência visto que a derivada temporal do ângulo  $\phi(t)$  varia linearmente com os parâmetros [39]. Se  $B$  é a largura máxima para o sistema LPV, calculada assumindo parâmetros estáticos,  $\Theta(t) = \Theta^{\text{Opt}}$  que maximiza o raio espectral, seja

$$\Delta F \triangleq \max \left\{ \frac{\dot{\phi}}{2\pi} \right\} - \min \left\{ \frac{\dot{\phi}}{2\pi} \right\} , \quad (3.24)$$

com os limites de (3.22) e (3.23).

O valor acima é utilizado em conjunto com a regra de Carson [40], que propõe

$$B_T = 2(\Delta F + B) \quad (3.25)$$

como estimativa para a largura de banda ocupada pelo sinal modulado em frequência.

Para o sistema LPV plano com  $\bar{A}(\Theta(t))$ , sabe-se que  $B = \Im\{\bar{\lambda}\} = \omega$ . Para um sistema de ordem superior com  $A(\Theta(t))$  genérica, porém, adota-se  $B = \rho(A(\Theta^{\text{Opt}}))$ : isso equivaleria a adotar  $B = |\bar{\lambda}|$  para  $\bar{A}(\Theta(t))$ . Fora esse conservadorismo, pode-se argumentar que o limite calculado pela regra de

Carson a rigor só valeria para  $\Theta(t) = \Theta^{\text{opt}}$ . Pode ser interessante no futuro investigar uma aproximação menos conservadora, mas potencialmente mais custosa do ponto de vista computacional, que a apresentada na Seção 3.1.3 para determinar o ponto exato onde aplicar a aproximação de Carson.

### 3.4 Exemplo

A fim de ilustrar os conceitos e resultados obtidos neste capítulo, o sistema autônomo (3.10) com

$$A(\Theta(t)) = A_0 + \sum_{k=1}^3 A_k \theta_k(t), \quad (3.26)$$

$$A_0 = \begin{bmatrix} -1 & \frac{7}{4} \\ -\frac{1}{4} & -1 \end{bmatrix}, A_1 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -\frac{3}{2} \\ -\frac{3}{2} & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$$

será estudado. Adaptado de [9], este sistema é instável para uma determinada família de trajetórias dos parâmetros, a saber

$$\Theta^*(t) = \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \end{bmatrix} = \begin{bmatrix} \cos^2(\omega t) \\ \frac{1}{2} + \cos(\omega t) \sin(\omega t) \\ \sin^2(\omega t) \end{bmatrix}, \quad (3.27)$$

com

$$\omega \in \left[ 1 - \frac{1}{\sqrt{2}}, 1 + \frac{1}{\sqrt{2}} \right]. \quad (3.28)$$

A possibilidade de se detectar uma situação como esta justifica o esforço envolvido na estimação dos parâmetros.

#### 3.4.1 Análise Preliminar

Substituindo  $\Theta^*(t)$  definido em (3.27) na matriz  $A(\Theta(t))$  de (3.26), tem-se

$$\begin{aligned} A(\Theta^*(t)) = A(t) &= \begin{bmatrix} -1 + \frac{3}{2} \cos^2(\omega t) & 1 - \frac{3}{2} \cos(\omega t) \sin(\omega t) \\ -1 - \frac{3}{2} \cos(\omega t) \sin(\omega t) & -1 + \frac{3}{2} \sin^2(\omega t) \end{bmatrix} \\ &= \begin{bmatrix} -1 + 3 \cos(2\omega t) & 4 - 3 \sin(2\omega t) \\ -4 - 3 \sin(2\omega t) & -1 - 3 \cos(2\omega t) \end{bmatrix} / 4. \end{aligned} \quad (3.29)$$

Os autovalores da matriz acima podem ser calculados diretamente,

$$\det(\lambda I - A(t)) = 0 \quad \Rightarrow \quad \lambda_{1,2} = \frac{-1 \pm j\sqrt{7}}{4},$$

e como  $\Re\{\lambda_{1,2}\} < 0$ , vê-se que  $A(t)$  é Hurwitz. Uma análise que desprezasse os efeitos da variação no tempo poderia levar a crer que o sistema LPV é estável. A razão da instabilidade com  $\Theta^*(t)$  será estudada em breve, mas a compatibilidade entre a análise acima e os resultados obtidos na Seção 3.2.1 será antes verificada.

### 3.4.2 Autovalores

Conforme já mencionado, os autovalores de  $A(\Theta(t))$  são fixos para  $\Theta(t) = \Theta^*(t)$ ; isso não é verdade no caso geral, portanto é interessante verificar como as características instantâneas dos autovalores podem mudar de acordo com a variação do determinante de  $A(\Theta(t)) \in \mathbb{R}^{2 \times 2}$  no tempo, que é uma função quadrática dos parâmetros.

De (3.12), tem-se que

$$M_{\Delta} = \begin{bmatrix} -\frac{7}{4} & 0 & -\frac{21}{2} & 0 & 0 \\ 0 & \frac{9}{4} & 0 & -\frac{9}{4} & 0 \\ \frac{3}{2} & 0 & 9 & 0 & 0 \\ 0 & -\frac{9}{4} & 0 & \frac{9}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Os autovalores da matriz acima são  $\lambda = \{\frac{29}{4}, \frac{9}{2}, 0, 0, 0\}$ , portanto  $M_{\Delta}$  é positiva semidefinida. Como nela estão incluídos os pseudo-parâmetros  $\theta_0(t) = 1$  e  $\theta_4(t) = 0$ , esse último com o coeficiente  $A_4 = 0$ , faz-se necessário descontar os efeitos desses elementos. O particionamento abaixo, como em (3.13), permite fazer isso:

$$\begin{aligned} \Delta(\Theta^+(t)) &= \Theta^+(t)^T M_{\Delta} \Theta^+(t) \\ &= \begin{bmatrix} 1 & \bar{\Theta}(t)^T & 0 \end{bmatrix} \begin{bmatrix} m_{1,1} & \bar{m}_{1,k}^T & m_{1,5} \\ \bar{m}_{k,1} & \bar{M}_{\Delta} & \bar{m}_{k,5} \\ m_{5,1} & \bar{m}_{5,k}^T & m_{5,5} \end{bmatrix} \begin{bmatrix} 1 \\ \bar{\Theta}(t) \\ 0 \end{bmatrix} \\ &= m_{1,1} + \bar{\Theta}(t)^T \bar{m}_{k,1} + \bar{m}_{1,k}^T \bar{\Theta}(t) + \bar{\Theta}(t)^T \bar{M}_{\Delta} \bar{\Theta}(t), \end{aligned}$$

com

$$\begin{aligned}
 m_{1,1} &= -\frac{7}{4} \\
 \bar{\Theta}(t) &= [\theta_1(t) \quad \theta_2(t) \quad \theta_3(t)]^T, \\
 \bar{m}_{k,1} &= \left[0 \quad \frac{3}{2} \quad 0\right]^T, \\
 \bar{m}_{1,k} &= \left[0 \quad -\frac{21}{2} \quad 0\right], \\
 \bar{M}_\Delta &= \begin{bmatrix} \frac{9}{4} & 0 & -\frac{9}{4} \\ 0 & 9 & 0 \\ -\frac{9}{4} & 0 & \frac{9}{4} \end{bmatrix}.
 \end{aligned}$$

A matriz  $\bar{M}_\Delta$  acima é, novamente, positiva semidefinida ( $\lambda = \{9, \frac{9}{2}, 0\}$ ). No entanto, se os parâmetros forem assumidos normalizados,  $0 \preceq \bar{\Theta}(t) \preceq 1$ , então  $\Delta(\Theta(t)) < 0$  para qualquer trajetória  $\Theta(t)$ , e os autovalores de  $A(\Theta(t))$  serão sempre conjugados complexos.

### 3.4.3 Coordenadas Polares

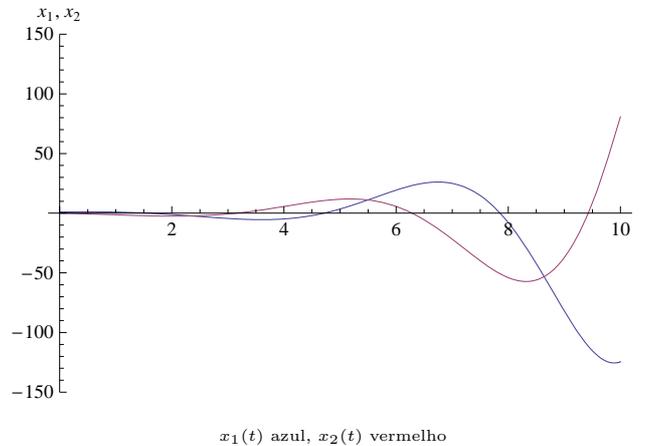


Figura 3.1: Resposta do sistema autônomo,  $\omega = 1$ , à condição inicial  $x(0) = [1, 0]^T$ .

As Figuras 3.1 e 3.2 ilustram a resposta do sistema autônomo com  $A(t)$  de (3.29),  $\omega = 1$ , à condição inicial  $x(0) = [1, 0]^T$ . A instabilidade pode ser facilmente entendida se o sistema for reescrito em coordenadas polares:

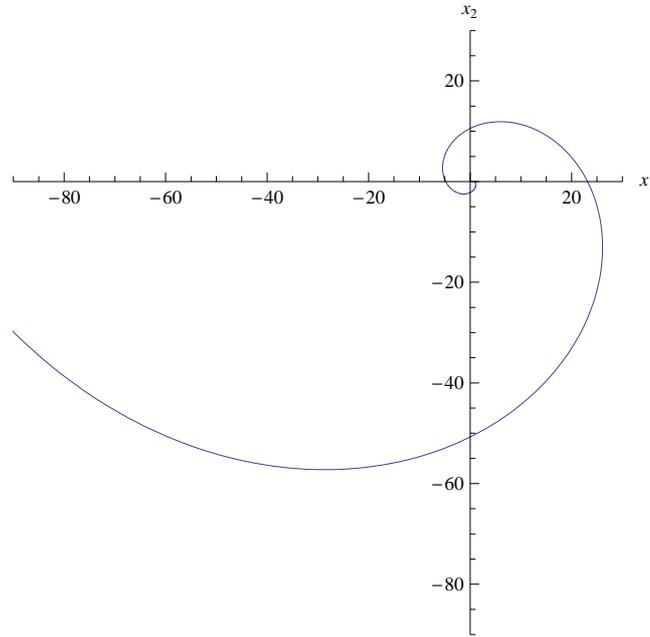


Figura 3.2: Resposta do sistema autônomo,  $\omega = 1$ , à condição inicial  $x(0) = [1, 0]^T$ , no espaço de estados.

aplicando (3.19) a (3.26), as equações diferenciais tornam-se

$$\begin{aligned} \dot{v}_{\ell\phi}(t) &= A_{\phi}(t)\Theta(t) \\ \Rightarrow \begin{bmatrix} \dot{\ell}(t) \\ \dot{\phi}(t) \end{bmatrix} &= \begin{bmatrix} \frac{3}{2} \cos(\phi(t)) \sin(\phi(t)) - \cos^2(\phi(t)) - \sin^2(\phi(t)) & \dots \\ -\frac{1}{4} \cos^2(\phi(t)) - \frac{7}{4} \sin^2(\phi(t)) & \dots \\ \dots & \frac{3}{2} \cos^2(\phi(t)) & -3 \cos(\phi(t)) \sin(\phi(t)) & \dots \\ \dots & -\frac{3}{2} \cos(\phi(t)) \sin(\phi(t)) & -\frac{3}{2} \cos^2(\phi(t)) + \frac{3}{2} \sin^2(\phi(t)) & \dots \\ \dots & \frac{3}{2} \sin^2(\phi(t)) & & \\ \dots & \frac{3}{2} \cos(\phi(t)) \sin(\phi(t)) & & \end{bmatrix} \begin{bmatrix} 1 \\ \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \end{bmatrix}, \end{aligned}$$

que podem ser rearrajadas como

$$\begin{aligned} \begin{bmatrix} \dot{\ell}(t) \\ \dot{\phi}(t) \end{bmatrix} &= \begin{bmatrix} \frac{3}{4} \theta_1(t) - \frac{3}{4} \theta_3(t) & -\frac{3}{2} \theta_2(t) + \frac{3}{4} & \dots \\ -\frac{3}{2} \theta_2(t) + \frac{3}{4} & -\frac{3}{4} \theta_1(t) + \frac{3}{4} \theta_3(t) & \dots \\ \dots & \frac{3}{4} \theta_1(t) + \frac{3}{4} \theta_3(t) - 1 & \\ \dots & & -1 & \end{bmatrix} \begin{bmatrix} \cos(2\phi(t)) \\ \sin(2\phi(t)) \\ 1 \end{bmatrix}. \end{aligned} \quad (3.30)$$

A segunda equação acima corresponde a (3.21), e  $\alpha(\Theta)$ ,  $\beta(\Theta)$  e  $\gamma(\Theta)$  podem ser lidas na segunda linha da matriz. Isso permitirá o cálculo de (3.22) e de (3.23), adiante.

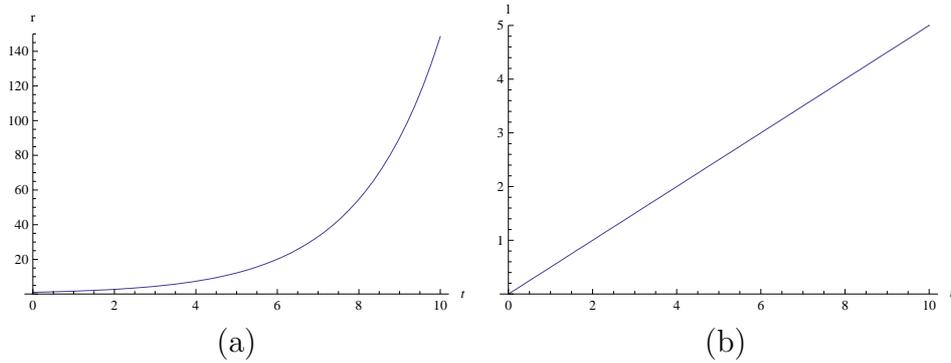


Figura 3.3: Magnitude  $r(t)$ , com crescimento exponencial, em (a), e o respectivo logaritmo  $\ell(t)$ , linear, em (b).

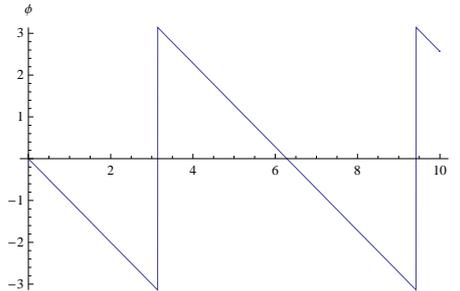


Figura 3.4: Fase  $\phi(t) = -\omega t$ , linear, mapeada entre  $-\pi$  e  $\pi$  ( $\omega = 1$ ).

Se os parâmetros  $\Theta^*(t)$  forem substituídos em (3.30) (lembrando que este é um exemplo teórico, na prática os parâmetros seriam desconhecidos), têm-se

$$\begin{aligned} \begin{bmatrix} \dot{\ell}(t) \\ \dot{\phi}(t) \end{bmatrix} &= \begin{bmatrix} \frac{3}{4} \cos(2\omega t) & \frac{3}{4} \sin(2\omega t) & -\frac{1}{4} \\ \frac{3}{4} \sin(2\omega t) & -\frac{3}{4} \cos(2\omega t) & -1 \end{bmatrix} \begin{bmatrix} \cos(2\phi(t)) \\ \sin(2\phi(t)) \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{4} + \frac{3}{4} \cos(2(\omega t + \phi(t))) \\ -1 - \frac{3}{4} \sin(2(\omega t + \phi(t))) \end{bmatrix}. \end{aligned}$$

É possível encontrar uma solução  $\phi(t)$  que satisfaz  $\dot{\phi}(t)$  acima por inspeção. No caso geral, se os parâmetros  $\theta_k(t)$  forem conhecidos será sempre possível

calcular uma solução numérica para essa equação. Neste exemplo, com as condições iniciais

$$x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \ell(0) \\ \phi(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} ,$$

a solução é

$$\begin{aligned} \phi(t) = -\omega t &\Rightarrow \dot{\phi}(t) = -1 - \frac{3}{4} \sin(2(\omega t + (-\omega t))) = -1 \\ &\Rightarrow \dot{\ell}(t) = -\frac{1}{4} + \frac{3}{4} \cos(2(\omega t + (-\omega t))) = 0.5 \\ &\Rightarrow r(t) = e^{0.5t} > 0 . \end{aligned}$$

O resultado acima, com  $r(t)$  crescente, explica a instabilidade do sistema. As Figuras 3.3 e 3.4 ilustram  $r(t)$  e  $\phi(t)$ , para  $\omega = 1$ , respectivamente.

### 3.4.4 Raios Espectrais

Passando agora à questão do raio espectral, os limites para seus valores mínimo e máximo serão calculados conforme descrito nas seções 3.1.1 e 3.1.2. Assumindo que os parâmetros sejam normalizados, uma análise de factibilidade é feita *a priori*: das  $2^4 = 16$  possíveis combinações de sinais dos quatro elementos  $a_{i,j}$ ,  $i, j \in \{1, 2\}$  da matriz  $A(\Theta(t))$ , somente quatro combinações são viáveis. Isso é ilustrado na Tabela 3.1.

Tabela 3.1: Combinações factíveis de sinais dos elementos da matriz.

Elementos da Matriz e Sinais				
Região	$a_{1,1} = -1 + \frac{3}{2}\theta_1$	$a_{1,2} = \frac{7}{4} - \frac{3}{2}\theta_2$	$a_{2,1} = -\frac{1}{4} - \frac{3}{2}\theta_2$	$a_{2,2} = -1 + \frac{3}{2}\theta_3$
<b>i</b>	$\geq 0$	$\geq 0$	$\leq 0$	$\geq 0$
<b>ii</b>	$\leq 0$	$\geq 0$	$\leq 0$	$\geq 0$
<b>iii</b>	$\geq 0$	$\geq 0$	$\leq 0$	$\leq 0$
<b>iv</b>	$\leq 0$	$\geq 0$	$\leq 0$	$\leq 0$

Juntamente com os limites  $\theta_k \in [0, 1]$ ,  $k = 1, 2, 3$ , as regiões **i** – **iv** da Tabela 3.1 podem ser expressas<sup>7</sup> em termos de  $\Theta = [\theta_1 \ \theta_2 \ \theta_3]^T$ :

$$\begin{aligned} \Theta_{\mathbf{i}} &= \left[ \left[ \frac{2}{3}, 1 \right] \ [0, 1] \ \left[ \frac{2}{3}, 1 \right] \right]^T & \Theta_{\mathbf{ii}} &= \left[ \left[ 0, \frac{2}{3} \right] \ [0, 1] \ \left[ \frac{2}{3}, 1 \right] \right]^T \\ \Theta_{\mathbf{iii}} &= \left[ \left[ \frac{2}{3}, 1 \right] \ [0, 1] \ \left[ 0, \frac{2}{3} \right] \right]^T & \Theta_{\mathbf{iv}} &= \left[ \left[ 0, \frac{2}{3} \right] \ [0, 1] \ \left[ 0, \frac{2}{3} \right] \right]^T \end{aligned}$$

<sup>7</sup>Notação:  $[a, b]$  = intervalo fechado  $a \leq \theta_n \leq b$ .

Não é possível afirmar no caso geral que o particionamento seja sempre simples como acima, portanto essas simplificações não foram usadas nos cálculos a seguir.

Os problemas das seções 3.1.1 e 3.1.2 podem ser escritos como:

$$\begin{aligned}
 & \min_{z, \theta_k} \quad z & (3.31) \\
 \text{s. a.} \quad & p_1 z \geq \pm(-1 + \frac{3}{2}\theta_1)p_1 + (\frac{7}{4} - \frac{3}{2}\theta_2)p_2 \\
 & p_2 z \geq -(-\frac{1}{4} - \frac{3}{2}\theta_2)p_1 \pm (-1 + \frac{3}{2}\theta_3)p_2 \\
 & 0 \leq \pm(-1 + \frac{3}{2}\theta_1) \\
 & 0 \leq (\frac{7}{4} - \frac{3}{2}\theta_2) \\
 & 0 \leq -(-\frac{1}{4} - \frac{3}{2}\theta_2) \\
 & 0 \leq \pm(-1 + \frac{3}{2}\theta_3) \\
 & 0 \leq \theta_k \leq 1, \quad k = 1, 2, 3,
 \end{aligned}$$

e

$$\begin{aligned}
 & \max_{z, \theta_k} \quad z & (3.32) \\
 \text{s. a.} \quad & p_1 z \leq \pm(-1 + \frac{3}{2}\theta_1)p_1 + (\frac{7}{4} - \frac{3}{2}\theta_2)p_2 \\
 & p_2 z \leq -(-\frac{1}{4} - \frac{3}{2}\theta_2)p_1 \pm (-1 + \frac{3}{2}\theta_3)p_2 \\
 & 0 \leq \pm(-1 + \frac{3}{2}\theta_1) \\
 & 0 \leq (\frac{7}{4} - \frac{3}{2}\theta_2) \\
 & 0 \leq -(-\frac{1}{4} - \frac{3}{2}\theta_2) \\
 & 0 \leq \pm(-1 + \frac{3}{2}\theta_3) \\
 & 0 \leq \theta_k \leq 1, \quad k = 1, 2, 3.
 \end{aligned}$$

Os sinais  $\pm$  acima dependem da região sendo considerada. Em cada uma delas ambos os problemas foram resolvidos pelo método simplex, mantendo  $p_1$  e  $p_2$  variáveis simbólicas estritamente positivas. Durante o cálculo hipóteses extras, como por exemplo  $p_1 \geq p_2\sqrt{7}$  ou vice-versa, precisaram ser introduzidas em alguns passos para garantir que os *tableaux* intermediários permanecessem factíveis: a seleção cuidadosa de *pivots* é capaz de evitar essa

necessidade até certo ponto, mas bifurcações dos cálculos devido a esse tipo de consideração são inevitáveis, visto que muitas soluções não são únicas. De qualquer forma os resultados do método simplex são ainda funções dos  $p$ 's, portanto o próximo passo é resolver problemas de maximização ou minimização não-lineares, mais simples. Por exemplo, os resultados da maximização na região **iii** foram<sup>8</sup>:

$$\begin{aligned}
 z &= \frac{p_1^2 + 4p_1p_2 + 2p_2^2}{2(p_1^2 + p_2^2)}, \\
 s_{1,1} &= \frac{1}{2}, & s_{\ell 1} &= \frac{p_1 s_{\ell 2}}{p_2}, & s_{\ell 2} &= 0, \\
 s_{1,2} &= \frac{p_1(4p_1 + p_2)}{2(p_1^2 + p_2^2)}, & \theta_1 &= 1, & s_{\theta 1} &= 0, \\
 s_{2,1} &= \frac{p_2(-p_1 + 4p_2)}{2(p_1^2 + p_2^2)}, & \theta_2 &= -\frac{p_1^2 + 2p_1p_2 - 7p_2^2}{6(p_1^2 + p_2^2)}, & s_{\theta 2} &= -\frac{1}{6} + \frac{p_1(4p_1 + p_2)}{(3(p_1^2 + p_2^2))}, \\
 s_{2,2} &= 1, & \theta_3 &= 0, & s_{\theta 3} &= 1.
 \end{aligned}$$

Aqui, maximizando  $z$  em função de  $p_1$  e  $p_2$  mais as restrições  $s_{i,j} \geq 0$  e  $0 \leq \theta_2 \leq 1$  resulta em  $z = \frac{17+5\sqrt{17}}{17+\sqrt{17}}$  e  $\Theta = [1 \ \frac{1}{2} \ 0]$ , para  $[p_1 \ p_2] = [1 \ \frac{1+\sqrt{17}}{4}]$ . O real valor do raio espectral em  $\Theta = [1 \ \frac{1}{2} \ 0]$  é  $\rho = \frac{1}{\sqrt{2}}$ , e uma análise rápida mostra que a solução alcançada não é de fato ótima, pois o valor da função objetivo pode ser aumentado reduzindo  $\theta_1$  ou aumentando  $\theta_3$ . Por causa disso os cálculos são repetidos para os novos valores de  $p_1$  e  $p_2$ , desta vez alcançando o resultado correto  $z = 1$  para  $\Theta = [\frac{2}{3} \ \frac{1}{2} \ [0, \frac{2}{3}]]$  ou  $[[\frac{2}{3}, 1] \ \frac{1}{2} \ \frac{2}{3}]$ .

O processo acima foi aplicado às quatro regiões factíveis, minimização e maximização. Em todos os casos as soluções finais foram identificadas em pontos extremos – vértices, cantos ou faces – das regiões correspondentes, politopos em função de  $\Theta$ . Neste exemplo em particular, ambas as desigualdades de linha foram verificadas ativas em cada caso, portanto não houve necessidade de repetir cálculos com uma delas omitida. A Tabela 3.2 apresenta os resultados<sup>9</sup>. Os dois melhores limites mostraram-se sem folgas, isto é,  $\rho_{\min} = \frac{1+\sqrt{2}}{4}$  e  $\rho_{\max} = \sqrt{2}$ . A partição simples do cubo unitário  $\Delta\Theta$  pelos planos  $\theta_1 = \frac{2}{3}$  e  $\theta_3 = \frac{2}{3}$  permite visualizar facilmente as soluções: as

<sup>8</sup>Notação: as variáveis  $s_{i,j}$ 's são folgas para os  $a_{i,j}$ 's,  $s_{\theta k}$ 's para os  $\theta_k$ 's, e  $s_{\ell 1}$ ,  $s_{\ell 2}$  para as desigualdades de linhas 1 e 2

<sup>9</sup>Notação:  $\{a, b\}$  = conjunto de valores,  $\theta_n = a$  ou  $\theta_n = b$ ;  $[a, b]$  = intervalo fechado  $a \leq \theta_n \leq b$ , como antes.

Figuras 3.5 (a) e (b) são, respectivamente, o subnível  $\rho(\Theta) \leq \frac{\sqrt{7}}{4}$  e o supernível  $\rho(\Theta) \geq 1$ . A família de subníveis de 3.5 (a) pode ser “encolhida” fazendo  $\frac{\sqrt{7}}{4} \rightarrow \frac{1+\sqrt{2}}{4}$ , revelando a localização dos mínimos que conferem com os resultados obtidos. O mesmo pode ser feito com 3.5 (b), com a superfície correspondente a  $\sqrt{2}$  colapsando no ponto  $\Theta = [0 \ \frac{1}{2} \ 0]^T$ , conforme calculado.

Tabela 3.2: Limites máximos e mínimos dos raios espectrais, para cada combinação factível dos sinais dos elementos da matriz.

Região	Mínimo	Máximo
i	$\frac{\sqrt{7}}{4}$ para $\Theta = \begin{cases} \left[ \frac{2}{3} \ \{0, 1\} \ \left[ \frac{2}{3}, 1 \right]^T \right. \\ \left. \left[ \left[ \frac{2}{3}, 1 \right] \ \{0, 1\} \ \frac{2}{3} \right]^T \right. \end{cases}$ ou	$\frac{\sqrt{5}}{2}$ para $\Theta = [1 \ \frac{1}{2} \ 1]^T$
ii	$\frac{1+\sqrt{2}}{4}$ para $\Theta = [0 \ \{0, 1\} \ 1]^T$	1 para $\Theta = \begin{cases} \left[ \frac{2}{3} \ \frac{1}{2} \ \left[ \frac{2}{3}, 1 \right]^T \right. \\ \left. \left[ \left[ 0, \frac{2}{3} \right] \ \frac{1}{2} \ \frac{2}{3} \right]^T \right. \end{cases}$ ou
iii	$\frac{1+\sqrt{2}}{4}$ para $\Theta = [1 \ \{0, 1\} \ 0]^T$	1 para $\Theta = \begin{cases} \left[ \frac{2}{3} \ \frac{1}{2} \ \left[ 0, \frac{2}{3} \right]^T \right. \\ \left. \left[ \left[ \frac{2}{3}, 1 \right] \ \frac{1}{2} \ \frac{2}{3} \right]^T \right. \end{cases}$ ou
iv	$\frac{\sqrt{7}}{4}$ para $\Theta = \begin{cases} \left[ \frac{2}{3} \ \{0, 1\} \ \left[ 0, \frac{2}{3} \right]^T \right. \\ \left. \left[ \left[ 0, \frac{2}{3} \right] \ \{0, 1\} \ \frac{2}{3} \right]^T \right. \end{cases}$ ou	$\sqrt{2}$ para $\Theta = [0 \ \frac{1}{2} \ 0]^T$

### 3.4.5 Amostragem

O raio espectral máximo  $\rho_{\max} = \sqrt{2}$  é alcançado para  $\Theta^{\text{Opt}} = [0 \ \frac{1}{2} \ 0]^T$ . As soluções de (3.22) e de (3.23) são

$$\dot{\phi}_{\max} = \frac{3\sqrt{2}}{4} - 1, \quad \begin{cases} \Theta = [0 \ 0 \ 1]^T, \\ \phi = \frac{\pi}{8} \end{cases},$$

$$\dot{\phi}_{\min} = -\frac{3\sqrt{2}}{4} - 1, \quad \begin{cases} \Theta = [1 \ 0 \ 0]^T, \\ \phi = \frac{\pi}{8} \end{cases},$$

(note que, dada a periodicidade de  $\sin()$  e  $\cos()$ , os valores de  $\Theta$  e de  $\phi$  acima não são únicos.) Assim, (3.24) e (3.25), essa última com  $B = 2\rho_{\max} = 2\sqrt{2}$ , resultam em:

$$\Delta F = \frac{\left(\frac{3\sqrt{2}}{4} - 1\right) - \left(-\frac{3\sqrt{2}}{4} - 1\right)}{2\pi} = \frac{3\sqrt{2}}{4\pi} \approx 0.3376$$

$$\therefore B_T = 2\left(\frac{3\sqrt{2}}{4\pi} + 2\sqrt{2}\right) \approx 6.3321 ,$$

$$f_S = \frac{B_T}{2\pi} \approx 1.008 .$$

Visto que  $\omega = 1 + \frac{1}{\sqrt{2}}$  corresponde à frequência mais alta encontrada na resposta do sistema autônomo, a saber o ciclo limite com período  $T = \frac{2\pi}{\omega} \approx 3.681$ , demonstra-se que a frequência de amostragem  $f_S$  é suficiente (mais que duas amostras por período).

Embora o resultado acima seja final, os cálculos de (3.22) e (3.23) podem ser repetidos com  $\Theta(t) = \Theta^*(t)$ . Os cálculos a seguir são puramente acadêmicos, visto que assumem conhecimento dos parâmetros. A idéia é meramente demonstrar que o conhecimento extra pode levar a um resultado menos conservador. De fato,

$$\dot{\phi}_{\max} = \max -1 - \frac{3}{4} \sin(2(\omega t + \phi(t))) = -\frac{1}{4} \text{ e}$$

$$\dot{\phi}_{\min} = \min -1 - \frac{3}{4} \sin(2(\omega t + \phi(t))) = -\frac{7}{4} ,$$

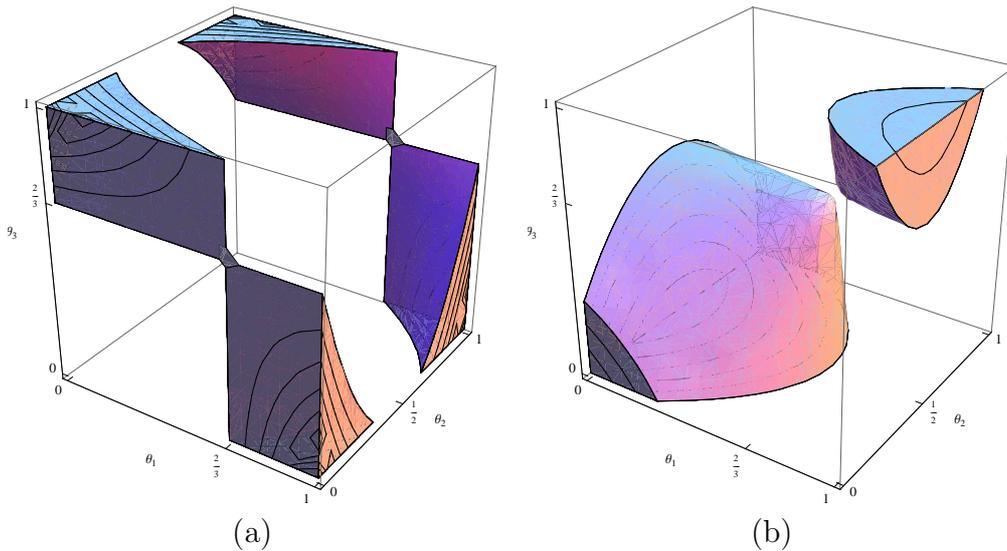


Figura 3.5: Subníveis  $\rho(A(\Theta)) \leq \frac{7}{4}$  (a), e superníveis  $\rho(A(\Theta)) \geq 1$  (b).

resultam em  $\Delta F = \frac{3}{4\pi}$ ,  $B_T = \frac{3}{2\pi} + 4\sqrt{2} \approx 6.134$  e  $f_S = \frac{B_T}{2\pi} \approx 0.9763$ , ainda suficiente. Por outro lado, (3.25) com  $\Delta F = \frac{3\sqrt{2}}{4\pi}$  e  $B = 2\mathfrak{S}\{\lambda(A(\Theta^*(t)))\} = \frac{\sqrt{7}}{2}$  resulta em  $B_T \approx 3.321$ ,  $f_S \approx 0.5286 < 0.5434 \approx 2\frac{\omega}{2\pi}$ , insuficiente (por pouco) para amostrar o sistema no ciclo limite. Assim, embora introduza conservadorismo, justifica-se neste caso usar o dobro do raio espectral como  $B$  em (3.25).

## Capítulo 4

# Estimação de Parâmetros via Otimização

Assim como os problemas de identificação de sistemas e de estimação de estados, a estimação de parâmetros de sistemas LPV pode ser interpretada como a minimização de uma medida de erro, neste caso a diferença entre a saída do sistema observado e o aquilo que o modelo, juntamente com uma estimativa preliminar dos valores dos parâmetros, predizem. A estimativa é então progressivamente refinada até atingir um valor ótimo, que minimiza a medida de erro adotada. Abordagens estatísticas, como a da informação estocástica em estado estacionário, PRBS ou o filtro de Kalman adaptado conforme descrito na Seção 2.3, fazem uso também de propriedades como propagação de covariâncias dos estados e parâmetros, bem como dos ruídos de processo e de medição. Esses últimos servem ainda para excitar o modelo e como artifício matemático para explicar e processar o erro.

Este capítulo, por outro lado, enfoca apenas a minimização do erro: assume-se que o estado seja mensurável, mas não se conhecem sua covariância inicial ou a covariância do ruído de medição. Assume-se também que o sistema seja autônomo, *i.e.* sem entradas exógenas, embora a inclusão dessas mesmas, quando conhecidas, seja trivial. As incertezas nas entradas ou no modelo são absorvidas pelo algoritmo de otimização, e não pelo ruído de processo. A abordagem aproveita a estrutura LPV para obter um problema simples e convexo mas, mesmo com a amostragem apropriada, em alguns casos o resultado é indeterminado – felizmente tais situações são detectáveis, *cf.* crescimento das covariâncias nas abordagens estatísticas.

Quanto à teoria de otimização, citam-se as referências [41] e [42]. Em termos de implementação, o pacote *Mathematica* utilizado no capítulo anterior para cálculos simbólicos foi também usado aqui para a parte numérica.

O *Optimization Toolbox* do MATLAB teria sido igualmente adequado.

## 4.1 Função Objetivo

Dado o sistema dinâmico (1.5), suponham-se conhecidos os valores do estado e do vetor de parâmetros num determinado instante  $t = t_0$ . Seja ainda  $\Delta t$  um valor menor que o intervalo mínimo de amostragem para o sistema LPV com a matriz  $A(\Theta(t))$ , conforme calculado no capítulo anterior. Se os parâmetros forem considerados constantes  $\Theta(t) \approx \Theta(t_0)$ ,  $\forall t \in [t_0, t_0 + \Delta t)$ , uma aproximação LTI poderia ser usada para calcular o valor do estado ao final desse intervalo. Tal valor seria

$$x(t_0 + \Delta t) = e^{\Delta t A(\Theta(t_0))} x(t_0) + \int_{t_0}^{t_0 + \Delta t} e^{(t_0 + \Delta t - \tau) A(\Theta(t_0))} B(\Theta(t_0)) u(\tau) d\tau .$$

Para simplificar o desenvolvimento, a entrada exógena será considerada nula: se  $u(t) = 0$  (sistema autônomo), elimina-se o termo com a integral. Uma segunda simplificação é escrever a exponencial  $\exp(\Delta t A(\Theta(t_0))) \approx I + \Delta t A(\Theta(t_0))$ , como uma aproximação de primeira ordem. Assim,

$$\tilde{x}(t_0 + \Delta t) \approx (I + (\Delta t) A(\Theta(t_0))) x(t_0) . \quad (4.1)$$

Comparando (4.1) acima com o valor real do estado,  $x(t_0 + \Delta t)$ , assumido mensurável, a diferença  $e(t_0 + \Delta t) = x(t_0 + \Delta t) - \tilde{x}(t_0 + \Delta t)$  é, a menos de ruídos ou do erro introduzido pela integração aproximada, função de  $\Theta(t_0)$ . A partir deste ponto, seja  $t = t_0 + \Delta t$ ; sejam também  $e(t)$  a diferença acima, e  $f(t)$  o quadrado de sua norma:

$$\begin{aligned} e(t) &\triangleq x(t) - \tilde{x}(t) \\ &= x(t) - x(t - \Delta t) - (\Delta t) A(\Theta(t - \Delta t)) x(t - \Delta t) \\ &= \Delta x(t) - (\Delta t) A(\Theta(t - \Delta t)) x(t - \Delta t) , \end{aligned} \quad (4.2)$$

com  $\Delta x(t) \triangleq x(t) - x(t - \Delta t)$  implícito acima, e

$$\begin{aligned} f(t) &\triangleq \|e(t)\|^2 = e(t)^T e(t) \\ &= \|\Delta x(t)\|^2 - (\Delta t) [(\Delta x(t))^T A(\Theta(t - \Delta t)) x(t - \Delta t) + \\ &\quad (x(t - \Delta t))^T A(\Theta(t - \Delta t)) (\Delta x(t))] + \\ &\quad (\Delta t)^2 (A(\Theta(t - \Delta t)) x(t - \Delta t))^T (A(\Theta(t - \Delta t)) x(t - \Delta t)) . \end{aligned}$$

Em outra simplificação notacional, a dependência do tempo será deixada implícita:  $A(\Theta(t - \Delta t))$  passa a ser escrita como  $A(\Theta)$ ,  $f(t) = f(t, \Theta(t))$  passa a ser  $f(\Theta)$ , e  $x = x(t - \Delta t)$ . Dessa forma,

$$f(\Theta) = \|\Delta x\|^2 - (\Delta t)((\Delta x)^T A(\Theta)x + x^T A(\Theta)(\Delta x)) + (\Delta t)^2(A(\Theta)x)^T(A(\Theta)x) . \quad (4.3)$$

Substituindo  $A(\Theta)$  dada por (1.12) em (4.3), verifica-se que  $f(\Theta)$  é quadrática em  $\Theta$ . O gradiente  $\nabla f(\Theta)$  e o Laplaciano  $\nabla^2 f(\Theta) = \nabla \cdot \nabla f(\Theta)$  em relação a  $\Theta$  podem ser calculados introduzindo (1.6) ou (1.12) em (4.3) e derivando. Os termos  $A_k^T A_k$  e  $A_k^T A_\ell + A_\ell^T A_k$  ( $k, \ell = 0, \dots, p$ ,  $\ell \neq k$ ) que aparecem nas expressões expandidas são positivos definidos ou semi-definidos, portanto  $f(\Theta)$  é convexa.

## 4.2 Implementação

Propõe-se aqui que  $\Theta^{\text{Opt}} = \arg \min_{\Theta} f(\Theta)$  seja a estimativa de  $\Theta(t - \Delta t)$ . Tal cálculo pode ser efetuado *on-line*, com um atraso  $\Delta t$ . Não há, em princípio, necessidade de incluir restrições extras; assumindo porém que os parâmetros sejam normalizados, é recomendável que isso seja feito. Conforme aludido no capítulo anterior, em determinados instantes a influência de alguns parâmetros pode desaparecer da resposta, *e.g.* elementos de  $M_\phi$  e de  $v_\phi$  (equação (3.17)) tendendo a zero para determinados valores de  $\phi(t)$ , no estudo em coordenadas polares (Seção 3.2.2). Isso, mais a possibilidade de o valor inicial para o algoritmo de otimização estar fora da região factível (no caso de parâmetros com escalas conhecidas) tornam recomendável o uso de restrições. O exemplo a seguir ilustra algumas dessas questões.

### 4.2.1 Exemplo

Continuando o exemplo do capítulo anterior, Seção 3.4, os parâmetros são dados por  $\Theta^*(t)$  em (3.27). Seja  $\omega = 1 + \frac{1}{\sqrt{2}}$ , valor máximo de (3.28). Nessas condições, o estado inicial  $x(0) = [1 \ 0]^T$  resulta no ciclo-limite ilustrado pela Figura 4.1, que também mostra a trajetória dos componentes de  $\Theta^*(t)$  a serem estimados, tanto no tempo quanto no espaço de parâmetros.

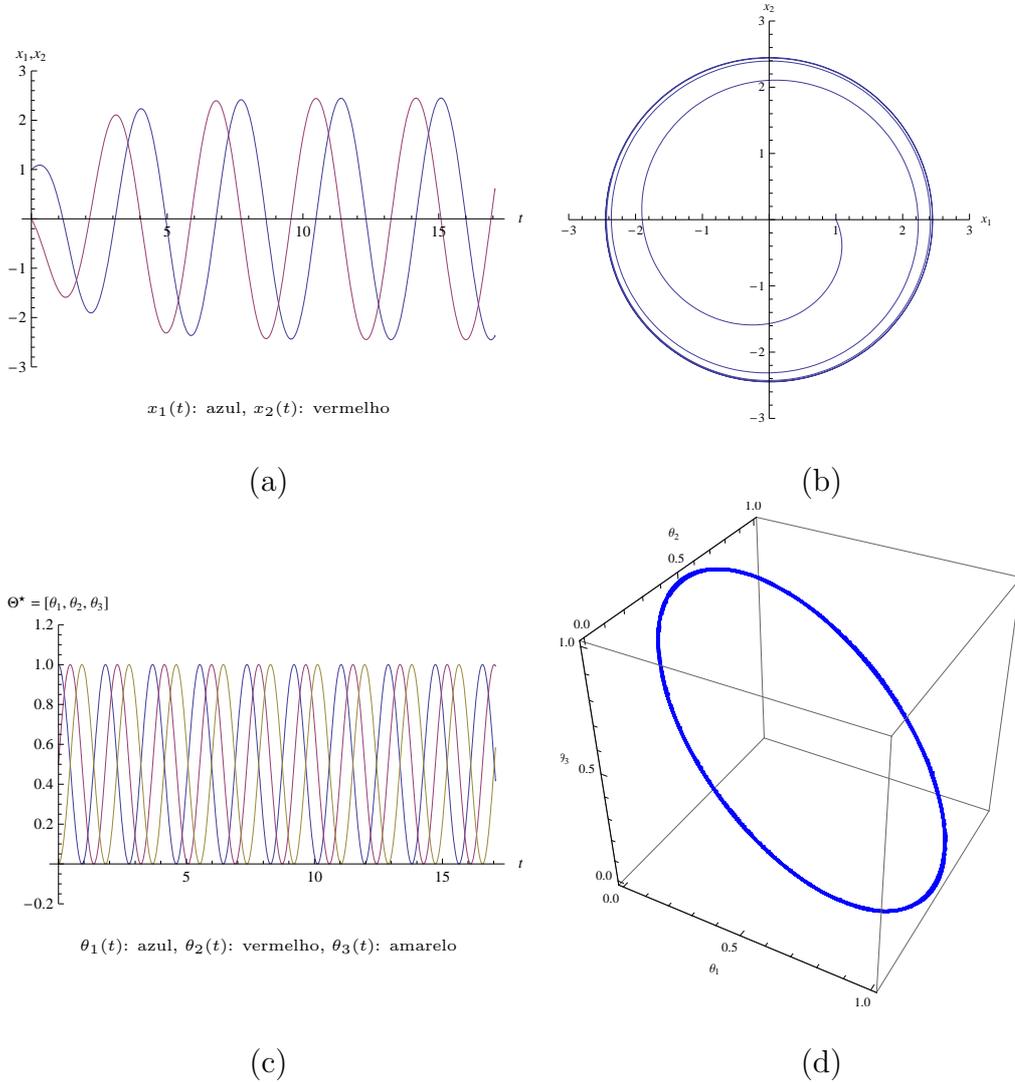


Figura 4.1: Resposta do sistema autônomo, exemplo da Seção 3.4, com parâmetros  $\Theta(t) = \Theta^*(\omega t) = [\cos^2(\omega t), \frac{1}{2} + \cos(\omega t) \sin(\omega t), \sin^2(\omega t)]^T$  e  $\omega = 1 + \frac{1}{\sqrt{2}}$ , para a condição inicial  $x(0) = [1, 0]^T$ , em coordenadas cartesianas (a) e polares (b). Os parâmetros são mostrados em (c) e (d).

Minimizando  $f$ , os parâmetros foram estimados desde  $t = 0$  até  $t = 17$ , em intervalos de duração  $\Delta t = 0.17$ . O problema de otimização  $\min f(\Theta)$  foi resolvido independentemente a cada instante, sem restrições. A estimativa inicial<sup>1</sup> a cada ponto foi  $\Theta_0 = [0.5, 0.5, 0.5]^T$ . A Figura 4.2 mostra os resultados.

<sup>1</sup>Estimativa inicial para o problema de otimização, não para a estimativa dos parâmetros. De fato, no primeiro ponto se espera que a estimativa convirja de  $\Theta_0 = [0.5, 0.5, 0.5]^T \rightarrow \Theta^*(0) = [1, 0.5, 0]^T$ , e assim por diante.

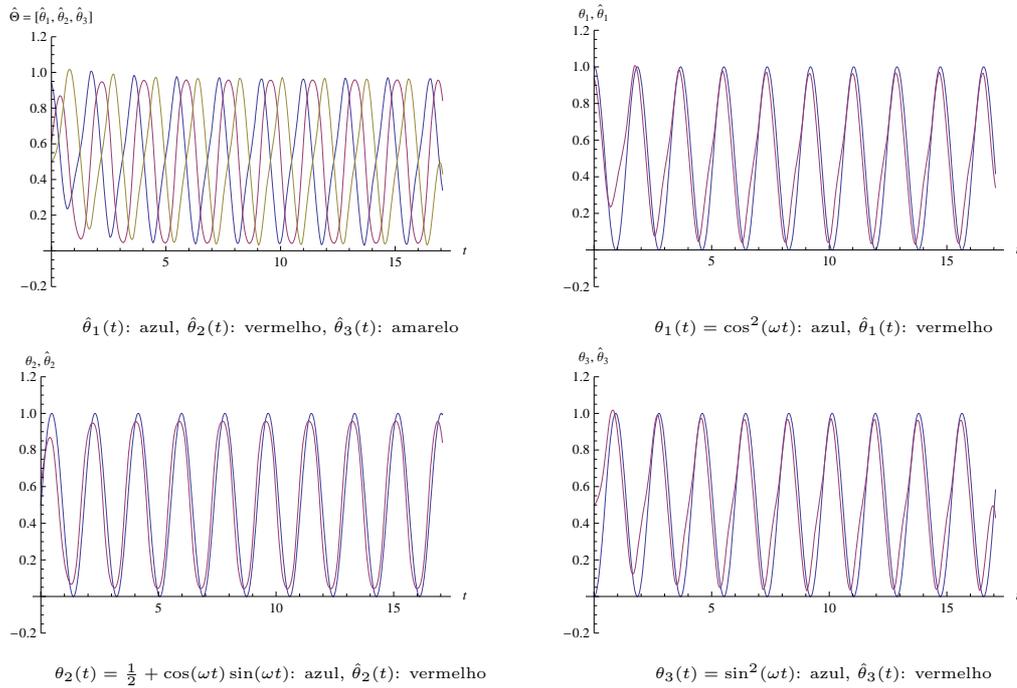


Figura 4.2: Estimativa dos parâmetros, otimização sem restrições, partindo de  $\Theta_0 = [0.5, 0.5, 0.5]^T$  a cada instante.

Embora os resultados obtidos via otimização sem restrições pareçam<sup>2</sup> razoáveis, tal solução pode ser numericamente frágil. Repetindo os mesmos cálculos com a estimativa inicial  $\Theta_0 = [0, 0, 0]^T$  os resultados são diferentes, como mostra a Figura 4.3.

<sup>2</sup>Neste trabalho, a análise será apenas qualitativa. Embora fosse possível quantificar o erro computando, por exemplo, seu valor quadrático médio, tal procedimento poderia ser enganador. Nos exemplos deste capítulo não se observam casos onde a estimativa de um ou mais parâmetros se torna indeterminada porque sua influência na saída do sistema desaparece, mesmo que momentaneamente. Isso é observado no próximo capítulo, podendo-se argumentar que um eventual erro de estimação de parâmetros nesses casos seria inconseqüente para fins de posterior utilização no controle LPV, embora relevante se a estimativa é considerada um processo de medição do parâmetro.

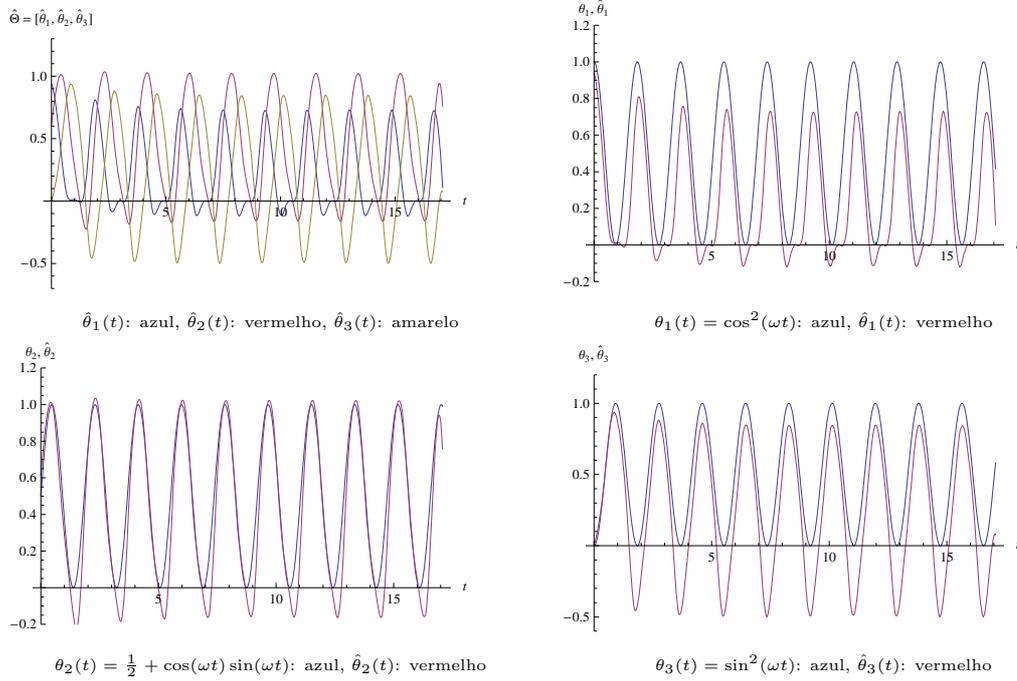


Figura 4.3: Estimativa dos parâmetros, otimização sem restrições, partindo de  $\Theta_0 = [0, 0, 0]^T$  a cada instante.

Tal comportamento pode ser melhor entendido observando os valores dos componentes do gradiente de  $f$  ao longo da trajetória  $\Theta^*(t)$ , no espaço dos parâmetros, ilustrados na Figura 4.4. A cor em 4.4 (b) é tal que vermelho corresponde ao valor absoluto da derivada parcial em relação a  $\theta_1(t)$ , verde a derivada parcial em relação a  $\theta_2(t)$ , e azul a derivada parcial em relação a  $\theta_3(t)$ : assim, por exemplo, o trecho da trajetória em verde puro indica uma forte sensibilidade a  $\theta_2(t)$ , e baixa quanto aos outros dois parâmetros. Outro ponto a se notar é que, por construção,  $\theta_1^*(t) + \theta_3^*(t) = \cos^2(\omega t) + \cos^2(\omega t) = 1, \forall t$ . Isso se reflete na tendência de erros na estimativa de um desses parâmetros causarem também erros na estimativa do outro. Por fim, a Figura 4.5 mostra um mapa de contorno da derivada parcial em relação a  $\theta_1(t)$ ,  $\frac{\partial f}{\partial \theta_1}$ , com  $\theta_2(t) = \theta_2^*(t)$  e  $\theta_3(t) = \theta_3^*(t)$  (a), e as linhas  $\frac{\partial f}{\partial \theta_1} = 0$  sobrepostas à estimativa do parâmetro. Pode-se perceber a piora da estimativa nas periódicas descontinuidades da superfície de (a).

Melhores estimativas podem ser obtidas via otimização com restrições, refletindo a hipótese de limites conhecidos para os valores dos parâmetros. No exemplo em questão esses últimos são normalizados, e o centro da região factível é exatamente o ponto de partida utilizado antes,  $\Theta_0 = [0.5, 0.5, 0.5]^T$ . Ademais, as funções-barreiras impedem que as estimativas de um parâmetro divirjam incorretamente na tentativa de compensar erros nas estimativas de outros parâmetros. A Figura 4.6 ilustra os resultados.

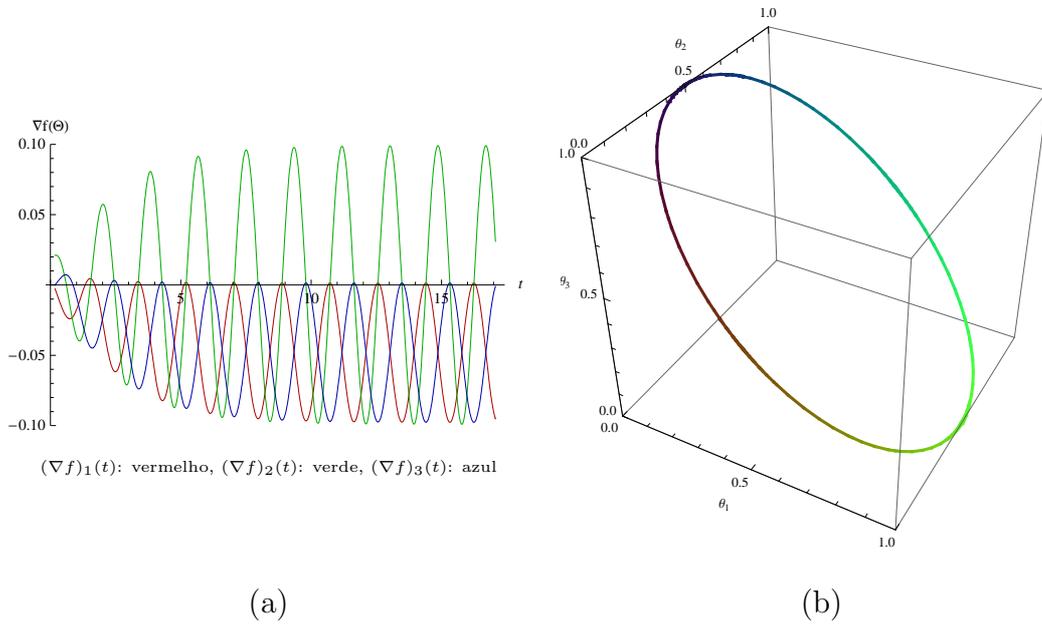


Figura 4.4: Componentes do gradiente  $\nabla f(\Theta)$  ao longo de  $\Theta^*(t)$ , em coordenadas cartesianas (a) e no espaço dos parâmetros (b) (informação de (a) em cores sobreposta à trajetória  $\Theta^*(t)$ ).

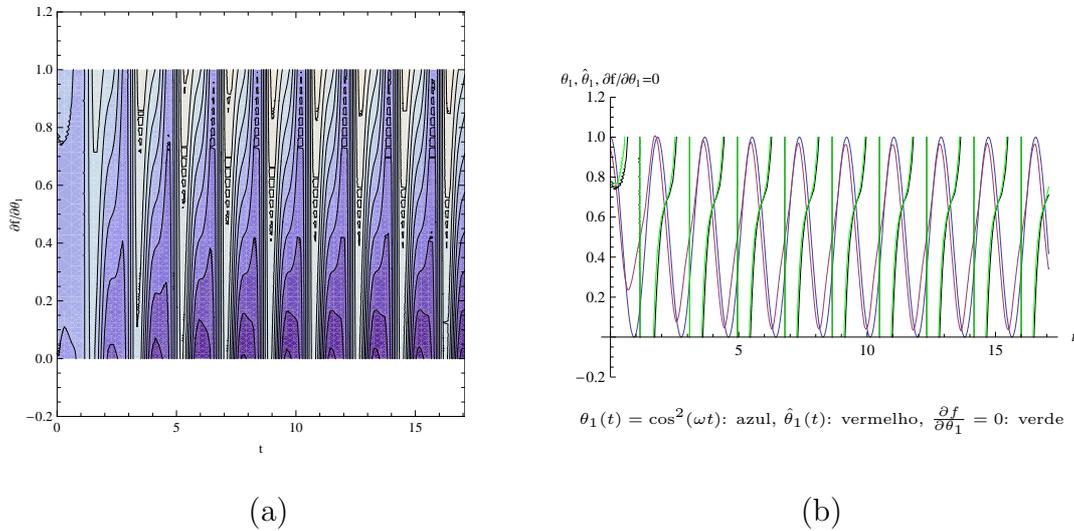


Figura 4.5: Mapa de contorno  $\frac{\partial f}{\partial \theta_1}$ , em (a) (valor absoluto proporcional à intensidade da cor azul), e sobreposição de  $\frac{\partial f}{\partial \theta_1} = 0$  (uma das famílias de curvas de nível em (a)) com o valor exato de  $\theta_1(t)$  e a estimativa desse parâmetro,  $\hat{\theta}_1(t)$ , em (b).

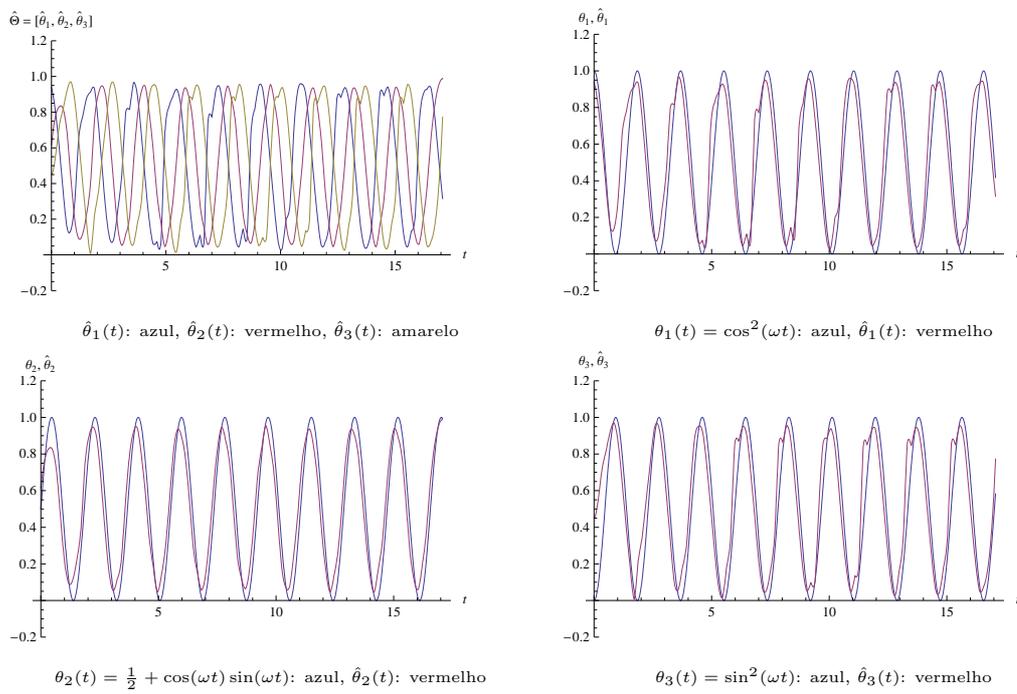


Figura 4.6: Estimativa dos parâmetros, restritos a  $0 \leq \hat{\theta}_k(t) \leq 1$ ,  $k = 1, 2, 3$ .

# Capítulo 5

## Estimação de Parâmetros por Métodos Escalares

Neste capítulo, a estimação de parâmetros é transformada em um problema escalar. O tratamento inicial supõe estimativa de um único parâmetro, resultando em métodos denominados de malha aberta e de malha fechada. Tal nomenclatura diz respeito à estrutura do estimador, e não à do sistema LPV como um todo. Na seqüência, discute-se como esses métodos podem ser estendidos para estimação simultânea de múltiplos parâmetros, e as condições onde isso é ou não é de fato possível. O conteúdo é também sumarizado em [43, 44].

### 5.1 Premissas e Definições

Conforme mencionado acima, o problema de estimação de um único parâmetro será tratado primeiro. A fim de manter a discussão direta, algumas premissas serão adotadas:

- O parâmetro é uma função limitada e contínua por partes<sup>1</sup>. Essa hipótese também vale para múltiplos parâmetros, e os valores são normalizados por conveniência,  $\theta_i(t) \in [0, 1]$ ;
- A entrada  $u(t)$  é conhecida;
- As condições iniciais  $x(0)$  do sistema são nulas ou conhecidas *a priori*;
- A matriz  $B$  independe do parâmetro. Caso contrário, seria necessário incorporar o mesmo em  $A(\Theta(t))$ , sob a forma de um pseudo-estado extra;

---

<sup>1</sup>*i.e.*, em intervalos finitos de tempo, há um número finito de descontinuidades. Esta condição é enunciada aqui unicamente por rigor matemático, visto que na prática parâmetros correspondentes a grandezas físicas não causariam problemas com integração/diferenciação.

- O estado é mensurável, *i.e.*  $C(\Theta(t)) = C = I$ . Essa hipótese pode ser eventualmente relaxada, desde que haja informação suficiente para estimar o parâmetro mesmo com conhecimento parcial do estado. Estimar primeiramente todo ou parte do estado para depois estimar o parâmetro só seria viável se o estimador de estado fosse independente do parâmetro e se sua dinâmica fosse mais rápida que a do estimador do parâmetro;
- Não há transmissão direta da entrada para a saída, *i.e.*  $D(\Theta) = 0$ . Lidar com  $D(\Theta) = D \neq 0$  constante não seria problema, pois supõe-se que a entrada  $u(t)$  seja conhecida. Por outro lado, se  $D(t) = D(\Theta(t))$  o parâmetro apareceria diretamente na saída a menos de um fator conhecido, portanto sua estimação seria trivial.

Dadas as hipóteses acima, o sistema com um parâmetro se torna

$$\begin{aligned}\dot{x}(t) &= (A_0 + A_1\theta_1(t))x(t) + Bu(t) , \\ y(t) &= x(t) .\end{aligned}\tag{5.1}$$

### 5.1.1 Função Escalar do Estado

O desenvolvimento que se segue faz uso da seguinte função escalar do vetor de estado  $x(t)$ . Seja

$$A_i = U_i \Sigma_i V_i^T \tag{5.2}$$

a decomposição em valores singulares (*singular value decomposition*, SVD) da matriz-coeficiente do parâmetro  $\theta_i(t)$ ,  $A_i$ . As matrizes  $U_i$  e  $V_i$  são unitárias, e

$$\Sigma_i = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$$

é a matriz diagonal  $n \times n$  que traz os  $n$  valores singulares<sup>2</sup>  $\sigma_s$  de  $A_i$ ,  $r \leq n$  dos quais não-nulos.

Definindo

$$\bar{\Sigma}_i \triangleq \frac{1}{\sum_{s=1}^r \sigma_s} [ \underbrace{1, \dots, 1}_r, \underbrace{0, \dots, 0}_{n-r} ] , \tag{5.3}$$

seja

$$\begin{aligned}f_i(t) &\triangleq \bar{\Sigma}_i U_i^T A_i x(t) \\ &= \bar{\Sigma}_i U_i^T U_i \Sigma_i V_i^T x(t) \\ &= \bar{\Sigma}_i \Sigma_i V_i^T x(t)\end{aligned}\tag{5.4}$$

---

<sup>2</sup>Raízes quadradas dos autovalores de  $A_i A_i^T$  ou de  $A_i^T A_i$ .

a função que se deseja definir aqui<sup>3</sup>. Ela funciona da seguinte maneira:  $V_i$  é unitária, portanto a pré-multiplicação de  $x(t)$  por  $V_i^T$  é uma mudança de coordenadas, especificamente uma rotação; a próxima multiplicação, pela matriz diagonal  $\Sigma_i$ , altera a escala relativa dos componentes do vetor de acordo com os valores singulares de  $A_i$ , e a multiplicação final pelo vetor  $\bar{\Sigma}_i$  resulta numa média ponderada desses componentes. Assim, quando o produto  $A_i x(t)$  é multiplicado por  $\theta_i(t)$ , pode-se notar que a influência do parâmetro em  $x(t)$  está restrita a um subespaço  $r$ -dimensional do espaço de estados original, e a função  $f_i(t)$  normaliza o ganho introduzido por  $A_i$ .

## 5.2 Desenvolvimento

Os métodos de estimação a seguir foram desenvolvidos começando de uma estrutura simples em malha aberta, que posteriormente evoluiu para estruturas mais robustas, em malha fechada. Conforme dito antes, o problema com um único parâmetro ( $\theta_1$ , sem perda de generalidade) será estudado primeiro.

### 5.2.1 Método em Malha Aberta

Tomando a primeira equação em (5.1),

$$\dot{x}(t) = (A_0 + A_1\theta_1(t))x(t) + Bu(t) , \quad (5.5)$$

seja  $\dot{\tilde{x}}(t) = A_0\tilde{x}(t) + Bu(t)$ . Com as mesmas condições iniciais e entrada de (5.5), essa equação forneceria a resposta do sistema se o parâmetro fosse identicamente nulo. Substituindo  $\tilde{x}(t)$  pelo o estado  $x(t)$  do sistema original, mensurável, no segundo membro da equação, ela se torna

$$\dot{\tilde{x}}(t) \triangleq A_0x(t) + Bu(t) , \quad (5.6)$$

e define-se  $\tilde{x}(t)$  como sua solução.

Seja então

$$\begin{aligned} \dot{e}(t) &\triangleq \bar{\Sigma}U^T(\dot{x}(t) - \dot{\tilde{x}}(t)) \\ &= \bar{\Sigma}U^T A_1\theta_1(t)x(t) \\ &= f_1(t)\theta_1(t) \\ \therefore e(t) &= \int f_1(\tau)\theta_1(\tau) d\tau , \end{aligned} \quad (5.7)$$

---

<sup>3</sup>Na verdade, a função surgiu originalmente no desenvolvimento de (5.7). Devido a sua maior generalidade, é introduzida antes.

com  $f_1(t)$  definida como em (5.4). Visto que  $f_1(t)$  pode ser calculada a partir do estado, conhecido, a integração acima pode ser feita em qualquer intervalo: em particular,  $t_0 \leq t < t_0 + \Delta t$ . Se a duração  $\Delta t$  for pequena o bastante (mais sobre isso em breve), é possível aproximar o parâmetro pelo seu valor médio nesse intervalo<sup>4</sup>. Seja  $\bar{\theta}_1$  o valor médio de  $\theta_1(t)$  para  $t \in [t_0, t_0 + \Delta t)$ , tal que, de (5.7),

$$e(t_0 + \Delta t) - e(t_0) = \int_{t_0}^{t_0 + \Delta t} f_1(\tau) \theta_1(\tau) d\tau \approx \bar{\theta}_1 \int_{t_0}^{t_0 + \Delta t} f_1(\tau) d\tau . \quad (5.8)$$

Assim como a integral de  $f_1(\tau)$ , o “erro”  $e(t)$  pode ser calculado de forma independente pois é função de  $\tilde{x}(t)$ , que por sua vez depende apenas de  $x(t)$  e de  $u(t)$ . A fim de simplificar a notação, sejam

$$\Delta e(t) \triangleq e(t) - e(t - \Delta t) \quad e \quad (5.9)$$

$$\Delta F_1(t) \triangleq \int_0^t f_1(\tau) d\tau - \int_0^{t-\Delta t} f_1(\tau) d\tau . \quad (5.10)$$

A duração do intervalo,  $\Delta t$ , torna-se implícita, e (5.8) pode ser reescrita como

$$\bar{\theta}_1(t) = \frac{\Delta e(t)}{\Delta F_1(t)} . \quad (5.11)$$

A equação acima poderia, em princípio, ser utilizada para calcular  $\bar{\theta}_1$ : as Figuras 5.1 e 5.2 sumarizam a computação. Particionando o domínio do tempo em intervalos de duração  $\Delta t$  e calculando  $\bar{\theta}_1$  para cada um deles, é possível definir uma função contínua por partes  $\bar{\theta}_1(t)$ , estimativa do parâmetro<sup>5</sup>. Antes de prosseguir nessa linha, há duas questões que devem ser discutidas.

O primeiro problema com (5.11) é a possibilidade de divisão por zero. Quando  $x(t) \rightarrow 0$ , *e.g.* um sistema estável retorna ao ponto de equilíbrio  $x = 0$ , ou  $A_1 x(t) \rightarrow 0$ , *e.g.*  $x(t) \neq 0 \in \mathcal{N}(A_1)$ , a trajetória do sistema entra no espaço nulo de  $A_1$ ,  $\Delta F_1(t) \rightarrow 0$  e torna-se impossível<sup>6</sup> estimar  $\theta_1(t)$  a partir de (5.11). Tal limitação é inerente aos métodos sendo desenvolvidos aqui, mas há um contraponto: se a estimativa  $\bar{\theta}_1(t)$  estiver sendo calculada para fins de controle, então o conhecimento do parâmetro não é necessário quando o mesmo não afeta o sistema. Do ponto de vista computacional, tais situações podem ser detectadas e tratadas de forma *ad hoc*, por exemplo limitando  $\bar{\theta}_1(t) \in [0, 1]$  quando a divisão falha, ou alertando o usuário quando  $|f_1(t)|$  é menor que algum valor predefinido. Ademais,

<sup>4</sup>O estado  $x(t)$  e a função  $f_1(t)$  são diferenciáveis, portanto contínuos.  $\theta_1(t)$  é assumida contínua por partes portanto, na pior das hipóteses, será aproximada pelo seu valor médio em eventuais descontinuidades (número finito) no intervalo. Para detalhes técnicos, vide *p.ex.* [45].

<sup>5</sup>Assume-se aqui alimentar a seqüência de “amostras”  $\bar{\theta}_1$  a um *hold* de ordem zero. Outras formas de interpolação seriam também possíveis.

<sup>6</sup>Divisão por zero, se  $\Delta e(t) \neq 0$ , ou forma indefinida  $0/0$ , se não.

cabe notar aqui que a presente discussão omite os ruídos de processo e de medição, que poderiam prover excitação adicional do sistema para manter  $\Delta F_1(t) \neq 0$  na maioria dos casos. A contrapartida seriam possíveis erros introduzidos na estimativa.

O segundo problema com (5.11) está relacionado à largura de banda: o valor de  $\Delta t$  funciona como um intervalo de amostragem. Sem tentar aqui produzir uma prova matemática rigorosa, mas apenas para fins de argumentação, seja a transformada de Laplace ( $\mathcal{L}$ ) da Equação (5.8), abaixo (com  $f_1(0) = 0$ , por hipótese):

$$\begin{aligned} \bar{\theta}_1(t) \left( \int_0^t f_1(\tau) d\tau - \int_0^{t-\Delta t} f_1(\tau) d\tau \right) &\approx \int_0^t f_1(\tau) \theta_1(t) d\tau \\ &\quad - \int_0^{t-\Delta t} f_1(\tau) \theta_1(t) d\tau \\ &\quad \mathcal{L} \downarrow \uparrow \mathcal{L}^{-1} \\ \bar{\Theta}_1(s) * \left( \frac{1 - e^{-(\Delta t)s}}{s} F_1(s) \right) &\approx \frac{1 - e^{-(\Delta t)s}}{s} (\Theta(s) * F_1(s)) . \end{aligned} \quad (5.12)$$

O primeiro membro de (5.12) pode ser interpretado como a convolução entre a transformada da estimativa  $\bar{\Theta}_1(s)$  e  $F_1(s)$ , essa última precedida pelo *hold* de ordem zero  $(1 - e^{-(\Delta t)s})/s$ . No segundo membro, a convolução entre a transformada do valor real do parâmetro  $\Theta_1(s)$  e  $F_1(s)$  ocorre antes de o resultado ser filtrado por aquele *hold*. A parte linear do modelo LPV normalmente funciona como um filtro passa-baixas para o parâmetro, e o lugar das raízes de  $\det(sI - (A_0 + A_1\theta_1)) = 0$  com  $\theta_1 \in [0, 1]$  poderia ser usado para determinar  $\Delta t$  a grosso modo, *i.e.* assumindo que o parâmetro seja estático ou varie lentamente. A análise correta passa pela aplicação dos resultados do Capítulo 3, caso contrário problemas como introdução de erros devido a *aliasing* ou atraso excessivo na estimativa, proporcional a  $\Delta t$ , podem acontecer.

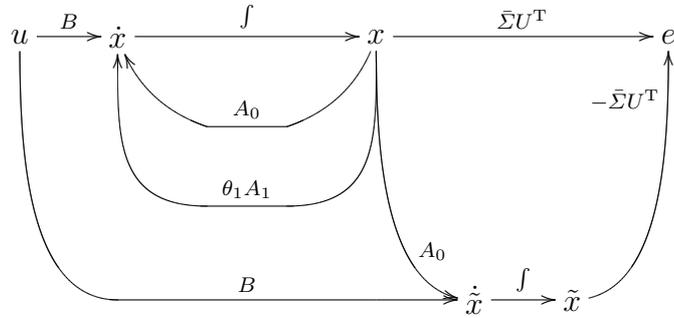
## 5.2.2 Métodos em Malha Fechada

A configuração em malha aberta descrita na seção anterior não usa realimentação para refinar a estimativa, o que poderia ajudar a eliminar erros e rejeitar distúrbios. Buscando essas características adicionais, uma nova estimativa em malha fechada, denotada  $\hat{\theta}_1(t)$ , é proposta aqui. Assumindo inicialmente que tal estimativa já fosse conhecida, sejam

$$e_\theta(t) \triangleq \theta_1(t) - \hat{\theta}_1(t) \quad (5.13)$$

e

$$\hat{x}(t) \triangleq x(0) + \int_0^t [(A_0 + \hat{\theta}_1(\tau)A_1)x(\tau) + Bu(\tau)] d\tau \quad (5.14)$$



$$e(t) = \int_0^t \theta_1(\tau) f_1(\tau) d\tau \approx e(t - \Delta t) + \bar{\theta}_1(t - \Delta t) \int_{t-\Delta t}^t f_1(\tau) d\tau$$

Figura 5.1: Cálculo do erro, método em malha aberta, um único parâmetro.

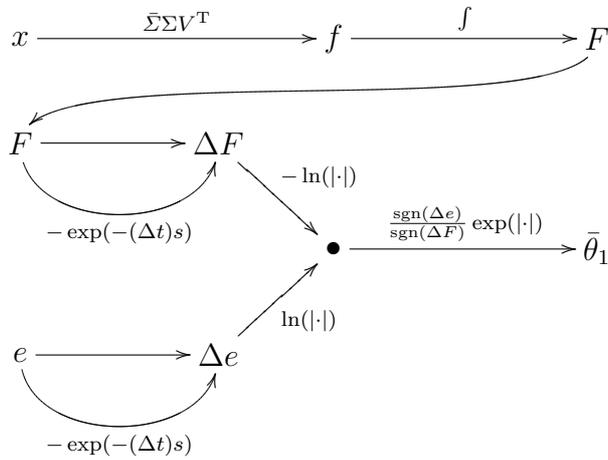


Figura 5.2: Cálculo da estimativa, método em malha aberta, um único parâmetro (incluindo o artifício numérico, ocasionalmente usado, de calcular a divisão por meio da diferença de logaritmos).

o erro de estimação e a resposta do sistema com o parâmetro estimado, respectivamente. Procedendo como em (5.7) com  $\hat{x}(t)$  no lugar de  $\tilde{x}(t)$ ,  $e(t)$  é redefinido como:

$$\begin{aligned}
 e(t) &\triangleq \bar{\Sigma}_1 U_1^T (x(t) - \hat{x}(t)) \\
 &= \bar{\Sigma}_1 U_1^T \int_0^t A_1 (\theta_1(\tau) - \hat{\theta}_1(\tau)) x(\tau) d\tau \\
 &= \int_0^t f_1(\tau) (\theta_1(\tau) - \hat{\theta}_1(\tau)) d\tau \\
 &= \int_0^t f_1(\tau) e_\theta(\tau) d\tau .
 \end{aligned} \tag{5.15}$$

Dado  $e(t)$  acima, o desenvolvimento é semelhante ao da seção anterior com uma maior diferença: antes, assumia-se que  $\theta_1(t)$  poderia ser aproximado pelo valor médio no intervalo em que  $\Delta e(t)$  era calculado, para que pudesse ser “extraído” da integral; a nova hipótese é que a dinâmica de  $e_\theta(t)$  seja mais rápida que a de  $f_1(t)$ . Uma vez que  $\hat{\theta}_1(t)$  deva acompanhar  $\theta_1(t)$  mesmo se  $x(t)$  e  $f_1(t)$  variarem rapidamente, o erro deve ser ainda mais rápido que essas duas grandezas. Isso informalmente justifica a hipótese. Assim, em intervalos de tempo curtos o bastante,  $f_1(t)$  é aproximada pelo seu valor médio e extraída da integral em (5.15):

$$\begin{aligned}
 e(t) - e(t - \Delta t) &= \int_{t-\Delta t}^t f_1(\tau) e_\theta(\tau) d\tau \\
 &\approx \bar{f}_1 \int_{t-\Delta t}^t e_\theta(\tau) d\tau .
 \end{aligned} \tag{5.16}$$

Na expressão acima,  $\bar{f}_1$  pode ser substituído pelo seu valor real  $f_1(t)$ , conhecido<sup>7</sup>. Reutilizando a definição (5.9), e definindo ainda

$$\Delta E_\theta(t) \triangleq \int_0^t e_\theta(\tau) d\tau - \int_0^{t-\Delta t} e_\theta(\tau) d\tau , \tag{5.17}$$

a variação do erro pode ser escrita como

$$\Delta E_\theta(t) = \frac{\Delta e(t)}{f_1(t)} . \tag{5.18}$$

Alguns problemas surgem, como antes, se o valor de  $f_1(t)$  se aproxima de zero:  $\Delta e(t)$  também tende a zero nesse caso. Uma vez mais, a implementação computacional deverá estar preparada para lidar com isso.

---

<sup>7</sup>O argumento de aproximação pela média foi utilizado apenas para justificar a “extração” de  $\bar{f}_1$  da integral. Utilizar depois o valor real, *i.e.*, desprezar a aproximação depois que ela cumpriu seu papel, não seria problema. Contrastando com isso, a aproximação que se segue em (5.19) (“derivação” aproximada) permanece válida até o final.



O procedimento acima é justificado ao reconhecer que o problema original pode ser reinterpretado como uma modulação em amplitude da portadora  $f_1(t)$  pelo sinal  $\theta_1(t)$  e que, em princípio, um demodulador síncrono (vide, por exemplo, [46]) poderia ser usado para recuperar o parâmetro. Na modulação de amplitude clássica, o sinal, centrado numa banda base, é multiplicado por uma portadora de alta frequência, usualmente uma senóide: o conteúdo espectral do primeiro é meramente translado para a frequência dessa última. A demodulação consiste em uma segunda multiplicação pela portadora, gerando um sinal com componentes tanto em torno da banda base quanto em frequências próximas ao dobro da portadora: essas são eliminadas por um filtro passa-baixas, e o que resta é o sinal original, recuperado. No problema de estimação de parâmetros aqui estudado, em contraste, a portadora  $f_1(t)$  não é necessariamente uma senóide, e seu espectro quase certamente se sobrepõe ao de  $\theta_1(t)$ . A função de transferência  $H(s)$  funciona como o filtro passa-baixas do demodulador clássico, mas enquanto esse último é configurado em malha aberta, aqui a configuração em malha fechada é essencial para rejeição de erros, permitindo eliminar a informação espúria de  $f_1(t)$  da estimativa.

### 5.3 Extensão para Múltiplos Parâmetros

Até este ponto a discussão esteve restrita à estimação de um único parâmetro. Os métodos apresentados podem ser estendidos para estimação simultânea de múltiplos parâmetros de forma relativamente direta, mas há restrições. Sem perda de generalidade, seja um sistema LPV com dois parâmetros,  $\theta_1(t)$  e  $\theta_2(t)$ , coeficientes  $A_1$  e  $A_2$ . Do ponto de vista dos métodos apresentados neste capítulo, só é possível estimar  $\theta_1(t)$ , por exemplo, se for possível isolar no estado informação que dependa apenas desse parâmetro e não de  $\theta_2(t)$ ; obviamente, a recíproca também vale para  $\theta_2(t)$ .

Para avaliar se existe interferência entre parâmetros, uma primeira providência é tentar reestruturar  $A(\Theta(t))$  numa forma diagonal por blocos ou, se possível, numa forma canônica de Jordan. A análise dos elementos ou blocos na diagonal revela quais são os modos naturais do sistema, e quais dos parâmetros afetam cada um desses modos; no caso de dois parâmetros, há quatro possibilidades: modos em que o único parâmetro é  $\theta_1(t)$ , modos em que o único parâmetro é  $\theta_2(t)$ , modos onde os dois parâmetros atuam, e modos que independem deles. Modos com parâmetro único são os mais simples de se lidar: se o parâmetro é  $\theta_1(t)$ , por exemplo, basta fazer com que os elementos de  $\bar{\Sigma}_1$  correspondentes aos modos afetados por  $\theta_2(t)$  sejam zero para que  $f_1(t)$  seja ortogonal a esse segundo parâmetro, *i.e.* não seja afetada por ele. Nesse caso,  $\theta_1(t)$  pode ser estimado como se fosse único. Tal condição pode ser expressa em termos de  $\text{range}(A_1) - \text{range}(A_2) \neq \emptyset$ , e a recíproca vale para o segundo parâmetro. Por outro lado, completa superposição

tal que  $\text{range}(A_1) = \text{range}(A_2)$  impede que qualquer um dos dois parâmetros seja estimado. Casos como  $\text{range}(A_1) \supsetneq \text{range}(A_2)$  recaem na primeira situação em termos de  $\theta_1(t)$ , mas na segunda para  $\theta_2(t)$ ;  $\theta_1(t)$  pode ser estimado normalmente e, se sua estimativa  $\hat{\theta}_1(t)$  for descontada daquilo que se obtém dos modos em que  $\theta_1(t)$  e  $\theta_2(t)$  se misturam, torna-se possível estimar  $\hat{\theta}_2(t)$  para o segundo parâmetro. O sucesso desse procedimento depende de que a dinâmica da estimativa de  $\theta_1(t)$  seja mais rápida que aquela da mistura, o que nem sempre é verdade. Para uma análise precisa é conveniente voltar ao Capítulo 3, repetindo os cálculos para o sistema só com  $\theta_1(t)$  (fazendo  $\theta_2(t) = 0$ ), *cf.* o sistema com a mistura de parâmetros.

Para encerrar esta seção, cabe comentar que as idéias discutidas no parágrafo anterior também valem para sistemas com três ou mais parâmetros. As possibilidades de superposição aumentam combinatorialmente, podendo ser pensadas da seguinte forma: um sistema de ordem  $n$  pode ter, no máximo,  $n$  modos independentes; se o número de parâmetros for maior que  $n$ , alguns dos modos certamente misturarão parâmetros<sup>8</sup>. Ainda assim, a existência de exatamente  $n$  parâmetros não impede que seja possível estimar todos por eliminação recursiva. Um número maior que  $n$  certamente implica que nem todos poderão ser estimados. Por outro lado, conforme visto antes, um número de parâmetros menor que  $n$ , mesmo tão pouco quanto apenas dois, não garante que não possa haver problemas. Cada caso deve ser analisado individualmente.

## 5.4 Exemplos

A fim de ilustrar a utilização e o desempenho dos métodos desenvolvidos acima, três exemplos serão apresentados aqui. O primeiro é uma versão não-autônoma (*i.e.*, com entrada  $u(t)$  exógena) do sistema introduzido na Seção 3.4, o segundo é original, e o terceiro é a conclusão daquele iniciado na Seção 1.2.1.

### 5.4.1 Primeiro Exemplo

Seja o sistema

$$\dot{x}(t) = A(\Theta(t))x(t) + Bu(t) ,$$

com  $A(\Theta(t))$  como em (3.26), repetida abaixo por conveniência,

$$A(\Theta(t)) = A_0 + \sum_{k=1}^3 A_k \theta_k(t) ,$$

$$A_0 = \begin{bmatrix} -1 & \frac{7}{4} \\ -\frac{1}{4} & -1 \end{bmatrix} , A_1 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 0 \end{bmatrix} , A_2 = \begin{bmatrix} 0 & -\frac{3}{2} \\ -\frac{3}{2} & 0 \end{bmatrix} , A_3 = \begin{bmatrix} 0 & 0 \\ 0 & \frac{3}{2} \end{bmatrix} ,$$

<sup>8</sup>Invoca-se aqui o “princípio dos buracos de pombos” (*pigeonhole*), também conhecido como princípio de Dirichlet. Vide, *p. ex.*, [47].

e  $B = [1 \ 1]^T$ . As condições iniciais são nulas e a excitação do sistema é feita pela entrada  $u(t)$ , uma onda quadrada de amplitude unitária e período 10.

A Figura 5.4 mostra a resposta do sistema para  $\theta_1(t) = \cos^2(t)$ ,  $\theta_2(t) = \theta_3(t) = 0$ , e os resultados da estimação desse primeiro parâmetro através dos três métodos: malha aberta, malha fechada e malha fechada simplificado. O valor de  $\Delta t$  adotado foi 0.05, e ambos os métodos de malha fechada usaram o mesmo filtro  $G(z) = H(z) = \frac{z-0.1}{z-0.9}$  (determinado empiricamente). A decomposição da matriz  $A_1$  é trivial,  $A_1 = U_1 \Sigma_1 V_1^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1.5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ , tal que  $\bar{\Sigma}_1 = \frac{1}{1.5+0} [1 \ 0]$  e  $f_1(t) = [1 \ 0] x(t)$ .

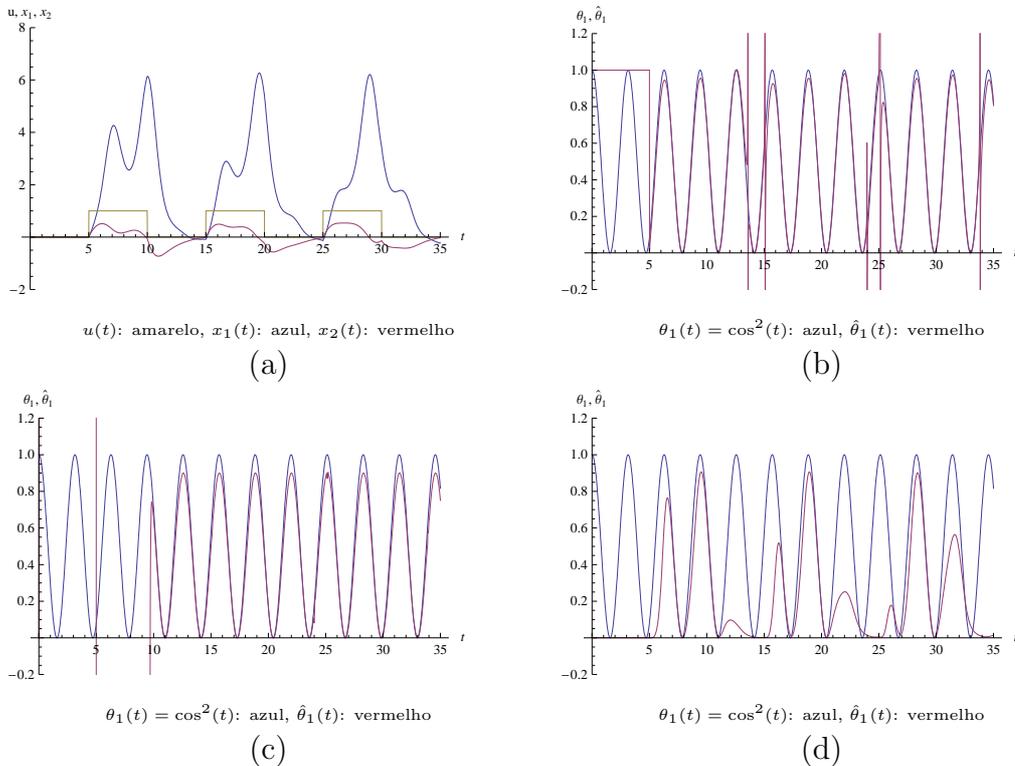


Figura 5.4: Primeiro exemplo, estimação de  $\theta_1(t)$ : entrada e saída do sistema (a), método em malha aberta (b), método em malha fechada (c) e método em malha fechada simplificado (d).

Quanto à aparente<sup>9</sup> exatidão dos resultados, nota-se que o melhor método é o em malha aberta, com os métodos em malha fechada regular e simplificado em segundo e terceiro lugares, respectivamente. Em defesa deste último pode-se argumentar que seu desempenho é proporcional à amplitude da resposta, enquanto que os dois outros métodos seriam numericamente mais frágeis. De fato, na Figura 5.5 podem ser vistos os resultados da mesma simulação com ruído branco,  $\sigma^2 = 0.01$ , adicionado ao estado (simulando ruído de medição): os métodos em malha aberta e em malha fechada divergem, enquanto o método em malha fechada simplificado continua funcionando.

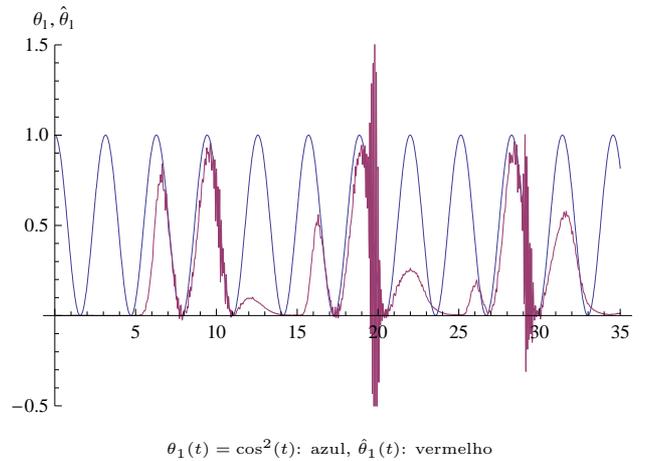
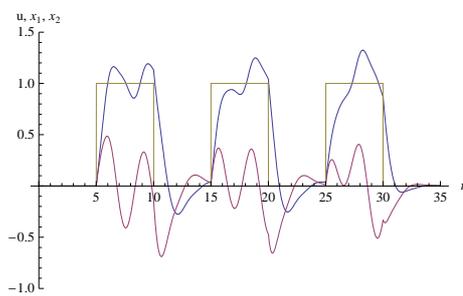


Figura 5.5: Primeiro exemplo, estimação de  $\theta_1(t)$  via o método em malha fechada simplificado, com ruído.

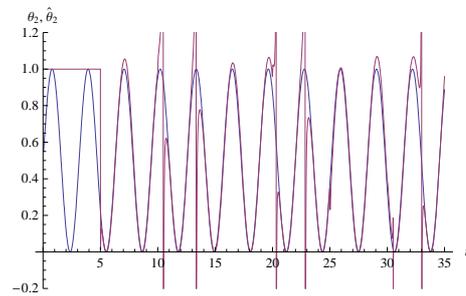
A estimação dos outros dois parâmetros, individualmente com  $\Theta(t) = [0, \frac{1}{2} + \cos(t)\sin(t), 0]^T$  e  $\Theta(t) = [0, 0, \sin^2(t)]^T$  para  $\theta_2(t)$  e  $\theta_3(t)$  são mostradas nas Figuras 5.6 e 5.7. As decomposições são, respectivamente,  $A_2 = U_2 \Sigma_2 V_2^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \therefore \bar{\Sigma}_2 = \frac{1}{1.5+1.5} [1 \ 1] \Rightarrow f_2(t) = [-0.5 \ -0.5] x(t)$  e  $A_3 = U_3 \Sigma_3 V_3^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1.5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \therefore \bar{\Sigma}_3 = \frac{1}{1.5+0} [1 \ 0] \Rightarrow f_3(t) = [0 \ 1] x(t)$ . Quase todos os comentários acima se aplicam, com duas novas observações: observando 5.6 (d) e 5.7 (d), vê-se que o estimador precisaria de um ganho maior; por outro lado, 5.7 (c) mostra uma estimativa que não converge. Em ambos os casos seria necessário adotar outros filtros  $G(z)$  e  $H(z)$ , conforme verificado experimentalmente (simulação).

<sup>9</sup>Mais uma vez, medida qualitativa. Do ponto de vista utilitarista (*vs.* ponto de vista de medida), uma estimativa de parâmetro imprecisa em instantes onde a influência do mesmo sobre o comportamento do sistema seja pequena ou desapareça não seria problema.



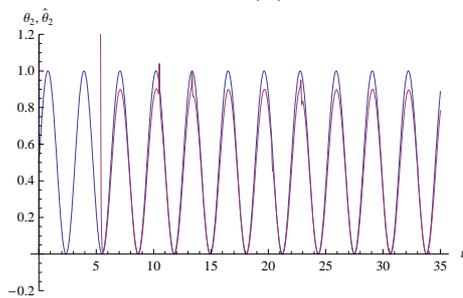
$u(t)$ : amarelo,  $x_1(t)$ : azul,  $x_2(t)$ : vermelho

(a)



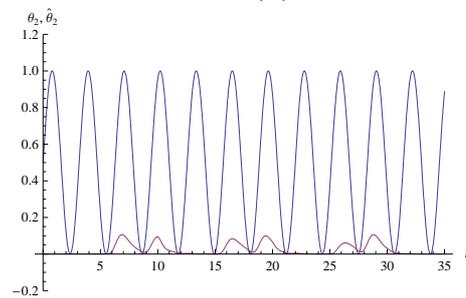
$\theta_2(t) = \frac{1}{2} + \cos(t) \sin(t)$ : azul,  $\hat{\theta}_2(t)$ : vermelho

(b)



$\theta_2(t) = \frac{1}{2} + \cos(t) \sin(t)$ : azul,  $\hat{\theta}_2(t)$ : vermelho

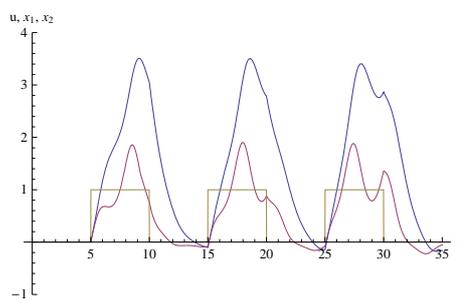
(c)



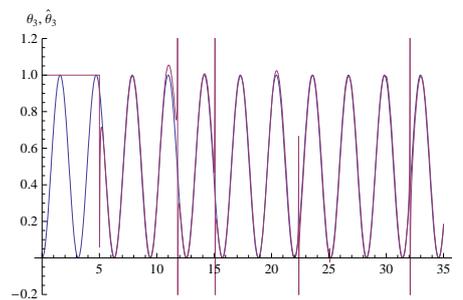
$\theta_2(t) = \frac{1}{2} + \cos(t) \sin(t)$ : azul,  $\hat{\theta}_2(t)$ : vermelho

(d)

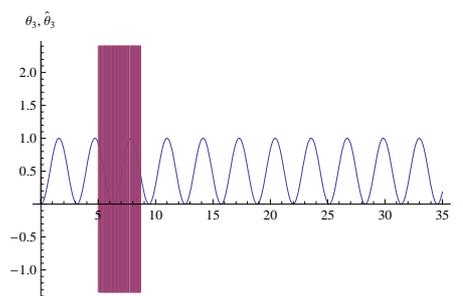
Figura 5.6: Primeiro exemplo,  $\theta_2(t)$ : entrada e saída (a), malha aberta (b), malha fechada (c) e malha fechada simplificado (d).



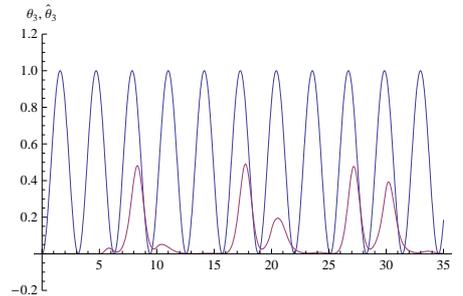
$u(t)$ : amarelo,  $x_1(t)$ : azul,  $x_2(t)$ : vermelho  
(a)



$\theta_3(t) = \sin^2(t)$ : azul,  $\hat{\theta}_3(t)$ : vermelho  
(b)



$\theta_3(t) = \sin^2(t)$ : azul,  $\hat{\theta}_3(t)$ : vermelho  
(c)



$\theta_3(t) = \sin^2(t)$ : azul,  $\hat{\theta}_3(t)$ : vermelho  
(d)

Figura 5.7: Primeiro exemplo,  $\theta_3(t)$ : entrada e saída (a), malha aberta (b), malha fechada (c) e malha fechada simplificado (d).

Uma outra questão relevante é a estimação simultânea de parâmetros, a ser ilustrada pelo exemplo a seguir.

### 5.4.2 Segundo Exemplo

Este exemplo explora um sistema não-autônomo relativamente simples, onde as matrizes são “misturadas” por uma transformação de similaridade. No exemplo anterior a superposição das influências dos parâmetros era clara, enquanto aqui ela deixa de ser óbvia. Tal sistema é

$$\begin{aligned} \dot{x}(t) &= A(\Theta(t))x(t) + Bu(t) \\ &= T \begin{bmatrix} \frac{1}{4}(-6 + \theta_1(t)) & 0 & 0 \\ 0 & \frac{1}{4}(-2 + \theta_2(t)) & 1 + \theta_3(t) \\ 0 & -1 - \theta_3(t) & \frac{1}{4}(-2 + \theta_2(t)) \end{bmatrix} T^{-1}x(t) + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u(t), \end{aligned} \quad (5.22)$$

com

$$T = \begin{bmatrix} 0.817389 & 0.111420 & 0.789526 \\ 0.187803 & 0.241361 & 0.0657388 \\ 0.542247 & 0.231155 & 0.396006 \end{bmatrix}.$$

A matriz  $T$  foi gerada aleatoriamente, tendo como única precondição ser invertível: seu único propósito é tornar o exemplo não trivial. As matrizes  $A_k$  como em (1.3) são obtidas a partir de  $A(\Theta(t))$  acima. As decomposições dessas últimas em valores singulares fornecem  $U_k$ ,  $\Sigma_k$  e  $V_k$ , definindo também os vetores  $\bar{\Sigma}_k$  e finalmente as funções  $f_k(t)$ . Durante esses cálculos, observa-se que  $A_1$  tem *rank* (posto) 1 enquanto  $A_2$  e  $A_3$  são de *rank* 2: conclui-se assim que as estimações de  $\theta_2(t)$  e de  $\theta_3(t)$  não são independentes. Essa mesma conclusão poderia ser obtida inspecionando (5.22), mas na prática não é necessário ter  $A(\Theta(t))$  numa forma de Jordan.

A Figura 5.8 compara o desempenho dos três métodos ao estimar  $\theta_1(t)$  em  $\Theta(t) = [\cos^2(t) \ 0 \ 0]^T$ . A entrada do sistema é  $u(t) = \sin(2\pi t/10)$ , e as condições iniciais são nulas. Foram adotados  $\Delta t = 0.02$  e os filtros discretos  $G(z) = 0.2(z - 0.001)/(z - 1)$  para o método em malha fechada original, e  $H(z) = 500(z - 0.001)/(z - 0.99)$  para o simplificado. (Os filtros foram determinados empiricamente.)

Uma vez mais, a melhor estimativa aparente é a do método em malha aberta, Figura 5.8 (a). Percebe-se que a segunda melhor estimativa, método em malha fechada na Figura 5.8 (b), apresenta um ligeiro atraso, esperado, e problemas numéricos quando  $f_1(t) \rightarrow 0$ . Apesar do ganho alto, o método simplificado sofre

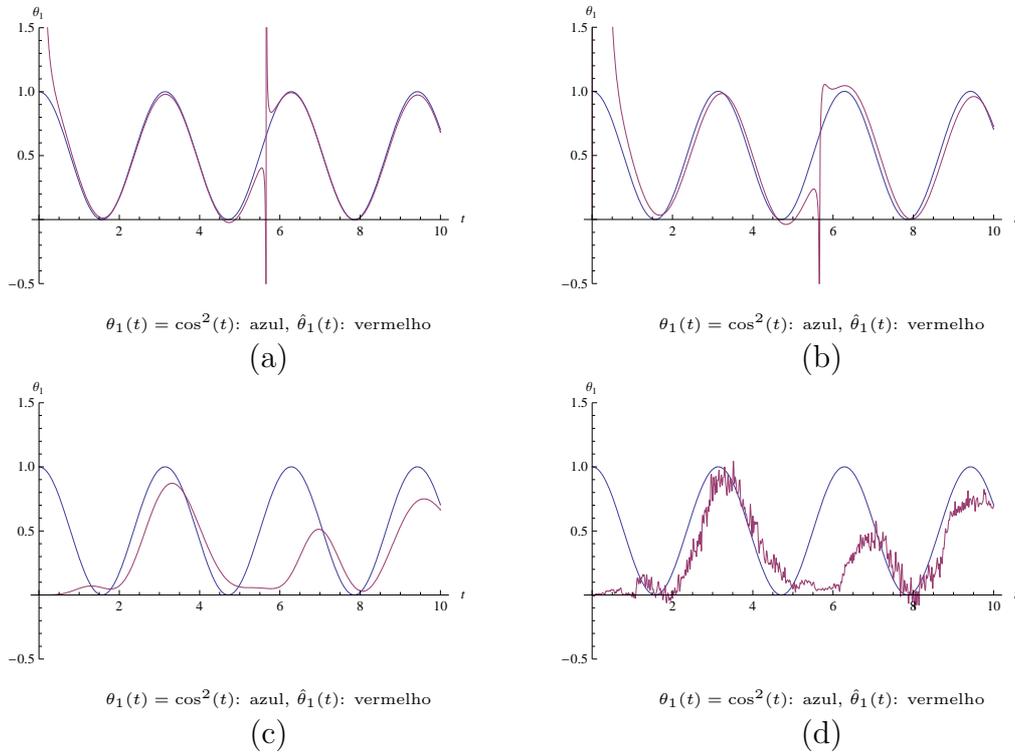


Figura 5.8: Segundo exemplo, estimação de um único parâmetro: métodos em malha aberta (a), malha fechada (b), e malha fechada simplificado com (c) e sem (d) ruído  $\sim \mathcal{N}(0, 0.01)$ .

ainda mais com esse último problema, Figura 5.8 (c), mas por outro lado sua robustez na presença de ruído, Figura 5.8 (d), e sua simplicidade de implementação compensam isso.

As Figuras 5.9 e 5.10 ilustram estimação simultânea de dois parâmetros:  $\theta_1(t) = \cos^2(t)$  e  $\theta_2(t) = 1/2 + \sin(t) \cos(t)$  na primeira, e  $\theta_2(t) = 1/2 + \sin(t) \cos(t)$  e  $\theta_3(t) = \sin^2(t)$  na segunda. Em ambas o método usado é o em malha aberta, e a excitação do sistema se dá pela condição inicial  $x(0) = [1 \ 1 \ 1]^T$  visto que a entrada exógena é  $u(t) = 0$ .

Na Figura 5.9 observa-se que, embora os resultados para  $\theta_2(t)$  sejam qualitativamente razoáveis, o mesmo não pode ser daqueles para  $\theta_1(t)$ . Na Figura 5.10 a estimativa  $\hat{\theta}_2(t)$  está incorreta, e  $\hat{\theta}_3(t)$  um pouco menos. Para contornar esses problemas é preciso tentar separar os parâmetros em diferentes modos.

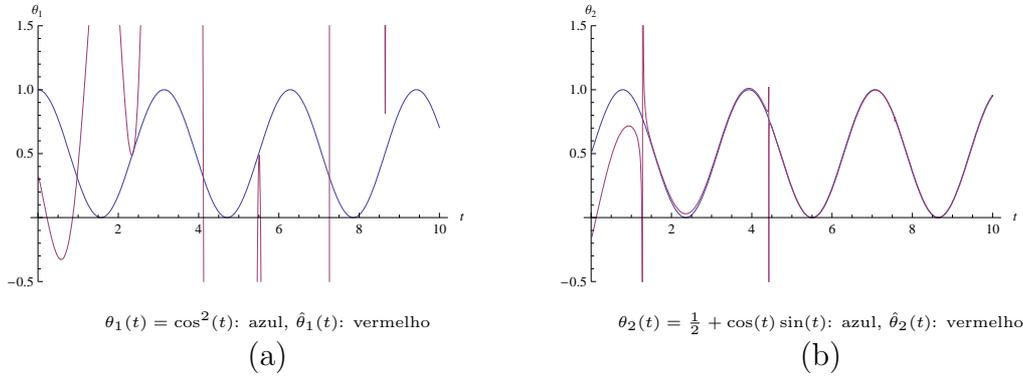


Figura 5.9: Segundo exemplo, estimação simultânea de  $\theta_1(t)$  (a) e de  $\theta_2(t)$  (b).

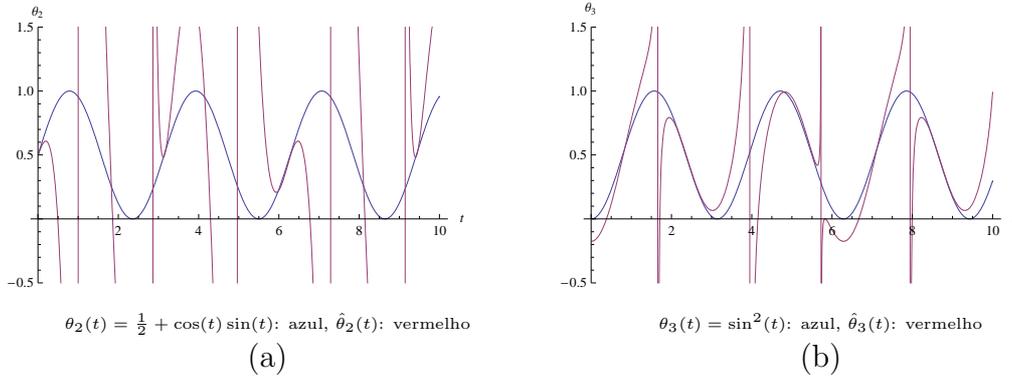


Figura 5.10: Segundo exemplo, estimação simultânea de  $\theta_2(t)$  (a) e de  $\theta_3(t)$  (b).

Seja a transformação de similaridade

$$\tilde{A}(\Theta(t)) = PA(\Theta(t))P^{-1} ,$$

com

$$P = \begin{bmatrix} -8.57155 & -14.7557 & 19.5388 \\ -0.22976 & 1 & 0 \\ -0.663388 & 0 & 1 \end{bmatrix} .$$

Essa mudança de variáveis torna  $A(\Theta(t))$  diagonal por blocos, o que já é suficiente para a estimação: não é preciso recuperar a forma em (5.22). Utilizando  $\tilde{A}(\Theta(t))$  e  $\tilde{x}(t) = Px(t)$  do sistema transformado, obtém-se a Figura 5.11: estimação simultânea e bem-sucedida de  $\theta_1(t)$  e de  $\theta_2(t)$  (método de malha aberta). Por outro lado, não há transformação capaz de separar  $\theta_2(t)$  e  $\theta_3(t)$ , o que ilustra a limitação fundamental da idéia de se separar parâmetros.

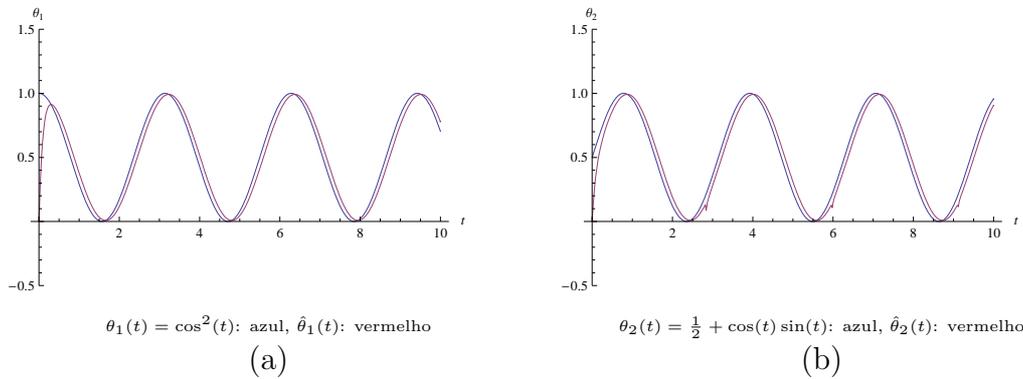


Figura 5.11: Segundo exemplo, estimação simultânea e bem-sucedida de  $\theta_1(t)$  (a) e de  $\theta_2(t)$  (b).

### 5.4.3 Terceiro Exemplo (Conclusão de 1.2.1)

Neste exemplo final retoma-se o sistema (1.5) com matrizes (1.15), conforme o exemplo da Seção 1.2.1. O mesmo controlador é usado aqui, mas o parâmetro foi mudado para  $\theta_1(t) = \sin(t)$  ao invés de  $\cos(y(t))$ , tão somente para mostrar que um parâmetro externo arbitrário pode ser estimado e usado tão facilmente como a saída medida e calculada do artigo original.

A Figura 5.12 compara a resposta do sistema LPV com o controlador usando o parâmetro exato, em (a), e a resposta do sistema com o controlador usando a estimativa  $\bar{\theta}_1(t)$ , calculada como em (5.11), em (b). Os erros (diferenças entre as respostas dos sistemas controlados e a referência) correspondentes a (a) e (b) são mostrados em (c) e (d), respectivamente. Apesar de a diferença de escala mostrar que o erro em (c) é menor, na prática o desempenho é o mesmo. A Figura 5.13 mostra a estimativa  $\bar{\theta}_1(t)$  usada para gerar a Figura 5.12 (b). O valor estimado é visto em (a), enquanto (b) é o respectivo erro (diferença entre estimativa e o valor  $\sin(t)$ ). À medida que o erro da Figura 5.12 (d) tende a zero a estimativa se torna menos precisa, mas isso não afeta a saída do sistema ou seu controle.

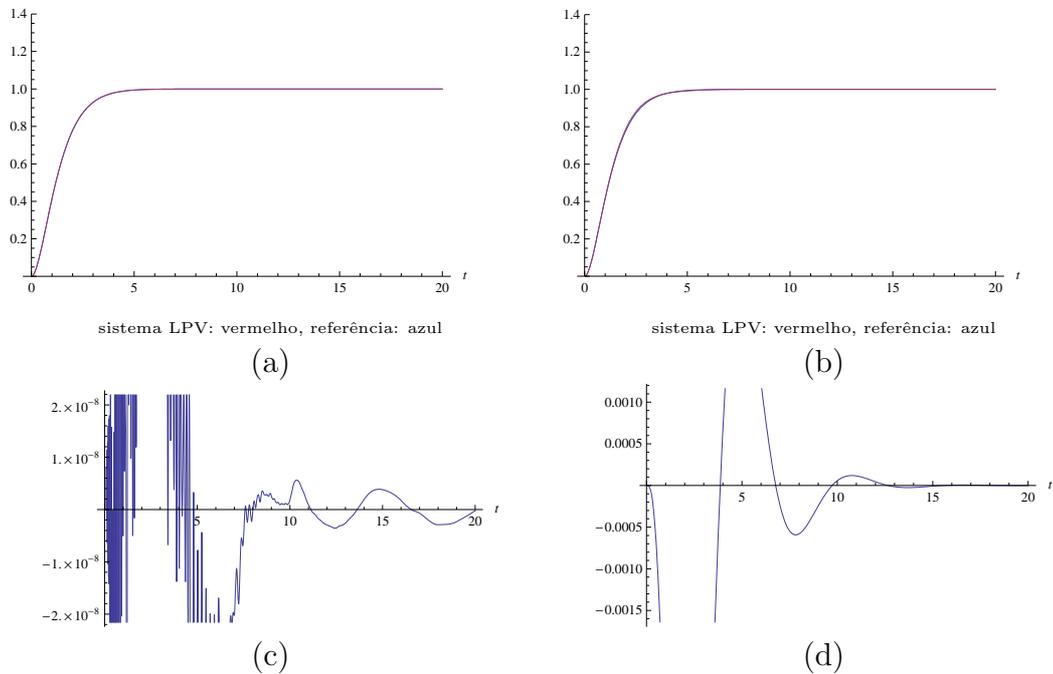


Figura 5.12: Terceiro exemplo, resposta ao degrau, controlador de [2] com o valor real do parâmetro (a), e respectivo erro (c). Em (b) o mesmo experimento é repetido com o controlador usando a estimativa mostrada na Figura 5.13, obtida simultaneamente; (d) é o erro correspondente a (b).

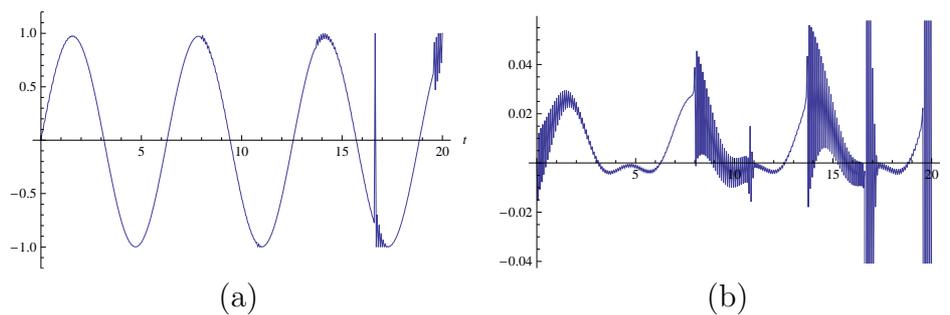


Figura 5.13: Terceiro exemplo, estimativa de  $\theta_1(t)$  (a) e erro (b), para sistema e controlador descritos em [2], com a diferença que o valor estimado é que parametriza o controlador.

# Capítulo 6

## Conclusões e Trabalhos Futuros

### 6.1 Conclusões

A primeira contribuição deste trabalho foi adaptar o resultado apresentado em [31] para o cálculo de limites para o raio espectral de matrizes linearmente dependentes de parâmetros. O raio espectral em si tem aplicações importantes em áreas como análise da estabilidade de sistemas discretos ou conectividade de grafos. O uso do limite superior desse raio para estimar a largura de banda máxima de sistemas LPV, conforme a aplicação / aproximação proposta no Capítulo 3, fornece o embasamento teórico necessário para a amostragem de entradas e saídas desses sistemas. O uso dessa informação amostrada na estimação de parâmetros via otimização, Capítulo 4, e a escolha do intervalo  $\Delta t$  nos métodos contínuos do Capítulo 5 ficam assim fundamentados.

Como segunda contribuição do trabalho, citam-se os métodos de estimação de parâmetros propriamente ditos, desenvolvidos nos Capítulos 4 e 5. O método de otimização seria uma aplicação básica de mínimos quadrados ao problema, abrindo caminho para o uso de algoritmos e técnicas de análise mais sofisticadas. Os métodos escalares, por sua vez, indicam como o problema de estimação de parâmetros pode ser resolvido através de estruturas simples, com filtros SISO (*single-input, single-output*). No caso de múltiplos parâmetros, a decomposição proposta também determina quais desses parâmetros podem ou não ser de fato estimados. O projeto formal dos filtros SISO utilizados, *cf.* a metodologia empírica aqui adotada, poderá ser explorada em trabalhos posteriores.

Por fim, há a questão de aplicações onde parâmetros estimados contribuem para o controle de sistemas LPV. O exemplo iniciado na Seção 1.2.1 e concluído na Seção 5.4.3 é apenas uma amostra do que pode ser explorado em trabalhos futuros. Outros temas para pesquisa serão discutidos um pouco mais a seguir.

## 6.2 Perspectivas de Trabalhos Futuros

O refinamento dos métodos de estimação propostos e a adaptação de projetos de controladores para sistemas LPV preexistentes tal que a informação estimada possa ser utilizada já foram mencionados na seção anterior. Há também uma vertente teórica motivada pelo problema de estimação de parâmetros, que seria o estudo mais profundo de “observabilidade” e de “controlabilidade” para sistemas LPV. No sentido convencional, essas duas palavras transmitem as idéias de se determinar o estado de um sistema e de capacidade de se levar esse estado, por meio de entradas de controle, até um valor desejado. Os parâmetros podem afetar esses dois processos, assumindo valores que podem tornar um ou mais estados do sistema indetectáveis, ou valores que tornem o sistema incontrollável / não-estabilizável. Além disso, uma vez que os parâmetros podem ser vistos como parte do estado do sistema LPV quando o mesmo é tratado como não-linear de dimensão  $\mathbb{R}^{n+p}$ , seria possível se falar num conceito de “observabilidade” dos parâmetros. Essa metáfora de “observabilidade” formalizaria a idéia de que valores do estado do sistema, em  $\mathbb{R}^n$ , podem tornar um ou mais parâmetros indetectáveis ( $\nabla f(\Theta) \rightarrow 0$  no Capítulo 4 ou  $f(x(t)) \rightarrow 0$  no Capítulo 5).

Também em termos de teoria, outra possibilidade interessante seria reinterpretar os resultados obtidos no Capítulo 3 e aspectos dos métodos de estimação de parâmetros do ponto de vista da geometria diferencial. O apêndice A apresenta alguns ponteiros nessa direção, com foco na solução de sistemas variantes no tempo através da série de Magnus. Um exemplo seria o uso de métodos de Runge-Kutta de ordens crescentes para se calcular a função exponencial da matriz do sistema, a fim de se obter a matriz de transição de estados. Em sistemas LPV, uma aproximação capaz de aproveitar a linearidade (afinidade) da matriz em relação aos parâmetros poderia ser mais adequada para computar a solução das equações diferenciais, e talvez uma representação dessa resposta mais adequada para a estimação dos parâmetros.

Do ponto de vista de controle de sistemas LPV, além da adaptação de controladores parametrizados projetados *a priori* e alimentados com os parâmetros estimados, outra idéia seria a computação de controladores *on-line*, integrando estimação de parâmetros e controle. Isso viabilizaria estratégias onde a incerteza sobre uma estimativa instantânea dos parâmetros possa ser usada no projeto de um controlador robusto, válido para aquele instante. A mesma idéia de estimação *on-line* e incerteza variante no tempo poderia também ser usada para se obter melhores aproximações LPV para sistemas não lineares, ou no problema de estimação de estados. Incluiriam-se nesse último caso o refinamento de métodos de Kalman robusto [48, 49], por exemplo: a estimação dos parâmetros poderia ser usada para refinar a estimação de estados e vice-versa, num processo iterativo.

# Apêndice A

## Teoria de Magnus

Este apêndice traz uma revisão sobre a série de Magnus, um método alternativo para solução de equações diferenciais variantes no tempo. Embora não se desenvolva no presente trabalho nenhum método de estimação de parâmetros diretamente baseado na teoria aqui apresentada, tal possibilidade fica em aberto. Além disso, é interessante mencioná-la não só por completeza, mas também pela nova perspectiva que se obtém em relação ao restante do trabalho.

### A.1 Solução de Sistemas Lineares Variantes no Tempo

A fim de se obter uma melhor perspectiva sobre os sistemas LPV, o caso mais geral dos sistemas LTV deve ser estudado antes: em particular, a solução temporal de (1.2). O material apresentado na primeira parte desta seção é padrão, baseado em [50]; embora seja uma boa introdução ao problema, essa teoria pode não ser a melhor escolha para se explorar a estimação de parâmetros LPV. Assim, a segunda parte desta seção apresenta o trabalho de Magnus [51], que parece ser um ponto de partida viável para se obter tanto resultados práticos quanto teóricos sobre sistemas LPV.

#### A.1.1 Solução Padrão

O primeiro passo para se resolver (1.2) é obter a solução do caso homogêneo, onde a entrada exógena  $u(t)$  é zero e a dinâmica do sistema é excitada apenas por uma condição inicial  $x(t_0)$ . Quatro casos distintos serão analisados, desde o caso escalar com coeficientes constantes até o matricial, variante no tempo:

$$\begin{aligned} \dot{x}(t) &= ax(t), & \dot{x}(t) &= a(t)x(t), \\ \dot{x}(t) &= Ax(t), & \dot{x}(t) &= A(t)x(t). \end{aligned}$$

Nas equações acima,  $a$  é uma constante escalar;  $a(t)$  é uma função escalar do tempo;  $A$  é uma matriz constante;  $A(t)$  é uma matriz variante no tempo. Assume-se que, para um dado estado inicial  $x(t_0) = x_0$  (sem perda de generalidade, seja  $t_0 = 0$ ), a solução seja única: para os casos variantes no tempo, é suficiente assumir que  $a(t)$  ou os elementos de  $A(t)$  sejam funções contínuas do tempo.

A solução mais simples é aquela de  $\dot{x}(t) = ax(t)$ : se a condição inicial é  $x_0$ , então

$$x(t) = e^{at} x_0 .$$

Verificando,

$$\dot{x}(t) = \frac{d}{dt} (e^{at} x_0) = e^{at} a x_0 = a (e^{at} x_0) = ax(t) ,$$

portanto a equação é satisfeita.

O caso matricial  $\dot{x}(t) = Ax(t)$  é quase tão direto quanto acima: a solução

$$x(t) = e^{At} x_0 ,$$

satisfaz

$$\dot{x}(t) = \frac{d}{dt} (e^{At} x(t_0)) = e^{At} A x_0 = A (e^{At} x_0) = Ax(t) .$$

Embora multiplicação de matrizes não seja, em geral, comutativa, essa propriedade vale no caso acima pois

$$e^{At} = I + tA + \frac{t^2}{2!} A^2 + \dots = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k ,$$

portanto pré- ou pós-multiplicar  $e^{At}$  por  $A$  resulta na mesma expressão.

A solução de  $\dot{x}(t) = a(t)x(t)$  é

$$x(t) = \exp \left( \int_0^t a(\tau) d\tau \right) x_0 ,$$

pois

$$\begin{aligned} \dot{x}(t) &= \frac{d}{dt} \left( \exp \left( \int_0^t a(\tau) d\tau \right) x_0 \right) = \exp \left( \int_0^t a(\tau) d\tau \right) a(t) x_0 \\ &= a(t) \left( \exp \left( \int_0^t a(\tau) d\tau \right) x_0 \right) = a(t) x(t) . \end{aligned}$$

A comutatividade das matrizes é, uma vez mais, essencial.

Infelizmente, a solução acima não pode ser estendida a  $\dot{x}(t) = A(t)x(t)$  pois, nesse caso, a matriz  $A(t)$  e sua exponencial não se comutam:

$$\exp\left(\int_0^t A(\tau)d\tau\right) = I + \int_0^t A(\tau)d\tau + \frac{1}{2!}\left(\int_0^t A(\tau)d\tau\right)\left(\int_0^t A(\sigma)d\sigma\right) + \dots,$$

e então

$$\frac{d}{dt}\left(\exp\left(\int_0^t A(\tau)d\tau\right)\right) = 0 + A(t) + \frac{1}{2}A(t)\left(\int_0^t A(\sigma)d\sigma\right) + \frac{1}{2}\left(\int_0^t A(\tau)d\tau\right)A(t) + \dots \neq A(t)\exp\left(\int_0^t A(\tau)d\tau\right).$$

Para se resolver a equação matricial LTV depende-se da uma hipótese extra que, se os elementos de  $A(t)$  forem funções contínuas do tempo, então para cada estado inicial  $x^{(i)}(t_0)$  deve haver uma solução única  $x^{(i)}(t)$ . Se  $x(t) \in \mathbb{R}^n$  e  $A(t) \in \mathbb{R}^{n \times n}$ , então deve ser possível selecionar  $n$  vetores iniciais  $x^{(i)}(t_0)$  linearmente independentes (LI), e então construir uma matriz não-singular  $X(t_0) = [x^{(1)}(t_0) \ x^{(2)}(t_0) \ \dots \ x^{(n)}(t_0)]$ . Assumindo que as soluções  $x^{(i)}(t)$  LI correspondentes a tais condições iniciais possam ser de alguma forma obtidas, então uma (não única) *matriz fundamental* pode ser definida:

$$X(t) \triangleq [x^{(1)}(t) \ x^{(2)}(t) \ \dots \ x^{(n)}(t)] . \tag{A.1}$$

Tal matriz satisfaz, por construção, a equação diferencial

$$\dot{X}(t) = A(t)X(t) . \tag{A.2}$$

A matriz fundamental é não-singular para todo  $t$ , caso contrário haveria algum vetor não-nulo  $v$  tal que  $x(t_1) = X(t_1)v = 0$ , mas então  $x(t) = X(t)v = 0$  para todo  $t$ , em particular para  $t = t_0$ : essa última afirmação contrariaria a hipótese de independência linear dos  $x^{(i)}(t_0)$ .

Embora a matriz fundamental  $X(t)$  não seja única, a *matriz de transição de estados* definida como:

$$\Phi(t, t_0) \triangleq X(t)X^{-1}(t_0) \tag{A.3}$$

é de fato a solução única para

$$\dot{\Phi}(t, t_0) = A(t)\Phi(t, t_0) \tag{A.4}$$

com a condição inicial  $\Phi(t_0, t_0) = I$ . Algumas propriedades importantes de  $\Phi$  são:

$$\begin{aligned}\Phi(t, t) &= I, \\ \Phi^{-1}(t, t_0) &= \Phi(t_0, t), \text{ e} \\ \Phi(t, t_0) &= \Phi(t, t_1)\Phi(t_1, t_0), \forall t, t_0, t_1.\end{aligned}$$

Dada (A.4), a solução da equação homogênea  $\dot{x}(t) = A(t)x(t)$  se torna

$$x(t) = \Phi(t, t_0)x(t_0), \quad (\text{A.5})$$

pois

$$\begin{aligned}\dot{x}(t) &= \dot{\Phi}(t, t_0)x(t_0) \\ &= A(t)\Phi(t, t_0)x(t_0) \\ &= A(t)x(t).\end{aligned}$$

A resposta completa de (1.2) é a soma da solução homogênea acima, (A.5), e a solução forçada pela entrada  $u(t)$ . Essa última é obtida pela convolução de  $\Phi(t, t_0)$  e  $B(t)u(t)$ , resultando em:

$$\begin{aligned}x(t) &= \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \tau)B(\tau)u(\tau)d\tau \\ &= \Phi(t, t_0) \left( x(t_0) + \int_{t_0}^t \Phi(t_0, \tau)B(\tau)u(\tau)d\tau \right) \\ y(t) &= C(t) \left( \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \tau)B(\tau)u(\tau)d\tau \right) + D(t)u(t)\end{aligned} \quad (\text{A.6})$$

A expressão acima tem valor limitado do ponto de vista computacional, pois o cálculo de  $\Phi(t, t_0)$  não é trivial exceto em alguns casos especiais; ainda assim, (A.6) tem seu valor do ponto de vista teórico.

Considerando agora a equação discreta

$$\begin{aligned}x[k+1] &= A[k]x[k] + B[k]u[k] \\ y[k] &= C[k]x[k] + D[k]u[k],\end{aligned} \quad (\text{A.7})$$

a solução para uma dada condição inicial  $x[k_0]$  e entrada  $u[k]$ , conhecida em  $k \geq k_0$ , pode ser calculada recursivamente. Seja  $\Phi[k, k_0]$  a matriz de transição, solução de

$$\Phi[k+1, k_0] = A[k]\Phi[k, k_0] \quad (\text{A.8})$$

com  $\Phi[k_0, k_0] = I$ . Apesar da analogia com (A.4), o cálculo de  $\Phi$  no caso discreto é bastante simples:

$$\Phi[k, k_0] = \begin{cases} A[k-1]A[k-2] \dots A[k_0], & k > k_0 \\ I, & k = k_0 \end{cases} \quad (\text{A.9})$$

Entretanto,  $\Phi[k, k_0]$  pode ser singular; além disso, só é válida para propagação no sentido positivo do tempo:

$$\Phi[k, k_0] = \Phi[k, k_1]\Phi[k_1, k_0], \quad \forall k \geq k_1 \geq k_0.$$

Com (A.9), a solução de (A.7) pode ser escrita como:

$$\begin{aligned} x[k] &= \Phi[k, k_0]x[k_0] + \sum_{m=k_0}^{k-1} \Phi[k, m+1]B[m]u[m] \\ y[k] &= C[k] \left( \Phi[k, k_0]x[k_0] + \sum_{m=k_0}^{k-1} \Phi[k, m+1]B[m]u[m] \right) + D[k]u[k]. \end{aligned} \quad (\text{A.10})$$

### A.1.2 Solução via Série de Magnus

Apesar de (A.6) teoricamente resolver (1.2) e portanto seu caso especial (1.3), o problema prático e não-trivial de se calcular a matriz de transição de estados permanece. Embora seja possível obter soluções numéricas via simulação, tais resultados sozinhos não podem ser base para o desenvolvimento da teoria. Felizmente a série de Magnus e a teoria dos grupos de Lie provêm não só resultados úteis para análise [52, 53, 54], quanto ainda uma alternativa do ponto de vista computacional [55, 56].

Começando com a equação diferencial

$$Y'(t) = A(t)Y(t), \quad t \geq 0, \quad Y(0) = Y_0 \in G, \quad (\text{A.11})$$

onde  $G$  é um *grupo de Lie* – um *manifold*<sup>1</sup> diferenciável, dotado de uma estrutura de grupo contínua em relação à topologia subjacente do *manifold* – e  $A(t) : \mathbb{R}^+ \rightarrow \mathfrak{g}$  é contínua no sentido de Lipschitz, *i.e.*, para todos  $t_1, t_2$  no domínio de  $A(t)$  há uma constante  $K \geq 0$  tal que  $|A(t_1) - A(t_2)| \leq K |t_1 - t_2|$ .  $\mathfrak{g}$  é a *álgebra de Lie*

<sup>1</sup>Em português, “variedade”. O termo em inglês, cujo uso também é comum, será mantido.

de  $G$ , a saber o espaço tangente a  $G$  – Conjunto de todos os possíveis valores de  $\gamma'(0)$ , onde  $\gamma(t) \in G$  é uma curva suave e  $\gamma(0) = I$ , o valor da identidade em  $G$ .

Da geometria diferencial, dado que a condição inicial de (A.11) pertence a  $G$ , sua solução deverá permanecer em  $G$  para todo  $t \geq 0$  [52]. Particularmente, se  $Y_0 = I$ , a identidade, então a solução pode ser dada por

$$Y(t) = \exp(\Omega(t)), \quad (\text{A.12})$$

com  $\Omega(t)$  a ser definido em breve. Pode-se argumentar que a expressão acima seja uma generalização razoável das equações estudadas na seção anterior: (A.4) é apenas (A.11) com a matriz de transição de estados na forma de uma exponencial, por exemplo.

Antes de continuar, é preciso introduzir alguma notação extra: defina-se  $\text{ad}^k(\cdot, \cdot)$  como

$$\text{ad}^k(p, q) \triangleq \begin{cases} p, & k = 0 \\ [\text{ad}^{k-1}(p, q), q], & k \geq 1, \end{cases}$$

onde  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  é o *Lie bracket* ou *comutador*

$$[A, B] \triangleq AB - BA.$$

Suas propriedades relevantes são:

- linearidade em cada componente,
- antissimetria:  $[A, B] = -[B, A]$ , e
- identidade de Jacobi:  $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$  .

Para que (A.12) seja de fato solução de (A.11),  $\Omega(t)$  com  $\Omega(0) = 0$  deverá satisfazer a equação diferencial implícita

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} \text{ad}^k(\Omega', \Omega) = A.$$

Magnus [51] reconheceu que a expressão acima poderia ser reescrita explicitamente como

$$\Omega' = A + \frac{1}{2} \text{ad}(A, \Omega) + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} \text{ad}^{2k}(A, \Omega),$$

onde  $B_m$  são números de Bernoulli.  $\Omega(t)$  pode então ser escrita concisamente em termos de  $A(t)$ :

$$\begin{aligned} \Omega(t) = & \int_0^t A(\kappa) d\kappa + \frac{1}{2} \int_0^t \left[ A(\kappa), \int_0^\kappa A(\xi) d\xi \right] d\kappa + \\ & \frac{1}{4} \int_0^t \left[ A(\kappa), \int_0^\kappa \left[ A(\xi), \int_0^\xi A(\eta) d\eta \right] d\xi \right] d\kappa + \\ & \frac{1}{12} \int_0^t \left[ \left[ A(\kappa), \int_0^\kappa A(\xi) d\xi \right], \int_0^\kappa A(\eta) d\eta \right] d\kappa + \dots \end{aligned} \quad (\text{A.13})$$

A implementação prática de (A.13) e a solução de (A.11) via (A.12) estão sujeitas a três fontes de erro:

- Truncamento da expansão infinita;
- Discretização das integrais multivariáveis;
- Computação das exponenciais de matrizes.

Em [55] os autores utilizam teoria de grafos para lidar com o primeiro desses problemas e analisar o segundo. Quanto às exponenciais de matrizes, embora reconheçam que o problema existe e deva ser estudado mais a fundo, na prática assumem que computação até os limites de precisão de máquina é suficiente. Resultados numéricos de um método de quarta ordem baseado na série de Magnus são comparados com aqueles obtidos via Runge–Kutta e Gauss–Legendre. De qualquer forma, tal discussão foge do escopo deste trabalho: não se pretende usar aqui a série de Magnus para simulação de sistemas dinâmicos, mas sim para obter limites de erro na análise de sistemas LPV – tal discussão se encontra a seguir.

## A.2 Sistemas LPV e a Série de Magnus

Seja a matriz

$$A(t) = \sum_{i=1}^m X_i \theta_i(t) ,$$

e seja  $\mathcal{L}$  a álgebra de Lie gerada pelos  $X_i$  acima e pelo produto de comutação  $[X_j, X_k]$ . Sejam ainda  $\ell$  a dimensão (finita) de  $\mathcal{L}$ , e  $\{X_1, X_2, \dots, X_\ell\}$  uma base para  $\mathcal{L}$ .

Considere novamente a equação (A.11), desta vez com a matriz identidade como condição inicial:

$$\dot{Y}(t) = A(t)Y(t) , \quad Y(0) = I .$$

A solução por Magnus é da forma

$$Y(t) = \exp \left( \sum_{i=1}^{\ell} \omega_i(t) X_i \right),$$

mas só é válida em uma vizinhança da origem. Por outro lado, Wei e Norman [57] demonstraram que

$$Y(t) = \prod_{i=1}^{\ell} \exp(\phi_i(t) X_i) \quad (\text{A.14})$$

pode ser, com a base correta, uma solução global para todas as álgebras de Lie resolvíveis, e para quaisquer equações em  $\mathbb{R}^{2 \times 2}$ . Como esses autores mencionam, a principal vantagem de (A.14) é possibilitar a investigação das propriedades de  $Y(t)$  a partir das propriedades espectrais de cada um dos  $X_i$ , individualmente.

A equação diferencial

$$\sum_{k=1}^{\ell} \theta_k(t) X_k = \sum_{i=1}^{\ell} \sum_{k=1}^{\ell} \phi'_k(t) \xi_{ki} X_k \quad (\text{A.15})$$

onde os  $\xi_{ki}$  são funções analíticas dos  $\phi_i$ 's, permite relacionar as funções  $\phi_i(t)$  aos parâmetros  $\theta_i(t)$ . Uma vez que os operadores  $X_k$  são linearmente independentes, então é possível escrever

$$\Theta = \Xi \Phi' \text{ ou,}$$

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_\ell \end{bmatrix} = \begin{bmatrix} \xi_{11} & \cdots & \xi_{1\ell} \\ \vdots & & \vdots \\ \xi_{\ell 1} & \cdots & \xi_{\ell\ell} \end{bmatrix} \begin{bmatrix} \phi'_1 \\ \phi'_2 \\ \vdots \\ \phi'_\ell \end{bmatrix}, \quad \Phi(0) = 0,$$

onde, uma vez mais, os elementos  $\xi_{ki}$  da matriz de transformação  $\Xi$  são funções analíticas dos  $\phi_i$ 's. Por causa disso, o determinante  $\det(\Xi)$  deve também ser uma função analítica dos  $\phi_i$ 's e, visto que  $\Xi(0) = I$ , deve haver pelo menos alguma vizinhança  $N_0$  de  $t = 0$  em que  $\det(\Xi) \neq 0$ . Assim,

$$\frac{d}{dt} \Phi = f(\Theta, \Phi) = \Xi^{-1} \Theta, \quad \Phi(0) = 0, \quad t \in N_0 \quad (\text{A.16})$$

Como mencionado antes, se  $\mathcal{L}$  é solucionável, uma escolha adequada de base pode tornar a solução acima global, *i.e.*  $N_0$  pode ser arbitrariamente grande. A condição seria a existência de uma cadeia de *ideais*  $0 \subset L_\ell \subset L_{\ell-1} \subset \cdots \subset L_1 = \mathcal{L}$  onde cada  $L_m$  de dimensão  $\ell - m + 1$  seja gerado por  $X_m, \dots, X_\ell$ .

Neste ponto, a discussão será interrompida. Restaria encontrar a matriz  $\Xi$  que relaciona  $\Theta$  e  $\Phi$ , para que a relação entre os limites de banda dos parâmetros e do sistema possa ser determinada.

# Apêndice B

## Transformações de Equivalência e de Lyapunov

Transformações de similaridade são uma ferramenta muito útil no estudo de sistemas lineares invariantes no tempo, pois a mudança certa de variáveis pode decompor um problema de ordem elevada em um conjunto de equações diferenciais de menor ordem, desacopladas. Embora transformações estáticas também tenham sua utilidade para sistemas variantes no tempo, as chamadas *transformações de equivalência* abrem maiores possibilidades de simplificação. Este apêndice traz uma revisão do assunto, baseada em [50]. Do ponto de vista deste trabalho, um aspecto extra deve ser observado: se um ou mais parâmetros do sistema LPV entram na transformação, as respectivas derivadas em relação ao tempo também entram. Isso pode ser um problema do ponto de vista da estimação, mas é consistente com o uso de tais derivadas em estratégias de controle LPV encontradas na literatura citada anteriormente nesta tese.

Considere (1.2), sem ruídos, com  $x(t) \in \mathbb{R}^n$ , e seja  $P(t) \in \mathbb{R}^{n \times n}$  uma matriz não-singular contínua com derivada  $\dot{P}(t)$  também contínua para todo  $t$ . Se

$$\bar{x}(t) \triangleq P(t)x(t), \quad (\text{B.1})$$

então o sistema

$$\begin{aligned} \dot{\bar{x}}(t) &= \bar{A}(t)\bar{x}(t) + \bar{B}(t)u(t) \\ y(t) &= \bar{C}(t)\bar{x}(t) + \bar{D}(t)u(t) \end{aligned} \quad (\text{B.2})$$

com

$$\bar{A}(t) = \left( P(t)A(t) + \dot{P}(t) \right) P^{-1}(t) \quad \bar{B}(t) = P(t)B(t) \quad (\text{B.3})$$

$$\bar{C}(t) = C(t)P^{-1}(t) \quad \bar{D}(t) = D(t) \quad (\text{B.4})$$

é *algebricamente equivalente* a (1.2). Isto pode ser verificado substituindo  $\bar{x}(t) = P(t)x(t)$  e  $\dot{\bar{x}}(t) = \dot{P}(t)x(t) + P(t)\dot{x}(t)$  em (1.2).

Se  $X(t)$  em (A.1) é uma matriz fundamental<sup>1</sup> de (1.2), então

$$\bar{X}(t) = P(t)X(t) \quad (\text{B.5})$$

é uma matriz fundamental de (B.2). Lembrando que tanto  $X(t)$  quanto  $P(t)$  são não-singulares (esta última por hipótese), então  $\bar{X}(t)$  também o é. Escrevendo-se:

$$\begin{aligned} \dot{\bar{X}}(t) &= \dot{P}(t)X(t) + P(t)\dot{X}(t) \\ &= \dot{P}(t)X(t) + P(t)A(t)X(t) \\ &= \left( \dot{P}(t) + P(t)A(t) \right) (P^{-1}(t)P(t)) X(t) \\ &= \bar{A}(t) (P(t)X(t)) = \bar{A}(t)\bar{X}(t) , \end{aligned}$$

verifica-se (B.5).

A matriz transformada  $\bar{A}(t)$  pode ser feita constante ou mesmo nula. A fim de verificar isso, dado que  $X(t)$  é não-singular, escreve-se

$$X^{-1}(t)X(t) = I ,$$

e deriva-se a expressão acima quanto a  $t$ :

$$\begin{aligned} \dot{X}^{-1}(t)X(t) + X^{-1}(t)\dot{X}(t) &= 0 \\ \Rightarrow \dot{X}^{-1}(t) &= -X^{-1}(t)\dot{X}(t)X^{-1}(t) \\ &= -X^{-1}(t)A(t)X(t)X^{-1}(t) \quad (\text{de A.2}) \\ &= -X^{-1}(t)A(t). \end{aligned} \quad (\text{B.6})$$

Se  $\bar{A}(t) = A_c$ , então  $\bar{X}(t) = e^{A_c t}$  é uma matriz fundamental de  $\dot{\bar{x}}(t) = \bar{A}(t)\bar{x}(t) = A_c\bar{x}(t)$ . De (B.5),

$$P(t) = \bar{X}(t)X^{-1}(t) = e^{A_c t} X^{-1}(t) ,$$

assim,

$$\begin{aligned} \bar{A}(t) &= \left( P(t)A(t) + \dot{P}(t) \right) P^{-1}(t) \\ &= \left( e^{A_c t} X^{-1}(t)A(t) + A_c e^{A_c t} X^{-1}(t) + e^{A_c t} \dot{X}^{-1}(t) \right) X(t)e^{-A_c t} \\ &= \left( e^{A_c t} X^{-1}(t)A(t) + A_c e^{A_c t} X^{-1}(t) - e^{A_c t} X^{-1}(t)A(t) \right) X(t)e^{-A_c t} \\ &= A_c e^{A_c t} X^{-1}(t)X(t)e^{-A_c t} \\ &= A_c. \end{aligned}$$

---

<sup>1</sup>Vide Seção A.1.1

Note que se  $A_c = 0$  então  $P(t) = X^{-1}(t)$  e  $\bar{A}(t) = 0$ , isto é, é possível eliminar a realimentação de estado em (1.2). Isto não é possível no caso invariante no tempo, *i.e.* com uma transformação de similaridade; por outro lado, se  $P(t)$  é uma constante  $P_c$ , então equivalência algébrica se reduz de fato a similaridade.

Se  $P(t)$  é não-singular, e  $P(t)$  e  $\dot{P}(t)$  são ambas contínuas e limitadas para todo  $t$ , então (B.1), (B.2), (B.3) definem uma chamada *transformação de Lyapunov*. Sob tais condições mais estritas, nem sempre é possível fazer  $\bar{A}(t)$  constante, menos ainda zero. Somente sistemas periódicos (*e.g.*  $A(t + T) = A(t)$ ) garantidamente têm transformações de Lyapunov que resultam em matrizes  $A$  constantes (teoria de Floquet).

# Referências Bibliográficas

- [1] J. E. Marsden, M. J. Hoffman, Elementary Classical Analysis, 2nd Edition, W. H. Freeman and Company, 1993.
- [2] A. Abdullah, M. Zribi, Model reference control of LPV systems, Journal of the Franklin Institute 346 (9) (2009) 854–871, doi:10.1016/j.jfranklin.2009.04.006.
- [3] D. J. Stilwell, W. J. Rugh, Stability and  $\mathcal{L}_2$  gain properties of LPV systems, Automatica 38 (2002) 1601–1606.
- [4] F. Blanchini, S. Miani, Stabilization of LPV systems: State feedback, state estimation, and duality, SIAM Journal on Control and Optimization 42 (1) (2003) 76–97.
- [5] C. E. de Souza, A. Trofino, J. de Oliveira, Parametric Lyapunov function approach to  $\mathcal{H}_2$  analysis and control of linear parameter–dependent systems, IEE Proceedings – Control Theory & Applications 150 (5) (2003) 501–508.
- [6] L. Xie, C. E. de Souza, Robust  $\mathcal{H}_\infty$  control for linear systems with norm–bounded time–varying uncertainty, IEEE Transactions on Automatic Control 37 (8) (1992) 1188–1191.
- [7] C. E. de Souza, K. A. Barbosa, A. Trofino Neto, Robust  $\mathcal{H}_\infty$  filtering for discrete–time linear systems with uncertain time–varying parameters, IEEE Transactions on Signal Processing 54 (6) (2006) 2110–2118.
- [8] W. Xie, T. Eisaka, Design of LPV control systems based on Youla parametrization, IEE Proceedings – Control Theory & Applications 151 (4) (2004) 465–472.
- [9] H. K. Khalil, Nonlinear Systems, 3rd Edition, Prentice Hall, 2002.
- [10] J. S. Shamma, M. Athans, Analysis of gain scheduled control for nonlinear plants, IEEE Transactions on Automatic Control 35 (8) (1990) 898–907.
- [11] D. J. Stilwell, W. J. Rugh, Interpolation of observer state feedback controllers for gain scheduling, IEEE Transactions on Automatic Control 44 (6) (1999) 1225–1229.

- [12] D. J. Stilwell, W. J. Rugh, Stability preserving interpolation methods for the synthesis of gain scheduled controllers, *Automatica* 36 (2000) 665–671.
- [13] T. S. Kuhn, *The Structure of Scientific Revolutions*, 3rd Edition, University Of Chicago Press, 1996.
- [14] D. Simon, *Optimal State Estimation: Kalman,  $\mathcal{H}_\infty$ , and Nonlinear Approaches*, Wiley, 2006.
- [15] G. F. Franklin, J. D. Powell, A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 3rd Edition, Addison–Wesley, 1995.
- [16] G. F. Franklin, J. D. Powell, M. L. Workman, *Digital Control of Dynamic Systems*, 3rd Edition, Addison–Wesley, 1998.
- [17] L. Ljung, *System Identification: A Theory for the User*, 2nd Edition, Prentice Hall, 1998.
- [18] L. A. Aguirre, *Introdução À Identificação de Sistemas: Técnicas Lineares e Não–Lineares Aplicadas a Sistemas Reais*, segunda Edition, Editora UFMG, 2004.
- [19] L. A. Aguirre, B. O. S. Teixeira, L. A. B. Torres, Using data–driven models and the unscented Kalman filter to estimate unobserved variables of nonlinear systems, *Physical Review E - Statistical Physics, Plasmas, Fluids and Related Interdisciplinary Topics* 72 (2) (2005) Paper 026226.
- [20] B. O. S. Teixeira, L. A. B. Torres, L. A. Aguirre, P. H. A. Iscold, Flight path reconstruction using the unscented Kalman filter algorithm, in: *International Congress of Mechanical Engineering 18th COBEM*, 2005.
- [21] B. O. S. Teixeira, L. A. B. Tôrres, L. A. Aguirre, D. S. Bernstein, On unscented Kalman filtering with state interval constraints, *Journal of Process Control* 20 (2010) 45–57.
- [22] L. C. A. Souza, R. M. Palhares, Spectral radius and sample rate bounding for linear parameter varying systems (A ser submetido).
- [23] F. Wirth, On Lipschitz continuity of the top Lyapunov exponent of linear parameter varying and linear switching systems, *Stochastics and Dynamics* 4 (3) (2004) 461–481.
- [24] F. Wirth, A converse Lyapunov theorem for linear parameter-varying and linear switching systems, *SIAM Journal on Control and Optimization* 44 (1) (2005) 210–239.
- [25] O. Toker, On the complexity of the robust stability problem for linear parameter varying systems, *Automatica* 33 (11) (1997) 2015–2017.

- [26] V. D. Blondel, J. N. Tsitsiklis, Complexity of stability and controllability of elementary hybrid systems, *Automatica* 35 (3) (1999) 479–489.
- [27] V. D. Blondel, J. N. Tsitsiklis, The boundedness of all products of a pair of matrices is undecidable, *Systems & Control Letters* 41 (2) (2000) 135–140.
- [28] V. D. Blondel, Y. Nesterov, J. Theys, On the accuracy of the ellipsoid norm approximation of the joint spectral radius, *Linear Algebra and its Applications* 394 (2005) 91–107.
- [29] A. Ghosh, S. Boyd, Upper bounds on algebraic connectivity via convex optimization, *Linear Algebra and its Applications* 418 (2006) 693–707.
- [30] I. M. Gelfand, Normierte Ringe, *Rec. Math. [ Mat. Sbornik ] N. S.* 9 (51) (1941) 3–24.
- [31] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [32] C. E. Shannon, Communication in the presence of noise, *Proceedings of the IRE* 37 (1) (1949) 10–21.
- [33] H. Nyquist, Certain topics in telegraph transmission theory, *Transactions of the AIEE* 47 (1928) 617–644.
- [34] M. Unser, Sampling - 50 years after Shannon, *Proceedings of the IEEE* 88 (4) (2000) 569–587.
- [35] D. M. Grobman, Homeomorphisms of systems of differential equations, *Doklady Akademii Nauk SSSR* 128 (1959) 880–881.
- [36] P. Hartman, A lemma in the theory of structural stability of differential equations, *Proceedings of the American Mathematical Society* 11 (4) (1960) 610–620.
- [37] P. Hartman, On local homeomorphisms of euclidean spaces, *Boletín Sociedad Matemática Mexicana* 5 (1960) 220–241.
- [38] S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control*, Springer, 1999.
- [39] L. W. Couch II, *Digital and Analog Communication Systems*, sixth Edition, Prentice Hall, 2001.
- [40] J. R. Carson, Notes on the theory of modulation, *Proceedings of the IRE* 10 (1) (1922) 57–64.
- [41] D. P. Bertsekas, *Nonlinear Programming*, 2nd Edition, Athena Scientific, 1999.

- [42] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [43] L. C. A. Souza, R. M. Palhares, Estimação de parâmetros em sistemas lineares variantes no tempo, in: XVII Congresso Brasileiro de Automática, 2008.
- [44] L. C. A. Souza, R. M. Palhares, Parameter estimation on linear time-varying systems, *Journal of the Franklin Institute* (Submetido em 2009).
- [45] K. A. Ross, *Elementary Analysis: The Theory of Calculus*, reprint Edition, Springer, 2000.
- [46] A. V. Oppenheim, A. S. Willsky, I. T. Young, *Signals and Systems*, Prentice Hall, 1983.
- [47] K. H. Rosen, *Discrete Mathematics and Its Applications*, 4th Edition, WCB/McGraw-Hill, 1999.
- [48] L. Xie, Y. C. Soh, C. E. de Souza, Robust Kalman filtering for uncertain discrete-time systems, *IEEE Transactions on Automatic Control* 39 (6) (1994) 1310–1314.
- [49] Y. K. Foo, Y. C. Soh, Robust Kalman filtering for uncertain discrete-time systems with probabilistic parameters bounded within a polytope, *Systems & Control Letters* 57 (2008) 482–488.
- [50] C.-T. Chen, *Linear System Theory and Design*, 3rd Edition, Oxford University Press, 1999.
- [51] W. Magnus, On the exponential solution of differential equations for a linear operator, *Communications on Pure and Applied Mathematics* 7 (4) (1954) 649–673.
- [52] A. Iserles, S. P. Nørsett, On the solution of linear differential equations in Lie groups, *Philosophical Transactions of the Royal Society of London Series A – Mathematical Physical and Engineering Sciences* 357 (1754) (1999) 983–1019.
- [53] E. Celledoni, A. Iserles, Approximating the exponential from a Lie algebra to a Lie group, *Mathematics of Computation* 69 (232) (2000) 1457–1480.
- [54] R. Suárez, L. Sáenz, Lie algebras associated with the exponential solutions of nonautonomous linear differential equations, *Journal of Mathematical Physics* 42 (9) (2001) 4582–4605.
- [55] A. Iserles, A. Marthinsen, S. P. Nørsett, On the implementation of the method of Magnus series for linear differential equations, *BIT Numerical Mathematics* 39 (2) (1999) 281–304.

- [56] C. Moan, J. Niesen, Convergence of the Magnus series, arXiv:math/0609198v2.
- [57] J. Wei, E. Norman, On global representations of the solutions of linear differential equations as a product of exponentials, Proceedings of the American Mathematical Society 15 (2) (1964) 327–334.