



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-14620-TDI/1199

**USO DE CONTROLE GEOMÉTRICO EM SINCRONIZAÇÃO
APLICADO À ESTIMAÇÃO DE PARÂMETROS DE
SISTEMAS DINÂMICOS**

Ubiratan Santos Freitas

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Elbert Einstein Nehrer Macau, aprovada em 11 de dezembro de 2006.

INPE
São José dos Campos
2007

Publicado por:

esta página é responsabilidade do SID

Instituto Nacional de Pesquisas Espaciais (INPE)

Gabinete do Diretor – (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 – CEP 12.245-970

São José dos Campos – SP – Brasil

Tel.: (012) 3945-6911

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

**Solicita-se intercâmbio
We ask for exchange**

Publicação Externa – É permitida sua reprodução para interessados.



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-14620-TDI/1199

**USO DE CONTROLE GEOMÉTRICO EM SINCRONIZAÇÃO
APLICADO À ESTIMAÇÃO DE PARÂMETROS DE
SISTEMAS DINÂMICOS**

Ubiratan Santos Freitas

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Elbert Einstein Nehrer Macau, aprovada em 11 de dezembro de 2006.

INPE
São José dos Campos
2007

681.511.4

Freitas, U. S.

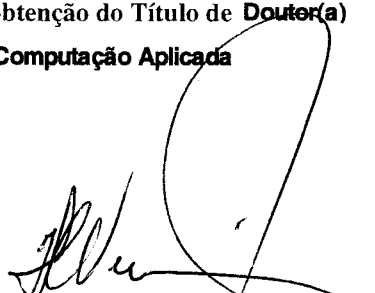
Uso de controle geométrico em sincronização aplicado à
estimação de parâmetros de sistemas dinâmicos / Ubiratan
Santos Freitas. – São José dos Campos: INPE, 2006.

116p. ; (INPE-14620-TDI/1199)

1.Sincronismo. 2.Sistemas lineares. 3.Identificação de
parâmetros. 4.Teoria de controle. 5.Identificação de sistemas.
I.Título.


Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de **Doutor(a)** em
Computação Aplicada

Dr. Haroldo Fraga de Campos Velho



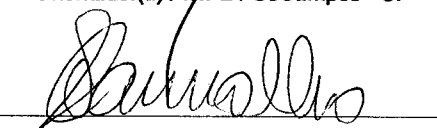
Presidente / INPE / SJC Campos - SP

Dr. Elbert Einstein Nehrer Macau



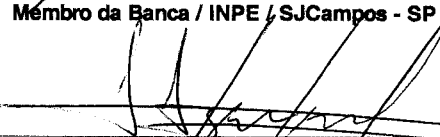
Orientador(a) / INPE / SJC Campos - SP

Dr. Solon Venâncio de Carvalho



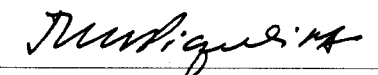
Membro da Banca / INPE / SJC Campos - SP

Dr. Luís Antônio Aguirre



Convidado(a) / UFMG / Belo Horizonte - MG

Dr. José Roberto Castilho Piqueira



Convidado(a) / USP / São Paulo - SP

Aluno (a): **Ubiratan Santos Freitas**

São José dos Campos, 11 de Dezembro de 2006

AGRADECIMENTOS

Diversas pessoas contribuíram para que este trabalho pudesse existir e chegar a termo. Eu gostaria, portanto, de lhes registrar aqui o meu agradecimento.

O primeiro agradecimento vai para os meus pais que, com seu esforço, fizeram com que eu tivesse a oportunidade de chegar até aqui. Um saudoso obrigado eu deixo a minha mãe, Lucília, que viu o começo desse trabalho mas não viveu para ver a sua conclusão. A meu pai, João Eudes, eu devo o meu interesse pela ciência. Obrigado a vocês dois.

O próximo agradecimento vai para o meu orientador, Elbert, por sua paciência e perseverança. Sem ele, este trabalho não teria chegado ao fim.

Gostaria também de agradecer a minha companheira, Juliana, por partilhar comigo o caminho da vida, nos melhores e nos piores momentos.

Sou grato também aos inúmeros amigos e colegas, alguns de longa data, outros nem tanto, que me apoiaram e animaram durante o período deste trabalho. Em particular, gostaria de lembrar dos amigos de BH, da Eng. Elétrica da UFMG, alguns agora em S. J. dos Campos, dos amigos e colegas do INPE, e dos amigos de república. A todos vocês o meu muito obrigado.

Um agradecimento importante vai para a FAPESP pelo profissionalismo e pelo apoio financeiro via o processo número 01/14089-6, e ao CNPq.

RESUMO

Este trabalho trata da proposição de um método de estimação de parâmetros para sistemas dinâmicos não lineares baseado no uso de técnicas de controle não linear e sincronização. A idéia principal do método é sincronizar o modelo com o sistema e, concomitantemente, estimar os parâmetros do sistema. A sincronização é obtida através de técnicas de controle não linear geométrico e os parâmetros são estimados usando um método adaptativo. As condições para ocorrência da sincronização são estudadas e a viabilidade do método é avaliada usando exemplos de simulação com os sistemas caóticos de Lorenz e Rössler.

USE OF GEOMETRIC CONTROL IN SYNCHRONIZATION APPLIED TO PARAMETER ESTIMATION OF DYNAMICAL SYSTEMS

ABSTRACT

This work proposes a parameter estimation method for nonlinear dynamical systems based on nonlinear control techniques and synchronization. The central idea of the method is to synchronize the model with the system, and, in parallel, to estimate the system's parameters. Synchronization is achieved by the use of nonlinear geometric control techniques, and the parameters are estimated using an adaptive method. The conditions for synchronization are studied, and the method's feasibility is evaluated using simulation examples with the Lorenz and Rössler chaotic systems.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

CAPÍTULO 1 - INTRODUÇÃO	19
1.1 - Motivação	22
1.1.1 - Sincronização na estimação de parâmetros	24
1.2 - Estrutura do texto	25
CAPÍTULO 2 - IDENTIFICAÇÃO DE SISTEMAS	27
2.1 - Introdução	27
2.2 - Definição	28
2.2.1 - Conceitos importantes	28
2.2.2 - Definição do problema de identificação de sistemas	29
2.3 - Procedimento de identificação	30
2.3.1 - Coleta de dados	31
2.3.2 - Escolha da representação	32
2.3.3 - Determinação da estrutura e estimação de parâmetros	33
2.3.4 - Validação do modelo	36
2.4 - Identificação de sistemas como um problema inverso	38
2.5 - Problemas da identificação de sistemas tradicional	40
2.6 - Conclusão	44
CAPÍTULO 3 - SISTEMAS DINÂMICOS CAÓTICOS E SIN- CRONIZAÇÃO	47
3.1 - Introdução	47
3.2 - Sistemas dinâmicos: definições	48
3.3 - Sincronização de sistemas caóticos	53
3.3.1 - Substituição completa	55
3.3.2 - Observadores como forma de sincronização	60
3.3.3 - Sincronização como um problema de controle	62
3.4 - Conclusão	62

CAPÍTULO 4 - SISTEMAS DE CONTROLE NÃO LINEAR	63
4.1 - Introdução	63
4.2 - Controle geométrico	64
4.2.1 - Linearização por realimentação de estado	67
4.3 - Controle Adaptativo	71
CAPÍTULO 5 - MÉTODO DE ESTIMAÇÃO DE PARÂMETROS	75
5.1 - Introdução	75
5.2 - Representação	76
5.3 - Versão monoparamétrica	77
5.4 - Versão multiparamétrica	80
5.4.1 - Controladores para a estimação de parâmetros	84
CAPÍTULO 6 - RESULTADOS	89
6.1 - Versão monoparamétrica	89
6.2 - Versão multiparamétrica	96
6.2.1 - Um estado observado — observador alto ganho	96
6.2.2 - Todos os estados —Controlador de alto ganho	101
CAPÍTULO 7 - CONCLUSÕES E PROPOSTAS DE CONTINUI- DADE	107
7.1 - Conclusões	107
7.2 - Propostas de continuidade	108
REFERÊNCIAS BIBLIOGRÁFICAS	109

LISTA DE FIGURAS

	<u>Pág.</u>
1.1 Esquema de circuito a ser modelado	20
2.1 Problema direto na identificação de sistemas. As causas são associadas à estrutura e aos parâmetros do modelo. Os efeitos são a evolução temporal do sistema, ou seja, os dados.	39
3.1 Órbita com a condição inicial $\mathbf{x}_0 = [1, 1, 1]^T$ para o sistema (3.2).	49
3.2 Trajetória ou órbita no espaço de estados.	50
3.3 Componente x_1 de duas órbitas não caóticas com condições iniciais distintas.	51
3.4 Projeção no plano x_1x_2 de duas órbitas não caóticas com condições iniciais distintas.	52
3.5 Evolução temporal das distâncias no espaço de estados de duas órbitas não caóticas com condições iniciais distintas.	52
3.6 Componente x_1 de duas órbitas caóticas com condições iniciais distintas.	53
3.7 Projeção no plano x_1x_2 de duas órbitas caóticas com condições iniciais distintas.	54
3.8 Evolução temporal das distâncias no espaço de estados de duas órbitas caóticas com condições iniciais distintas.	54
3.9 Representação gráfica do esquema de sincronização mestre-escravo para o sistema de Lorenz com substituição completa.	59
3.10 Projeções no plano (x_1, x_3) dos osciladores mestre e escravo para o início da simulação.	59

3.11	Projeções no plano (x_1, x_3) dos osciladores mestre e escravo para a simulação completa.	60
4.1	Pêndulo simples	65
4.2	Simulação do pêndulo utilizando controle. O ângulo está representado em vermelho, em azul está a velocidade angular e em verde a ação de controle.	66
4.3	Esquema do controle via linearização por realimentação de estados. CL é o controlador projetado para o sistema linear e SNL representa o sistema não linear.	71
4.4	Esquema de um controlador adaptativo do tipo auto-ajustável.	72
5.1	Diagrama em blocos do sistema de controle para a estimação monoparamétrica.	80
6.1	Variável x em função do tempo para o sistema original e para o modelo.	95
6.2	Projeção no plano xz dos atratores do sistema e do modelo.	95
6.3	Ação de controle.	96
6.4	Evolução temporal do erro de sincronização entre o modelo e o sistema, usando apenas a variável x_1 como observável.	97
6.5	Evolução temporal dos parâmetros estimados usando apenas a variável x_1 como observável.	98
6.6	Evolução temporal do erro de estimação dos parâmetros usando apenas a variável x_1 como observável.	98
6.7	Evolução temporal da variável x do modelo com os parâmetros estimados.	99
6.8	Projeção no plano xy do atrator do modelo com os parâmetros estimados.	100
6.9	Evolução temporal dos parâmetros θ_2 e θ_8	103
6.10	Evolução temporal dos parâmetros θ_5 e θ_9	105

6.11 Projeções planas dos atratores do sistema original (esquerda) e do modelo com parâmetros estimados (direita).	105
---	-----

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Causas e efeitos do PDC e do PDR	40
2.2 Parâmetros estimados usando diferentes estratégias.	42
2.3 Expoentes de Lyapunov para o sistema e os modelos identificados	43
6.1 Parâmetros estimados para o sistema de Lorenz usando apenas x_1 como observável.	98
6.2 Valores finais dos parâmetros	102
6.3 Valores finais dos parâmetros para o sistema de Rössler	105

CAPÍTULO 1

INTRODUÇÃO

A relevância deste trabalho está intimamente ligada à grande importância que os modelos, principalmente os matemáticos, possuem em ciência e tecnologia. Um modelo é um análogo de um sistema, que preserva algumas de suas características enquanto descarta outras (AGUIRRE, 2000a). Este trabalho concentra-se em modelos matemáticos para sistemas dinâmicos (MONTEIRO, 2002). A importância dos modelos se deve ao fato de que se pode substituir um sistema pelo seu modelo na hora de se fazer experimentos, tomados os devidos cuidados.

Essa aplicação dos modelos pode ser melhor descrita através de um exemplo. Seja o processo de construir um elevador. Tal elevador deve cumprir um conjunto de requisitos como suportar uma determinada carga, deslocar-se a uma certa velocidade, parar num determinado andar com uma certa precisão, entre outros. Como saber se um determinado projeto de elevador será capaz de cumprir esses requisitos? Uma alternativa é construir o elevador e depois testá-lo. Se o projeto estiver adequado, o trabalho está concluído. Se não, desmonta-se o elevador, altera-se o projeto, e constrói-se outro elevador. Essa abordagem apresenta desvantagens. Ela é demorada e cara, pois em cada ciclo um novo elevador deve ser construído e testado, gastando recursos e tempo. Ela é arriscada, pois durante um teste, uma falha pode levar a danos pessoais ou materiais. O uso de um modelo para o elevador projetado pode simplificar o processo. Ao invés de testar um elevador real, pode-se simular o modelo e verificar se ele cumpre os requisitos. Dessa forma, é possível detectar e corrigir alguns problemas no projeto sem que seja necessário construir o sistema.

O que os modelos fornecem é uma certa capacidade de prever o futuro. Em outras palavras, é possível usar modelos para *simular* um sistema. “Determinada estrutura será capaz de suportar determinada carga?” “Qual corrente passará neste circuito quando ligado?” “Esse sistema de controle é estável?” Esses são exemplos de perguntas que podem ser respondidas por simulação, ao invés de se realizar experimentos diretamente com os sistemas em questão. Naturalmente, a qualidade dessas respostas, ou seja, quão bem as respostas obtidas pela análise dos modelos refletem as respostas que seriam encontradas num experimento real, depende da qualidade do próprio modelo.

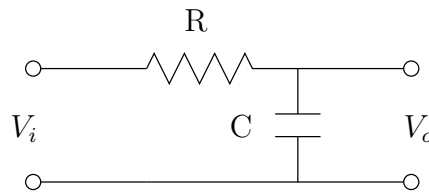


FIGURA 1.1 - Esquema de circuito a ser modelado

Atualmente, é inconcebível a construção de um sistema que apresente um mínimo de complexidade sem se realizar algum tipo de simulação antes. Um fator de grande importância para essa situação é a popularização e o desenvolvimento do computador pessoal, que permitiu a execução, de forma acessível, de simulações outrora inviáveis.

No entanto, todas essas possibilidades de análise partem do princípio da disponibilidade de um modelo adequado. Existem basicamente duas formas de obtenção de modelos, dependendo da origem das informações usadas para elaborá-los: a *identificação de sistemas* (LJUNG, 1987; AGUIRRE, 2000a) e a *modelagem por primeiros princípios* (GARCIA, 1997).

A identificação de sistemas é uma forma empírica de obtenção de modelos, baseada em experimentos e coleta de dados, enquanto a modelagem por primeiros princípios é um procedimento teórico, baseado em conhecimento científico já adquirido. O que difere as duas abordagens é, principalmente, a origem das informações necessárias para construir o modelo. No caso da modelagem por primeiros princípios, essas informações vêm de conhecimento científico prévio. A partir do conhecimento das partes do sistema a ser modelado e da forma como essas partes se relacionam, constrói-se o modelo. As informações, neste caso, são de caráter geral, aplicáveis a um conjunto de sistemas semelhantes.

No caso da identificação de sistemas, as informações usadas na geração do modelo são, em sua maior parte, oriundas do próprio sistema. Elas são obtidas em experimentos, o que confere o caráter empírico a essa abordagem. Ao contrário da modelagem por primeiros princípios, as informações são de caráter particular, específicas ao sistema sob estudo e a sua forma particular de operação.

Por exemplo, seja o problema de criar um modelo que relacione a tensão de saída $V_o(t)$ com a tensão de entrada $V_i(t)$ do circuito RC da Figura 1.1.

A modelagem por primeiros princípios permite criar um modelo a partir do esquema da Figura 1.1, de equações que descrevam a relação entre tensão e corrente nos terminais de resistores e capacitores e das leis de Kirchhoff. Um circuito real não é necessário, apenas o seu diagrama é suficiente. Além do mais, o modelo criado será função de parâmetros mensuráveis dos componentes do circuito como resistências e capacitâncias, e será válido para outros circuitos que difiram deste apenas nesses valores.

A identificação, por outro lado, requer um circuito real para poder gerar modelos. O circuito passará por experimentos nos quais serão colocadas tensões conhecidas em V_i e medidas as tensões V_o correspondentes. Os dados coletados são, então, usados na confecção dos modelos. Estes serão específicos do circuito identificado e, em geral, não o representariam corretamente se os parâmetros de seus componentes fossem alterados. No entanto, a princípio, pouca ou nenhuma hipótese precisa ser feita sobre o sistema a ser identificado. A modelagem, pelo contrário, costuma fazer algumas hipóteses importantes. Por exemplo, é comum considerar ideais os componentes do circuito. Se essa hipótese não for válida, o modelo deixa de representar adequadamente o sistema. A identificação, por não usar tais hipóteses, tem menos chances de apresentar tais problemas.

Em geral, o método preferido para obter modelos é a modelagem por primeiros princípios. Os modelos assim obtidos são usualmente válidos para uma classe inteira de sistemas e não somente para um em particular, como no caso empírico. Além do mais, a modelagem por primeiros princípios descreve o funcionamento interno do sistema, inclusive as relações entre os seus subsistemas. A identificação, por outro lado, gera modelos apenas para o que pode ser apreendido pelos dados, sendo muitas vezes cega para os detalhes internos do sistema em investigação.

Entretanto, embora preferível, a modelagem por primeiros princípios nem sempre é possível. Muitas vezes acontece da ciência não ser capaz de fornecer informações precisas o suficiente para que se construa um modelo adequado. Em outros casos, essa abordagem pode levar a modelos excessivamente complexos, sem que se encontrem simplificações que os tornem tratáveis. No entanto, se se puder coletar dados no sistema, a identificação é alternativa de escolha.

1.1 Motivação

Identificar modelos para sistemas não lineares, principalmente sistemas caóticos, é uma tarefa difícil, quando comparada com a identificação de sistemas dinâmicos lineares. Representações matemáticas não lineares possuem uma ampla gama de estruturas possíveis e o problema de detecção de estrutura é mais sério nessas representações (AGUIRRE, 2000a). Por exemplo, modelos com uma estrutura exageradamente complexa costumam exibir comportamentos dinâmicos espúrios (AGUIRRE; BILLINGS, 1995).

Existem diferentes técnicas de identificação de sistemas não lineares, mas todas englobam o ajuste de uma estrutura flexível aos dados dos sistemas. Esse ajuste é feito através da estimação de parâmetros ou através de métodos não paramétricos (LJUNG, 1987; AGUIRRE, 2000a). A diferença entre essas abordagens é que os métodos paramétricos limitam *a priori* a busca por um modelo adequado a um espaço de dimensão finita (LJUNG, 1987), o espaço de parâmetros, enquanto os métodos não paramétricos fazem essa busca num espaço de funções¹. Os métodos não paramétricos não são abordados nesse trabalho.

Um método de identificação paramétrico é uma função que associa os dados de identificação a um ponto específico do espaço de parâmetros (LJUNG, 1987). Existem várias formas de se fazer esse mapeamento. Uma forma muito comum é escolher os parâmetros de maneira que o erro de predição um passo à frente seja mínimo. Essa abordagem é usada tanto em representações lineares (LJUNG, 1987; SÖDERSTRÖM; STOICA, 1989), quanto em representações não lineares (PHAM; LIU, 1995; HAYKIN, 2001).

Representações de tempo contínuo usam outras formas de estimar parâmetros. Por exemplo, a abordagem via *poisson moment functional* (SAHA; RAO, 1983), a expansão em polinômios ortogonais com relação a medida invariante (GIONA *et al.*, 1991), aproximação via mínimos quadrados após o cálculo de derivadas (GOUESBET; LETELLIER, 1994; AGUIRRE *et al.*, 2001), técnicas *multiple shooting* (BAAKE *et al.*, 1992; TIMMER *et al.*, 2000), método das trajetórias (PERONA *et al.*, 2000). Esses métodos, com a exceção dos dois últimos, não levam explicitamente em conta a evolução temporal do sistema a ser identificado, fato importante no desempenho dos

¹Na prática, os métodos não paramétricos acabam recaindo num problema de dimensão finita, mas essa restrição não é feita *a priori*.

modelos identificados, conforme mostrado no Capítulo 2. Já os métodos *multiple shooting* e das trajetórias recaem em complexos problemas de otimização não linear na estimação dos parâmetros.

O método apresentado neste trabalho procura resolver o problema da estimação de parâmetros de uma outra forma. O que motivou o desenvolvimento desse método de estimação foi, em parte, o conceito de *qualidade de modelo*. Um modelo deve ser capaz de reproduzir características desejadas do sistema que ele representa. No caso do tipo de modelo produzido aqui, supõe-se que a principal característica desejada é o comportamento temporal. Em outras palavras, espera-se que o modelo seja capaz de reproduzir numa simulação a evolução temporal de grandezas de interesse do sistema. No caso de sistemas não lineares, principalmente sistemas que apresentam comportamentos dinâmicos complexos como caos, esse é um requisito um tanto difícil. Por exemplo, um sistema caótico possui como uma de suas características mais básicas a sensibilidade às condições iniciais (ALLIGOOD *et al.*, 1994). Devido a ela, a evolução temporal de um sistema caótico, para uma determinada condição inicial, diverge exponencialmente, em média, de outra solução com uma condição inicial distinta porém arbitrariamente próxima. Como consequência, a simulação de um modelo não é capaz de reproduzir exatamente o comportamento temporal de um sistema caótico devido à presença de ruído experimental e à precisão limitada a que uma simulação digital está sempre sujeita. Da mesma forma, um usuário de técnicas de identificação não espera que uma simulação de um modelo obtido para um sistema caótico reproduza as mesmas séries temporais observadas no sistema. Em outras palavras, não se espera que um modelo seja capaz de reproduzir, instante à instante, sinais observados no sistema, e a qualidade de modelos para sistemas caóticos é normalmente avaliada usando outras propriedades (AGUIRRE; BILLINGS, 1994), e não no domínio do tempo.

No entanto, apesar da sensibilidade às condições iniciais, existem situações nas quais dois sistemas caóticos podem exibir o mesmo comportamento temporal, mesmo que tenham sido iniciados em condições distintas. Esse fenômeno é conhecido como *sincronização*. O tema da sincronização de sistemas caóticos é abordado com profundidade no Capítulo 3. Em resumo, é possível, em certas ocasiões, fazer com que dois ou mais sistemas caóticos exibam o mesmo comportamento temporal, mesmo que seja um regime caótico. Em outras palavras, se se deseja que dois sistemas caóticos se sincronizem, deve-se fazer com que eles apresentem o mesmo comportamento

temporal.

Nesse ponto, é possível notar que “mesmo comportamento temporal” é um componente presente, de algum modo, tanto em identificação de sistemas quanto em sincronização. É natural, portanto, indagar se existe alguma relação que com utilidade prática entre esses dois ramos da ciência. Particularmente, se é possível empregar de alguma forma a sincronização na identificação de sistemas como uma forma de estimação de parâmetros.

Esse assunto já é investigado na literatura científica. A presente Tese é uma contribuição nessa área.

1.1.1 Sincronização na estimação de parâmetros

O uso de sincronização na estimação de parâmetros é um tópico presente na literatura científica recente.

Em [Parlitz *et al.* \(1996\)](#) o erro de sincronização é usado em um esquema de otimização não linear para estimar parâmetros desconhecidos de sistemas caóticos. A sincronização é feita usando uma decomposição do sistema em uma parte ativa e outra passiva. O uso de otimização para minimizar o erro de sincronização também é a estratégia adotada em [Timmer *et al.* \(2000\)](#) e em [Zhang *et al.* \(2006b\)](#), com a diferença que neste último, a sincronização é obtida através de um controle linear. Em [Zhang *et al.* \(2006a\)](#), um método parecido é usado, empregando um método de otimização mais simples.

Em [Marino e Miguez \(2006b\)](#) e [Marino e Miguez \(2006a\)](#), um interessante método de estimação usa derivadas dos sinais de erro de sincronização em um método de otimização para estimar os parâmetros e sincronizar os sistemas. No entanto, não são fornecidas as condições de convergência para o método.

Um classe mais interessante de métodos é baseada no uso de controle adaptativo junto com sincronização de sistemas. Um trabalho de fundamental importância nesse sentido é [Fradkov e Pogromsky \(1996\)](#), onde é apresentado um método de controle adaptativo que permite a estimação de parâmetros a partir da sincronização de sistemas. Existem outros trabalhos nessa linha ([MAYBHATE; AMRITKAR, 1999](#); [MAYBHATE; AMRITKAR, 2000](#); [HUANG; GUO, 2004](#); [KONNUR, 2005](#)), mas eles não fornecem um método de aplicação genérica ou provam sua convergência.

O presente trabalho procura estender o método do gradiente de velocidade (FRADKOV; POGROMSKY, 1996) com técnicas de controle não linear para criar uma ferramenta geral de estimação de parâmetros.

1.2 Estrutura do texto

O Capítulo 2 apresenta o problema de identificação de sistemas, juntamente com outros conceitos importantes. Nele, as estratégias convencionais são agrupadas numa estrutura geral e essa estrutura é analisada, evidenciando seus problemas. No Capítulo 3 são apresentados os sistemas caóticos e o fenômeno da sincronização. O Capítulo 4 descreve a estratégia de controle não linear usada no método de identificação proposto, descrito no Capítulo 5. Os resultados obtidos com esse método são mostrados no Capítulo 6 e as conclusões encerram o texto no Capítulo 7.

CAPÍTULO 2

IDENTIFICAÇÃO DE SISTEMAS

“Muito melhor uma resposta aproximada à pergunta certa, que é freqüentemente vaga, do que uma resposta exata à pergunta errada, que sempre pode ser tornada precisa.”

(John W. Tukey, 1962)

2.1 Introdução

A identificação de sistemas é um assunto central desse trabalho e é também uma área de conhecimento bastante dinâmica, com extensa literatura a respeito (HSIA, 1977; SAHA; RAO, 1983; LJUNG, 1987; SÖDERSTRÖM; STOICA, 1989; JOHANSSON, 1993; AGUIRRE, 2000a). Embora variados, os diversos métodos de identificação possuem uma estrutura comum, que pode ser descrita de uma forma unificada, pelo menos no que tange aos métodos mais usados.

O objetivo deste Capítulo é descrever essa estrutura comum e apontar seus principais problemas, bem como indicar possíveis soluções. De particular interesse nesse sentido é a formulação do problema de identificação como um problema inverso. Os problemas descritos mostram as limitações das técnicas tradicionais de identificação, sendo em parte, a motivação para o desenvolvimento do método apresentado nesta Tese.

Quanto à organização do Capítulo, inicialmente a definição do problema de identificação de sistemas é apresentada na Seção 2.2, junto com outros conceitos importantes. A seguir, a Seção 2.3 apresenta a estrutura básica dos métodos mais usados de identificação. A Seção 2.4 descreve a identificação de sistemas como um problema inverso. Os problemas do método tradicional de identificação são mostrados na Seção 2.5 e, finalmente, a Seção 2.6 conclui o Capítulo.

2.2 Definição

2.2.1 Conceitos importantes

Antes de definir identificação de sistemas, alguns outros conceitos devem ser abordados. O primeiro conceito importante é o de *sistema dinâmico*.

Uma boa definição de sistema é dada por Monteiro (2002, p. 39):

Um *sistema* pode ser definido como um conjunto de objetos agrupados por alguma interação ou interdependência, de modo que existam relações de causa e efeito nos fenômenos que ocorrem com os elementos desse conjunto. São exemplos de sistemas: o circuito elétrico de um rádio-telescópio, Júpiter e seus satélites, o sistema nervoso de um canguru, a situação financeira de uma família, o ecossistema de um mangue.

Um sistema *dinâmico* é aquele que possui memória, ou seja, a evolução temporal das grandezas que o definem depende de alguma forma dos valores passados dessas grandezas. Aos sistemas dinâmicos opõem-se os sistemas instantâneos, que não possuem essa dependência. Como exemplo de sistema instantâneo, pode-se citar uma lâmpada incandescente. A lâmpada está acesa se o interruptor estiver ligado. Do contrário, ela está apagada. Ou seja, para saber o *estado* da lâmpada, basta olhar para o interruptor. Um exemplo de sistema dinâmico é uma panela com água no fogão. Não é possível saber se a água está quente ou fria apenas verificando se a chama está acesa ou apagada. É possível que a água esteja fria estando a chama acesa. Basta que a panela tenha acabado de ser colocada no fogo. Em outras palavras, é necessário conhecer o passado da água para saber o estado da sua temperatura. Somente a informação sobre a chama não é suficiente para concluir sobre a temperatura.

Um sistema dinâmico é dito *determinístico* se o seu estado futuro é função somente do seu passado e do comportamento de eventuais entradas. Um sistema determinístico se contrapõe a um sistema estocástico, no qual algum grau de aleatoriedade se faz presente.

O estado de um sistema dinâmico é o conjunto mínimo de valores, chamados *variáveis de estado* (MONTEIRO, 2002), necessários para, conjuntamente com as possíveis entradas, caracterizar a evolução temporal do sistema.

Neste trabalho serão abordados sistemas dinâmicos determinísticos, de agora em diante referidos apenas por *sistemas*, ou por *sistemas dinâmicos*.

Outro conceito de grande importância neste trabalho é o de *modelo*. Um modelo,

conforme apresentado no Capítulo 1 é um análogo de um sistema, que preserva algumas de suas características. Por exemplo, uma maquete é um modelo para um edifício. Esse modelo preserva algumas de suas características, como as proporções de suas dimensões, enquanto descarta outras, como os materiais usados na construção. Outros exemplos de modelos são uma foto, como modelo para um ser humano e um mapa como modelo para um terreno.

Um *modelo matemático* para um sistema é um modelo que descreve matematicamente uma ou mais características desse sistema. Um modelo matemático permite uma análise matemática das características do sistema. Um exemplo de um modelo matemático é a descrição do movimento de uma partícula de massa m num eixo x , sob a ação de uma força F , conforme a segunda Lei de Newton

$$\frac{d^2x}{dt^2} = \ddot{x} = \frac{F}{m} . \quad (2.1)$$

Modelos matemáticos para sistemas dinâmicos podem ter diferentes representações. Neste trabalho, duas representações são usadas, dependendo se a variável independente, normalmente o tempo, é discreta ou contínua.

No caso de *sistemas de tempo contínuo*, usam-se sistemas de equações diferenciais ordinárias, como a equação (2.1). Em todos esses casos é usado o ponto para indicar derivada em relação ao tempo como em \ddot{x} .

No caso de *sistemas de tempo discreto*, são usadas as equações à diferença, nas quais o valor de uma grandeza em um determinado instante de tempo é colocado em função do valor dessa grandeza, e/ou outras grandezas como entradas, em outros instantes discretos de tempo. Por exemplo,

$$y(k) = \mu y(k-1) (1 - y(k-1)) , \quad (2.2)$$

onde $k \in \mathbb{N}$ representa o instante discreto de tempo e μ é um parâmetro.

2.2.2 Definição do problema de identificação de sistemas

Nesse ponto é possível definir a identificação de sistemas. Ela consiste no procedimento que permite obter, a partir de dados experimentais, modelos matemáticos para sistemas quaisquer (HSIA, 1977; SÖDERSTRÖM; STOICA, 1989; JOHANSSON,

1993; AGUIRRE, 2000a). O resultado da identificação é, pois, um modelo matemático que representa (bem ou mal) algumas características de um determinado sistema.

Essa definição é um tanto geral e permite que vários procedimentos possam ser classificados como identificação de sistemas. No entanto, o assunto tratado nesse trabalho é mais restrito. Apenas a identificação de sistemas *dinâmicos* é abordada. Também, nem todo tipo possível de modelo matemático é estudado. Esse trabalho se concentra em modelos matemáticos *paramétricos*, ou seja, modelos que possuem parâmetros numéricos como μ em (2.2).

Outro ponto importante é que tipo de modelo se busca, ou melhor, quais características do sistema esperam-se que o modelo seja capaz de reproduzir. Isso depende do uso a que se destina o modelo. No entanto, neste trabalho, supõe-se que a principal característica dos modelos é que eles possuam, tanto quanto possível, o mesmo *comportamento dinâmico* do sistema, ou seja o mesmo comportamento temporal.

2.3 Procedimento de identificação

Existem diversos procedimentos disponíveis para um eventual usuário que queira identificar um modelo a partir de dados coletados. Apesar dessa multiplicidade, os métodos mais usados partilham a mesma essência e podem ser apresentados em uma forma geral comum.

Nesta Seção, esta forma geral é apresentada, com o objetivo mostrar o mecanismo de identificação e, principalmente, apontar os problemas que esses métodos de identificação compartilham.

Segundo Aguirre (2000a), processo de identificação pode ser dividido em 5 etapas principais:

- a) Coleta de dados;
- b) Escolha da representação matemática a ser usada;
- c) Determinação da estrutura do modelo;
- d) Estimação de parâmetros;
- e) Validação do modelo.

A seguir, cada uma dessas etapas é descrita.

2.3.1 Coleta de dados

A coleta de dados constitui o ponto de partida do processo de identificação. Esses dados são grandezas que devem ser amostradas no sistema sob estudo e que quantificam características de interesse do mesmo.

Quanto à forma de obtenção dos dados, diversas situações podem acontecer, dependendo do controle que o usuário de identificação tenha sobre o processo de coleta de dados. Num extremo, o usuário pode projetar completamente o ensaio de coleta de dados, determinando questões como formas de onda das entradas e período de amostragem adequado. No outro extremo, o usuário pode ter que partir apenas dos dados, sobre os quais ele não teve nenhuma influência no processo de obtenção, como acontece com frequência na análise de séries temporais (LJUNG, 1987).

A importância da coleta de dados reside no fato de que toda a informação que o método de identificação precisa sobre o sistema deve estar nos dados¹. Em outras palavras, se não houver informação sobre uma determinada característica do sistema nos dados, não há como garantir que o modelo preserve essa característica. Logo, o ensaio de coleta de dados deve procurar garantir que haja informação suficiente sobre as características de interesse nos dados e que essa informação esteja disponível para o processo de identificação.

Um dos principais parâmetros da coleta de dados é o tempo ou período de amostragem T_s . Este deve ser compatível com as escalas de tempo características do sistema. Um período de amostragem grande demais não será capaz de acomodar informação sobre os comportamentos dinâmicos mais rápidos do sistema. Por outro lado, razões práticas como capacidade de armazenamento de dados e limitações do sistema de coleta de dados impedem que se utilize tempos de amostragem arbitrariamente curtos. Deve-se, portanto, empregar um T_s intermediário. Maiores detalhes sobre como determinar T_s podem ser obtidos em (FREITAS, 2001) e em (AGUIRRE, 2000a).

Outro ponto muito importante para o ensaio de coleta de dados é o sinal de excitação ou de entrada. Nos casos em que o sistema a ser identificado possui entradas,

¹Na verdade, existem outras formas de incluir informação sobre um sistema no processo de identificação, além dos dados (AGUIRRE, 2000b). No entanto, os dados são a principal fonte de informação.

muitas vezes é possível, e mesmo desejável, que sinais de excitação especialmente projetados para a identificação sejam utilizados. O objetivo é garantir que o máximo de informação possível sobre o sistema esteja presente nos dados a serem coletados. Na prática, na identificação de sistemas não lineares, isso se traduz em sinais de entrada de amplo espectro de frequências e de amplitudes variadas. Estratégias para a escolha de sinais de entrada podem ser vistas em [Aguirre \(2000a\)](#).

Uma vez de posse dos dados, é possível realizar um pré-processamento antes da identificação. Um dos objetivos dessa etapa é melhorar a qualidade dos dados. Por exemplo, os dados podem ser decimados para se aumentar o T_s ([AGUIRRE, 2000a](#)), no caso deste se mostrar pequeno demais. Pode-se ainda aplicar filtros para diminuir o efeito de ruído, computar funções de autocorrelação e correlação cruzada para se escolher entradas e saídas adequadas, entre outros.

Um procedimento muito comum consiste em realizar uma *reconstrução de espaço de estados*, no qual, a partir de séries temporais de baixa dimensão, geram-se séries de dimensão mais alta ([GIBSON et al., 1992](#)). Dessa forma está sendo criado um espaço de imersão para as trajetórias diferente do espaço de estados original, mas que é topologicamente equivalente ao espaço de estados original se a dimensão do espaço reconstruído for suficientemente grande ([TAKENS, 1980](#)). Dentre as diversas técnicas existentes para isso, a mais utilizada é a reconstrução por coordenadas de atraso. Nela, parte-se de uma série temporal $y(k)$, sendo $k = 1, 2, \dots, N$ os instantes discretos de tempo, e geram-se vetores $\mathbf{y}(k)$ dados por

$$\mathbf{y}(k) = [y(k), y(k - \tau), \dots, y(k - (d_e - 1)\tau)]^T, \quad (2.3)$$

sendo τ o atraso de reconstrução e d_e a dimensão do espaço de reconstrução. Tal procedimento é muito importante na identificação de sistemas de dimensão mais alta do que a dos dados coletados.

2.3.2 Escolha da representação

A escolha da representação matemática é uma etapa crucial. Essa escolha determina como serão as próximas etapas ([AGUIRRE, 2000a](#)). Nesse ponto, o usuário de identificação deve optar entre diversas possibilidades. Por exemplo, ele pode escolher entre modelos de tempo discreto ou de tempo contínuo e entre modelos lineares e não lineares, entre muitas outras decisões.

Dentre os fatores que influenciam essa escolha, pode-se citar a familiaridade do usuário com uma determinada representação matemática, algum conhecimento que se tenha *a priori* sobre o sistema a identificar (AGUIRRE, 2000b) e a aplicação a que se destina o modelo.

Entre as representações para modelos dinâmicos não lineares e de tempo discreto pode-se citar as redes MLP (*Multilayer Perceptron*) (ELSNER, 1992), redes RBF's (*Radial Basis Function*) (BROOMHEAD; LOWE, 1988), lógica nebulosa (TAKAGI; SUGENO, 1985), redes híbridas como a *Neo Fuzzy Neuron* (YAMAKWA *et al.*, 1992; CAMINHAS; GOMIDE, 2000), modelos NARMAX polinomiais (LEONTARITIS; BILLINGS, 1985a; LEONTARITIS; BILLINGS, 1985b) ou racionais (BILLINGS; CHEN, 1989), entre muitas outras. Cada uma dessas representações tem características distintas em relação à complexidade, flexibilidade, facilidade de treinamento, capacidade de generalização, custo computacional e outros quesitos. Essas representações, no entanto, apresentam um ponto em comum: elas são todas paramétricas. Ou seja, elas possuem parâmetros numéricos que precisam ser ajustados. São justamente esses parâmetros que conferem a essas estruturas a sua flexibilidade, característica fundamental na identificação de sistemas.

2.3.3 Determinação da estrutura e estimação de parâmetros

A seguir vêm as etapas de escolha de estrutura e estimação de parâmetros. Qualquer representação matemática útil para identificação de sistemas deve possuir uma flexibilidade que a permita se moldar a um sistema qualquer. Num exemplo didático, seja o problema de se aproximar uma função qualquer $f : \mathbb{R} \rightarrow \mathbb{R}$ por um polinômio $p(x) = \sum_{i=0}^g a_i x^i$. Para se determinar um polinômio, deve-se escolher, ou ajustar, tanto o seu grau g , quanto os seus coeficientes a_i . Essas duas “escolhas” são distintas em sua natureza.

A determinação do grau g deve ser feita num espaço discreto. A escolha do grau está relacionada com a flexibilidade do polinômio. Quanto maior o grau, maior a flexibilidade. Também, o grau determina o número de coeficientes que precisarão ser ajustados.

Já o ajuste dos coeficientes é uma busca num espaço contínuo. É ajustando ou estimando os coeficientes que o polinômio genérico $p(x)$ toma uma forma específica e pode ser usado no lugar da função que ele aproxima.

A escolha do grau na aproximação polinomial é o análogo da escolha da estrutura para a identificação de sistemas. Esta última é feita num espaço discreto e influencia a flexibilidade dos modelos a serem gerados. Essa escolha depende da representação utilizada. O ajuste dos coeficientes na aproximação polinomial é o análogo da estimação de parâmetros na identificação. Esta é uma busca num espaço contínuo e é responsável por colocar o modelo numa forma específica capaz de representar o sistema. Em outras palavras, a estrutura é um mapeamento entre um vetor de parâmetros e um conjunto de modelos no qual se busca o modelo particular que melhor se adequa aos dados (LJUNG, 1987).

A escolha da estrutura e a estimação de parâmetros são o cerne das técnicas de identificação, pois são esses procedimentos que geram os modelos para os sistemas.

Um modelo não linear dinâmico de tempo discreto pode ser descrito por

$$\mathbf{y}(k) = f(\mathbf{y}(k-1), \mathbf{u}(k-1)) \quad , \quad (2.4)$$

sendo $\mathbf{y}(k) \in \mathbb{R}^q$ o vetor de estados no instante de tempo discreto k , $\mathbf{u}(k) \in \mathbb{R}^m$ o vetor de entradas no instante k e $f : \mathbb{R}^q \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ uma função não linear qualquer. Os vetores $\mathbf{y}(k)$ e $\mathbf{u}(k)$ podem representar tanto vetores observáveis no sistema, quanto vetores de um espaço vetorial reconstruído, vide equação (2.3). A equação (2.4) permite obter, a partir do conhecimento do estado e das entradas no instante $k-1$, o estado no instante k . Ou seja, o modelo permite avançar no tempo o conhecimento do estado. Repetindo-se o processo para o instante $k+1$, usando o valor de $\mathbf{y}(k)$ obtido no passo anterior, pode-se conhecer o valor de $\mathbf{y}(k+1)$. Assim, por iteração, é possível obter a evolução temporal dos estados. A equação (2.4) é, *junto com o processo de iteração*, um modelo não linear de tempo discreto para um sistema dinâmico.

Na identificação de sistemas, escolhe-se para $f(\cdot)$ uma representação paramétrica qualquer, como por exemplo as já citadas redes *perceptron* multicamadas ou os modelos NARMAX polinomiais, entre outros. Nessa representação, $f(\cdot)$ toma uma forma que depende de uma estrutura $s \in S$ e de um vetor de parâmetros $\boldsymbol{\theta} \in \mathbb{R}^{p(s)}$, sendo S o conjunto de todas as estruturas possíveis. O número de parâmetros p depende da estrutura escolhida, enquanto o significado da estrutura e dos parâmetros depende da representação.

Como ilustração, seja o caso de um modelo com $q = 1$ e sem entradas. Uma possível representação para $f(\cdot)$ é a polinomial. Nesse caso, a estrutura s seria um conjunto específico de termos de um polinômio, por exemplo os termos com graus 1, 3 e 5. Por sua vez, S seria o conjunto de todos os conjuntos de termos possíveis. O vetor de parâmetros seria formado, então, pelos coeficientes desses termos.

Se fosse empregado o *perceptron* multicamadas como representação, o conjunto S seria o conjunto de *todas* as topologias de redes com uma entrada e uma saída. Uma estrutura particular s para o modelo seria então uma topologia específica, determinando quantas camadas escondidas, quantos neurônios em cada camada e quais as ligações entre eles. O vetor de parâmetros é o conjunto de todos os pesos e termos de polarização (*bias*) para uma determinada topologia.

Assim, escolher a estrutura é escolher um s particular de S e estimar os parâmetros é tomar um vetor particular θ . Ambas ações devem ter por objetivo gerar um modelo que capture as características desejadas de um determinado sistema. Para isso lança-se mão dos dados coletados.

O procedimento adotado na maior parte das técnicas de identificação de sistemas é o de se definir uma *função custo* que dependa do modelo e dos dados e de se utilizar algum *método de otimização*.

Normalmente, a função de custo empregada em identificação é definida sobre o *erro de predição de um passo à frente* (LJUNG, 1987; SÖDERSTRÖM; STOICA, 1989; AGUIRRE, 2000a)

$$\xi(k) = \mathbf{y}(k) - f(\mathbf{y}(k-1), \mathbf{u}(k-1), s, \theta) , \quad (2.5)$$

sendo que a dependência de $f(\cdot)$ com relação a estrutura s e ao vetor de parâmetros θ foi tornada explícita. De agora em diante, supõe-se que $\mathbf{y}(k)$ e $\mathbf{u}(k)$, para $k = 1, 2, \dots, N$ são formados pelos dados coletados, sejam vetores observados diretamente, sejam vetores oriundos de algum procedimento de reconstrução de espaço de estados.

A função de custo normalmente utilizada é o somatório dos erros quadráticos²

$$J = \sum_{k=1}^N \|\boldsymbol{\xi}(k)\|_2^2, \quad (2.6)$$

sendo $\|\cdot\|_2$ a norma euclidiana.

Essa função depende da estrutura s e do vetor de parâmetros $\boldsymbol{\theta}$, bem como dos dados coletados. Dessa forma, os problemas de detecção de estrutura e estimação de parâmetros se resumem em escolher s e $\boldsymbol{\theta}$ de modo que J seja mínimo. Esses problemas são normalmente resolvidos separadamente, escolhendo-se primeiro uma estrutura e depois estimando-se os parâmetros.

Existem técnicas automáticas³ de detecção de estrutura para diversas representações, como modelos NARMAX polinomiais (BILLINGS *et al.*, 1989), racionais (BILLINGS; CHEN, 1989; CORRÊA, 1997), redes neurais (REED, 1993), modelos contínuos (FREITAS, 2001; AGUIRRE *et al.*, 2001), entre outras. Tais técnicas permitem escolher a estrutura de um modelo baseada nos dados de identificação. No entanto, elas apenas se aproximam da estrutura ideal. Encontrar a estrutura ótima é, em geral, um problema muito difícil.

Uma estrutura pode também ser escolhida de forma arbitrária. Dessa forma, espera-se que apenas a estimação de parâmetros seja suficiente para gerar um modelo adequado.

Uma vez obtida a estrutura, utiliza-se alguma técnica de otimização para estimar os parâmetros. Dependendo da representação escolhida, o erro de predição um passo à frente pode ser linear ou não linear nos parâmetros. No primeiro caso, a função custo é convexa em $\boldsymbol{\theta}$ e o problema de otimização é mais simples de se resolver do que quando o erro de predição é não linear nos parâmetros. Neste último caso, há a possibilidade de existência de mínimos locais.

2.3.4 Validação do modelo

Uma vez de posse de um modelo identificado, procede-se a sua validação. A validação é a etapa na qual a qualidade dos modelos identificados é aferida. Nela são

²Outras funções de custo também são empregadas. Ver Ljung (1987).

³Na prática, algum grau de interferência do usuário sempre é necessário.

descartados os modelos que não foram capazes de representar adequadamente as características desejadas do sistema original. É importante observar que *a qualidade do modelo é um conceito relativo, pois depende do objetivo para o qual o modelo se destina* (AGUIRRE, 2000a). Idealmente, o modelo deveria ser capaz de se comportar exatamente como o sistema em estudo em quaisquer situações, mas isso, em geral, não é viável. Assim, durante a identificação, deve-se ter em mente o *uso a que se destina o modelo*. Um bom modelo para predição pode não ser tão bom quando utilizado para controle e, por sua vez, um péssimo modelo para predição pode ser suficiente para a extração de alguma informação sobre do sistema (AGUIRRE; SOUZA, 1998).

Ainda assim, um emprego muito freqüente para os modelos é a predição, ou seja, usar o modelo para saber como o sistema se comportaria frente a determinadas condições. Nesse caso, deseja-se que o modelo tenha um comportamento temporal o mais próximo possível daquele do sistema original.

Uma das formas mais simples de se fazer a validação nesse sentido é através da comparação entre o sinal coletado do sistema real e o sinal resultante da simulação do modelo identificado. Se o modelo acompanhar o sistema, a menos de um erro pequeno, então o modelo é “bom”. Preferencialmente, essa comparação é feita com dados diferentes daqueles usados para identificar o modelo, pois é possível que o modelo identificado capture alguma característica particular do conjunto de dados de identificação, como ruído, por exemplo (AGUIRRE, 2000a).

Porém, se o sistema a ser identificado é caótico, a comparação do sinal simulado com o sinal oriundo do modelo não é capaz de dizer se um modelo é adequado ou não. Isso ocorre devido às características próprias desse tipo de sistema. Uma trajetória simulada de um modelo irá, em geral, divergir de uma trajetória real de um sistema caótico, mesmo que elas tenham começado em pontos extremamente próximos e independentemente da qualidade do modelo identificado. Essa divergência é devido à sensibilidade às condições iniciais, a característica principal dos sistemas caóticos. Isso torna necessária a utilização, no caso de sistemas caóticos, de outras técnicas de validação que não a simples comparação de séries temporais.

Uma alternativa simples é a comparação visual do atrator do sistema com o atrator do modelo identificado. O atrator é o lugar geométrico, no espaço de estados, para o qual as trajetórias tendem quando o tempo tende ao infinito (ALLIGOOD

et al., 1994). Ele é uma característica de “regime permanente” e independente tanto das séries temporais como das condições iniciais. Dessa forma, mesmo que as séries temporais sejam diferentes, é de se esperar que o atrator do modelo e o atrator do sistema sejam parecidos, se o modelo for bom. Porém, essa abordagem não é possível para todos os sistemas. Por exemplo, sistemas não dissipativos não possuem atrator.

Comparar visualmente dois atratores traz um elemento subjetivo ao processo de validação. Em vez disso, pode-se comparar medidas objetivas desses atratores, como a dimensão de correlação. Pode-se também comparar grandezas invariantes, como expoentes de Lyapunov (AGUIRRE; BILLINGS, 1994). Por fim, quando uma validação mais cuidadosa for necessária, pode-se comparar mapas de primeiro retorno em seções de Poincaré e populações de órbitas periódicas instáveis (LETELLIER *et al.*, 1995). Porém, esses procedimentos de validação são de utilização mais complexa e requerem conhecimento de caracterização topológica de atratores.

O método aqui apresentado é a forma usual de se fazer identificação de sistemas, classificado como método do erro de predição (SÖDERSTRÖM; STOICA, 1989). Essa formulação tem problemas intrínsecos que dificultam a obtenção de bons modelos, principalmente para sistemas não lineares. A seguir, os principais problemas dessa abordagem são explorados. Mas, antes, é importante formular a identificação de sistemas como um problema inverso.

2.4 Identificação de sistemas como um problema inverso

A identificação de sistemas, ou mais precisamente, as etapas de detecção de estrutura e estimação de parâmetros, consiste em gerar modelos a partir de comportamentos temporais conhecidos de sistemas dinâmicos. Esse é o problema inverso ao de se determinar, a partir do conhecimento do modelo, o comportamento temporal do mesmo.

A literatura sobre problemas inversos é bastante vasta. Um problema inverso pode ser visto como o problema de se determinar as causas a partir de seus efeitos, enquanto o problema direto seria determinar os efeitos a partir do conhecimento das suas causas (ALIFANOV, 1979, citado por MUNIZ, 1999). Matematicamente, sejam W e Z espaços normados e $K : W \rightarrow Z$ uma função. O problema direto é, dado as causas $w \in W$, determinar os seus efeitos $z \in Z$ correspondentes tal que $z = K(w)$. O problema inverso é, a partir de um dado efeito $z \in Z$ encontrar suas causas $w \in W$

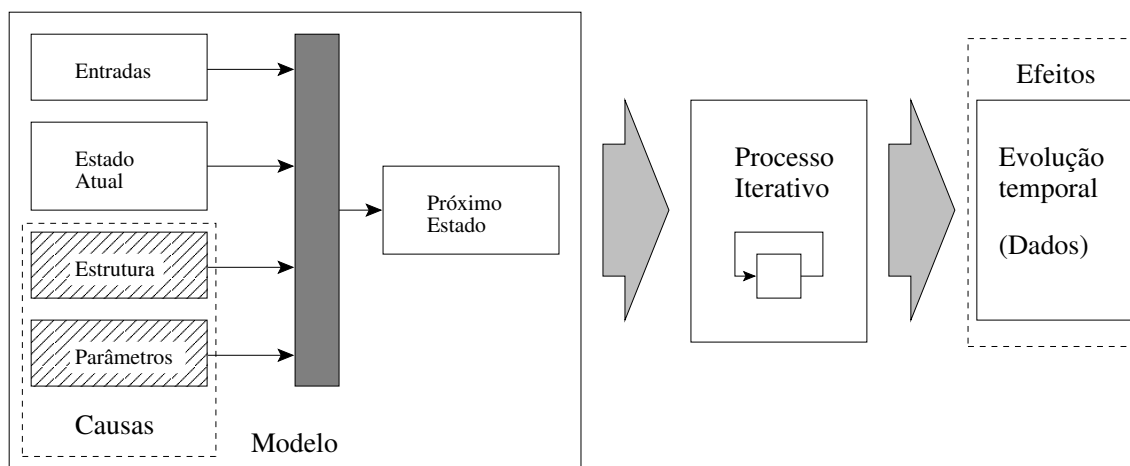


FIGURA 2.1 - Problema direto na identificação de sistemas. As causas são associadas à estrutura e aos parâmetros do modelo. Os efeitos são a evolução temporal do sistema, ou seja, os dados.

tal que $K(w) = z$.

No caso da identificação de sistemas, w representa o modelo, ou, mais especificamente, uma estrutura s particular junto com um vetor de parâmetros θ específico, supondo uma representação já escolhida, enquanto z representa o conjunto de dados e K é a função que toma essa estrutura e parâmetros como entrada e gera a evolução temporal do modelo. A Figura 2.1 ilustra esse problema direto. O *modelo*, na Figura, é um mecanismo que produz o próximo estado a partir das entradas e do estado atual. Esse mecanismo depende de uma estrutura particular e de um vetor de parâmetros. O modelo é então usado em um processo iterativo que produz a evolução temporal. Na identificação de sistemas, a estrutura s e o vetor de parâmetros θ são as *causas* (w) do problema direto, enquanto a evolução temporal representa os *efeitos* (z) e está associada aos dados coletados no sistema a ser identificado.

O procedimento de identificação é o problema inverso de, a partir dos dados, obter a estrutura s e os parâmetros θ . É importante notar que existe, no problema direto, um processo de iteração que deve ser levado em conta no momento de se atacar o problema inverso. Um modelo, tal qual o descrito na equação (2.4), apenas relaciona o próximo estado com o estado atual e as possíveis entradas. Para que esse modelo seja capaz de gerar sinais que evoluem no tempo, é necessário empregar esse processo de iteração. No caso de modelos de tempo contínuo, o análogo desse processo de iteração é a integração das equações diferenciais. Da mesma forma que nos modelos

de tempo discreto precisam do processo de iteração, os modelos contínuos em si não são capazes de gerar sinais sem a integração das suas equações.

Um modelo da forma (2.4) representa um outro problema direto, no qual as causas são o estado atual, as entradas, a estrutura e os parâmetros. O efeito é o próximo estado. Esse outro problema direto é referido, daqui por diante, por *Problema Direto Reduzido*, ou PDR. A sua contra-partida, o problema direto da Figura 2.1, é referido como *Problema Direto Completo*, ou PDC. Os problemas inversos associados são chamados *Problema Inverso Reduzido*, ou PIR, e *Problema Inverso Completo*, ou PIC, respectivamente. As características do PDC e do PDR estão listadas na Tabela 2.1.

TABELA 2.1 - Causas e efeitos do PDC e do PDR

	PDC	PDR
Causas (x)	Estrutura (s) Parâmetros (θ) Condição inicial	Estrutura (s) Parâmetros (θ) Estado atual ($\mathbf{y}(k-1)$)
	Entradas ($\mathbf{u}(k), k = 1, \dots, N$)	Entrada atual ($\mathbf{u}(k-1)$)
Efeitos	Evolução temporal ($\mathbf{y}(k), k = 1, \dots, N$)	Próximo estado ($\mathbf{y}(k)$)

Um problema inverso pode ser resolvido como um problema de otimização. Seja a função de custo dada por

$$J_{\text{pi}}(w) = \|z - K(w)\|, \quad (2.7)$$

sendo $\|\cdot\|$ uma norma em Z . Pode-se definir como solução do problema inverso o $w = w^*$ para o qual $J_{\text{pi}}(w)$ seja mínimo.

2.5 Problemas da identificação de sistemas tradicional

Nesse ponto, é possível mostrar um dos problemas do método de identificação apresentado na Seção 2.3. Como apresentado naquela Seção, as técnicas de identificação de sistemas, em geral, recaem em um problema de otimização. É possível supor, então, que tais técnicas implicitamente resolvem um problema inverso, embora sem postulá-lo. Observando a equação do erro de predição um passo à frente (2.5) e a função de custo J (2.6), nota-se que, embora J dependa de s e de θ , não há referência a nenhum processo de iteração. Em outras palavras, a função de custo J

depende do modelo que fornece o próximo estado a partir do atual, mas não depende do processo que usa esse modelo para gerar a evolução temporal característica dos sistemas dinâmicos.

O resultado disso é que o problema inverso que as técnicas apresentadas na Seção 2.3 resolvem não é aquele associado ao problema da Figura 2.1, pois elas *não fazem referência ao processo iterativo*. O problema que elas na verdade tratam é determinar qual modelo (estrutura e parâmetros) retorna o próximo estado a partir do estado atual e entradas contidos nos dados. Em resumo, elas resolvem o PIR ao invés do PIC, ou seja, elas *resolvem o problema errado*.

Embora a diferença entre o PIC e o PIR possa parecer sutil a princípio, ela pode comprometer seriamente o desempenho da identificação, pois a solução para o PIR não é, necessariamente, uma boa solução para o PIC, principalmente quando se trata de identificar modelos para sistemas não lineares. Por exemplo, um modelo solução para o PIR (um preditor um passo à frente para os dados) pode, quando colocado em um regime iterativo, apresentar um comportamento dinâmico completamente diferente daquele presente nos dados, podendo ser até mesmo instável. Tal modelo dificilmente poderia ser considerado como uma solução adequada para o PIC. Essa situação ocorre com frequência na identificação de sistemas não lineares (AGUIRRE; BILLINGS, 1995; AGUIRRE *et al.*, 2001; FREITAS, 2001).

Um exemplo pode ilustrar esse problema. Suponha que o sistema a ser identificado é dado pela equação logística, um clássico exemplo de sistema caótico discreto

$$y(k) = \lambda y(k-1) [1 - y(k-1)] , \quad (2.8)$$

sendo λ o seu parâmetro. O valor de λ determina o comportamento temporal do sistema. Vários regimes são possíveis, variando do estado estacionário ao caos, passando por regimes periódicos e intermitências (ARGYRIS *et al.*, 1994).

Para este exemplo, o sistema da equação (2.8) foi simulado com $\lambda = 3,97$ gerando 500 pontos. A estes, foi adicionado ruído gaussiano i.i.d. com desvio padrão de 0,02. O método de identificação tradicional foi então implementado, empregando-se uma representação polinomial de grau 2 e um modelo de primeira ordem, ou seja, a função $f(\cdot)$ em (2.4) foi tomada como sendo um polinômio

$$f(y(k-1)) = \theta_0 + \theta_1 y(k-1) + \theta_2 y^2(k-1) . \quad (2.9)$$

TABELA 2.2 - Parâmetros estimados usando diferentes estratégias.

	θ_0	θ_1	θ_2
Sistema original	0	3,97	-3,97
Tradicional	$1,4378 \times 10^{-2}$	3,8760	-3,8662
Mínimo J_2	$-9,7035 \times 10^{-3}$	4,0046	-3,9839
Mínimo J_3	$9,8315 \times 10^{-4}$	3,9694	-3,9698
Mínimo J_4	$3,8400 \times 10^{-3}$	3,9627	-3,9692
Mínimo J_5	$1,5804 \times 10^{-3}$	3,9682	-3,9669
Mínimo J_6	$1,3337 \times 10^{-3}$	3,9670	-3,9641
Mínimo J_7	$-3,0083 \times 10^{-3}$	3,9811	-3,9725
Mínimo J_8	$-1,5342 \times 10^{-3}$	3,9772	-3,9671
Mínimo J_9	$4,7746 \times 10^{-4}$	3,9720	-3,9668
Mínimo J_{10}	$-1,9323 \times 10^{-3}$	3,9803	-3,9739

Dessa forma, a estrutura foi fixada e possui um vetor de três parâmetros $\boldsymbol{\theta} = [\theta_0 \ \theta_1 \ \theta_2]^T$. Esses parâmetros foram então estimados usando duas metodologias: a tradicional, via minimização do erro de predição um passo à frente (2.6), que resolve o PIR e uma que emprega a minimização dos erros até n passos à frente, que se aproxima de resolver o PIC. O erro de predição n passos à frente é dado por

$$\boldsymbol{\xi}_n(k+n) = \mathbf{y}(k+n) - f^{[n]}(\mathbf{y}(k)) \ , \quad (2.10)$$

sendo $f^{[n]}(\cdot) = f(f^{[n-1]}(\cdot))$, com $f^{[1]}(\cdot) = f(\cdot)$, a n -ésima iteração de $f(\cdot)$. A função de custo adotada é o somatório dos erros até o passo n (JAEGER; KANTZ, 1996)

$$J_n = \sum_{k=1}^N \sum_{i=1}^n \|\boldsymbol{\xi}_i(k)\|_2^2 \ . \quad (2.11)$$

Se $n = N$, ou seja, o erro envolver todos os dados coletados, a função de custo (2.11) está associada ao PIC. Na prática, $n \ll N$, mas qualquer $n > 1$ deveria ser melhor do que o erro de predição um passo à frente.

Como estratégias alternativas à identificação tradicional, foram estimados modelos que minimizam J_i para $i = 2, \dots, 10$. Os modelos obtidos foram então simulados. Os parâmetros estimados usando a estratégia convencional (erro um passo à frente) e as funções de custo (2.11) são mostrados na Tabela 2.2, juntamente com os parâmetros do sistema original.

Pode-se notar que todas as estratégias alternativas apresentam coeficientes mais pró-

TABELA 2.3 - Expoentes de Lyapunov para o sistema e os modelos identificados

Método	Expoente
Sistema original	0,59116
Tradicional	0,51312
Mínimo J_2	0,60886
Mínimo J_3	0,59432
Mínimo J_4	0,59998
Mínimo J_5	0,60437
Mínimo J_6	0,59289
Mínimo J_7	0,60558
Mínimo J_8	0,60022
Mínimo J_9	0,60466
Mínimo J_{10}	0,61156

ximos do sistema original que os do método convencional, sendo o modelo de mínimo J_3 o que mais se aproxima do sistema. No entanto, mais importante que os valores dos parâmetros em si é o comportamento temporal do modelo identificado. Para mensurar esse comportamento temporal, foram computados os expoentes de Lyapunov do sistema original e dos modelos identificados. Os resultados se encontram na Tabela 2.3.

O melhor modelo identificado, segundo o critério do expoente de Lyapunov, é aquele que minimiza J_6 , enquanto o modelo de “melhores” coeficientes é o de mínimo J_3 . Isso é uma indicação de que, sob o ponto de vista dinâmico, os coeficientes em si talvez não sejam tão importantes. Outro ponto interessante é que *todos* os modelos que minimizam de J_2 a J_{10} possuem os expoentes de Lyapunov mais próximos dos do sistema original.

Existe, ainda, uma outra dificuldade decorrente de se resolver o PIR. Observando a Tabela 2.1, nota-se que os dados coletados do sistema a ser identificado aparecem listados como causas no PDR, além de serem também efeitos. Isso se torna um problema quando os dados se encontram contaminados com ruído, como normalmente acontece em casos práticos.

A solução para um problema inverso baseada na minimização da função custo (2.7) supõe que as incertezas, ou ruídos, estejam contidas somente nos dados y e não nas causas x . Especificamente falando, essa abordagem não é adequada para casos em que há ruído contaminando x . Esse problema leva à polarização da solução do pro-

blema inverso (JAEGER; KANTZ, 1996). A polarização, ou viés, é uma característica normalmente indesejável em identificação de sistemas. Ela significa que as soluções obtidas para o problema inverso são sistematicamente erradas, ou seja, sua média não corresponde com o valor “ótimo” desejado.

Para que a abordagem via otimização não gere soluções polarizadas é necessário que os ruídos nos dados sejam ruídos dinâmicos e brancos. Para maiores detalhes ver (AGUIRRE, 2000a). Se há ruído aditivo, uma situação bastante comum, é necessário usar um tratamento específico para evitar a polarização. Uma possibilidade é usar uma função de custo modificada mostrada em (JAEGER; KANTZ, 1996). Outra alternativa é usar algumas das técnicas mostradas em (AGUIRRE, 2000a), como o estimador de mínimos quadrados estendidos. De qualquer forma, o processo de otimização se torna mais complexo.

Tal problema ocorre devido aos dados aparecerem como causa no PDR. Isso não acontece no PDC pois as suas causas não dependem dos dados⁴, o que diminui a chance da solução do PIC ser polarizada.

2.6 Conclusão

Neste Capítulo, a estrutura básica dos métodos de identificação de sistemas tradicionais é apresentada. O problema da identificação de sistemas é formulado explicitamente como um problema inverso. Tal problema inverso consiste na determinação da estrutura e dos parâmetros de um modelo a partir de dados da evolução temporal de um sistema. Observa-se que as técnicas tradicionais de identificação não atacam esse problema, chamado aqui de Problema Inverso Completo, ou PIC, mas sim resolvem um problema diferente, o Problema Inverso Reduzido, ou PIR.

Essa é uma das causas do mau desempenho dos métodos tradicionais muitas vezes observado em casos práticos, principalmente quando se trata de sistemas não lineares.

Existem outros métodos de identificação que contornam essa dificuldade e procuram resolver o PIC. Por exemplo, em Timmer *et al.* (2000) é usada uma estratégia mais próxima do PIC, levando a resultados bastante interessantes. Porém, essa abordagem

⁴A rigor, no PDC elas podem depender dos dados e serem contaminadas com ruído se o modelo tiver entradas. Geralmente, no entanto, as entradas são conhecidas com precisão pois foram projetadas pelo usuário de identificação. Por outro lado, se elas precisam ser medidas como os outros sinais, a solução para o PIC pode sofrer de polarização da mesma forma que a do PIR.

leva a um problema de otimização não linear sujeito a restrições bastante complexo.

O presente trabalho procura ser uma forma alternativa para a solução do PIC, construída usando por base as características dinâmicas do sistema sob estudo, e, ao mesmo tempo, evitando um problema de otimização muito complexo.

CAPÍTULO 3

SISTEMAS DINÂMICOS CAÓTICOS E SINCRONIZAÇÃO

3.1 Introdução

A dinâmica caótica em sistemas físicos (GLEICK, 1991; FIEDLER-FERRARA; PRADO, 1994) é um tópico de estudo bastante ativo. Um sistema caótico, apesar de determinístico, apresenta um comportamento dinâmico muito complexo, aperiódico. Depois que a atenção da comunidade científica foi atraída para o assunto, o caos foi observado em uma grande variedade de sistemas físicos (GLEICK, 1991).

Uma das características principais dos sistemas caóticos é a chamada *sensibilidade às condições iniciais*: a trajetória de um sistema caótico iniciada em um determinado ponto diverge exponencialmente, em média, de uma outra trajetória iniciada em um ponto próximo, porém distinto do anterior.

Sincronização de sistemas caóticos é um campo de estudos recente, porém bastante dinâmico. O início do assunto se dá a partir do trabalho de Yamada e Fujisaka (1984) que mostra como dois sistemas caóticos podem ser sincronizados. Uma referência muito conhecida é Pecora e Carroll (1990).

Essa possibilidade de sincronismo é uma das características mais interessantes dos sistemas caóticos. É também, de certa forma, surpreendente, pois, devido à sensibilidade às condições iniciais, dois sistemas caóticos idênticos e independentes, partindo de condições iniciais próximas, apresentam soluções que se distanciam rapidamente e terminam gerando trajetórias não correlacionadas.

A sincronização de sistemas também é um assunto central neste trabalho, pois ela é uma das bases da ferramenta de identificação aqui desenvolvida.

A questão da sincronização foi abordada de várias formas na literatura, levando a resultados diferentes. O objetivo deste Capítulo é introduzir a sincronização de sistemas caóticos bem como rever as abordagens que mais interessam para este trabalho.

A Seção 3.2 apresenta algumas definições sobre sistemas dinâmicos que serão usadas neste Capítulo. Na Seção 3.3 o problema da sincronização é apresentado e a literatura a respeito é revista.

3.2 Sistemas dinâmicos: definições

Como apresentado no Capítulo 2, um sistema dinâmico pode ser representado por uma equação diferencial ordinária. Usualmente, utilizam-se *sistemas* de equações diferenciais ordinárias de primeira ordem na análise de sistemas dinâmicos. Um sistema de equações diferenciais tem a seguinte forma:

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), t), \quad (3.1)$$

sendo $\mathbf{x} \in \mathbb{R}^n$ o vetor de estados, $\mathbf{u} \in \mathbb{R}^p$ o vetor de entradas e t é o tempo. Nesta equação, $f : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^1 \rightarrow \mathbb{R}^n$ é uma função não linear genérica que determina as derivadas temporais das variáveis de estado. O valor n é a *ordem* do sistema.

Um *problema de valor inicial* consiste em determinar qual a evolução temporal dos estados $\mathbf{x}(t)$ que satisfaz (3.1), sujeito a um vetor de entradas $\mathbf{u}(t)$ e que passa pelo estado \mathbf{x}_0 no instante $t = 0$. Tal evolução é a *solução* do problema de valor inicial e também é chamada de *órbita*. Esta evolução temporal representa o comportamento do sistema para a entrada $\mathbf{u}(t)$ e condição inicial \mathbf{x}_0 .

Por exemplo, seja o seguinte sistema de equações diferenciais ordinárias, o sistema de Lorenz (LORENZ, 1963)

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1) \\ \dot{x}_2 = rx_1 - x_2 - x_1x_3 \\ \dot{x}_3 = -bx_3 + x_1x_2 \end{cases} \quad (3.2)$$

sendo que r , σ e b são parâmetros fixos.

Este é um sistema de terceira ordem, autônomo (sem dependência explícita do tempo) e sem entradas. Nesse caso, as derivadas temporais dos estados são determinadas apenas pelos próprios estados e para se conhecer o comportamento temporal do sistema, basta um estado inicial. Uma órbita para o sistema (3.2) com a condição inicial $\mathbf{x}_0 = [1, 1, 1]^T$ e $(r, \sigma, b) = (28, 10, \frac{8}{3})$ foi calculada numericamente e é mostrada na Figura 3.1.

Uma ferramenta de análise muito usada é o *espaço de estados*. Este é um espaço euclidiano de dimensão igual à ordem do sistema em estudo e cujas coordenadas são as variáveis de estado (MONTEIRO, 2002). Uma órbita de um sistema é represen-

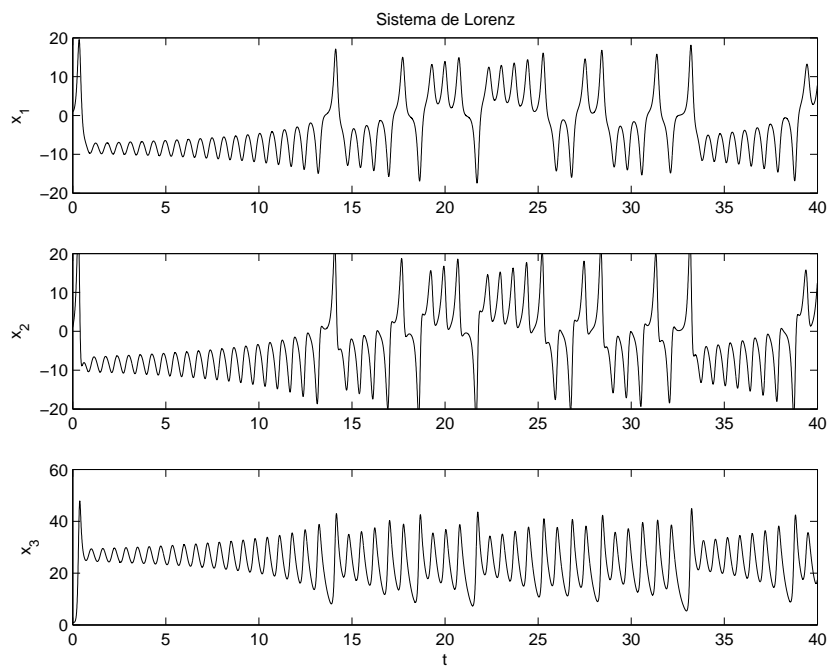


FIGURA 3.1 - Órbita com a condição inicial $\mathbf{x}_0 = [1, 1, 1]^T$ para o sistema (3.2).

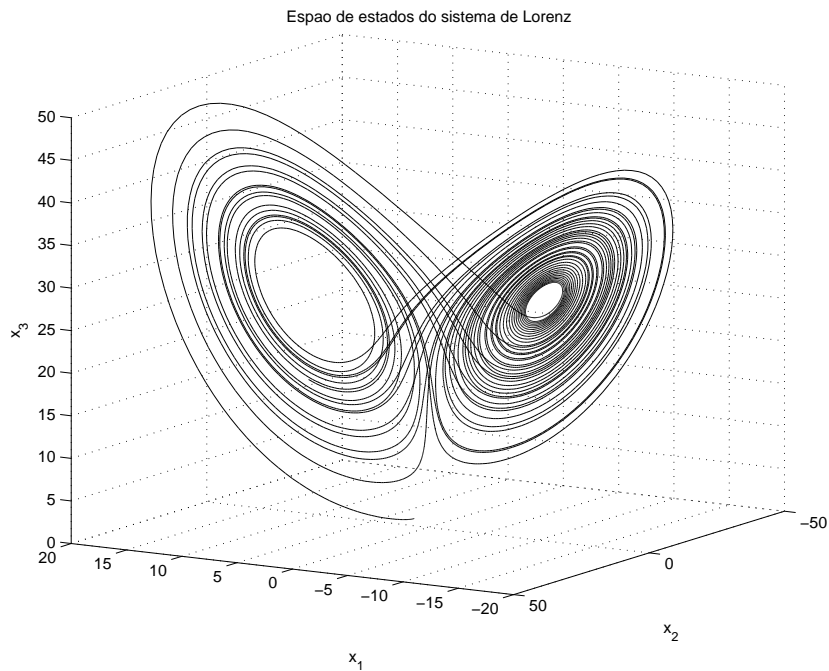


FIGURA 3.2 - Trajetória ou órbita no espaço de estados.

tada por uma trajetória no espaço de estados. Esta trajetória é uma curva que é parametrizada pelo tempo.

A Figura 3.2 mostra a mesma órbita da Figura 3.1 no espaço de estados.

O sistema de Lorenz (3.2), para os parâmetros $(r, \sigma, b) = (28, 10, \frac{8}{3})$, apresenta comportamento caótico.

O significado usual da palavra *caos* é desordem. É uma palavra de origem grega que representava o vazio absoluto que reinava antes do surgimento do universo, ou seja, do *cosmo*, na mitologia grega (FERREIRA, 1975). No entanto, o significado de caos usado neste trabalho é diferente. Caos aqui designa um tipo de comportamento que ocorre com certa frequência em sistemas dinâmicos e cujas principais características são a aperiodicidade e a sensível dependência a condições iniciais.

O comportamento caótico se parece, superficialmente, com o comportamento aleatório, mas não poderia ser conceitualmente mais distinto. Enquanto o comportamento aleatório ou estocástico é resultado de um processo que envolve acaso, o comportamento caótico é o resultado de um processo *determinístico*.

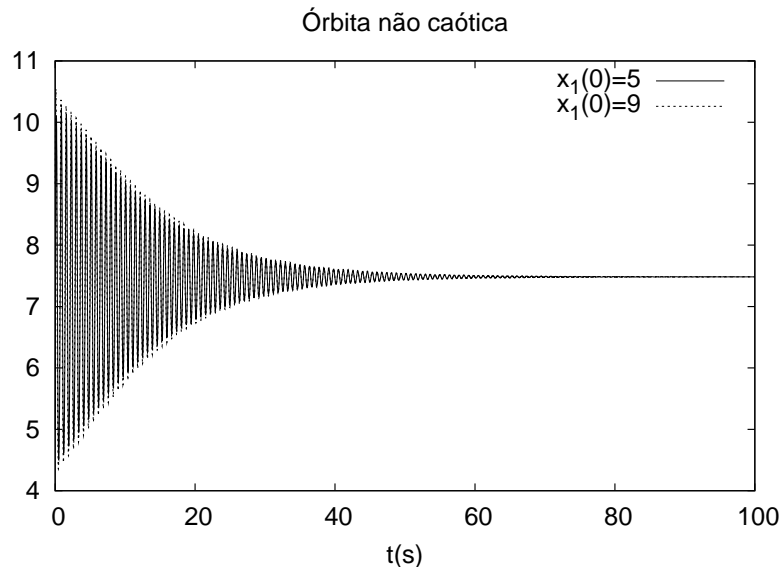


FIGURA 3.3 - Componente x_1 de duas órbitas não caóticas com condições iniciais distintas.

Um sistema caótico é um sistema dinâmico (como definido no Capítulo 2) que apresenta órbitas caóticas. Uma órbita caótica é limitada, não-periódica e diverge em média exponencialmente de outras órbitas próximas (MONTEIRO, 2002).

A principal característica de um sistema caótico é a sensibilidade às condições iniciais. A rigor, as órbitas de qualquer sistema dinâmico dependem das condições iniciais. No entanto, num sistema caótico, essa dependência é muito grande. Num sistema não caótico limitado, órbitas de condições iniciais distintas tendem a se aproximar com o tempo, se manter a distâncias fixas ou se afastar linearmente umas das outras. Um exemplo típico de órbitas não caóticas pode ser encontrado no próprio sistema de Lorenz (3.2). Para os valores de parâmetros $(r, \sigma, b) = (22, 10, \frac{8}{3})$, esse sistema apresenta órbitas não caóticas. A Figura 3.3 mostra a componente x_1 de duas órbitas, uma com a condição inicial $(x_1, x_2, x_3) = (9, 11, 18)$ e a outra com $(x_1, x_2, x_3) = (5, 11, 18)$. A Figura 3.4 mostra a projeção no plano x_1x_2 do espaço de estados dessas mesmas órbitas enquanto que a Figura 3.5 mostra a evolução temporal da distância dessas órbitas no espaço de estados. As órbitas se aproximam praticamente de forma exponencial.

O comportamento de órbitas caóticas é bastante diferente. O mesmo sistema de Lorenz, agora com os parâmetros nos valores $(r, \sigma, b) = (28, 10, \frac{8}{3})$, exibe comportamento caótico. Duas órbitas, uma com condição inicial $(x_1, x_2, x_3) =$

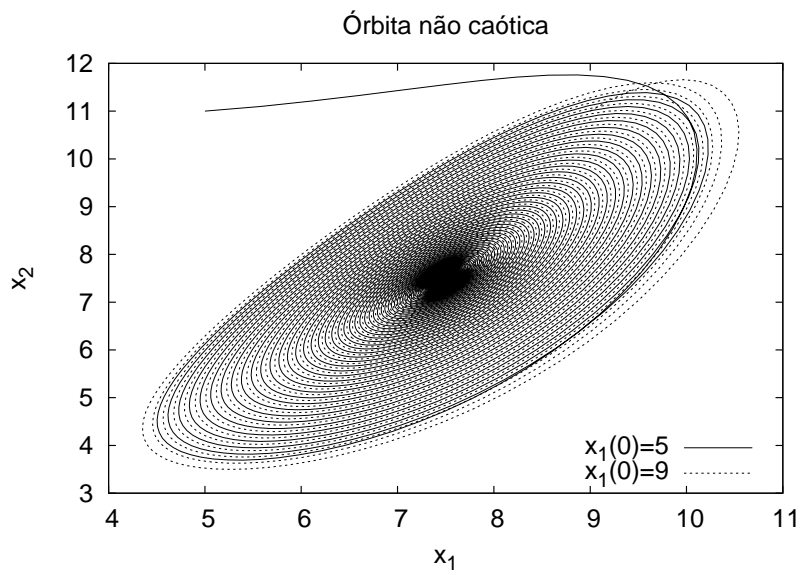


FIGURA 3.4 - Projeção no plano x_1x_2 de duas órbitas não caóticas com condições iniciais distintas.

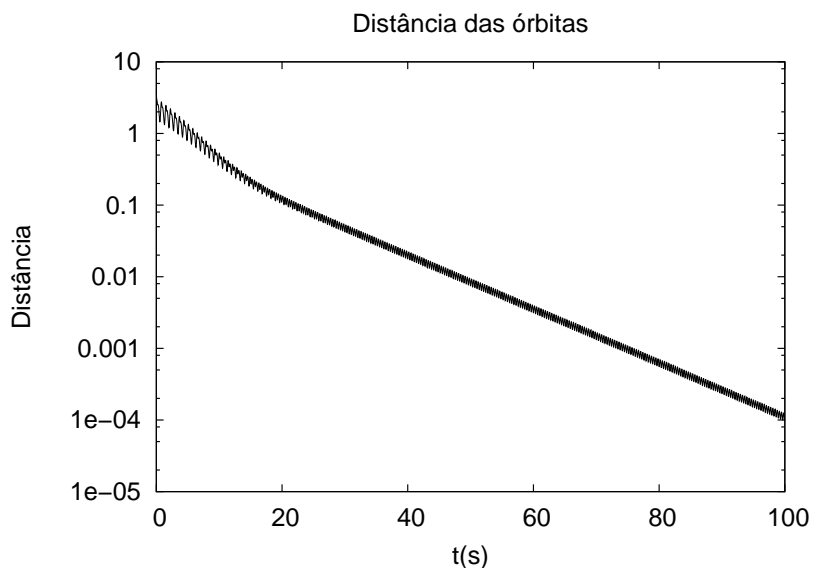


FIGURA 3.5 - Evolução temporal das distâncias no espaço de estados de duas órbitas não caóticas com condições iniciais distintas.

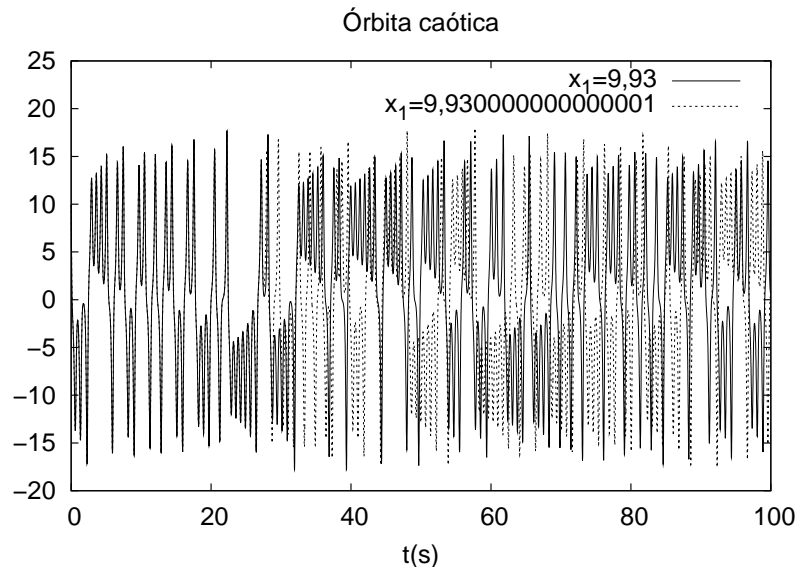


FIGURA 3.6 - Componente x_1 de duas órbitas caóticas com condições iniciais distintas.

$(9, 93; -0, 76; 38, 30)$ e a outra com $(x_1, x_2, x_3) = (9, 9300000000000001; -0, 76; 38, 30)$ foram simuladas e são mostradas na Figuras 3.6, 3.7 e 3.5. Ao contrário do exemplo anterior, a distância inicial, embora mais de 14 ordens de grandeza menor que no caso não caótico, *avumenta* com o passar do tempo.

As Figuras de 3.6 a 3.8 ilustram as principais características de uma órbita caótica: a aperiodicidade e a sensibilidade às condições iniciais. Na Figura 3.6, nota-se que as órbitas não se repetem, embora permaneçam confinadas a um intervalo limitado. Na Figura 3.8, observa-se que as duas órbitas que, inicialmente estão próximas, se afastam quase exponencialmente até o limite do tamanho do atrator.

3.3 Sincronização de sistemas caóticos

Um dos fatos mais marcantes sobre sistemas caóticos é a sua possibilidade de sincronização. Esse fato é marcante porque, num certo sentido, os sistemas caóticos *desafiam* a sincronização por possuírem sensibilidade às condições iniciais (BOCCALETTI *et al.*, 2002). As órbitas de dois sistemas caóticos, mesmo que sejam iguais e comecem de condições iniciais arbitrariamente próximas, divergem exponencialmente. Para que ocorra a sincronização entre dois sistemas caóticos é necessário alguma forma de acoplamento entre eles.

Existem duas maneiras possíveis de se acoplar dois sistemas (BOCCALETTI *et al.*,

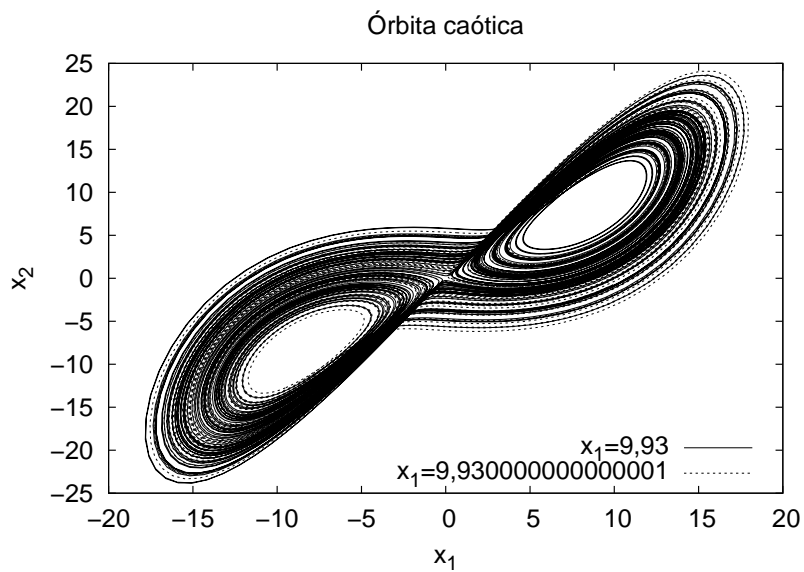


FIGURA 3.7 - Projeção no plano x_1x_2 de duas órbitas caóticas com condições iniciais distintas.

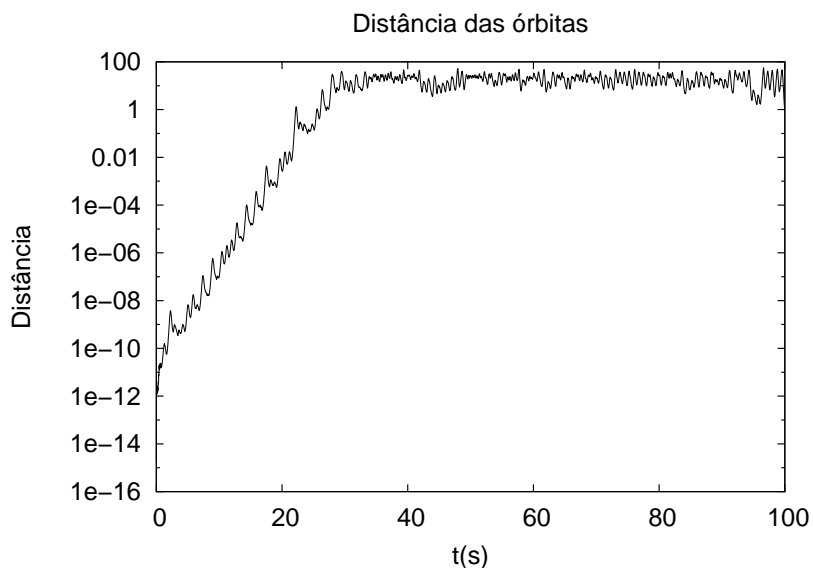


FIGURA 3.8 - Evolução temporal das distâncias no espaço de estados de duas órbitas caóticas com condições iniciais distintas.

2002). Uma possibilidade é cada sistema influenciar o outro, num acoplamento bidirecional. A outra é um sistema evoluir livremente enquanto “perturba” o outro. Neste caso, o acoplamento é unidirecional. Esta segunda forma de acoplamento é a estudada neste trabalho. O motivo disso é que, no método de identificação aqui proposto, deseja-se sincronizar o modelo ao sistema sob estudo. Dessa forma, é mais flexível supor que apenas os dados coletados do sistema estão disponíveis e que não há meios de se fazer um acoplamento bidirecional entre este e o modelo.

Esse acoplamento unidirecional também é chamado de acoplamento *mestre-escravo*, sendo o sistema mestre aquele que evolui livremente e o escravo aquele que o “segue”.

Existem várias definições de sincronismo para sistemas caóticos, como *sincronismo idêntico*, de *fase* e *generalizado* (BOCCALETTI *et al.*, 2002). Dentre essas diversas possibilidades, o tipo de sincronismo que interessa para esse trabalho é o sincronismo idêntico. Nesse tipo de sincronização, os estados do sistema escravo se tornam uma cópia dos do sistema mestre. Ou seja, suponha que $\mathbf{x}_1(t)$ seja o vetor de estados do do sistema mestre e $\mathbf{x}_2(t)$ seja o vetor de estados do sistema escravo. Define-se como sincronização idêntica a condição

$$\lim_{t \rightarrow \infty} [\mathbf{x}_1(t) - \mathbf{x}_2](t) = 0 . \quad (3.3)$$

A ocorrência ou não de sincronismo entre dois sistemas caóticos depende da natureza dos sistemas e da forma de acoplamento entre eles. Há na literatura várias formas de se abordar o problema de sincronização. A seguir, algumas dessas formas são revistas.

3.3.1 Substituição completa

O princípio desse método é substituir algumas das variáveis de estado do sistema escravo, que neste caso é uma cópia do sistema mestre, por suas correspondentes do sistema mestre (PECORA *et al.*, 1997). As variáveis \mathbf{v} formam então o acoplamento entre os dois sistemas.

Seja o sistema mestre um oscilador autônomo não linear de ordem n , possivelmente caótico, representado por¹

$$\dot{\mathbf{x}} = f(\mathbf{x}) . \quad (3.4)$$

¹Supõe-se aqui que o sistema admite solução única para uma condição inicial qualquer.

Esse sistema é então dividido arbitrariamente em dois subsistemas \mathbf{v} e \mathbf{w} de forma que

$$\mathbf{x} = [\mathbf{v}, \mathbf{w}]^T \quad (3.5)$$

e

$$\begin{aligned} \dot{\mathbf{v}} &= g(\mathbf{v}, \mathbf{w}) \\ \dot{\mathbf{w}} &= h(\mathbf{v}, \mathbf{w}) \end{aligned} \quad (3.6)$$

sendo $\mathbf{v} = (x_1, \dots, x_m)^T$, $g = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T$, $\mathbf{w} = (x_{m+1}, \dots, x_n)^T$ e $h = (f_{m+1}(\mathbf{x}), \dots, f_n(\mathbf{x}))^T$.

Em seguida, o subsistema \mathbf{w} é copiado, identicamente, gerando o subsistema \mathbf{w}' , chamado sistema escravo. Este último é então acrescentado ao sistema original gerando

$$\begin{aligned} \dot{\mathbf{v}} &= g(\mathbf{v}, \mathbf{w}) \\ \dot{\mathbf{w}} &= h(\mathbf{v}, \mathbf{w}) \\ \dot{\mathbf{w}}' &= h(\mathbf{v}, \mathbf{w}') \end{aligned} \quad (3.7)$$

Ao final, tem-se um sistema mestre composto por um oscilador não linear autônomo e uma cópia idêntica de um subsistema do sistema mestre, o sistema escravo. Nesse arranjo, o subsistema escravo se comporta como se fosse um outro oscilador, idêntico ao mestre, do qual algumas variáveis de estado foram substituídas pelas correspondentes do sistema mestre. Essas variáveis são a informação que é transmitida do mestre para o escravo.

Dependendo do sistema em questão e da escolha do subsistema, a substituição completa leva à sincronização idêntica do subsistema escravo com o subsistema correspondente do mestre, ou seja

$$\lim_{t \rightarrow \infty} \Delta \mathbf{w} = \lim_{t \rightarrow \infty} \mathbf{w} - \mathbf{w}' = 0 \quad (3.8)$$

A ocorrência ou não de sincronização, dado um certo par mestre escravo, depende de alguns fatores.

Inicialmente, deve existir um hiperplano de sincronização que seja invariante segundo a dinâmica. Para um sistema² descrito pelas equações (3.7) o hiperplano de sincronização é dado por $\mathbf{w} = \mathbf{w}'$. O hiperplano ser invariante segundo a dinâmica significa

²Nesse caso o sistema é formado pela união do mestre com o escravo.

que, para qualquer condição inicial no hiperplano, a solução do sistema permanece no hiperplano. Para o sistema (3.7), a invariância é trivial. Como os subsistemas \mathbf{w} e \mathbf{w}' são idênticos, se em algum instante de tempo t , os estados $\mathbf{w}(t)$ e $\mathbf{w}'(t)$ forem iguais eles permanecerão iguais a partir de então. Isso é uma decorrência direta do teorema da existência e unicidade das soluções de equações diferenciais.

Uma outra condição necessária para a sincronização é que esse hiperplano seja localmente estável, ou seja, pequenas perturbações para fora do hiperplano devem ser amortecidas. Uma forma de analisar essa condição utiliza a seguinte mudança de coordenadas (PECORA *et al.*, 1997):

$$\begin{aligned} \mathbf{v} &= \mathbf{v} \\ \mathbf{w}_{\parallel} &= \mathbf{w} + \mathbf{w}' \quad . \\ \mathbf{w}_{\perp} &= \mathbf{w} - \mathbf{w}' \end{aligned} \quad (3.9)$$

Essa mudança de coordenadas leva a um sistema do qual $2n - m$ variáveis de estado (\mathbf{v} e \mathbf{w}_{\parallel}) se encontram no hiperplano de sincronização e $n - m$ variáveis (\mathbf{w}_{\perp}) se encontram no espaço perpendicular ao hiperplano. A estabilidade do hiperplano de sincronização depende da estabilidade do subsistema \mathbf{w}_{\perp} .

A equação para \mathbf{w}_{\perp} é

$$\dot{\mathbf{w}}_{\perp} = h(\mathbf{v}, \mathbf{w}) - h(\mathbf{v}, \mathbf{w}') \quad (3.10)$$

que é aproximada no limite $\mathbf{w}_{\perp} \rightarrow 0$ por

$$\dot{\mathbf{w}}_{\perp} = Dh(\mathbf{v}, \mathbf{w})\mathbf{w}_{\perp} \quad , \quad (3.11)$$

sendo $Dh(\mathbf{v}, \mathbf{w})$ a matriz jacobiana de $h(\cdot)$ calculada em (\mathbf{v}, \mathbf{w}) . Se $Dh(\mathbf{v}, \mathbf{w})$ não depender de (\mathbf{v}, \mathbf{w}) e o sistema for hiperbólico, pode-se calcular os seus autovalores e determinar a estabilidade local do hiperplano a partir deles. Porém, esse não é o caso para a maioria dos sistemas não lineares. Assim é necessária uma outra forma de abordar o problema. Pode-se utilizar nesse caso os expoentes de Lyapunov. Mais especificamente os expoentes de Lyapunov restritos ao subsistema \mathbf{w}_{\perp} (PECORA *et al.*, 1997).

Os expoentes de Lyapunov medem as taxas médias, em diferentes direções, com as

quais trajetórias próximas se afastam ou se aproximam³. Normalmente os expoentes de Lyapunov são calculados para o sistema inteiro, mas o interessante aqui é calcular os expoentes associados somente ao subsistema \mathbf{w}_\perp . Se estes expoentes forem negativos, o hiperplano de sincronização é localmente estável. É importante ressaltar que a condição de estabilidade local é apenas necessária, e não suficiente, para ocorrer o sincronismo.

Os expoentes de Lyapunov de um sistema qualquer dependem da sua matriz jacobiana. Como a matriz jacobiana para o subsistema \mathbf{w}_\perp é a mesma para o subsistema \mathbf{w} , basta calcular os expoentes de Lyapunov para esse último. Esse cálculo é feito da forma tradicional (ECKMANN; RUELLE, 1985), porém tendo como sistema somente o subsistema desejado.

A seguir é mostrado um exemplo de sincronização de sistemas caóticos usando o sistema de Lorenz 3.2.

Esse sistema é passível de ser sincronizado via substituição completa se se fizer $\mathbf{v} = x_2$ e $\mathbf{w} = (x_1, x_3)^\top$. Também existe a possibilidade de se transmitir x_1 e usar $\mathbf{w} = (x_2, x_3)^\top$. A outra opção (transmitir x_3) não é localmente estável.

Neste exemplo foi implementado a opção com $\mathbf{v} = x_2$. Os parâmetros usados foram $(r, \sigma, b) = (60, 10, \frac{8}{3})$. Na figura 3.9 está representado esquematicamente o sistema, com o oscilador mestre, o escravo e o sinal transmitido.

O sistema completo — o mestre junto com escravo — foi simulado em computador, utilizando o programa Octave. Como condição inicial foi usado $(x_1, x_2, x_3, y_2, x_3) = (15, 0, 60, -15, 60)$.

Na figura 3.10 é mostrado o início da simulação quando o oscilador escravo ainda não havia sincronizado. Nota-se que, mesmo havendo começado do outro lado do atrator, o sistema escravo se aproxima rapidamente do mestre.

A figura 3.11 mostra a simulação inteira. Pode-se ver que apenas um traço é visível a maior parte do tempo, o que indica que o estado sincronizado se mantém.

³Para maiores detalhes ver Ott (1993)

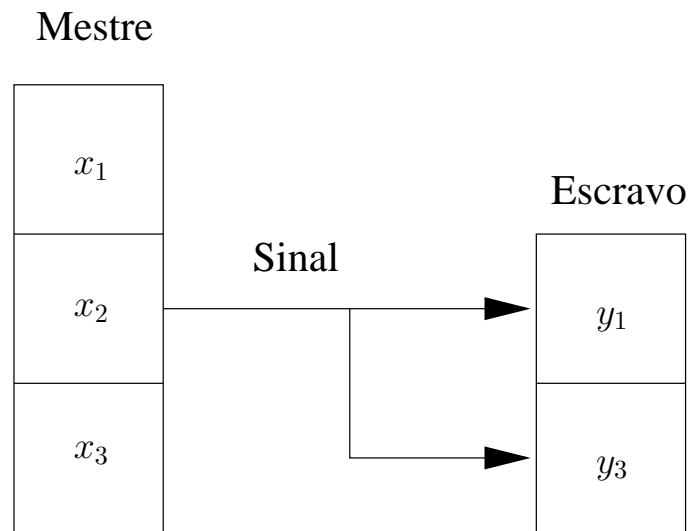


FIGURA 3.9 - Representação gráfica do esquema de sincronização mestre-escravo para o sistema de Lorenz com substituição completa.

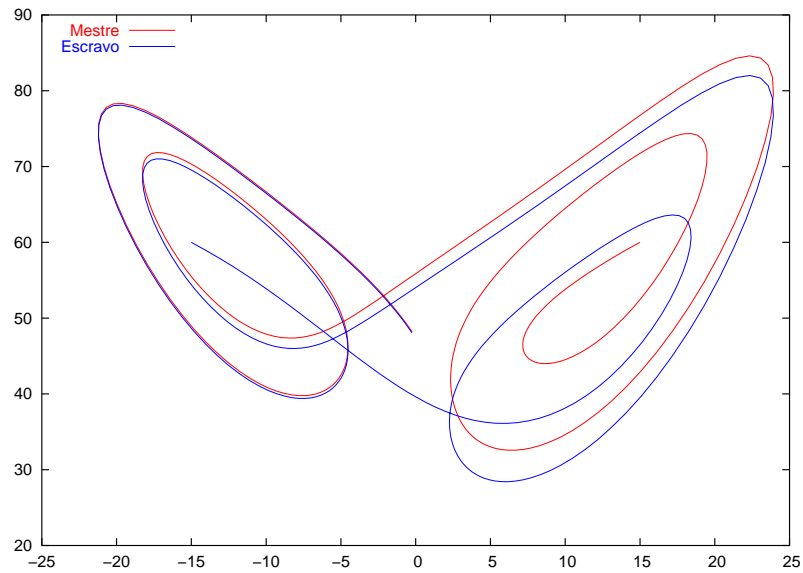


FIGURA 3.10 - Projeções no plano (x_1, x_3) dos osciladores mestre e escravo para o início da simulação.

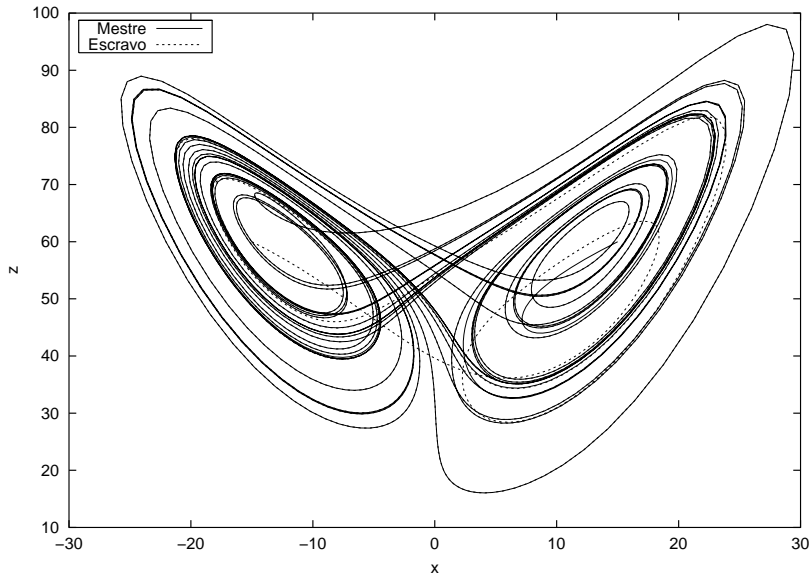


FIGURA 3.11 - Projeções no plano (x_1, x_3) dos osciladores mestre e escravo para a simulação completa.

3.3.2 Observadores como forma de sincronização

O problema de observação de um sistema é clássico na teoria de controle linear. Em resumo, ele consiste em reproduzir o estado de um sistema a partir de observações do mesmo e do conhecimento de seu modelo.

O problema linear é da seguinte forma. Seja um sistema linear dado por

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} \\ \mathbf{y} = C\mathbf{x} \end{cases}, \quad (3.12)$$

sendo \mathbf{x} o vetor de estado do sistema, \mathbf{y} o vetor de saída. A e C são matrizes, respectivamente $n \times n$ e $p \times n$. O observador linear é então construído como uma cópia do sistema original acrescida de um termo de correção

$$\dot{\hat{\mathbf{x}}} = A\hat{\mathbf{x}} + K(\mathbf{y} - C\hat{\mathbf{x}}), \quad (3.13)$$

sendo $\hat{\mathbf{x}}$ o vetor de estados estimado e K uma matriz de ganho cujo cálculo depende de A e C , entre outros fatores.

Dependendo do sistema, é possível escolher K de forma que o estado estimado se

aproxime assintoticamente do estado do sistema original, ou seja,

$$\lim_{t \rightarrow \infty} [\mathbf{x}(t) - \hat{\mathbf{x}}(t)] = 0 . \quad (3.14)$$

Para que isso ocorra o sistema (3.12) deve ser *observável*. A observabilidade é uma propriedade que pode ser verificada para um sistema linear e depende das suas matrizes A e C (CHEN, 1993).

Comparando 3.14 com 3.3 percebe-se que o que um observador faz é exatamente sincronizar com o sistema para recuperar o seu estado.

Seguindo nessa filosofia, a sincronização de sistemas caóticos pode ser vista como um problema de observação. Tem-se um oscilador não linear representado por

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}) \\ \mathbf{y} = g(\mathbf{x}) \end{cases} . \quad (3.15)$$

O problema agora é criar um estimador para o vetor de estados \mathbf{x} , baseando-se no modelo do sistema, e que funcione a partir dos dados de saída $\mathbf{y}(t)$. Essa abordagem é filosoficamente diferente da abordagem via substituição completa. O foco agora é determinar se o sistema é observável e, se possível, projetar um sistema escravo que seja capaz de o sincronizar. É uma metodologia mais voltada para a síntese. Normalmente o sistema escravo é uma cópia do sistema mestre, mas isso não é obrigatório. Qualquer sistema que seja capaz de recuperar o estado do mestre é válido.

Uma boa referência sobre essa abordagem é Nijmeijer e Mareels (1997). A abordagem por observadores para o problema da sincronização se mostrou uma idéia produtiva. Muitos trabalhos foram feitos na área (SO *et al.*, 1994; GRASSI; MILLER, 2002; MILLERIOUX; DAAFOUZ, 2001; LIAO; HUANG, 2001; BOUTAYEB *et al.*, 2002; WANG; WANG, 1998; WU; LEE, 2000).

No caso de sistemas não lineares, o problema ainda se encontra em aberto. Em Nijmeijer e Mareels (1997) alguns resultados interessantes são apresentados.

3.3.3 Sincronização como um problema de controle

Uma outra forma de se abordar o problema de sincronização é sob a óptica de sistemas de controle (NIJMEIJER, 2001; TÔRRES, 2001). Enquanto a abordagem via substituição completa é mais focada no sistema mestre e a abordagem via observadores se preocupa com o projeto do sistema escravo, o objetivo da abordagem de controle é projetar o acoplamento. Neste caso, tanto o oscilador mestre quanto o escravo são dados *a priori*.

Seja o oscilador mestre dado por (3.15) e o oscilador escravo dado por

$$\dot{\hat{\mathbf{x}}} = h(\hat{\mathbf{x}}, \mathbf{u}) , \quad (3.16)$$

sendo \mathbf{u} o vetor de controle. O problema de sincronismo se torna agora projetar a lei de controle de forma que o erro de sincronização $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ tenda para zero independentemente da condição inicial. Uma possibilidade é a lei de controle ser função da observação \mathbf{y} e do estado do estimador $\hat{\mathbf{x}}$

$$u = \alpha(\hat{\mathbf{x}}, \mathbf{y}) . \quad (3.17)$$

A idéia principal nessa abordagem é que o sistema mestre passa a ser visto como a *referência* para o sistema escravo. O objetivo da lei de controle é garantir que o sistema escravo siga essa referência.

Resolver o problema de projetar a lei de controle $\alpha(\cdot)$ pode ser difícil ou mesmo impossível em alguns casos, assim como na abordagem via observadores. No entanto, essa é justamente a abordagem escolhida para tratar do problema de identificação usando sincronização.

As ferramentas para o desenvolvimento dos controladores usados neste trabalho são apresentadas no Capítulo 4.

3.4 Conclusão

Este Capítulo apresentou o fenômeno da sincronização de sistemas caóticos e revê algumas das abordagens usadas no seu estudo.

CAPÍTULO 4

SISTEMAS DE CONTROLE NÃO LINEAR

4.1 Introdução

A teoria de controle, juntamente com a identificação de sistemas e a sincronização de sistemas caóticos, forma a base deste trabalho. A teoria de controle linear é bastante conhecida e usada em aplicações práticas. Ela envolve conhecer e estudar as relações entre as entradas e os estados de um sistema, bem como as relações entre os estados e as saídas. De fundamental importância para esse estudo são os conceitos de *controlabilidade* e *observabilidade* introduzidos por Kalman.

Um sistema é dito controlável se, a partir de um determinado estado dado em um instante inicial, existe um controle (ou entrada) que aplicado ao sistema o leva a um estado final em um instante final também dados¹. O conceito de observabilidade está relacionado com a possibilidade de se reconstruir os estados a partir das saídas do sistema (OGATA, 1993).

A teoria de sistemas lineares fornece as condições necessárias e suficientes para a controlabilidade e a observabilidade, mas somente para sistemas lineares. No caso de sistemas não lineares, essas propriedades são muito mais difíceis de serem verificadas, pelo menos de forma genérica e global. Existem resultados locais e globais, porém restritos a alguns tipos de sistemas.

A teoria de controle geométrico procura definir esses conceitos para sistemas não lineares e é uma ferramenta fundamental neste trabalho devido à natureza não linear dos sistemas aqui tratados.

Outra ferramenta importante são as técnicas de controle adaptativo. O controle adaptativo trata de problemas onde o sistema a ser controlado possui parâmetros que são desconhecidos e é necessário projetar um controlador para tal sistema. A abordagem adaptativa é projetar leis de controle que se “adaptam” ao sistema de alguma forma (SLOTINE; LI, 1991).

O objetivo deste Capítulo é introduzir os principais conceitos da teoria de controle geométrico e controle adaptativo que são usados no Capítulo 5, sendo que o controle

¹A rigor, o tempo final não precisa ser dado, mas deve ser finito.

geométrico é abordado na Seção 4.2 e o controle adaptativo na Seção 4.3.

4.2 Controle geométrico

A teoria de sistemas de controle geométrico é um assunto profundo e razoavelmente recente, sendo que os seus principais resultados foram produzidos a partir da década de 1970 (RESPONDEK, 2001). Uma revisão a fundo dessa teoria está além dos objetivos deste trabalho. Dessa forma, esta Seção apenas apresenta os conceitos que são usados no método aqui proposto.

Uma idéia importante no controle geométrico é a *linearização via realimentação*. Nela, um sistema de controle não linear dado por

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}) \quad (4.1)$$

é transformado, via a mudança de coordenadas

$$\begin{aligned} \tilde{\mathbf{x}} &= \Phi(\mathbf{x}) \\ \tilde{\mathbf{u}} &= \Psi(\mathbf{x}, \mathbf{u}) \end{aligned} \quad (4.2)$$

no sistema

$$\dot{\tilde{\mathbf{x}}} = A\tilde{\mathbf{x}} + B\tilde{\mathbf{u}} . \quad (4.3)$$

Essa abordagem é diferente da linearização do sistema em torno de um ponto de operação, procedimento normalmente empregado em problemas não lineares (OGATA, 1993). O sistema transformado (4.3) é o próprio sistema original, mas em um outro sistema de coordenadas. Por outro lado, a linearização em torno de um ponto de operação é uma aproximação do sistema real no mesmo espaço de coordenadas.

As transformações (4.2) são inversíveis e diferenciáveis por definição. Mais precisamente são difeomorfismos (ISIDORI, 1995). Elas transformam trajetórias do sistema original biunivocamente em trajetórias de (4.3). Isso permite resolver o problema de controle no espaço de (4.3) usando a teoria de sistemas lineares e depois usar as transformações inversas para voltar ao espaço inicial.

A teoria de controle geométrico lida com as questões de quando essas transformações existem e como encontrá-las. Para isso lança mão de ferramentas da geometria diferencial, daí o nome.

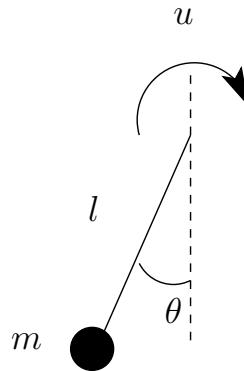


FIGURA 4.1 - Pêndulo simples

Para ilustrar essas idéias, é mostrado a seguir um exemplo, extraído de [Respondek \(2001\)](#).

Seja um pêndulo simples não linear conforme a Figura 4.1.

Esse sistema pode ser descrito por

$$\begin{aligned}\dot{\theta} &= \omega \\ \dot{\omega} &= -\frac{g}{l}\text{sen}\theta + \frac{u}{ml^2},\end{aligned}\tag{4.4}$$

sendo m a massa do pêndulo, l o seu comprimento, g a aceleração da gravidade e u o torque externo aplicado ao pêndulo. As variáveis de estado são o ângulo θ e a velocidade angular ω .

O sistema (4.4) pode representar, por exemplo, um atuador robótico com u sendo o torque fornecido pelo motor e θ o ângulo do braço.

Suponha que a ação de controle u seja substituída pela função

$$u = ml^2\tilde{u} + ml\text{gsen}\theta\tag{4.5}$$

na qual \tilde{u} é uma nova função de entrada. Assim substituindo (4.5) em (4.4) obtém-se

$$\begin{aligned}\dot{\theta} &= \omega \\ \dot{\omega} &= \tilde{u},\end{aligned}\tag{4.6}$$

que é um sistema linear. Para o sistema (4.6) pode-se usar a teoria de controle linear para se projetar um controlador.

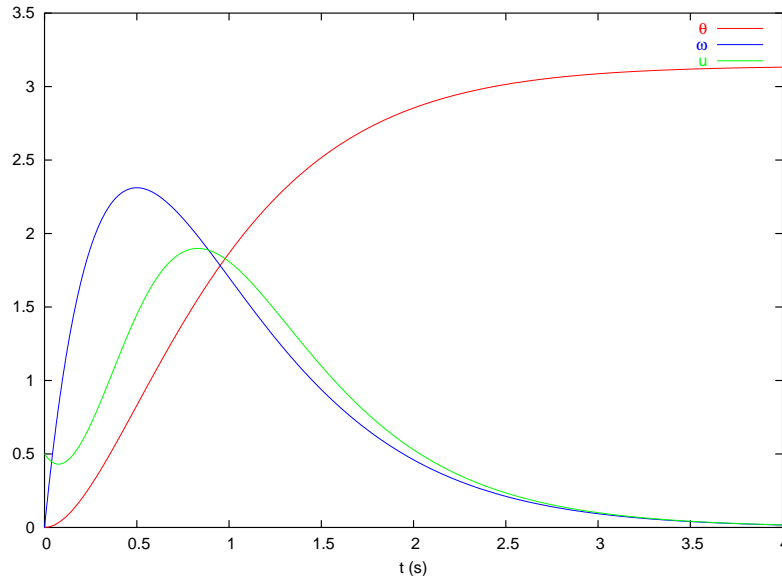


FIGURA 4.2 - Simulação do pêndulo utilizando controle. O ângulo está representado em vermelho, em azul está a velocidade angular e em verde a ação de controle.

Por exemplo, uma possível abordagem para o projeto do controlador é via realimentação de estados

$$\tilde{u} = k_1(\theta - \theta_r) + k_2\omega, \quad (4.7)$$

em que os k_i são valores a serem determinados em função da resposta que se deseja para o sistema e θ_r é a referência. Como (4.6) é controlável (no sentido linear), os seus pólos podem ser alocados teoricamente de forma arbitrária. Supondo que os pólos desejados para o sistema realimentado sejam λ_1 e λ_2 , então os valores dos ganhos do controlador devem ser $k_1 = -\lambda_1\lambda_2$ e $k_2 = \lambda_1 + \lambda_2$.

O sistema (4.4) foi simulado com a substituição (4.5) e o controlador linear (4.7). Como parâmetros usou-se $m = 1\text{kg}$, $l = 0,2\text{m}$ e $g = 9,81\text{m/s}^2$. A referência usada foi $\theta_r = \pi$ (deseja-se estabilizar o ponto de equilíbrio instável) e os parâmetros do controlador foram escolhidos de forma que $\lambda_1 = \lambda_2 = -2$.

O resultado da simulação é mostrado na Figura 4.2.

Pode-se notar a resposta criticamente amortecida do ângulo do pêndulo. Outro ponto importante é a forma da ação de controle. Na Figura 4.2, é mostrada a ação de controle no espaço não linear, ou seja $u(t)$. Ao contrário, $\tilde{u}(t)$ é a ação de controle calculada para um sistema linear.

Esse exemplo ilustra adequadamente alguns conceitos da linearização por realimentação.

Inicialmente, o exemplo mostra um fato comum em controle geométrico: o *cancelamento de dinâmica* (JACKSON; GROSU, 1995; TÔRRES, 2001). A transformação de coordenadas da equação (4.5) efetivamente cancela a parte não linear do sistema (4.4). Com um sistema da forma da equação (4.4) esse cancelamento é trivial de ser obtido. Não é necessário usar as ferramentas do controle geométrico. Porém, se o sistema fosse um pouco diferente, por exemplo, se a não linearidade estivesse na primeira equação de (4.4) ao invés de na segunda, a linearização poderia não ser tão facilmente obtida.

Um cancelamento dessa forma nem sempre é possível. A teoria de controle geométrico procura responder quando essa transformação é exequível e como obtê-la.

A seguir, são formulados conceitos matemáticos que podem ser usados para projetar sistemas de controle de uma única entrada.

4.2.1 Linearização por realimentação de estado

Esta Seção introduz os conceitos de controle não linear que são usados neste trabalho. A descrição detalhada dessa teoria pode ser encontrada em outros textos (NIJMEIJER; SCHAFT, 1990; SLOTINE; LI, 1991; ISIDORI, 1995; JAKUBCZYK, 2001; RESPONDEK, 2001).

Nas definições a seguir, em geral, existe a restrição de suavidade sobre as funções. Uma função f ser suave quer dizer que ela é infinitamente derivável (C^∞). No entanto, na prática, essa restrição pode ser relaxada para apenas um certo número de derivadas contínuas, ou seja, C^p , com p suficientemente grande.

O primeiro conceito importante é a *derivada de Lie* de uma função escalar suave h com respeito a um campo vetorial f , denotada por $L_f h$. Ela é uma função escalar definida como o produto interno do gradiente de h pelo campo vetorial f . É simplesmente a derivada direcional de h na direção de f (ISIDORI, 1995). Uma outra forma de ver a derivada de Lie é considerar que um campo vetorial f define um operador diferencial de primeira ordem no espaço das funções suaves (JAKUBCZYK, 2001).

Em coordenadas, $L_f h$ é definido como

$$L_f h(\mathbf{x}) = \sum_{i=1}^n f_i(\mathbf{x}) \frac{\partial}{\partial x_i} h(\mathbf{x}) , \quad (4.8)$$

sendo f_i as componentes do campo vetorial $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, h uma função escalar $h : \mathbb{R}^n \rightarrow \mathbb{R}$ e $x \in \mathbb{R}^n$.

A ferramenta de capital importância para a análise e síntese de sistemas de controle não linear é o *colchete de Lie*. É uma operação em dois campos vetoriais que produz um terceiro campo vetorial. O colchete de Lie dos campos vetoriais suaves f e g é denotado por $[f, g]$ e pode ser calculado, num sistema de coordenadas, segundo (ISIDORI, 1995)

$$[f, g] = \nabla g f - \nabla f g , \quad (4.9)$$

sendo ∇f a matriz jacobiana de f .

Uma forma alternativa de representar o colchete de Lie de f e g é $\text{ad}_f g$.

Outro conceito importante é o *difeomorfismo*. Uma função $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um difeomorfismo se e somente se ela for suave e sua inversa Φ^{-1} existir e for suave (ISIDORI, 1995). O difeomorfismo pode ser visto como uma mudança de coordenadas e não precisa estar definido em todo o \mathbb{R}^n . É comum haver difeomorfismos válidos somente em uma região aberta Ω de \mathbb{R}^n . Tais difeomorfismos são chamados *locais* e os válidos em todo o espaço, *globais*.

Conforme já mencionado, o objetivo da linearização por realimentação é transformar um sistema não linear genérico (4.1) em um sistema linear controlável através de mudanças de coordenadas da forma (4.2). O problema genérico é um tanto difícil. Nesta seção, será abordado o problema restrito a sistemas com uma entrada e do tipo *afim*

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})u , \quad (4.10)$$

sendo $f(\cdot)$ e $g(\cdot)$ campos vetoriais suaves em \mathbb{R}^n , \mathbf{x} o vetor de estados e u o sinal de entrada.

A forma (4.10) é bastante genérica e descreve uma ampla classe de sistemas. A maior restrição é o fato de se ter somente uma entrada.

Um sistema da forma (4.10) é dito *linearizável por realimentação*, ou controlável, se existir, em uma região aberta Ω em \mathbb{R}^n , um difeomorfismo

$$(\tilde{\mathbf{x}}, \tilde{u}) = \chi(\mathbf{x}, u) = (\Phi(\mathbf{x}), \psi(\mathbf{x}, u)) ,$$

que transforme o sistema (4.10) em um sistema linear

$$\dot{\tilde{\mathbf{x}}} = A\tilde{\mathbf{x}} + b\tilde{u} , \quad (4.11)$$

com

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} .$$

Essa forma particular do sistema (4.11), chamada *forma canônica de Brunovský* (RESPONDEK, 2001), é especialmente interessante para o controle. O sistema (4.11) pode também ser descrito por

$$\begin{cases} \dot{\tilde{x}}_i = \tilde{x}_{i+1} , & i = 1, \dots, n-1 \\ \dot{\tilde{x}}_n = \tilde{u} . \end{cases}$$

Um sistema da forma (4.10) é (localmente) controlável se existir uma região aberta Ω em \mathbb{R}^n onde (RESPONDEK, 2001):

- os campos vetoriais $\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-1} g\}$ sejam linearmente independentes e
- o conjunto $\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-2} g\}$ seja *involutivo*.

A primeira condição é um análogo da condição de controlabilidade para sistemas lineares. A segunda condição está ligada à existência do difeomorfismo $\tilde{\mathbf{x}} = \Phi(\mathbf{x})$.

A independência linear de campos vetoriais é semelhante àquela dos vetores. Um

conjunto de m campos vetoriais é independente se a única solução para

$$\sum_{i=1}^m \alpha_i(\mathbf{x}) f_i = 0$$

em termos das funções escalares α_i for $\alpha_i = 0, i = 1, \dots, m$.

Um conjunto de campos vetoriais é involutivo se o colchete de Lie de quaisquer dois elementos seus forma um conjunto linearmente dependente com os demais campos vetoriais (JAKUBCZYK, 2001).

Normalmente essas condições não são satisfeitas em todo o espaço. Nesse caso, o controlador é válido somente em uma região aberta do espaço de estados. Uma vez verificado se as condições de controlabilidade foram satisfeitas, procede-se à síntese do controlador.

O primeiro passo é obter a transformação $\tilde{\mathbf{x}} = \Phi(\mathbf{x})$. Ela é dada por (SLOTINE; LI, 1991)

$$\Phi(\mathbf{x}) = \begin{bmatrix} \phi(\mathbf{x}) \\ L_f \phi(\mathbf{x}) \\ \vdots \\ L_f^{n-1} \phi(\mathbf{x}) \end{bmatrix} \quad (4.12)$$

e a função $\phi(\cdot)$ é uma solução para o sistema de equações diferenciais parciais

$$\begin{cases} L_g \phi(\mathbf{x}) = 0 \\ L_g L_f \phi(\mathbf{x}) = 0 \\ \vdots \\ L_g L_f^{n-2} \phi(\mathbf{x}) = 0 \\ L_g L_f^{n-1} \phi(\mathbf{x}) \neq 0. \end{cases} \quad (4.13)$$

A existência da função $\phi(\cdot)$ como solução para esse sistema é uma consequência da condição de involutividade (SLOTINE; LI, 1991).

A outra etapa é a função de realimentação $u = \Psi(\mathbf{x}, \tilde{u})$. Ela é dada por (SLOTINE; LI, 1991)

$$\Psi(\mathbf{x}, \tilde{u}) = \frac{\tilde{u} - \alpha(\mathbf{x})}{\beta(\mathbf{x})}, \quad (4.14)$$

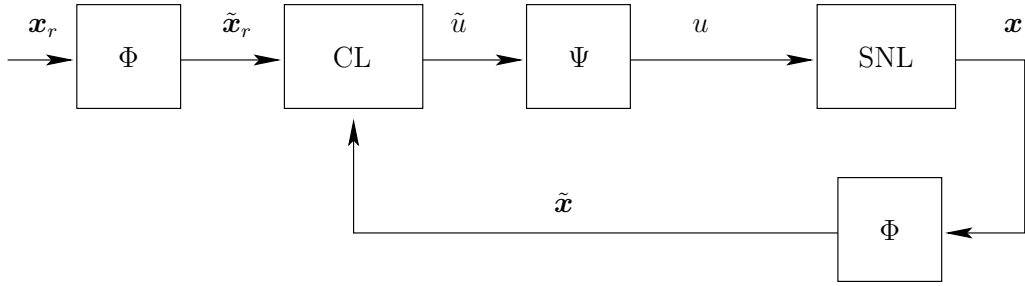


FIGURA 4.3 - Esquema do controle via linearização por realimentação de estados. CL é o controlador projetado para o sistema linear e SNL representa o sistema não linear.

sendo

$$\alpha(\mathbf{x}) = L_f^n \phi(\mathbf{x}) \quad (4.15)$$

e

$$\beta(\mathbf{x}) = L_g L_f^{n-1} \phi(\mathbf{x}) . \quad (4.16)$$

De posse dessas transformações é possível projetar um controlador através da linearização por realimentação de estados. A estratégia é usar as funções (4.12) e (4.14) para transformar o sistema a ser controlado (4.10) num sistema linear da forma (4.11). A seguir, projeta-se um controlador linear para esse sistema linear. Esse controlador segue uma referência no espaço do sistema linear $\tilde{\mathbf{x}}_r$ dada por

$$\tilde{\mathbf{x}}_r = \Phi(\mathbf{x}_r) , \quad (4.17)$$

sendo \mathbf{x}_r a referência no espaço original. O controlador, usando o estado transformado $\tilde{\mathbf{x}}$, gera então uma ação de controle \tilde{u} no espaço transformado. Essa ação é transformada de volta para o espaço original usando $\Psi(\cdot)$. Este arranjo é ilustrado na Figura 4.3.

4.3 Controle Adaptativo

Existem duas abordagens principais para o controle adaptativo: o método do modelo de referência e o método do auto-ajuste (SLOTINE; LI, 1991). Ambos possuem duas partes essenciais: um controlador, que atua no sistema a ser controlado, e um *mecanismo de adaptação*, que atua num conjunto de parâmetros.

O objetivo do controle adaptativo é manter um desempenho consistente do controlador face a incertezas ou variações desconhecidas nos parâmetros do sistema

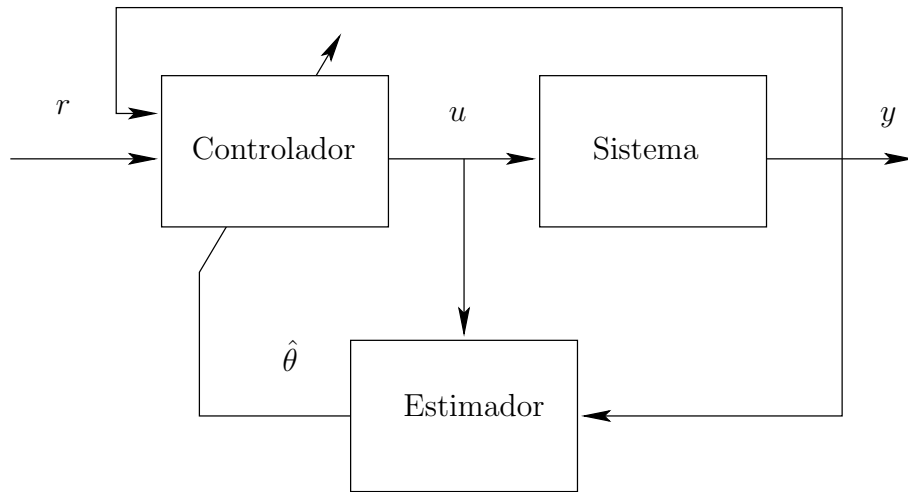


FIGURA 4.4 - Esquema de um controlador adaptativo do tipo auto-ajustável.

controlado (SLOTINE; LI, 1991).

A abordagem adaptativa usada neste trabalho é o método do auto-ajuste. Mas especificamente o método conhecido como *gradiente de velocidade* (FRADKOV; POGROMSKY, 1996).

Um método do tipo auto-ajuste pode ser representado pelo esquema da Figura 4.4. Um controlador recebe a referência r e a saída y do sistema e gera uma ação de controle u . Este controlador também depende de um vetor de parâmetros estimados $\hat{\theta}$. Esses parâmetros são ajustados por um estimador que recebe como entrada os sinais u e y .

O problema de síntese, nessa abordagem, é formado pelo projeto do controlador e da lei de adaptação responsável pelo ajuste dos parâmetros.

A seguir apresenta-se uma importante técnica para o projeto de leis de adaptação: o gradiente de velocidade (FRADKOV; POGROMSKY, 1996).

Seja o seguinte sistema de controle

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\xi}) , \quad (4.18)$$

sendo $\mathbf{x} \in \mathbb{R}^n$ o vetor de estados, $\mathbf{u} \in \mathbb{R}^m$ o vetor de entradas e $\boldsymbol{\xi}$ um vetor de parâmetros.

Suponha que a meta de controle seja definida como

$$\lim_{t \rightarrow \infty} Q(\mathbf{x}, t) = 0, \quad (4.19)$$

sendo $Q : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ uma função escalar definida positiva. Em outras palavras, o objetivo da lei de controle

$$\mathbf{u}(t) = U(\mathbf{x}, \boldsymbol{\xi}, t) \quad (4.20)$$

é garantir (4.19).

O método do gradiente de velocidade fornece uma lei de adaptação que permite tornar um controlador que resolva o problema (4.19) “imune” ao desconhecimento dos parâmetros $\boldsymbol{\xi}$. Ou seja, se for possível encontrar uma lei de controle (4.20) que alcance a meta (4.19) *globalmente* para um vetor de parâmetros particular de (4.18), então existe uma lei de adaptação que garante que o controlador atingirá a meta para *qualquer* outro valor do vetor de parâmetros.

Suponha que uma lei de controle que garanta (4.19) para um valor de $\boldsymbol{\xi}$ tenha sido escolhida. Substitua $\boldsymbol{\xi}$ pelo vetor de parâmetros $\boldsymbol{\theta}(t)$ em (4.20) e, a seguir, substitua (4.20) em (4.18) levando a

$$\dot{\mathbf{x}} = F(\mathbf{x}, \boldsymbol{\theta}, t). \quad (4.21)$$

A derivada de $Q(\cdot)$ em relação a t , ao longo de uma trajetória de (4.21), (a “velocidade” de $Q(\cdot)$) é dada por

$$\dot{Q} = \omega(\mathbf{x}, \boldsymbol{\theta}, t) = \frac{\partial Q}{\partial t} + L_F Q. \quad (4.22)$$

A idéia do método do gradiente de velocidade é atualizar o vetor de parâmetros $\boldsymbol{\theta}$ ao longo do gradiente de $\omega(\mathbf{x}, \boldsymbol{\theta}, t)$ em relação a $\boldsymbol{\theta}$. A forma mais geral dessa atualização é (FRADKOV; POGROMSKY, 1996)

$$\frac{d}{dt}(\boldsymbol{\theta} + \varphi(\mathbf{x}, t)) = -\Gamma \nabla_{\boldsymbol{\theta}} \omega(\mathbf{x}, \boldsymbol{\theta}, t), \quad (4.23)$$

sendo $\Gamma = \Gamma^T \geq 0$ uma matriz de ganho definida positiva e $\varphi(\mathbf{x}, t)$ alguma função

vetorial que satisfaça a condição de pseudo-gradiente

$$\varphi^T \nabla_{\boldsymbol{\theta}} \omega(\mathbf{x}, \boldsymbol{\theta}, t) \geq 0. \quad (4.24)$$

Um teorema para a estabilidade global do método foi provado em [Fradkov e Pogromsky \(1996\)](#). Considere os sistemas (4.18), (4.23) sujeitos às seguintes condições:

A1 $F(\mathbf{x}, \boldsymbol{\theta}, t)$ e $\nabla_{\boldsymbol{\theta}} \omega(\mathbf{x}, \boldsymbol{\theta}, t)$ e suas derivadas parciais são limitadas uniformemente em qualquer conjunto limitado dos estados $(\mathbf{x}, \boldsymbol{\theta})$ para $t \geq 0$.

A2 $\omega(\mathbf{x}, \boldsymbol{\theta}, t)$ é convexo em $\boldsymbol{\theta}$.

A3 Existem um $\boldsymbol{\theta}_* \in \mathbb{R}^n$ e uma função escalar $\rho(\mathbf{x}) \geq 0$, $\rho(0) = 0$ uniformemente contínua em cada região limitada tais que $\omega(\mathbf{x}, \boldsymbol{\theta}_*, t) \leq -\rho(\mathbf{x})$ vale para todo $\mathbf{x} \in \mathbb{R}^n$.

A4 $Q(\mathbf{x}, t)$ ser limitada implica que $\mathbf{x}(t)$ é limitado.

Então cada trajetória $(\mathbf{x}(t), \boldsymbol{\theta}(t))$ é limitada e $\lim_{t \rightarrow \infty} \rho(\mathbf{x}(t)) = 0$.

Adicionalmente, se $\rho(\mathbf{x})$ depende explicitamente apenas de $Q(\mathbf{x}, t)$ então a meta (4.19) é atingida ([FRADKOV; POGROMSKY, 1996](#)).

O método do gradiente de velocidade é uma importante ferramenta de controle. No entanto, é importante frisar algumas de suas limitações. Primeiro, as condições A1-A4 *não garantem* a convergência dos parâmetros $\boldsymbol{\theta}(t)$ para os parâmetros “ideais” $\boldsymbol{\theta}_* = \boldsymbol{\xi}$ ([FRADKOV; POGROMSKY, 1996](#)). Segundo, o método depende da existência de um controlador que garanta a meta *globalmente* (condição A3) para um determinado conjunto de parâmetros. Esse método do gradiente de velocidade não fornece esse controlador, que deve então ser projetado por outros meios.

A idéia por trás do método (e da prova do teorema da estabilidade) é que a lei de adaptação (4.23) permite a criação de uma segunda função de Lyapunov que inclui (4.19) e o vetor de parâmetros. A lei (4.23) permite que a derivada temporal desta nova função de Lyapunov seja uma função semidefinida negativa ([FRADKOV; POGROMSKY, 1996](#)).

CAPÍTULO 5

MÉTODO DE ESTIMAÇÃO DE PARÂMETROS

5.1 Introdução

A principal contribuição desta Tese é a proposição de um novo método de estimação de parâmetros para sistemas não lineares. Esse método se baseia no uso de controle (Capítulo 4) e sincronismo (Capítulo 3) como uma forma alternativa de se fazer estimação de parâmetros de forma que o Problema Inverso Completo (PIC) seja resolvido (Capítulo 2).

A idéia central do método é sincronizar o modelo com os dados oriundos do sistema a ser identificado. Isso é uma forma de sincronismo tipo mestre-escravo, onde o modelo (o escravo) segue o sistema (o mestre). O acoplamento é necessariamente unidirecional, pois supõe-se que os dados já foram coletados, de forma que não haveria como o modelo interferir no sistema.

Outro ponto importante do método é o acoplamento entre o sistema mestre e o escravo. Esse acoplamento é feito por um sistema de controle não linear.

Essencialmente, o método proposto consiste no projeto de um controlador não linear cujo objetivo é controlar um sistema bem particular: o modelo a ser identificado e utilizando uma referência bem particular: os dados de identificação.

Duas versões do método são apresentadas. A primeira é uma versão que permite estimar apenas um parâmetro e a segunda permite estimar vários parâmetros. A versão monoparamétrica, apesar ter uma utilidade prática limitada, é apresentada aqui com o objetivo de introduzir o projeto de um controlador para *estimação de parâmetros*, com finalidades didáticas, e também porque o controlador desenvolvido para a versão monoparamétrica será utilizado na versão multiparamétrica.

Antes de descrever essas versões, a representação polinomial usada no trabalho é apresentada. Os controladores projetados nesse capítulo são particulares para essa representação. No entanto, nada impede que a seqüência de passos usada nesse projeto seja aplicada em outras representações.

5.2 Representação

A representação matemática usada em um método de identificação é quesito muito importante. Ela consiste na forma matemática particular escolhida para descrever o sistema a ser identificado. Escolher uma determinada representação implica assumir que o sistema a ser identificado pode ser bem descrito pela representação escolhida.

A representação usada no método de estimação de parâmetros proposto é de tempo contínuo, usando sistemas de equações diferenciais ordinárias de primeira ordem como a equação (3.1), reproduzida aqui

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t)) . \quad (5.1)$$

Além do mais, a função $f(\cdot)$ é parametrizada usando funções do tipo polinomial. *Parametrização* é definida como a *especificação de uma curva, superfície, etc. usando uma ou mais variáveis* (WEISSTEIN, 1999). Essas variáveis ou *parâmetros* podem ser ajustados livremente pelo método de identificação para criar uma função que se adeqüe ao sistema a ser identificado, num processo chamado *estimação de parâmetros* conforme mostrado no Capítulo 2.

A função $f : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ pode ser escrita como um vetor de n funções $f_i : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$

$$f(\mathbf{x}(t), \mathbf{u}(t)) = \begin{bmatrix} f_1(\mathbf{x}(t), \mathbf{u}(t)) \\ f_2(\mathbf{x}(t), \mathbf{u}(t)) \\ \vdots \\ f_i(\mathbf{x}(t), \mathbf{u}(t)) \\ \vdots \\ f_n(\mathbf{x}(t), \mathbf{u}(t)) \end{bmatrix} . \quad (5.2)$$

Cada uma das $f_i(\cdot)$ é um polinômio multivariável nas componentes de \mathbf{x} e \mathbf{u} . As funções $f_i(\cdot)$ são da forma

$$f_i(\mathbf{x}(t), \mathbf{u}(t)) = \sum_{j=1}^{\rho_i} \theta_{ij} m_{ij}(\mathbf{x}(t), \mathbf{u}(t)) , \quad (5.3)$$

sendo ρ_i o número de termos de $f_i(\cdot)$, θ_{ij} o *coeficiente* do termo $m_{ij}(\mathbf{x}(t), \mathbf{u}(t))$. Este

é dado por

$$m_{ij}(\mathbf{x}, \mathbf{u}) = \prod_{k=1}^n x_k^{\alpha_{ijk}} \prod_{k=1}^p u_k^{\beta_{ijk}}, \quad (5.4)$$

sendo $\alpha_{ijk}, \beta_{ijk} \in \mathbb{N}$ os expoentes das componentes individuais dos vetores \mathbf{x} e \mathbf{u} . A dependência em relação a t foi suprimida por motivo de clareza. A representação polinomial descrita nas equações (5.3) e (5.4), apesar de simples, é capaz de descrever uma grande variedade de sistemas não lineares e aproximar qualquer campo vetorial analítico. O sistema de Lorenz (3.2), por exemplo, é descrito na representação polinomial

$$\begin{aligned} \dot{x}_1 &= \theta_{11}m_{11}(\mathbf{x}) + \theta_{12}m_{12}(\mathbf{x}) = -\sigma x_1^1 x_2^0 x_3^0 + \sigma x_1^0 x_2^1 x_3^0 \\ \dot{x}_2 &= \theta_{21}m_{21}(\mathbf{x}) + \theta_{22}m_{22}(\mathbf{x}) + \theta_{23}m_{23}(\mathbf{x}) = r x_1^1 x_2^0 x_3^0 - x_1^0 x_2^1 x_3^0 - x_1^1 x_2^0 x_3^1 \\ \dot{x}_3 &= \theta_{31}m_{31}(\mathbf{x}) + \theta_{32}m_{32}(\mathbf{x}) = -b x_1^0 x_2^0 x_3^1 + x_1^1 x_2^1 x_3^0. \end{aligned} \quad (5.5)$$

5.3 Versão monoparamétrica

A versão monoparamétrica é um método de estimação de parâmetros bastante restrito, de utilidade mais didática do que prática. A primeira restrição importante é que esse método se aplica a sistemas sem entrada, ou seja, não há termos em \mathbf{u} . A segunda restrição é que apenas *um* dos coeficientes θ_{ij} é suposto desconhecido e deve ser estimado. Por isso essa versão é chamada monoparamétrica. Esse método foi publicado em Freitas *et al.* (2005).

O método consiste em isolar esse coeficiente e tratá-lo como uma entrada de controle. É importante não confundir essa entrada com as entradas \mathbf{u} em (5.2). Essas últimas representam as entradas físicas presentes no sistema a identificar e, por hipótese, não existem no caso da versão monoparamétrica, enquanto o coeficiente desconhecido é transformado numa entrada por um artifício matemático, visando a sua estimação.

A primeira etapa é escrever a equação (5.2) numa forma afim como em (4.10), de forma que o parâmetro desconhecido ocupe o lugar da entrada escalar

$$\dot{\mathbf{x}} = F(\mathbf{x}) + g(\mathbf{x})\theta_{ij}, \quad (5.6)$$

sendo que $g_k(\mathbf{x}) = 0$ para $k \neq i$ e $g_i(\mathbf{x}) = m_{ij}$ e $F(\mathbf{x}) = f(\mathbf{x}) - \theta_{ij}g(\mathbf{x})$. Por hipótese, o sistema a ser identificado é descrito por uma equação idêntica a (5.6), sendo que todos os parâmetros são conhecidos exceto θ_{ij} .

Se o sistema (5.6) for controlável conforme definido no Capítulo (4) então é possível projetar um controlador que, atuando em θ_{ij} , leve o sistema para qualquer estado. Em particular, é possível fazê-lo seguir o sistema a ser identificado, ou seja, é possível fazê-los sincronizar. Na prática, o que o modelo estará “seguindo” como referência são os dados de identificação. Para que o sistema (5.6) seja controlável os campos vetoriais $F(\cdot)$ e $g(\cdot)$ devem ser tais que exista uma região aberta Ω em \mathbb{R}^n onde:

- os campos vetoriais $\{g, \text{ad}_F g, \dots, \text{ad}_F^{n-1} g\}$ sejam linearmente independentes e
- o conjunto $\{g, \text{ad}_F g, \dots, \text{ad}_F^{n-2} g\}$ seja *involutivo*.

Essas condições devem ser verificadas caso-a-caso, pois dependem das formas particulares dos campos $F(\cdot)$ e $g(\cdot)$. Particularmente, a existência e o tamanho da região Ω dependem desses campos. Verificar essas condições pode ser trabalhoso do ponto de vista algébrico, mas é um procedimento direto e quase mecânico. Por esse motivo, foram desenvolvidas rotinas para um sistema de álgebra computacional que, dados um $F(\cdot)$ e um $g(\cdot)$ particulares, verificam as condições de controlabilidade.

Se os campos vetoriais satisfizerem as condições de controlabilidade, é possível encontrar uma transformação de coordenadas não linear $\tilde{\mathbf{x}} = \Phi(\mathbf{x})$ e uma transformação da entrada $\tilde{\theta}_{ij} = \Psi^{-1}(\mathbf{x}, \theta_{ij})$, que transforma o sistema (5.6) em um sistema linear da forma (4.11) reproduzido aqui,

$$\dot{\tilde{\mathbf{x}}} = A\tilde{\mathbf{x}} + \mathbf{b}\tilde{u} . \quad (5.7)$$

O que essa transformação faz é escrever o sistema (5.6) em um outro sistema de coordenadas, no qual ele tem a forma (5.7). As transformações $\Phi(\cdot)$ e $\Psi^{-1}(\cdot)$ mapeiam o sistema num espaço de estados particular onde o controle é muito mais simples do que no espaço original.

A mudança de coordenadas $\Phi(\cdot)$ é dada por

$$\Phi(\mathbf{x}) = \begin{bmatrix} \phi(\mathbf{x}) \\ L_F \phi(\mathbf{x}) \\ \vdots \\ L_F^{n-1} \phi(\mathbf{x}) \end{bmatrix} , \quad (5.8)$$

sendo $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ a solução do sistema de equações diferenciais parciais (ver Capítulo 4)

$$\begin{cases} L_g \phi(\mathbf{x}) & = 0 \\ L_g L_F \phi(\mathbf{x}) & = 0 \\ & \vdots \\ L_g L_F^{n-2} \phi(\mathbf{x}) & = 0 \\ L_g L_F^{n-1} \phi(\mathbf{x}) & \neq 0. \end{cases} \quad (5.9)$$

Este sistema sempre tem solução se os campos vetoriais satisfizerem as condições de controlabilidade, embora essa solução não seja única. A falta de unicidade, nesse caso, não é crítica, pois qualquer solução de (5.9) resolve o problema. Também foram implementadas funções num sistema de álgebra computacional para resolver o sistema (5.9) de forma semi-automática, com um mínimo de intervenção do usuário.

No novo sistema de coordenadas, o modelo se torna um sistema linear, que pode ser facilmente controlado. Dessa forma é possível projetar um sistema de controle linear de modo a garantir que o modelo siga uma referência nesse sistema de coordenadas. O objetivo é fazer com que o modelo siga o sistema a ser identificado.

Como $\Phi(\cdot)$ é um difeomorfismo, $\tilde{\mathbf{r}} = \tilde{\mathbf{x}}$ implica $\mathbf{r} = \mathbf{x}$, ou seja, esse método de controle é capaz de sincronizar o modelo com o sistema a ser identificado.

Uma vez que ocorra o sincronismo, quer dizer, que os estados do modelo e do sistema fiquem idênticos a partir de um certo instante de tempo, a saída do controlador converge para o valor do parâmetro desconhecido do sistema original. Isso pode ser mostrado pelo seguinte argumento. Seja o sistema a ser identificado dado por

$$\dot{\mathbf{y}} = F(\mathbf{y}) + g(\mathbf{y})\theta'_{ij} \quad (5.10)$$

sendo θ'_{ij} o parâmetro desconhecido. O erro de sincronização é definido por $\mathbf{e}(t) = \mathbf{y}(t) - \mathbf{x}(t)$ e a sua evolução temporal obedece a

$$\dot{\mathbf{e}}(t) = F(\mathbf{y}(t)) - F(\mathbf{x}(t)) + \theta'_{ij}g(\mathbf{y}(t)) - \theta_{ij}g(\mathbf{x}(t)). \quad (5.11)$$

Na variedade de sincronização, ou seja em $\mathbf{y}(t) = \mathbf{x}(t)$, a equação (5.11) se reduz a

$$\dot{\mathbf{e}}(t) = (\theta'_{ij} - \theta_{ij})g(\mathbf{y}(t)). \quad (5.12)$$

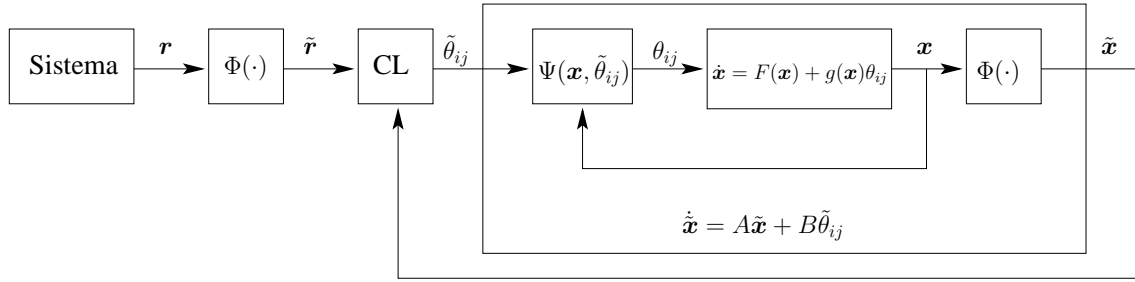


FIGURA 5.1 - Diagrama em blocos do sistema de controle para a estimação monoparamétrica.

Se a sincronização ocorrer e for estável, ou seja $e(t) = \dot{e}(t) = 0$ para algum $t \geq t_0$ tem-se que $g(\mathbf{y}(t)) = 0$ ou $\theta_{ij} = \theta'_{ij}$. A primeira possibilidade pode ser descartada pois implicaria que o sistema é independente do parâmetro a ser estimado, ou seja, qualquer que fosse o valor desse parâmetro o sistema se comportaria da mesma forma (teria o mesmo campo vetorial). Isso tornaria a estimação de tal parâmetro sem sentido. A possibilidade restante garante então a capacidade do método de estimar corretamente o parâmetro desconhecido.

O método monoparamétrico é ilustrado na Figura 5.1.

O sistema a ser identificado, representado pelo bloco “Sistema” na Figura 5.1, gera a referência $\mathbf{r} \in \mathbb{R}^n$. Essa referência é então transformada para o novo sistema de coordenadas usando o difeomorfismo $\Phi(\mathbf{r})$, dado pela equação (5.8), gerando a referência $\tilde{\mathbf{r}}$. Essa referência é usada pelo controlador linear “CL” que, usando o estado do modelo no espaço transformado, gera um sinal de controle $\tilde{\theta}_{ij}$ que faz com que $\tilde{\mathbf{r}}(t) - \tilde{\mathbf{x}}(t) \rightarrow 0$ quando $t \rightarrow \infty$. Isso é possível porque nesse espaço transformado o modelo é um sistema linear. Note que θ_{ij} é o parâmetro a ser identificado. Um outro ponto importante é que a “referência” seguida pelo modelo é formada pelos dados de identificação.

5.4 Versão multiparamétrica

Conforme mencionado anteriormente, a versão monoparamétrica é um método de estimação de parâmetros de utilidade limitada. A maior parte das aplicações de identificação possuem mais do que um parâmetro desconhecido. O que é necessário, então, é uma extensão do método para vários parâmetros.

As ferramentas para cumprir esse objetivo são as técnicas do Capítulo 4. Como

um item importante daquele Capítulo, a técnica de controle adaptativo (Seção 4.3) fornece o seguinte resultado: se um determinado controlador funcionar para um certo valor do vetor de parâmetros do sistema controlado, então é possível projetar uma lei de adaptação que faça com que esse controlador funcione para *qualquer* valor desse vetor de parâmetros.

Suponha que se deseje estimar os parâmetros de um sistema dado por

$$\dot{\mathbf{x}}_1 = f(\mathbf{x}_1, \boldsymbol{\xi}), \quad (5.13)$$

onde $\mathbf{x}_1 \in \mathbb{R}^n$ é o vetor de estados e $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um campo vetorial¹ que depende dos parâmetros $\boldsymbol{\xi} \in \mathbb{R}^p$.

O modelo é semelhante a (5.13) e é dado por

$$\dot{\mathbf{x}}_2 = f(\mathbf{x}_2, \boldsymbol{\theta}) + \mathbf{u}, \quad (5.14)$$

sendo $\mathbf{x}_2 \in \mathbb{R}^n$ o vetor de estados, $\boldsymbol{\theta} \in \mathbb{R}^p$ um vetor de parâmetros e $\mathbf{u} \in \mathbb{R}^n$ o vetor de entradas.

A estratégia é sincronizar o modelo (5.14) ao sistema (5.13). Para isso, utiliza-se um controlador atuando no modelo através das entradas \mathbf{u} usando como referência os estados do sistema, \mathbf{x}_1 , e uma lei de adaptação para os parâmetros $\boldsymbol{\theta}$ dada pelo método do gradiente de velocidade (vide Seção 4.3). O controlador atua então como um acoplamento entre o sistema e o modelo, tendo como objetivo anular o erro de sincronização, $\mathbf{e} = \mathbf{x}_1 - \mathbf{x}_2$. Com o sistema de sincronização colocado dessa forma, a derivada temporal do erro de sincronização, $\dot{\mathbf{e}}$, é dada por

$$\dot{\mathbf{e}} = \dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2 = f(\mathbf{x}_1, \boldsymbol{\xi}) - f(\mathbf{x}_2, \boldsymbol{\theta}) - \mathbf{u} = F(\mathbf{e}, \boldsymbol{\theta}, t). \quad (5.15)$$

Define-se a meta de controle como sendo

$$Q(\mathbf{e}, \boldsymbol{\theta}) = \frac{1}{2} \mathbf{e}^T \mathbf{e}, \quad (5.16)$$

¹Na Seção 5.2, $f(\cdot)$ depende de um vetor de entradas $\mathbf{u}(t) \in \mathbb{R}^p$, além dos estados. Na exposição que se segue, essa dependência foi suprimida para simplificar a argumentação e também para evitar ambigüidade com a ação de controle usada no método de estimação de parâmetros, que não tem relação nenhuma com esse vetor de entradas. No entanto, supondo que as entradas são conhecidas, como normalmente ocorre em identificação, pode-se considerar que $f(\mathbf{x}, \mathbf{u}(t), \boldsymbol{\xi}) = f(\mathbf{x}, t, \boldsymbol{\xi})$ e o método continua válido.

sendo que a dependência em relação a $\boldsymbol{\theta}$ é implícita.

A função $Q(\cdot)$ é definida positiva em \mathbf{e} , ou seja, $Q(\mathbf{e}, \boldsymbol{\theta}) \geq 0$ e $Q(\mathbf{e}, \boldsymbol{\theta}) = 0 \Leftrightarrow \mathbf{e} = 0$. O objetivo na sincronização é fazer $e(t) \rightarrow 0$ quando $t \rightarrow \infty$. Isso é equivalente a fazer $\lim_{t \rightarrow \infty} Q(e(t), \boldsymbol{\theta}) = 0$.

Suponha que seja possível projetar um controlador que garanta a meta (5.16) se $\boldsymbol{\theta}$ assumir um valor específico $\boldsymbol{\theta} = \boldsymbol{\xi}$. Então, se as condições para o método do gradiente de velocidade puderem ser satisfeitas, é possível usar uma lei de adaptação na forma (4.23) que garanta que a meta é alcançada para qualquer valor do vetor de parâmetros $\boldsymbol{\theta}$, ou seja, que este controlador, junto com a lei de adaptação, é capaz de sincronizar o modelo com o sistema *mesmo que os parâmetros sejam desconhecidos*.

A derivada temporal de $Q(\cdot)$ ao longo da trajetória de (5.15) (a “velocidade” de $Q(\cdot)$) é dada por

$$\omega(\mathbf{e}, \boldsymbol{\theta}) = \dot{Q} = L_F Q = \frac{\partial Q}{\partial \boldsymbol{\theta}} F. \quad (5.17)$$

Substituindo (5.15) e (5.16) em (5.17) tem-se

$$\omega(\mathbf{e}, \boldsymbol{\theta}) = \mathbf{e}^T f(\mathbf{x}_1, \boldsymbol{\xi}) - \mathbf{e}^T f(\mathbf{x}_2, \boldsymbol{\theta}) - \mathbf{e}^T \mathbf{u} \quad (5.18)$$

O gradiente de $\omega(\mathbf{e}, \boldsymbol{\theta})$ em relação a $\boldsymbol{\theta}$ fica então

$$\nabla_{\boldsymbol{\theta}} \omega(\mathbf{e}, \boldsymbol{\theta}) = -(\mathbf{e}^T \nabla_{\boldsymbol{\theta}} f(\mathbf{x}_2, \boldsymbol{\theta}))^T - (\mathbf{e}^T \nabla_{\boldsymbol{\theta}} \mathbf{u})^T, \quad (5.19)$$

sendo $\nabla_{\boldsymbol{\theta}} f(\mathbf{x}_2, \boldsymbol{\theta})$ a matriz jacobiana de $f(\mathbf{x}_2, \boldsymbol{\theta})$ em relação a $\boldsymbol{\theta}$.

Neste ponto, usa-se explicitamente o fato de que o modelo é uma função polinomial² da forma (5.3), com um certo conjunto de termos m_{ij} . Essa suposição é equivalente a assumir que o sistema (5.13) possa ser representado da mesma forma. Se isso não for verdade, não há nenhum valor para o vetor de parâmetros $\boldsymbol{\theta}$ que faça com que o modelo represente o sistema e a estimação de parâmetros não faz sentido. Neste caso, deve-se alterar a representação. Por esse motivo, é importante garantir que $f(\cdot)$ tenha flexibilidade suficiente para representar o sistema (JUDD; MEES, 1995).

²A argumentação também vale para uma função C^∞ linear nos parâmetros.

Assim

$$\begin{aligned} \frac{\partial}{\partial \theta_{ij}} f_k(\mathbf{x}_2, \boldsymbol{\theta}) &= m_{ij}, \quad k = i \\ \frac{\partial}{\partial \theta_{ij}} f_k(\mathbf{x}_2, \boldsymbol{\theta}) &= 0, \quad k \neq i. \end{aligned} \quad (5.20)$$

A lei de adaptação usada neste trabalho é um caso particular de (4.23), chamada *forma diferencial* (FRADKOV; MARKOV, 1997), definida como

$$\dot{\boldsymbol{\theta}} = -\Gamma \nabla_{\boldsymbol{\theta}} \omega(\mathbf{e}, \boldsymbol{\theta}), \quad (5.21)$$

onde Γ é uma matriz simétrica definida positiva.

O método proposto consiste no modelo (5.14) como um oscilador escravo que é sincronizado ao sistema (5.13) via um controlador, juntamente com a lei de atualização dos parâmetros (5.21).

Para que esse método seja uma forma válida de estimar parâmetros é necessário:

- a) existir um controlador que se enquadre nas hipóteses do método do gradiente de velocidade;
- b) verificar se o sistema (5.15) se enquadra nessas mesmas hipóteses e
- c) mostrar que se os itens a) e b) forem satisfeitos, então o vetor de parâmetros converge para o valor desejado.

A questão do projeto do controlador é tratada mais adiante. A seguir, verifica-se o item b).

A função $f(\mathbf{x}, \boldsymbol{\theta})$ é polinomial, e portanto, C^∞ . Assim, $f(\cdot)$ e suas derivadas são limitadas em qualquer conjunto $(\mathbf{x}, \boldsymbol{\theta})$ limitado. A função $Q(\mathbf{e})$ é definida positiva e ilimitada radialmente, logo $Q(\mathbf{e})$ ser limitada implica que \mathbf{e} é limitado (KHALIL, 1996). Isso, por sua vez, implica que \mathbf{x}_2 também é limitado, uma vez que \mathbf{x}_1 seja limitado, i.e. os estados do sistema são limitados (o sistema é caótico e evolui num atrator). Com isso, a condição A4 (Capítulo 4, p. 63) e parte da condição A1 são satisfeitas. Analisando (5.18) e (5.19), nota-se que se \mathbf{u} for independente de $\boldsymbol{\theta}$, então $\omega(\cdot)$ é convexo em $\boldsymbol{\theta}$ (é linear em $\boldsymbol{\theta}$) e $\nabla_{\boldsymbol{\theta}} \omega(\mathbf{e}, \boldsymbol{\theta})$ é C^∞ . Dessa forma, $\nabla_{\boldsymbol{\theta}} \omega(\mathbf{e}, \boldsymbol{\theta})$ é limitado e com derivadas parciais limitadas em qualquer conjunto limitado. Assim, as condições A1, A2 e A4 são satisfeitas. Se \mathbf{u} não for independente de $\boldsymbol{\theta}$,

então essas condições passam a depender somente da lei de controle, pois o resto do sistema (5.15) as cumpre.

Resta, agora, verificar em que condições o vetor de parâmetros converge para o valor desejado ξ . De acordo com o método do gradiente de velocidade, o sistema (5.15) converge para o conjunto de estados Ω tais que $\rho(\mathbf{x}_2) = 0$. Também, uma função definida positiva formada por $Q(\cdot)$ adicionada de uma função definida positiva dos parâmetros tem a sua derivada ao longo de (5.15), (5.21) dada por uma função semidefinida negativa (FRADKOV; POGROMSKY, 1996). Dessa forma, os estados convergem para o conjunto dos pontos onde $Q(\mathbf{e}, \boldsymbol{\theta}) = 0$ que seja também positivamente invariante (KHALIL, 1996). Admite-se aqui que isso é suficiente para a convergência para os valores desejados. Em Yu *et al.* (2006) argumenta-se que os parâmetros convergem como desejado para sistemas caóticos, mas uma prova não é apresentada. A seguir, são tratadas as possibilidades de projeto de controladores.

5.4.1 Controladores para a estimação de parâmetros

A essência do método multiparamétrico proposto é o uso simultâneo de um controlador não linear junto com o método adaptativo do gradiente de velocidade numa arquitetura de sincronismo com o objetivo de estimar parâmetros. Quando Fradkov e Pogromsky (1996) propuseram o método do gradiente de velocidade para a sincronização de sistemas caóticos, o objetivo era apenas tornar um esquema de sincronização robusto a incertezas nos parâmetros. O método do gradiente de velocidade requer a existência de um controlador que atinja a meta (5.16) globalmente para um determinado conjunto de parâmetros. Esses autores não mostraram como tal controlador genérico poderia ser construído.

Por outro lado, existem na literatura trabalhos que mostram como controladores com essas características podem ser projetados (GAUTHIER *et al.*, 1992; NIJMEIJER; MAREELS, 1997; NIJMEIJER, 2001). Novamente, a ênfase é a sincronização entre os sistemas.

A contribuição desta Tese é usar a sincronização para *estimar parâmetros* para os sistemas em estudo. Desta forma, a sincronização deixa de ser um fim para ser um meio para outro objetivo: a determinação de parâmetros. Nesse sentido, algumas possibilidades de projeto de controlador, algumas existentes na literatura, outras apresentadas aqui, são sugeridas para serem usadas em conjunto com o método do

gradiente de velocidade.

A primeira possibilidade é usar um controlador não linear geométrico, como o que foi empregado na Seção 5.3. Nesse caso, deve-se supor conhecida a estrutura de (5.2), ou seja, quais são os valores de α_{ijk} e β_{ijk} em (5.4). No entanto, essa alternativa apresenta algumas desvantagens. O controlador depende necessariamente do vetor de parâmetros θ . Assim, é necessário verificar as condições A1-A3 caso à caso, bem como a controlabilidade do sistema (5.14).

Uma outra possibilidade é o uso de *observadores*. Note que um controlador que cumpra a meta (5.16) é um observador para o sistema. A principal vantagem do uso de observadores é que eles permitem o uso do método de estimação de parâmetros aqui proposto para os casos onde o vetor de estados \mathbf{x}_1 não é amostrado diretamente, mas através de uma função escalar. Em Nijmeijer e Mareels (1997), algumas possibilidades de observadores são apresentadas. Por exemplo, sistemas do tipo *Lur'e*

$$\begin{aligned}\dot{\mathbf{x}}_1 &= A\mathbf{x}_1 + f(C\mathbf{x}_1, t) \\ y_1(t) &= C\mathbf{x}_1 ,\end{aligned}\tag{5.22}$$

sendo A e C um par de matrizes detectável. Para esse caso, o modelo é da forma

$$\begin{aligned}\dot{\mathbf{x}}_2 &= A\mathbf{x}_2 + f(y_1(t), t) + K(y_2(t) - y_1(t)) \\ y_2(t) &= C\mathbf{x}_2 ,\end{aligned}\tag{5.23}$$

sendo que K é tal que $A + KC$ é assintoticamente estável. Esse tipo de sistema apresenta uma dinâmica linear para o erro de sincronização. Dessa forma, o método do gradiente de velocidade pode ser usado para estimar os parâmetros de $f(\cdot)$, ou mesmo de A , se for possível escolher um K que faça $A + KC$ assintoticamente estável, sem conhecer alguns de seus parâmetros.

Outra possibilidade é o uso de um observador de alto ganho (GAUTHIER *et al.*, 1992; NIJMEIJER; MAREELS, 1997). Suponha que um sistema da forma

$$\begin{aligned}\dot{\mathbf{x}}_1(t) &= f(\mathbf{x}_1(t)) \\ y_1(t) &= h(\mathbf{x}_1(t))\end{aligned}\tag{5.24}$$

evolua em um conjunto $\Omega \subset \mathbb{R}^n$ compacto positivamente invariante sob a dinâmica e que $\mathbf{x}'_1 = \Phi(\mathbf{x}_1) = (h(\mathbf{x}_1), L_f h(\mathbf{x}_1), \dots, L_f^{n-1} h(\mathbf{x}_1))^T$ define um difeomorfismo em um conjunto aberto contendo Ω . Ou seja, esse sistema é *observável* (GAUTHIER *et*

al., 1992). Nas coordenadas \mathbf{x}'_1 o sistema (5.24) tem a forma

$$\begin{aligned}\dot{\mathbf{x}}'_1(t) &= F(\mathbf{x}'_1(t), \boldsymbol{\xi}) \\ y_1(t) &= x_1^{1'} ,\end{aligned}\tag{5.25}$$

sendo $x_1^{1'}$ a primeira coordenada de \mathbf{x}'_1 .

Uma possibilidade de observador para (5.25) é (GAUTHIER *et al.*, 1992)

$$\begin{aligned}\dot{\mathbf{x}}(1)'_2 &= F(\mathbf{x}'_2, \boldsymbol{\xi}) + K_\gamma(y_2(t) - y_1(t)) \\ y_2(t) &= x_2^{1'} ,\end{aligned}\tag{5.26}$$

sendo $K_\gamma = -S_\gamma^{-1}C^T$ e $S_\gamma = S_\gamma^T > 0$ a solução para

$$\gamma S_\gamma + A^T S_\gamma + S_\gamma A - C^T C = 0 ,\tag{5.27}$$

$$C = (1 \ 0 \ \dots \ 0) \in \mathbb{R}^{1 \times n}$$

e

$$A = \begin{pmatrix} 0_{(n-1) \times 1} & I_{n-1} \\ 0 & 0_{1 \times (n-1)} \end{pmatrix} ,$$

com $\gamma > 0$ suficientemente grande.

É importante notar que nessas duas últimas possibilidades a ação de controle \mathbf{u} é independente do vetor de parâmetros e, portanto, esses controladores cumprem as condições A1-A4.

Essas duas possibilidades surgiram originalmente para resolver o problema da observação. No entanto, se todas as variáveis de estado \mathbf{x}_1 forem conhecidas, o problema de observação não existe. Porém, como o objetivo neste trabalho é a estimação e não a observação, ainda é necessário projetar um controlador que atenda as necessidades do método do gradiente de velocidade neste caso.

A proposta deste trabalho é usar um controlador linear de alto ganho

$$\dot{\mathbf{x}}_2 = f(\mathbf{x}_2, \boldsymbol{\theta}) + A(\mathbf{x}_1 - \mathbf{x}_2) ,\tag{5.28}$$

sendo $A = \text{diag}(k_1 \ k_2 \ \dots \ k_n)$ uma matriz diagonal de ganhos positivos suficientemente grandes.

Para que este controlador seja válido, é necessário provar que ele atinge a meta (5.16) para algum valor do vetor de parâmetros. A prova, um dos resultados desse trabalho, é dada a seguir.

Repare que, para atingir esse objetivo, é suficiente provar que a meta (5.16) é uma função de Lyapunov para o sistema (5.15), já que o controlador de (5.28) é independente de θ . Suponha que $f(\mathbf{x}, \boldsymbol{\xi})$ seja uniformemente Lipschitz, ou seja,

$$\|f(\mathbf{x}, \boldsymbol{\xi}) - f(\mathbf{y}, \boldsymbol{\xi})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad (5.29)$$

Assim, substituindo (5.7) em (5.18) e fazendo $\theta = \boldsymbol{\xi}$ tem-se

$$\dot{Q}(\mathbf{e}) = \mathbf{e}^T(f(\mathbf{x}_1, \boldsymbol{\xi}) - f(\mathbf{x}_2, \boldsymbol{\xi})) - \mathbf{e}^T A \mathbf{e}.$$

Mas

$$\mathbf{e}^T(f(\mathbf{x}_1, \boldsymbol{\xi}) - f(\mathbf{x}_2, \boldsymbol{\xi})) \leq \|\mathbf{e}\| \|f(\mathbf{x}_1, \boldsymbol{\xi}) - f(\mathbf{x}_2, \boldsymbol{\xi})\| \leq L\|\mathbf{e}\| \|\mathbf{x}_1 - \mathbf{x}_2\| \leq L\|\mathbf{e}\|^2,$$

sendo que usou-se o fato que $f(\cdot)$ é Lipschitz. Dessa forma, se $k_i = k$ for escolhido de forma que $\mathbf{e}^T A \mathbf{e} > L\|\mathbf{e}\|^2$, $\dot{Q}(\mathbf{e})$ é definida negativa e o ponto fixo $\mathbf{e} = 0$ é globalmente assintoticamente estável.

A condição “*uniformemente Lipschitz*” não vale em geral para modelos polinomiais. No entanto, ela pode ser relaxada para “*localmente Lipschitz*” se os estados \mathbf{x}_1 e \mathbf{x}_2 forem restritos a um conjunto compacto e $f(\cdot)$ for contínua (KHALIL, 1996). Como já se supõe que isso é verdade para \mathbf{x}_1 e $f(\cdot)$ é polinomial, essa condição se aplica. Isso completa a prova.

CAPÍTULO 6

RESULTADOS

Neste Capítulo são apresentados resultados obtidos com o método de estimação de parâmetros proposto. Inicialmente é mostrado o resultado de simulação da versão monoparamétrica e, a seguir, os resultados da versão multiparamétrica.

6.1 Versão monoparamétrica

Nesta seção apresenta-se um exemplo de aplicação que ilustra como a estratégia de estimação baseada em controle e sincronismo de caos funciona no caso da estimação de apenas um parâmetro.

Suponha que o sistema a ser identificado seja o sistema de Lorenz (3.2) (LORENZ, 1963), reproduzido abaixo

$$\begin{cases} \dot{x}_1 = \sigma(y_1 - x_1) , \\ \dot{y}_1 = rx_1 - y_1 - x_1z_1 , \\ \dot{z}_1 = -bz_1 + x_1y_1 . \end{cases} \quad (6.1)$$

Suponha ainda que os parâmetros σ e r sejam conhecidos, bem como as equações (6.1) e que o único parâmetro desconhecido seja b . Mais ainda, que todas as variáveis de estado estão disponíveis para coleta e não há ruído. Isto é uma situação bastante idealizada, que provavelmente nunca ocorrerá na prática, mas serve como prova de conceito.

O modelo a ser identificado é idêntico a (6.1), exceto pelo parâmetro b que no modelo é substituído pela entrada de controle u

$$\begin{cases} \dot{x}_2 = \sigma(y_2 - x_2) , \\ \dot{y}_2 = rx_2 - y_2 - x_2z_2 , \\ \dot{z}_2 = -uz_2 + x_2y_2 . \end{cases} \quad (6.2)$$

O objetivo é estimar o parâmetro b via controle e sincronismo. Para isso, é necessário um controlador que seja capaz de, computando a entrada u , sincronizar o modelo (6.2) ao sistema original (6.1).

O primeiro passo é projetar o controlador, o que foi feito usando a teoria de controle geométrico. Inicialmente, o sistema a ser controlado deve ser colocado na forma (6.3)

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})u . \quad (6.3)$$

Para o sistema (6.2), tem-se¹

$$f(\mathbf{x}) = [\sigma(y - x), \quad rx - y - xz, \quad xy]^T \quad (6.4)$$

e

$$g(\mathbf{x}) = [0, \quad 0, \quad -z]^T . \quad (6.5)$$

A seguir, a controlabilidade do sistema deve ser avaliada. Para isso, é necessário testar duas condições. Primeiro, os campos vetoriais $\{g, \text{ad}_f g, \text{ad}_f^2 g\}$ devem ser linearmente independentes e, segundo, o conjunto $\{g, \text{ad}_f g\}$ deve ser involutivo. Essas condições devem ser verificadas em um subconjunto aberto de \mathbb{R}^3 , onde o controlador é válido.

O colchete de Lie $[f, g]$ é dado pela equação (4.9) que, aplicada a (6.4) e a (6.5), leva a

$$\begin{aligned} [f, g] = \nabla g f - \nabla f g &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \sigma(y - x) \\ rx - y - xz \\ xy \end{bmatrix} \\ &\quad - \begin{bmatrix} -\sigma & \sigma & 0 \\ r - z & -1 & -x \\ y & x & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -z \end{bmatrix} = \begin{bmatrix} 0 \\ -xz \\ -xy \end{bmatrix} . \quad (6.6) \end{aligned}$$

¹Os subscritos foram retirados por motivo de clareza.

Usando-a novamente obtém-se $\text{ad}_f^2 g$

$$\begin{aligned}
[f, [f, g]] &= \nabla_{\text{ad}_f g} f - \nabla f \text{ad}_f g \\
&= \begin{bmatrix} 0 & 0 & 0 \\ -z & 0 & -x \\ -y & -x & 0 \end{bmatrix} \begin{bmatrix} \sigma(y-x) \\ rx-y-xz \\ xy \end{bmatrix} - \begin{bmatrix} -\sigma & \sigma & 0 \\ r-z & -1 & -x \\ y & x & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -xz \\ -xy \end{bmatrix} \\
&= \begin{bmatrix} \sigma xz \\ (\sigma-1)xz - \sigma yz - 2x^2 y \\ (\sigma+1)xy + 2x^2 z - \sigma y^2 - rx^2 \end{bmatrix}. \quad (6.7)
\end{aligned}$$

Os campos vetoriais g , $\text{ad}_f g$ e $\text{ad}_f^2 g$ são linearmente independentes se $\sigma xz \neq 0$. Isso pode ser verificado notando que g tem os dois primeiros componentes nulos, $[f, g]$ tem o primeiro componente nulo e $\text{ad}_f^2 g$ não tem componentes nulos, se $\sigma xz \neq 0$. Ao contrário, se $\sigma xz = 0$, os campos g , $\text{ad}_f g$ e $\text{ad}_f^2 g$ formam um conjunto linearmente dependente.

Para testar se o conjunto $\{g, \text{ad}_f g\}$ é involutivo é suficiente verificar se $[g, \text{ad}_f g]$ pode ser escrito como uma combinação linear de g e $\text{ad}_f g$. Usando as equações (6.5) e (6.6) tem-se

$$\begin{aligned}
[g, \text{ad}_f g] &= \\
&= \begin{bmatrix} 0 & 0 & 0 \\ -z & 0 & -x \\ -y & -x & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -z \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -xz \\ -xy \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ xz \\ -xy \end{bmatrix} \quad (6.8)
\end{aligned}$$

que, por sua vez, é igual a

$$[g, \text{ad}_f g] = c_1(\mathbf{x})g + c_2(\mathbf{x})\text{ad}_f g, \quad (6.9)$$

sendo $c_2(\mathbf{x}) = -1$ e

$$c_1(\mathbf{x}) = \frac{2xy}{z}.$$

Dessa forma, o conjunto $\{g, \text{ad}_f g\}$ é involutivo para todo o \mathbb{R}^3 . Assim, conclui-se que o sistema da equação (6.2) é *controlável* e que o subconjunto do espaço de estados no qual isso ocorre é o próprio \mathbb{R}^3 menos os planos $x = 0$ e $z = 0$. Essa é uma conclusão importante. O sistema de Lorenz, com o parâmetro b visto como uma entrada, pode ser controlado de forma a seguir, com algumas restrições, qualquer trajetória. Particularmente, é possível controlá-lo de forma a seguir *outro* sistema de Lorenz, ou seja, o sistema a ser identificado.

Outro ponto interessante é que a condição $\sigma xz \neq 0$ pode ser derivada a partir de uma análise do próprio sistema de Lorenz. Se $\sigma = 0$ a primeira equação do sistema se torna $\dot{x} = 0$, o que quer dizer que a variável de estado x é constante e, portanto, desacoplada do resto do sistema, tornando-o não controlável. Se $z = 0$ a última equação do sistema, e o sistema como um todo, fica independente da ação de controle. Por fim, se $x = 0$, a segunda equação fica independente da última. Isso faz com que não seja possível levar informações da última equação para as outras duas, tornando o sistema não controlável.

A seguir, deve-se projetar o controlador. O difeomorfismo que transforma o espaço de estados é dado por

$$\tilde{\mathbf{x}} = \Phi(\mathbf{x}) = [\phi(\mathbf{x}), L_f\phi(\mathbf{x}), L_f^2\phi(\mathbf{x})]^T. \quad (6.10)$$

A função ϕ é obtida pela solução do seguinte sistema equações diferenciais parciais

$$\left\{ \begin{array}{l} \frac{\partial \phi}{\partial z} z = 0 \\ \frac{\partial \phi}{\partial y} xz + \frac{\partial \phi}{\partial z} xy = 0 \\ -\frac{\partial \phi}{\partial x} \sigma xz + \frac{\partial \phi}{\partial y} [(\sigma - 1)xz - \sigma yz - 2x^2y] \\ \quad + \frac{\partial \phi}{\partial z} [(\sigma + 1)xy + 2x^2z - \sigma y^2 - rx^2] \neq 0. \end{array} \right. \quad (6.11)$$

Uma solução simples é

$$\phi(\mathbf{x}) = [x, 0, 0]^T \quad (6.12)$$

que leva ao difeomorfismo

$$\tilde{\mathbf{x}} = \Phi(\mathbf{x}) = \begin{bmatrix} x \\ \sigma(y - x) \\ (\sigma r + \sigma^2)x - (\sigma + \sigma^2)y - \sigma xz \end{bmatrix}. \quad (6.13)$$

A lei de controle linear usada foi

$$\tilde{u} = \dot{\tilde{x}}_{d3} - k_0 e - k_1 \dot{e} - k_2 \ddot{e}, \quad (6.14)$$

sendo os k_i os ganhos do controlador. O sinal de erro e é definido como $\tilde{x}_1 - \tilde{z}_{d1}$. \tilde{x}_1 é a primeira coordenada de² $\tilde{\mathbf{x}} = \Phi(\mathbf{x}_2)$, enquanto que \tilde{x}_{d1} é a primeira coordenada de $\tilde{\mathbf{x}}_d = \Phi(\mathbf{x}_1)$ que é o estado do oscilador mestre (6.1) transformado para o espaço linear. Dessa forma, as subseqüentes derivadas de e são obtidas tomando $\dot{e} = \tilde{x}_2 - \tilde{x}_{d2}$ e $\ddot{e} = \tilde{x}_3 - \tilde{x}_{d3}$. O sinal $\dot{\tilde{x}}_{d3}$ é obtido derivando-se a última linha da equação (6.13) o que leva a

$$\begin{aligned} \dot{\tilde{x}}_{d3} = & (\sigma r + \sigma^2)\sigma(y_1 - x_1) \\ & - (\sigma + \sigma^2)(rx_1 - y_1 - x_1z_1) \\ & - \sigma [\sigma(y_1 - x_1)z_z + x_1(-bz_1 + x_1y_1)] . \end{aligned} \quad (6.15)$$

Os ganhos k_i foram escolhidos de modo a obter uma resposta estável, rápida e não oscilatória para o erro de sincronização. Os três pólos do sistema foram escolhidos reais e iguais a -20 o que leva aos valores $k_0 = 8000$, $k_1 = 1200$ e $k_2 = 60$.

O difeomorfismo $\Phi(\cdot)$, juntamente com a transformação da entrada a ser mostrada a seguir, é responsável por linearizar a dinâmica do erro de sincronismo. Na verdade, essa transformação coloca o sistema (6.2) em um sistema de coordenadas no qual ele é *linear*. Nesse espaço, pode-se usar a teoria de controle linear para projetar um controlador, como foi feito. No entanto, há ressalvas. Uma delas é que quaisquer referências que o sistema controlado deva seguir devem ser expressas no novo espaço de coordenadas. Por isso foi necessário transformar também os sinais do sistema mestre.

²Neste ponto, para evitar confusão, os sub-índices de (6.2) são usados novamente.

A última etapa do projeto do sistema de controle é o cálculo da função de transformação da entrada. Ela é dada por (vide equação (4.14))

$$u(\mathbf{x}) = \frac{\tilde{u} - \alpha(\mathbf{x})}{\beta(\mathbf{x})}, \quad (6.16)$$

sendo

$$\alpha(\mathbf{x}) = L_f^3 \phi(\mathbf{x}) \quad (6.17)$$

e

$$\beta(\mathbf{x}) = L_g L_f^2 \phi(\mathbf{x}). \quad (6.18)$$

Usando as equações (6.13) e (6.17) tem-se

$$L_f^3 \phi = [(\sigma r + \sigma^2) - \sigma z] \sigma(y - x) - (\sigma + \sigma^2)(rx - y - xz) - \sigma x^2 y \quad (6.19)$$

e a partir das equações (6.13) e (6.18) chega-se a

$$L_g L_f^2 \phi = \sigma x z. \quad (6.20)$$

É interessante notar que (6.16) é singular para $\sigma x z = 0$, o que é coerente com a condição de controlabilidade.

Uma vez de posse do projeto do controlador, partiu-se para a implementação. O sistema mestre (6.1), o sistema escravo (6.2) e o controlador foram simulados conjuntamente usando o Octave. Os parâmetros usados foram $(r, \sigma, b) = (60, 10, \frac{8}{3})$. Usou-se como condição inicial para o sistema um ponto do atrator, enquanto para o modelo foi usado um ponto a uma distância de 10 unidades do mesmo. Não foi incluído nenhum tipo de ruído.

Na Figura 6.1 aparece a variável x do sistema e do modelo. Projeções bidimensionais dos atratores do sistema e do modelo no plano xz são mostradas na Figura 6.2. Nota-se que o controlador é capaz de sincronizar perfeitamente o modelo ao sistema original. Não é possível, nas figuras, discernir o sistema do modelo após o transitório.

No entanto, o resultado mais interessante é o comportamento da ação de controle u . Ele é mostrado na Figura 6.3.

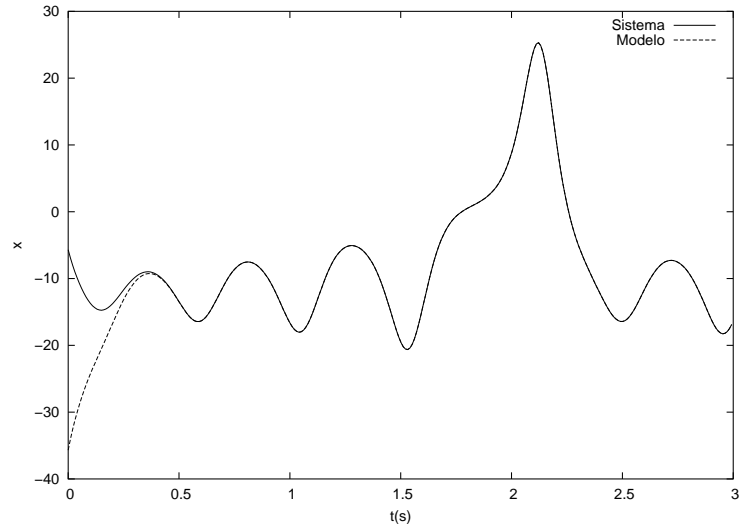


FIGURA 6.1 - Variável x em função do tempo para o sistema original e para o modelo.

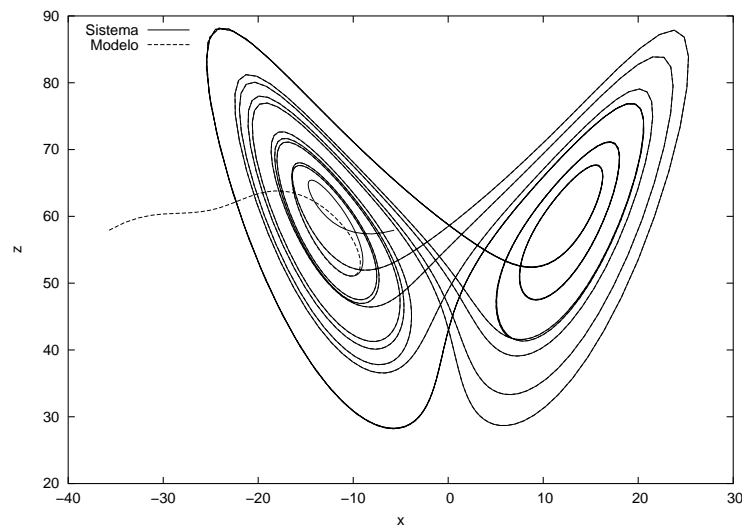


FIGURA 6.2 - Projeção no plano xz dos atratores do sistema e do modelo.

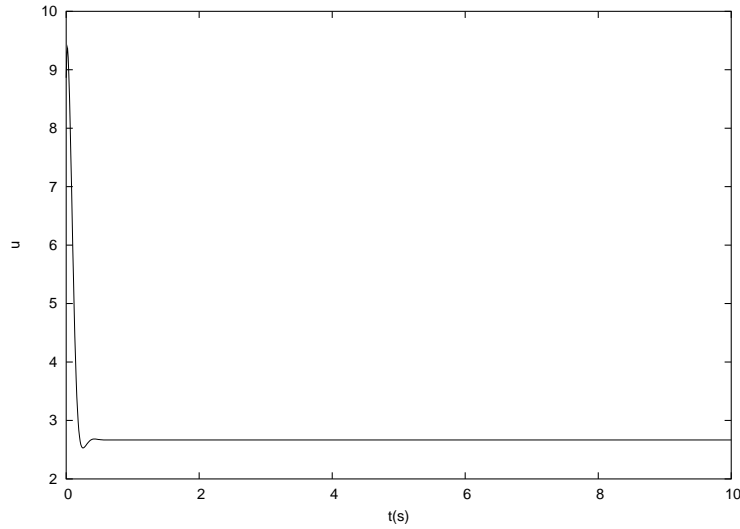


FIGURA 6.3 - Ação de controle.

À medida que ocorre o sincronismo, a ação de controle converge para o valor do parâmetro b do sistema original. Em outras palavras, a ação de controle é um estimador para o parâmetro b do sistema. Isso confirma que é possível, neste caso, usar sincronismo e controle para estimar um parâmetro desconhecido de um sistema caótico.

6.2 Versão multiparamétrica

Nesta Seção são mostrados os resultados do método de estimação de parâmetros na sua versão multiparamétrica.

6.2.1 Um estado observado — observador alto ganho

Neste exemplo, utiliza-se a técnica do observador de alto ganho (GAUTHIER *et al.*, 1992) como opção de controlador para se estimar dois parâmetros do sistema de Lorenz (6.1) sendo que é medido apenas o estado x_1 .

O sistema de Lorenz foi simulado com os parâmetros $(r, \sigma, b) = (28, 10, \frac{8}{3})$. O modelo usado na estimação de parâmetros foi

$$\begin{cases} \dot{x}_2 = \theta_1(y_2 - x_2) \\ \dot{y}_2 = \theta_2 x_2 - y_2 - x_2 z_2 + K_\gamma(x_1 - x_2) , \\ \dot{z}_2 = -b z_2 + x_2 y_2 \end{cases} \quad (6.21)$$

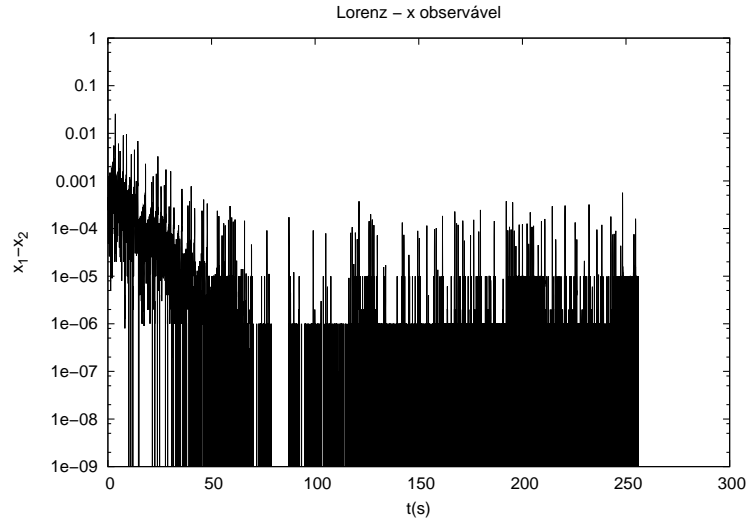


FIGURA 6.4 - Evolução temporal do erro de sincronização entre o modelo e o sistema, usando apenas a variável x_1 como observável.

sendo θ_1 e θ_2 os parâmetros a estimar e K_γ uma matriz de ganhos dada por $K_\gamma = -S_\gamma^{-1}[1\ 0\ 0]^T$. A matriz S_γ é a solução da equação (5.27) (vide p. 86). O valor de γ usado foi 150. Foram simulados 16000 pontos com passo de integração 0,016s. Foi usado o valor 0 como condição inicial para os parâmetros a serem estimados.

O controlador, conforme esperado, é capaz de sincronizar os dois sistemas. A evolução temporal do erro de sincronização é mostrada na Figura 6.4.

A evolução temporal dos parâmetros estimados é mostrada na Figura 6.5.

É possível notar que os parâmetros convergem para uma região próxima aos valores esperados. Isso pode ser melhor visualizado na Figura 6.6, onde é mostrado em escala logarítmica o valor absoluto do erro de estimação dos parâmetros. Inicialmente, o erro de estimação cai quase exponencialmente. No entanto, após esse período, o valor estimado fica oscilando próximo do valor real.

A média e o desvio padrão dos parâmetros estimados, descontando-se o período transitório, são mostrados na Tabela 6.1.

Como forma de validação, o modelo com os parâmetros estimados foi simulado infinitos passos à frente. Além do mais, ao invés de se usar os valores médios da Tabela 6.1 foram usados os últimos valores obtidos na simulação, a saber $\theta_1 = 10,04267$

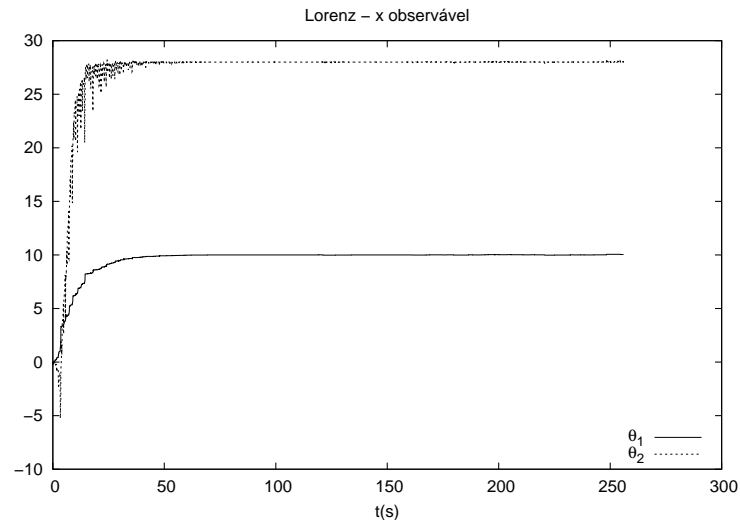


FIGURA 6.5 - Evolução temporal dos parâmetros estimados usando apenas a variável x_1 como observável.

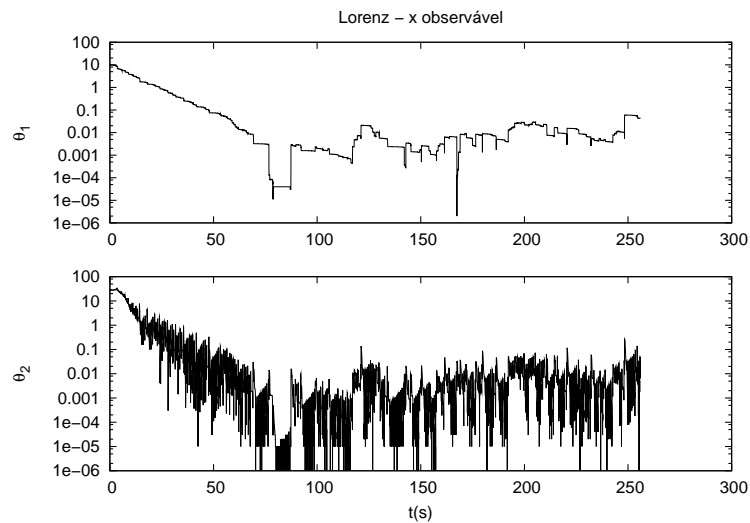


FIGURA 6.6 - Evolução temporal do erro de estimação dos parâmetros usando apenas a variável x_1 como observável.

TABELA 6.1 - Parâmetros estimados para o sistema de Lorenz usando apenas x_1 como observável.

Parâmetro	Média	Desvio	Valor esperado
θ_1	10,00501	0,01528	10
θ_2	28,00335	0,01739	28

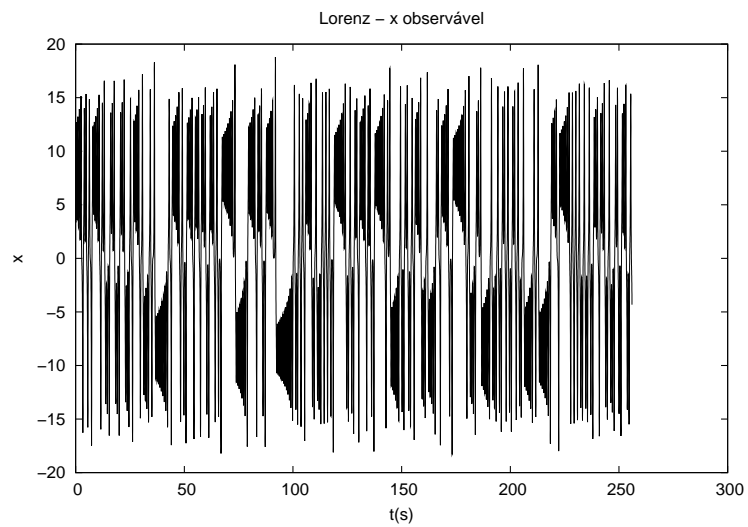


FIGURA 6.7 - Evolução temporal da variável x do modelo com os parâmetros estimados.

e $\theta_2 = 28,00646$. A evolução temporal da variável x do modelo é mostrada na Figura 6.7 e a projeção do atrator do modelo no plano xy na Figura 6.8.

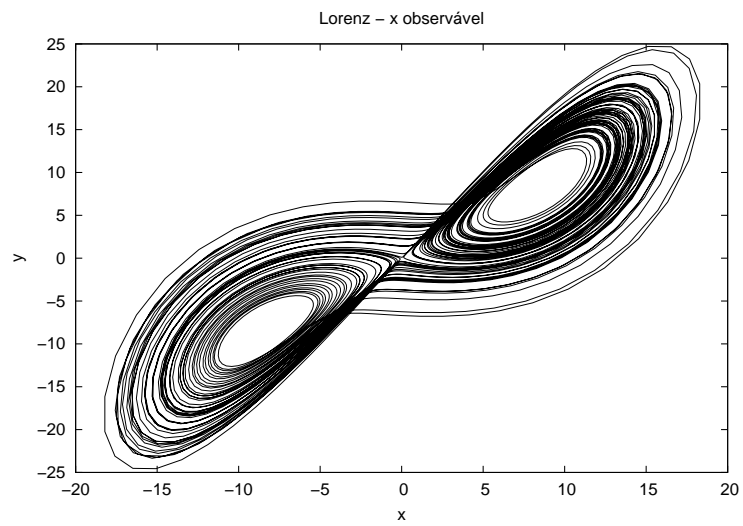


FIGURA 6.8 - Projeção no plano xy do atrator do modelo com os parâmetros estimados.

6.2.2 Todos os estados —Controlador de alto ganho

Para ilustrar a versão multiparamétrica com a observação de todos os estados, o sistema escravo (modelo) é construído de uma forma que aproxima a de uma identificação real. Ou seja, *todos* os coeficientes dos termos de (6.1) são supostos desconhecidos e alguns *termos espúrios* (que não estão presentes no sistema mestre) são adicionados ao modelo. O modelo é definido então como

$$\begin{cases} \dot{x}_2 = \theta_1 y_2 - \theta_4 x_2 + \theta_8 z_2 + \theta_9 x_2 y_2 + k_1(x_1 - x_2) , \\ \dot{y}_2 = \theta_2 x_2 - \theta_5 y_2 - \theta_6 x_2 z_2 + k_2(y_1 - y_2) , \\ \dot{z}_2 = -\theta_3 z_2 + \theta_7 x_2 y_2 + k_3(z_1 - z_2) , \end{cases} \quad (6.22)$$

sendo os θ_i os parâmetros do modelo, os k_i os coeficientes do controlador linear de alto ganho (vide (5.28) na p. 86), e $[x_2 \ y_2 \ z_2]^T$ o vetor de estados. Note que todos os termos do mestre (6.1) são reproduzidos no escravo com coeficientes desconhecidos e que termos com os coeficientes θ_8 e θ_9 não existem no sistema mestre. O erro de sincronização é definido como

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \\ z_1 - z_2 \end{bmatrix} , \quad (6.23)$$

e a meta de controle como

$$Q(\mathbf{e}, \boldsymbol{\theta}) = \frac{1}{2}(\mathbf{e}^T \mathbf{e}) . \quad (6.24)$$

A derivada de $Q(\mathbf{e}, \boldsymbol{\theta})$ ao longo da trajetória do sistema é dada por

$$\omega(\mathbf{e}, \boldsymbol{\theta}) = \frac{\partial Q}{\partial \mathbf{e}} \begin{bmatrix} \sigma(y_1 - x_1) - [\theta_1 y_2 - \theta_4 x_2 + \theta_8 z_2 + \theta_9 x_2 y_2 + k_1(x_1 - x_2)] \\ r x_1 - y_1 - x_1 z_1 - [\theta_2 x_2 - \theta_5 y_2 - \theta_6 x_2 z_2 + k_2(y_1 - y_2)] \\ -b z_1 + x_1 y_1 - [-\theta_3 z_2 + \theta_7 x_2 y_2 + k_3(z_1 - z_2)] \end{bmatrix} . \quad (6.25)$$

TABELA 6.2 - Valores finais dos parâmetros

Parâmetro	Valor	Esperado
θ_1	$9,9996 \times 10^0$	10
θ_2	$2,7998 \times 10^1$	28
θ_3	$2,6667 \times 10^0$	$\frac{8}{3}$
θ_4	$1,0006 \times 10^1$	10
θ_5	$1,0004 \times 10^0$	1
θ_6	$9,9994 \times 10^{-1}$	1
θ_7	$9,9998 \times 10^{-1}$	1
θ_8	$1,5324 \times 10^{-3}$	0
θ_9	$6,1248 \times 10^{-4}$	0

Assim, o gradiente de (6.25) é expresso por

$$\nabla_{\theta} \omega(\mathbf{e}, \boldsymbol{\theta}) = \text{diag} \left(\begin{bmatrix} y_2(x_1 - x_2) \\ x_2(y_1 - y_2) \\ -z_2(z_1 - z_2) \\ -x_2(x_1 - x_2) \\ -y_2(y_1 - y_2) \\ -x_2z_2(y_1 - y_2) \\ x_2y_2(z_1 - z_2) \\ z_2(x_1 - x_2) \\ x_2y_2(x_1 - x_2) \end{bmatrix} \right). \quad (6.26)$$

As equações diferenciais (6.1) e (6.22) foram resolvidas numericamente pelo método Runge-Kutta de quarta ordem, com passo de integração 0,002. Os parâmetros do modelo, θ , foram atualizados segundo o método da forma diferencial. Os parâmetros do sistema de Lorenz usados foram $(\sigma, r, b) = (10, 28, \frac{8}{3})$. Os ganhos usados no controlador foram $k_i = 100$, $i = 1, 2, 3$. A matriz Γ foi escolhida como uma matriz diagonal $\Gamma = \text{diag}([10 \ 10^3 \ 10 \ 10 \ 10 \ 1 \ 10 \ 10 \ 10]^T)$. As condições iniciais para os parâmetros foram todas nulas.

Todos os parâmetros convergiram para os valores esperados, inclusive θ_8 e θ_9 convergiram para valores pequenos, aproximadamente 3 ordens de grandeza menores que os demais parâmetros. Os valores finais dos parâmetros, após 240s de simulação, se encontram na tabela 6.2.

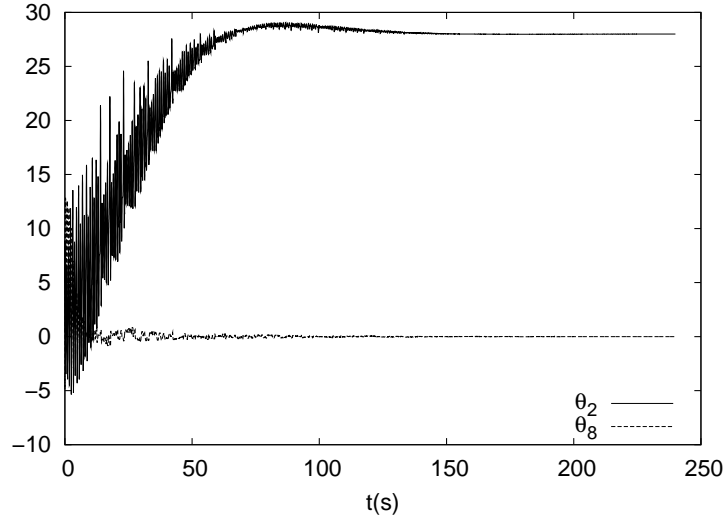


FIGURA 6.9 - Evolução temporal dos parâmetros θ_2 e θ_8 .

Na Figura 6.9 é mostrada a evolução temporal de dois parâmetros do modelo. Os outros parâmetros se comportam de maneira similar.

Um outro exemplo de aplicação usa o sistema de Rössler (RÖSSLER, 1976)

$$\begin{cases} \dot{x}_1 = -y_1 - z_1 , \\ \dot{y}_1 = x_1 + ay_1 , \\ \dot{z}_1 = b + z_1(x_1 - c) , \end{cases} \quad (6.27)$$

sendo (a, b, c) os parâmetros.

Novamente, é suposto que *todos* os coeficientes de (6.27) são desconhecidos, não apenas a tripla (a, b, c) . Dessa forma, o modelo a ter os parâmetros estimados é da forma

$$\begin{cases} \dot{x}_2 = \theta_1 y_2 + \theta_2 z_2 + \theta_8 x_2 + k_1(x_1 - x_2) , \\ \dot{y}_2 = \theta_3 x_2 + \theta_4 y_2 + k_2(y_1 - y_2) , \\ \dot{z}_2 = \theta_5 + \theta_6 z_2 x_2 + \theta_7 z_2 + \theta_9 y_2 + k_3(z_1 - z_2) , \end{cases} \quad (6.28)$$

sendo os θ_i os parâmetros do modelo, os k_i os coeficientes do controlador linear de alto ganho, e $[x_2 \ y_2 \ z_2]^T$ o vetor de estados, como no exemplo anterior. Da mesma forma, todos os termos do mestre (6.27) são reproduzidos no escravo com coeficientes desconhecidos os termos com os coeficientes θ_8 e θ_9 não existem no sistema mestre. O erro de sincronização e a meta de controle são definidos da mesma forma que no exemplo anterior.

Similarmente, a derivada de $Q(\mathbf{e}, \boldsymbol{\theta})$ ao longo da trajetória do sistema é dada por

$$\omega(\mathbf{e}, \boldsymbol{\theta}) = \frac{\partial Q}{\partial \mathbf{e}} \begin{bmatrix} -y_1 - z_1 - [\theta_1 y_2 + \theta_2 z_2 + \theta_8 x_2 + k_1(x_1 - x_2)] \\ x_1 + ay_1[\theta_3 x_2 + \theta_4 y_2 + k_2(y_1 - y_2)] \\ b + z_1(x_1 - c)[\theta_5 + \theta_6 z_2 x_2 + \theta_7 z_2 + \theta_9 y_2 + k_3(z_1 - z_2)] \end{bmatrix} \quad (6.29)$$

e o seu gradiente por

$$\nabla_{\boldsymbol{\theta}} \omega(\mathbf{e}, \boldsymbol{\theta}) = \text{diag} \left(\begin{bmatrix} y_2(x_1 - x_2) \\ z_2(x_1 - x_2) \\ x_2(y_1 - y_2) \\ y_2(y_1 - y_2) \\ (z_1 - z_2) \\ x_2 z_2(z_1 - z_2) \\ z_2(z_1 - z_2) \\ x_2(x_1 - x_2) \\ y_2(z_1 - z_2) \end{bmatrix} \right). \quad (6.30)$$

As equações (6.27) e (6.28) foram integradas numericamente pelo método Runge-Kutta de quarta ordem com passo de integração 0,002 e foi usando na atualização dos parâmetros $\boldsymbol{\theta}$ o método da forma diferencial. Os parâmetros do sistema de Rössler usados foram $(a, b, c) = (0, 398; 2; 4)$. Os ganhos usados no controlador foram $k_i = 10$, $i = 1, 2, 3$. A matriz *Gamma* foi escolhida como uma matriz diagonal $\Gamma = \text{diag}([1 \ 1 \ 1 \ 1 \ 10 \ 10 \ 10 \ 1 \ 10]^T)$. As condições iniciais para os parâmetros foram todas nulas.

Todos os parâmetros convergiram para os valores esperados, inclusive θ_8 e θ_9 convergiram para valores pequenos, pelo menos 3 ordens de grandeza menores que os demais parâmetros. Os valores finais dos parâmetros, após 1280s de simulação, se encontram na tabela 6.3.

Na Figura 6.10 é mostrada a evolução temporal de dois parâmetros do modelo. Os outros parâmetros se comportam de maneira similar ou convergem mais rápido. A Figura 6.11 mostra projeções planas dos atratores do sistema original e do modelo com os parâmetros estimados.

TABELA 6.3 - Valores finais dos parâmetros para o sistema de Rössler

Parâmetro	Valor	Esperado
θ_1	$-9,9999 \times 10^{-1}$	-1
θ_2	$-9,9999 \times 10^{-1}$	-1
θ_3	$1,0000 \times 10^0$	1
θ_4	$3,9800 \times 10^{-1}$	0,398
θ_5	$1,9982 \times 10^0$	2
θ_6	$9,9933 \times 10^{-1}$	1
θ_7	$-3,9972 \times 10^0$	-4
θ_8	$-3,5410 \times 10^{-6}$	0
θ_9	$-1,2868 \times 10^{-3}$	0

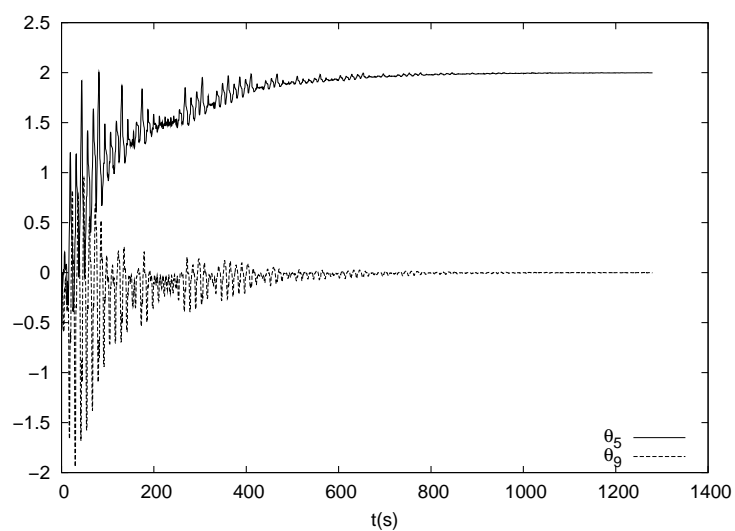


FIGURA 6.10 - Evolução temporal dos parâmetros θ_5 e θ_9 .

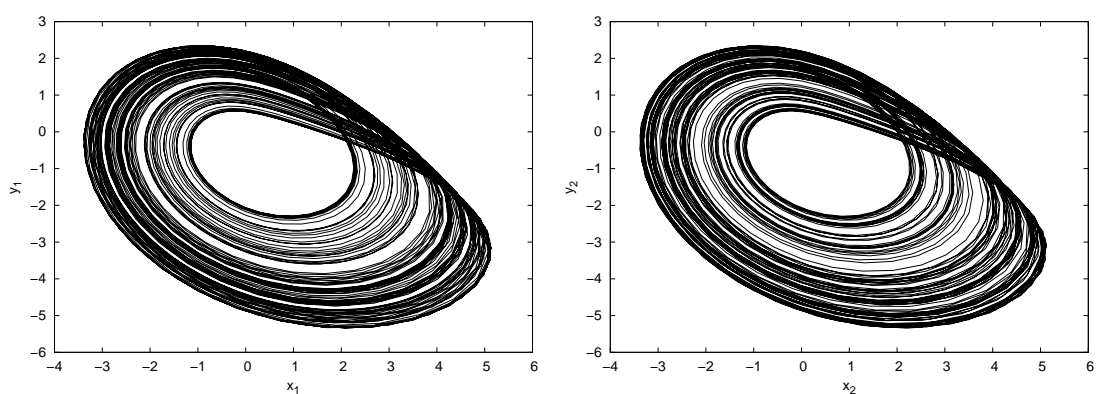


FIGURA 6.11 - Projeções planas dos atratores do sistema original (esquerda) e do modelo com parâmetros estimados (direita).

CAPÍTULO 7

CONCLUSÕES E PROPOSTAS DE CONTINUIDADE

Identificação de sistemas é uma forma de se obter modelos matemáticos. A relevância desse tema está ligada à importância que os modelos matemáticos possuem em ciências e tecnologia modernas.

Este trabalho apresenta um método de estimação de parâmetros que se baseia em sincronização e controle.

O método desenvolvido aqui utiliza o fenômeno da sincronização de sistemas como uma ferramenta de estimação de parâmetros. Para isso, a questão da sincronização é abordada sob o ponto de vista da teoria de controle não linear. Tal abordagem permitiu a criação de métodos de estimação de parâmetros para sistemas não lineares, tanto para problemas com um parâmetro desconhecido quanto para problemas multiparamétricos.

A versão monoparamétrica do método usa o controle não linear geométrico para atuar no parâmetro desconhecido do modelo de forma que este sincronize com o sistema sob estudo.

A versão multiparamétrica funciona usando técnicas de controle não linear em conjunto com um método de controle adaptativo para estimar os parâmetros desconhecidos. Três possibilidades são apresentadas: dois observadores e um controlador de alto ganho.

O método apresentado foi testado, via simulação, em dois sistemas caóticos bem estudados: o sistema de Lorenz e o sistema de Rössler.

7.1 Conclusões

O Capítulo 2 formula a identificação de sistemas como um problema inverso e descreve como uma classe de métodos de identificação ataca esse problema. Também é mostrado que existem dois problemas inversos associados à identificação de sistemas dinâmicos: o Problema Inverso Completo, PIC e o Problema Inverso Reduzido, PIR. Muitos métodos de identificação de sistemas tradicionais resolvem o PIR ao invés de resolver o PIC. Isso acontece com os métodos que usam o erro de predição um passo à frente.

A solução para o PIC, que é o problema inverso que melhor descreve o que se deseja em identificação de sistemas, recai num problema de otimização muito difícil, principalmente em se tratando da identificação de sistemas não lineares. No entanto, esse é o problema de deve ser abordado se se deseja que os modelos identificados reproduzam melhor o comportamento temporal dos sistemas sob estudo.

Para problemas monoparamétricos, mostrou-se que, na ausência de ruído, se um controlador atuando no parâmetro desconhecido é capaz de sincronizar o modelo com o sistema então a ação de controle converge para o valor do parâmetro do sistema.

Para o caso multiparamétrico, mostrou-se que a união de observadores para sistemas caóticos com o método adaptativo do gradiente de velocidade fornece uma ferramenta para a identificação de sistemas. No caso em que todos os estados do sistema são medidos, mostrou-se que é possível usar um controlador linear de alto ganho no método do gradiente de velocidade como uma ferramenta de identificação.

7.2 Propostas de continuidade

Um método de identificação, quando aplicado em casos práticos, enfrentará necessariamente o problema de ruído. Uma questão que deve ser então abordada é o comportamento do método aqui proposto face à presença de ruído nos dados de identificação.

Uma possibilidade interessante é verificar a viabilidade da extensão do método para modelos de tempo discreto.

Finalmente, deve-se aplicar o método a problemas práticos de identificação de sistemas e análise de séries temporais.

REFERÊNCIAS BIBLIOGRÁFICAS

AGUIRRE, L. A. **Introdução à identificação de sistemas**: técnicas lineares e não lineares aplicadas a sistemas reais. Belo Horizonte: Editora UFMG, 2000. [19](#), [20](#), [22](#), [27](#), [29](#), [30](#), [31](#), [32](#), [35](#), [37](#), [44](#)

_____. A nonlinear dynamical approach to system identification. **IEEE Circuits Syst. Newsletter**, v. 11, n. 2, p. 10–23, 2000. [31](#), [33](#)

AGUIRRE, L. A.; BILLINGS, S. A. Validating identified nonlinear models with chaotic dynamics. **Int. J. Bifurcation and Chaos**, v. 4, n. 1, p. 109–125, 1994. [23](#), [38](#)

_____. Dynamical effects of overparametrization in nonlinear models. **Physica D**, v. 80, n. 1,2, p. 26–40, 1995. [22](#), [41](#)

AGUIRRE, L. A.; FREITAS, U. S.; LETELLIER, C.; MAQUET, J. Structure selection techniques applied to continuous-time nonlinear models. **Physica D**, v. 158, n. 2, p. 1–18, 2001. [22](#), [36](#), [41](#)

AGUIRRE, L. A.; SOUZA, A. V. P. An algorithm for estimating fixed points of dynamical systems from time series. **Int. J. Bifurcation and Chaos**, v. 8, n. 11, p. 2203–2213, 1998. [37](#)

ALIFANOV, O. M. **Identification of heat transfer process of flying vehicles (an introduction to the theory of inverse heat transfer problems)**. Moscou: Mashinostroenie Publishing Academy, 1979. [38](#)

ALLIGOOD, K. T.; SAUER, T. D.; YORKE, J. A. **Chaos** — an introduction to dynamical systems. London: Springer Verlag, 1994. [23](#), [38](#)

ARGYRIS, J.; FAUST, G.; HAASE, M. **An exploration of chaos**. Amsterdam: North-Holland, 1994. [41](#)

BAAKE, E.; BAAKE, M.; BOCK, H. G.; BRIGGS, K. M. Fitting ordinary differential equations to chaotic data. **Phys. Rev. A**, v. 45, n. 8, p. 5524–5529, 1992. [22](#)

BILLINGS, S. A.; CHEN, S. Identification of nonlinear rational systems using a predictor-error estimation algorithm. **Int. J. Systems Sci.**, v. 20, n. 3, p. 467–494, 1989. [33](#), [36](#)

- BILLINGS, S. A.; CHEN, S.; KORENBERG, M. J. Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. **Int. J. Control**, v. 49, n. 6, p. 2157–2189, 1989. 36
- BOCCALETTI, S.; KURTHS, J.; OSIPOV, G.; VALLADARES, D. L.; ZHOU, C. S. The synchronization of chaotic systems. **Physics Reports**, v. 366, n. 1, p. 1–101, 2002. 53, 55
- BOUTAYEB, M.; DAROUACH, M.; RAFARALAHY, H. Generalized state-space observers for chaotic synchronization and secure communications. **IEEE Trans. Circuits Syst. I**, v. 49, n. 3, p. 345–349, 2002. 61
- BROOMHEAD, D. S.; LOWE, D. Multivariable functional interpolation and adaptive networks. **Complex Systems**, v. 2, p. 321–355, 1988. 33
- CAMINHAS, W. M.; GOMIDE, F. A. C. A fast learning algorithm for neofuzzy networks. In: IPMU, 8, 2000, Madrid, Spain. **Proceedings....** Madrid, Spain: IMPU, 2000. p. 1784–1790. 33
- CHEN, C. T. **Analog and digital control system design**. London: Saunders, 1993. 61
- CORRÊA, M. **Identificação de sistemas dinâmicos não lineares utilizando modelos NARMAX racionais** — aplicação a sistemas reais. Dissertação (Mestrado) — PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, 1997. 36
- ECKMANN, J. P.; RUELE, D. Ergodic theory of chaos and strange attractors. **Rev. Mod. Phys.**, v. 57, n. 3, p. 617–656, 1985. 58
- ELSNER, J. B. Predicting time series using a neural network as a method of distinguishing chaos from noise. **J. Phys. A: Math. Gen.**, v. 25, p. 843–850, 1992. 33
- FERREIRA, A. B. d. H. **Novo dicionário da língua portuguesa**. Primeira. Rio de Janeiro: Editora Nova Fronteira, 1975. 50
- FIEDLER-FERRARA, N.; PRADO, C. P. C. **Caos: Uma introdução**. São Paulo: Editora Edgard Blücher LTDA., 1994. 47

FRADKOV, A. L.; MARKOV, A. Y. Adaptive synchronization of chaotic systems based on speed gradient method and passification. **IEEE Trans. Circuits Syst. I**, v. 44, n. 10, p. 905–912, 1997. [83](#)

FRADKOV, A. L.; POGROMSKY, A. Y. Speed gradient control of chaotic continuous-time systems. **IEEE Trans. Circuits Syst. I**, v. 43, n. 11, p. 907–913, 1996. [24](#), [25](#), [72](#), [73](#), [74](#), [84](#)

FREITAS, U. S. **Uso de técnicas de detecção de estrutura na identificação de modelos dinâmicos não lineares contínuos polinomiais**. Dissertação (Mestrado) — PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, 2001. [31](#), [36](#), [41](#)

FREITAS, U. S.; MACAU, E. E. N.; GREBOGI, C. Using geometric control and chaotic synchronization to estimate an unknown model parameter. **Physical Review E**, v. 71, p. 047203, 2005. [77](#)

GARCIA, C. **Modelagem e simulação de processos industriais e de sistemas eletromecânicos**. São Paulo: EDUSP, 1997. [20](#)

GAUTHIER, J. P.; HAMMOURI, H.; OTHMAN, S. A simple observer for nonlinear systems, applications to bioreactors. **IEEE Transactions on Automatic Control**, v. 37, n. 6, p. 875–880, 1992. [84](#), [85](#), [86](#), [96](#)

GIBSON, J. F.; FARMER, J. D.; CASDAGLI, M.; EUBANK, S. An analytic approach to practical state space reconstruction. **Physica D**, v. 57, p. 1, 1992. [32](#)

GIONA, M.; LENTINI, F.; CIMAGALLI, V. Functional reconstruction and local prediction of chaotic time series. **Phys. Rev. A**, v. 44, n. 6, p. 3496–3502, 1991. [22](#)

GLEICK, J. **Caos e a criação de uma nova ciência**. Quarta. Rio de Janeiro: Editora Campus, 1991. [47](#)

GOUESBET, G.; LETELLIER, C. Global vector field reconstruction by using a multivariate polynomial L_2 approximation on nets. **Phys. Rev. E**, v. 49, n. 6, p. 4955–4972, 1994. [22](#)

GRASSI, G.; MILLER, D. A. Theory and experimental realization of observer-based discrete-time hyperchaos synchronization. **IEEE Trans. Circuits Syst. I**, v. 49, n. 3, 2002. [61](#)

- HAYKIN, S. **Redes neurais: princípios e prática**. Porto Alegre, RS: Bookman, 2001. 22
- HSIA, T. C. **System identification**. Lexington, Massachusetts: Lexington Books, 1977. 27, 29, 30
- HUANG, D.; GUO, R. Identifying parameter by identical synchronization between different systems. **Chaos**, v. 14, n. 1, p. 152–159, 2004. 24
- ISIDORI, A. **Nonlinear control systems**. 3rd. ed. London, England: Springer, 1995. 64, 67, 68
- JACKSON, E. A.; GROSU, I. An open-plus-closed-loop (OPCL) control of complex dynamic system. **Physica D**, v. 85, p. 1–9, 1995. 67
- JAEGER, L.; KANTZ, H. Unbiased reconstruction of the dynamics underlying a noisy chaotic time series. **Chaos**, v. 6, n. 3, p. 440–450, 1996. 42, 44
- JAKUBCZYK, B. **Introduction to geometric nonlinear control: controllability and Lie bracket**. Trieste, Itália, 2001. Lectures given at the Summer Schools on Mathematical Control Theory. Disponível em: <http://www.ist.uni-stuttgart.de/education/courses/respondek/Jakubczyk.pdf>. Acesso em: 16 nov. 2006. 67, 70
- JOHANSSON, R. **System modeling and identification**. Englewood Cliffs: Prentice-Hall, 1993. 27, 29, 30
- JUDD, K.; MEES, A. On selecting models for nonlinear time series. **Physica D**, v. 82, n. 4, p. 426–444, 1995. 82
- KHALIL, H. K. **Nonlinear systems**. Upper Saddle River, NJ: Prentice-Hall, 1996. 83, 84, 87
- KONNUR, R. Estimation of all model parameters of chaotic systems from discrete scalar time series measurements. **Physics Letters A**, v. 346, n. 4, p. 275–280, OCT 17 2005. 24
- LEONTARITIS, I. J.; BILLINGS, S. A. Input-output parametric models for nonlinear systems part I: deterministic nonlinear systems. **Int. J. Control**, v. 41, n. 2, p. 303–328, 1985. 33

_____. Input-output parametric models for nonlinear systems part II: stochastic nonlinear systems. **Int. J. Control**, v. 41, n. 2, p. 329–344, 1985. [33](#)

LETELLIER, C.; SCELLER, L. L.; DUTERTRE, P.; GOUESBET, G.; FEI, Z.; HUDSON, J. L. Topological characterization and global vector field reconstruction of an experimental electrochemical system. **Journal of Physical Chemistry**, v. 99, n. 18, p. 7016–7027, 1995. [38](#)

LIAO, T.-L.; HUANG, N.-S. An observer-based approach for chaotic synchronization with applications to secure communications. **IEEE Trans. Circuits Syst. I**, v. 46, n. 9, p. 1144–1150, 2001. [61](#)

LJUNG, L. **System identification, theory for the user**. New Jersey: Prentice Hall, 1987. [20](#), [22](#), [27](#), [31](#), [34](#), [35](#), [36](#)

LORENZ, E. Deterministic nonperiodic flow. **J. Atmos. Sci.**, v. 20, p. 282–293, 1963. [48](#), [89](#)

MARINO, I. P.; MIGUEZ, J. An approximate gradient-descent method for joint parameter estimation and synchronization of coupled chaotic systems. **Physics Letters A**, v. 351, n. 4-5, p. 262–267, MAR 6 2006. [24](#)

_____. On a recursive method for the estimation of unknown parameters of partially observed chaotic systems. **Physica D-Nonlinear Phenomena**, v. 220, n. 2, p. 175–182, AUG 15 2006. [24](#)

MAYBHATE, A.; AMRITKAR, R. E. Use of synchronization and adaptive control in parameter estimation from a time series. **Physical Review E**, v. 59, n. 1, p. 284–293, 1999. [24](#)

_____. Dynamic algorithm for parameter estimation and its applications. **Physical Review E**, v. 61, n. 6, p. 6461–6470, 2000. [24](#)

MILLERIOUX, G.; DAAFOUZ, J. Global chaos synchronization and robust filtering in noisy context. **IEEE Trans. Circuits Syst. I**, v. 48, n. 10, p. 1170–1176, 2001. [61](#)

MONTEIRO, L. H. A. **Sistemas dinâmicos**. São Paulo: Editora Livraria da Física, 2002. [19](#), [28](#), [48](#), [51](#)

- MUNIZ, W. B. **Um problema inverso em condução de calor utilizando métodos de regularização**. 79 p. Dissertação (Mestrado em Matemática Aplicada) — Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999. [38](#)
- NIJMEIJER, H. A dynamical control view on synchronization. **Physica D**, v. 154, p. 219–228, 2001. [62](#), [84](#)
- NIJMEIJER, H.; MAREELS, I. M. Y. An observer looks at synchronization. **IEEE Trans. Circuits Syst. I**, v. 44, n. 10, p. 882–890, 1997. [61](#), [84](#), [85](#)
- NIJMEIJER, H.; SCHAFT, A. J. van der. **Nonlinear dynamical control systems**. New York, NY: Springer-Verlag, 1990. [67](#)
- OGATA, K. **Engenharia de controle moderno**. São Paulo: Segunda Edição, Prentice Hall do Brasil, 1993. [63](#), [64](#)
- OTT, E. **Chaos in dynamical systems**. Cambridge, England: Cambridge University Press, 1993. [58](#)
- PARLITZ, U.; JUNGE, L.; KOCAREV, L. Synchronization-based parameter estimations from time series. **Physical Review E**, v. 54, n. 6, p. 6253–6259, 1996. [24](#)
- PECORA, L. M.; CARROLL, T. L. Synchronization in chaotic systems. **Phys. Rev. Letters**, v. 64, n. 8, p. 821–824, 1990. [47](#)
- PECORA, L. M.; CARROLL, T. L.; JOHNSON, G. A.; MAR, D. J.; HEAGY, J. F. Fundamentals of synchronization in chaotic systems, concepts, and applications. **Chaos**, v. 7, n. 4, p. 520–542, 1997. [55](#), [57](#)
- PERONA, P.; PORPORATO, A.; RIDOLFI, L. On the trajectory method for the reconstruction of differential equations from time series. **Nonlinear Dynamics**, v. 23, n. 1, p. 13–33, 2000. [22](#)
- PHAM, D. T.; LIU, X. **Neural networks for identification, prediction and control**. London: Springer-Verlag, 1995. [22](#)
- REED, R. Pruning algorithms — a survey. **IEEE Trans. Neural Networks**, v. 4, n. 5, p. 740–746, 1993. [36](#)

RESPONDEK, W. **Geometry of static and dynamic feedback**. Trieste, Itália, 2001. Lectures given at the Summer Schools on Mathematical Control Theory.

Disponível em: [http:](http://www.ist.uni-stuttgart.de/education/courses/respondek/Respondek.pdf)

[//www.ist.uni-stuttgart.de/education/courses/respondek/Respondek.pdf](http://www.ist.uni-stuttgart.de/education/courses/respondek/Respondek.pdf).

Acesso em: 16 nov. 2006. 64, 65, 67, 69

RÖSSLER, O. An equation for continuous chaos. **Physics Letters A**, v. 57, n. 5, p. 397–398, 1976. 103

SAHA, D. C.; RAO, G. P. **Identification of continuous dynamical systems: the poisson moment functional (PMF) approach**. Berlin: Springer-Verlag, 1983. 22, 27

SLOTINE, J.-J. E.; LI, W. **Applied nonlinear control**. Englewood Cliffs, NJ: Prentice Hall, 1991. 63, 67, 70, 71, 72

SO, P.; OTT, E.; DAYAWANSA, W. P. Observing chaos: Deducing and tracking the state of a chaotic system from limited observation. **Phys. Rev. E**, v. 49, n. 4, 1994. 61

SÖDERSTRÖM, T.; STOICA, P. **System Identification**. London: Prentice Hall International, 1989. 22, 27, 29, 30, 35, 38

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE Trans. Syst., Man, Cybern.**, v. 15, n. 1, p. 116–132, 1985. 33

TAKENS, F. Detecting strange attractors in turbulence. In: RAND, D. A.; YOUNG, L. S. (Ed.). **Dynamical systems and turbulence, Lecture Notes in Mathematics, vol. 898**. Berlin: Springer Verlag, 1980. p. 366–381. 32

TIMMER, J.; RUST, H.; HORBELT, W.; VOSS, H. U. Parametric, nonparametric and parametric modelling of a chaotic circuit time series. **Physics Letters A**, v. 274, p. 123–134, 2000. 22, 24, 44

TÔRRES, L. A. B. **Controle e sincronismo de osciladores caóticos**. Tese (Doutorado) — PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, 2001. 62, 67

WANG, X. F.; WANG, Z. Q. Synchronizing chaos and hyperchaos with any scalar transmitted signal. **IEEE Trans. Circuits Syst. I**, v. 45, n. 10, p. 1101–1103, 1998. 61

WEISSTEIN, E. W. Parameterization. In: **Mathworld**—a Wolfram web resource. Wolfram Research, Inc., 1999. Disponível em:

<http://mathworld.wolfram.com/Parameterization.html>. Acesso em: 20 jun. 2006. 76

WU, C.-J.; LEE, Y.-C. Observer-based method for secure communications of chaotic systems. **Electronics Letters**, v. 36, n. 22, p. 1842–1843, 2000. 61

YAMADA, T.; FUJISAKA, H. Stability theory of synchronized motion in coupled-oscillator systems .3. mapping model for continuous system. **Progress of Theoretical Physics**, v. 72, n. 5, p. 885–894, 1984. 47

YAMAKWA, T.; UCHINO, E.; MIKI, T.; KUSANAGI, T. A neo fuzzy neuron and its applications to system identification and predictions to system behavior. In: IIZUKA, 2, Japan, 1992. **Proceedings...** Japan: IIZUKA, 1992. p. 477–483. 33

YU, D.; DUANE, G. S.; KODAREV, L. Identical synchronization implies parameter estimation. **Phys. Rev. Lett.**, 2006. 84

ZHANG, Y.; TAO, C.; JIANG, J. J. Parameter estimation of an asymmetric vocal-fold system from glottal area time series using chaos synchronization. **Chaos**, v. 16, n. 2, jun. 2006. 24

_____. Theoretical and experimental studies of parameter estimation based on chaos feedback synchronization. **Chaos**, v. 16, n. 4, dez. 2006. 24

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programa de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. São aceitos tanto programas fonte quanto executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.