

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
ENGENHARIA ELÉTRICA

RAFAEL RUI

MODELOS COM
PARAMETRIZAÇÃO
POLINOMIAL:
IDENTIFICABILIDADE,
INFORMATIVIDADE E
IDENTIFICAÇÃO

Porto Alegre
2012

RAFAEL RUI

**MODELOS COM
PARAMETRIZAÇÃO
POLINOMIAL:
IDENTIFICABILIDADE,
INFORMATIVIDADE E
IDENTIFICAÇÃO**

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

ORIENTADOR: Prof. Dr. Alexandre Sanfelice
Bazanella

Porto Alegre
2012

RAFAEL RUI

**MODELOS COM
PARAMETRIZAÇÃO
POLINOMIAL:
IDENTIFICABILIDADE,
INFORMATIVIDADE E
IDENTIFICAÇÃO**

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

Orientador: _____
Prof. Dr. Alexandre Sanfelice Bazanella, UFRGS
Doutor pela Universidade Federal de Santa Catarina,
UFSC

Banca Examinadora:

Prof. Dr. Luis Antonio Aguirre, CPDEE – UFMG
Doutor pela Universidade de Sheffield – Sheffield, Inglaterra

Prof. Dr. Daniel Ferreira Coutinho, CTC – UFSC
Doutor pela Universidade Federal de Santa Catarina – Florianópolis, Brasil

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

Prof. Dra. Lucíola Campestrini, PPGEE – UFRGS
Doutora pela Universidade Federal do Rio Grande do Sul – Porto Alegre,
Brasil

Coordenador do Curso: _____
Prof. Dr. João Manoel Gomes da Silva Jr.

Porto Alegre, Novembro de 2012.

AGRADECIMENTOS

Aos meus pais por todo suporte nestes anos de estudo.

Aos colegas Diego, Lucíola e Manoel pelas muitas discussões que auxiliaram no desenvolvimento do trabalho.

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

Ao PPGEE e à CAPES pelo ensino de qualidade e concessão da bolsa.

RESUMO

Na modelagem caixa branca obtém-se um modelo para um processo a partir do equacionamento dos fenômenos físicos/químicos envolvidos. Estes modelos são parametrizados, mas os valores dos parâmetros utilizados muitas vezes são desconhecidos. Nestes casos é necessário efetuar um procedimento de identificação paramétrica o que representa um problema altamente desafiador, com muitas questões teóricas e práticas em aberto, quando os parâmetros aparecem de forma não linear no modelo. O objetivo deste trabalho é apresentar e estudar um método capaz de determinar se uma estrutura de modelo predeterminada pode ser identificada e que possa ser utilizado em conjunto com algum outro método de identificação, para identificar o sistema. O método que será apresentado é baseado em álgebra diferencial e é conhecido como algoritmo de Ritt. O algoritmo de Ritt transforma uma estrutura de modelo polinomial predeterminada em regressões lineares nos parâmetros a partir das quais pode-se utilizar os métodos dos mínimos quadrados ou variáveis instrumentais para identificar o sistema. Apresentaremos alguns estudos de caso e faremos a análise de identificabilidade para cada um deles. Em alguns casos identificaremos o sistema e estudaremos a consistência e precisão das estimativas.

Palavras-chave: Identificação de Sistemas, Identificabilidade, Sistemas Não-Lineares, Algoritmo de Ritt.

ABSTRACT

In white box modeling we obtain a model for a process from the equations of the physical/chemical phenomena involved. These models are parameterized, but the parameters used are often unknown. In these cases it is necessary to perform a parametric identification procedure which represents a highly challenging problem, with many theoretical and practical open questions when the parameters are nonlinear in the model. The aim of this work is to present and study a method able to determine whether a predetermined model structure can be identified and that can be used in conjunction with another identification method to identify the system. The method that will be presented is based on differential algebra and is known as Ritt algorithm. The Ritt's algorithm transforms a predetermined model structure in linear regression in the parameters from which one can use the least squares method or instrumental variables to identify the system. We will present some case studies and realise the analysis of identifiability for each case. For some cases we will identify the system, and then present a study for the consistency and precision of the estimates.

Keywords: System Identification, Analysis of Identifiability, Nonlinear Systems, Ritt's Algorithm.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	15
LISTA DE TABELAS	17
LISTA DE ABREVIATURAS	19
LISTA DE SÍMBOLOS	21
1 INTRODUÇÃO	23
1.1 Apresentação do trabalho	25
2 CONCEITOS DE IDENTIFICAÇÃO DE SISTEMAS	27
2.1 Introdução	27
2.2 Classe de Modelos	28
2.2.1 Estruturas de Modelos Padrão	30
2.3 Identificação por minimização do erro de predição	30
2.3.1 Estimador de Mínimos Quadrados	32
2.4 Incertezas nas estimativas dos parâmetros	33
2.4.1 Propriedades estatísticas do estimador de MQ	34
2.5 Variáveis instrumentais	35
2.6 Identificabilidade x Informatividade	36
2.6.1 Identificabilidade	37
2.6.2 Identificabilidade das estruturas padrões	37
2.6.3 Informatividade	38
2.7 Considerações Finais	39
3 ÁLGEBRA DIFERENCIAL	41
3.1 Introdução	41
3.2 Definições	41
3.3 Algoritmo de Ritt	46
3.3.1 Ritt Forma Alternativa	52
3.4 Algoritmo de Ritt Tempo Discreto	55
4 ESTUDO DE CASOS	59
4.1 Tempo Contínuo	59
4.1.1 Caso 1 – “Goodwin’s Napkin”	59
4.1.2 Caso 2 – Modelo não linear	61
4.1.3 Caso 3 – Modelo compartimental	62
4.1.4 Caso 4 – Três modelos Compartimentais	63

4.1.5	Caso 5	64
4.1.6	Caso 6 – Modelo Farmacocinético	66
4.1.7	Caso 7 – Modelo Compartimental	66
4.2	Tempo Discreto	69
4.3	Estudos numéricos	71
4.3.1	Resultados Sistema \mathcal{S}_1	71
4.3.2	Resultados Sistema \mathcal{S}_2	74
4.3.3	Resultados Sistema \mathcal{S}_3	74
4.4	Considerações Finais	81
5	CONCLUSÃO	83
	REFERÊNCIAS	85
	APÊNDICE A ROTINAS MUPAD	89
A.1	Rotina principal	89
A.2	Encontra lider	89
A.3	Conjunto	90
A.4	Qual reduzir	93
A.5	Reduções	94
A.6	Simplifica 1	97
A.7	Simplifica 2	98
A.8	Ordem	98
A.9	Ordem da Derivada	99
A.10	Rankink	100
A.11	Reduzido	101
A.12	Identificável	102

LISTA DE ILUSTRAÇÕES

4.1	Funções custo $V_N(\theta, Z^N)$ linha cheia e $J_N(\theta, Z^N)$ linha tracejada para o sistema \mathcal{S}_1	72
4.2	Sinal $y(t)$ para o sistema \mathcal{S}_1	73
4.3	Funções custo $V_N(\theta, Z^N)$ linha cheia e $J_N(\theta, Z^N)$ linha tracejada para o sistema \mathcal{S}_2	74
4.4	Função custo sistema \mathcal{S}_3	75
4.5	Função custo $J_N(\theta, Z^N)$ para sistema \mathcal{S}_3	76
4.6	Dispersão dos parâmetros utilizando o estimador de MQ para entrada $u_1(t)$	77
4.7	Dispersão dos parâmetros utilizando o estimador de MQ para entrada $u_2(t)$	78
4.8	Dispersão dos parâmetros utilizando o estimador de VI para entrada $u_1(t)$	79
4.9	Dispersão dos parâmetros utilizando o estimador de VI para entrada $u_2(t)$	80
4.10	Dispersão dos parâmetros utilizando o estimador de MQ e VI com ruído $\sigma_e = 0,6$	82

LISTA DE TABELAS

2.1	Estrutura de (2.9) comumente utilizadas.	30
2.2	Distribuição χ^2 para alguns valores de probabilidade α e diferentes graus de liberdade n	35
4.1	Média e desvio padrão para os estimadores, Sistema \mathcal{S}_1	73
4.2	Média e desvio padrão para os estimadores. Sistema \mathcal{S}_2	74

LISTA DE ABREVIATURAS

AR	<i>Auto Regressive</i>
ARMA	<i>Auto Regressive Moving Average</i>
ARX	<i>AutoRegressive with eXternal input</i>
ARMAX	<i>AutoRegressive Moving Average with with eXternal input</i>
BJ	<i>Box-Jenkings</i>
FIR	<i>Finite Impulse Response</i>
GPU	<i>Graphics Processing Unit</i>
MQ	Mínimos Quadrados
OE	<i>Output Error</i>
SISO	<i>Single-Input Single-Output</i>
VI	Variáveis Instrumentais

LISTA DE SÍMBOLOS

$[\cdot]^T$	Transposição de vetor ou matriz
\triangleq	Igual por definição
$f \equiv 0$	$f = 0 \forall t$
\sim	Mesmo rank
\prec	Ordenado abaixo de
$(\dot{\cdot})$	Operador diferenciação ou atraso temporal
A^T	Transposta da matriz A
$A \setminus B$	Diferença entre os conjuntos A e B , $\{x \mid x \in A \wedge x \notin B\}$
$\arg \min f(x)$	Valor de x que minimiza $f(x)$
$\epsilon(t, \theta)$	Vetor erro de predição
$e(t)$	Ruído ou perturbação do sistema
$\bar{E}[\cdot]$	Esperança matemática
\mathbb{F}	Campo \mathbb{F}
$\mathbb{F}\{\cdot\}$	Anel diferencial com coeficientes em \mathbb{F}
$G(q)$	Função de transferência de $u(t)$ para $y(t)$
$G(q, \theta)$	Função de transferência de uma estrutura, parametrizada por θ
$H(q)$	Função de transferência de $e(t)$ para $y(t)$
$H(q, \theta)$	Função de transferência de uma estrutura, parametrizada por θ
\mathcal{I}	Conjunto de polinômios diferenciais
I	Ideal diferencial
i	i -ésima componente de um vetor
$\rho(A)$	Posto da matriz A
\in	Pertence ao
\mathcal{M}	Estrutura do modelo
$M(\theta)$	Modelo correspondente ao valor de parâmetro θ
\mathbb{N}	Conjunto dos números naturais

θ	Vetor de parâmetros
$\hat{\theta}_N$	Estimativa de θ
P_θ	Matriz de correlação assintótica
p	Dimensão de θ
q^{-1}	Operador atraso, $q^{-1}u(t) = u(t - 1)$
\mathbb{R}	Conjunto dos números reais
\mathcal{S}	Sistema real
$\varphi(t)$	Vetor Regressor
$\zeta(t)$	Variável Instrumental
σ_e^2	Variância do ruído
$u(t)$	Sinal de entrada do sistema no instante t
$V_N(\theta, Z^N)$	Função custo
$\frac{\partial}{\partial \theta} V(y, u, \theta)$	Gradiente de $V(y, u, \theta)$ com respeito a θ
$\chi^2(n)$	Distribuição Qui-quadrado com n graus de liberdade
$y(t)$	Sinal de saída do sistema no instante t
$\hat{y}(t t - 1, \theta)$	Preditor um passo a frente
Z^N	Conjunto de dados do sistema

1 INTRODUÇÃO

A identificação de sistemas não lineares é um tópico muito encontrado na literatura e possui raízes em diferentes áreas: identificação de sistemas lineares na comunidade de controle (GEVERS, 2005), regressão não paramétrica na comunidade estatística (HARDLE, 1990), expansão de funções na comunidade de redes neurais, álgebra diferencial na comunidade biomédica (FRANÇOIS; FRANÇOIS, 2009), (AUDOLY et al., 2001) entre outras. Neste trabalho trataremos estruturas de modelos racionais muito utilizadas em sistemas biológicos (BEN-ZVI, 2010), (NEMCOVÁ, 2010), (CHIS; BANGA; BALSACANTO, 2011).

A identificação de sistemas tenta modelar sistemas dinâmicos a partir de um conjunto de dados. Uma suposição comum é que o sistema pode ser aproximado por uma estrutura parametrizada. A escolha dessa parametrização é altamente dependente da aplicação, onde o conhecimento do sistema é de vital importância para que se possa escolher uma parametrização adequada. Uma vez escolhida uma estrutura de modelo deve-se definir algum critério que seja capaz de avaliar se a estrutura de modelo escolhida é capaz de representar o conjunto de dados medidos do sistema. Neste trabalho utilizaremos a modelagem *caixa branca* ou *modelagem pela física do processo*, neste tipo de modelagem é necessário conhecer bem o sistema a ser modelado, bem como as leis físicas que descrevem o sistema a ser modelado (AGUIRRE, 2004). Os modelos obtidos através da modelagem *caixa branca* são parametrizados, e muitas vezes os parâmetros utilizados para parametrizar o modelo são desconhecidos.

Existem diversos métodos de estimação de parâmetros, onde dados uma estrutura de modelo parametrizada e um conjunto de dados dos sinais de entrada e saída, realizam-se uma série de procedimentos numéricos para obter uma estimativa para os parâmetros. No entanto, uma questão fundamental deve ser respondida antes de se utilizar tais métodos – a estrutura de modelo escolhida pode ser identificada?

A identificabilidade de uma estrutura é uma propriedade da estrutura do modelo. Esta propriedade é que garante que os parâmetros possam ser unicamente determinados a partir do conhecimento dos comportamentos da entrada e saída do sistema. Não é incomum que estruturas de modelos obtidas através de modelagens físicas e/ou químicas resultem em estruturas não identificáveis, ou seja, que existam infinitos valores para os parâmetros capazes de descrever os sinais de entrada e saída medidos do sistema.

A análise de identificabilidade de uma estrutura para modelos lineares é bastante conhecida e existem muitos métodos capazes de fazer essa análise (LJUNG, 1999) (AGUIRRE, 2004). Por outro lado, para modelos não lineares, racionais, existem alguns poucos métodos: métodos baseados em expansão em séries de Taylor

(POHJANPALO, 1978) (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001), mudança de similaridade (BEN-ZVI, 2010), teste direto (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001), álgebra diferencial (LJUNG, 1999) (AUDOLY et al., 2001) (MARGARIA et al., 2001), e mais recentemente métodos probabilísticos (KARLSSON; ANGUELOVA; JIRSTRAND, 2012).

Este trabalho tem como principal objetivo apresentar e estudar métodos de testar se uma estrutura predeterminada é identificável. O método utilizado é baseado em álgebra diferencial e é conhecido como Algoritmo de Ritt (RITT, 1950). O algoritmo de Ritt foi desenvolvido em 1950, de modo que podemos dizer que ele constitui um tema clássico. O uso do algoritmo de Ritt para analisar a identificabilidade de estruturas de modelos em tempo contínuo é conhecido pela comunidade de engenharia há no mínimo dezoito anos, (LJUNG; GLAD, 1994) e mais recentemente sua aplicação para a análise da identificabilidade de sistemas de tempo discreto foi apresentada em (LYZELL et al., 2011).

O algoritmo de Ritt é um algoritmo iterativo e tem como objetivo transformar uma estrutura de modelo parametrizada em regressões lineares nos parâmetros. Através de manipulações algébrico-diferenciais o algoritmo transforma uma estrutura de modelo predefinida em regressões lineares nos parâmetros, de tal modo que estas regressões lineares possuam a mesma solução da estrutura original. A partir dessa nova estrutura é possível, de forma simples, determinar se a estrutura original é identificável.

A vantagem de se utilizar métodos baseados em álgebra diferencial é que a solução das equações algébricas associadas à estrutura original nos dão informações precisas sobre a identificabilidade ou não identificabilidade da estrutura tratada. Porém conseguir escrever uma estrutura qualquer na forma de regressão linear tem um custo, este custo é a alta complexidade das expressões analíticas que tendem a aparecer na regressão linear equivalente. Não se encontram muitas aplicações do algoritmo de Ritt na literatura, mas sim algoritmos alternativos, também baseados em álgebra diferencial, capazes de determinar se uma estrutura é identificável, porém não de forma direta como o algoritmo de Ritt. Tais algoritmos utilizam, por exemplo, bases de Gröbner (MESHKAT; EISENBERG; III, 2009), cálculo de determinantes funcionais (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001), algoritmo de Buchberger (AUDOLY et al., 2001), etc. Neste trabalho também se propõe apresentar uma forma alternativa da utilização do algoritmo de Ritt para o estudo da identificabilidade de estruturas predeterminadas. Este método baseia-se na matriz de informação do modelo (GEVERS et al., 2009) (BAZANELLA; BOMBOIS; GEVERS, 2012).

Para que se possa realizar a identificação do sistema é necessário o conhecimento dos sinais de entrada e saída medidos no sistema que se deseja identificar. A qualidade do modelo obtido estará intimamente ligada com as características do sinal de excitação, ou seja, o modelo só reproduzirá as características do sistema se o sinal de excitação for suficientemente informativo. O algoritmo de Ritt consegue encontrar condições de persistência de excitação dos sinais medidos do sistema. Além disso, o algoritmo permite reescrever uma estrutura parametrizada não linear em uma estrutura linear nos parâmetros, tornando assim a identificação simples de ser implementada, diferentemente do que acontece quando se tenta identificar uma estrutura não linear diretamente.

1.1 Apresentação do trabalho

O principal objetivo deste trabalho é estudar e apresentar um método capaz de testar se uma estrutura predeterminada é identificável e o seu uso para a identificação de sistemas não lineares. O método que será apresentado é chamado de *algoritmo de Ritt* e se baseia em álgebra diferencial. Este método pode ser aplicado tanto em sistemas de tempo contínuo como em sistemas de tempo discreto.

Como mencionado anteriormente o algoritmo de Ritt é capaz de transformar uma estrutura de modelo parametrizada não linear em regressões lineares nos parâmetros. Essa nova estrutura em forma de regressões lineares possui a mesma solução que a estrutura original, ou seja, os sinais que satisfazem a estrutura original também irão satisfazer a estrutura resultante do algoritmo. Deste modo, a identificação de um sistema modelado por uma estrutura não linear pode ser feita através de uma estrutura equivalente parametrizada linearmente, onde através de métodos simples e de fácil implementação como mínimos quadrados ou variáveis instrumentais se pode realizar a identificação do sistema. Por esse motivo, começaremos apresentando, no capítulo 2, os conceitos para a identificação de sistemas lineares. Define-se o que é um modelo, uma estrutura de modelo, é apresentado o método de identificação por minimização do erro de predição, o estimador de mínimos quadrados e variáveis instrumentais. Então, apresentam-se aspectos gerais sobre as incertezas nos estimadores. São definidos também os conceitos de identificabilidade de uma estrutura de modelo paramétrica e informatividade do conjunto de dados.

No capítulo 3 são introduzidos conceitos gerais de álgebra diferencial necessários para a compreensão do algoritmo de Ritt, e então são apresentados o algoritmo de Ritt em tempo contínuo e tempo discreto. Também se apresenta uma proposta para uma forma alternativa da utilização do algoritmo de Ritt para a determinação da identificabilidade de estruturas de modelos.

No capítulo 4 apresentam-se os resultados obtidos através de uma série de estudos de casos, tanto em tempo contínuo quanto tempo discreto. Para sistemas de tempo contínuo são apresentados estudos da identificabilidade de estruturas utilizando o algoritmo de Ritt e a forma alternativa proposta no capítulo anterior sem considerar a presença de ruído no sistema. Para sistemas de tempo discreto acrescenta-se ruído ao sistema, e além do estudo da identificabilidade, apresenta-se um estudo numérico da identificação, consistência e precisão dos estimadores, utilizando as relações obtidas pelo algoritmo de Ritt.

Por fim, no capítulo 5, apresentam-se as conclusões do trabalho, e também algumas propostas de continuação do mesmo.

2 CONCEITOS DE IDENTIFICAÇÃO DE SISTEMAS

2.1 Introdução

Modelos matemáticos de sistemas são muito úteis em várias áreas do conhecimento. O modelo matemático de um sistema é um conjunto de equações que descrevem o seu comportamento dinâmico e estático. Existem diversas maneiras de classificar as técnicas de modelagem de sistemas. Uma delas agrupa os métodos em três categorias (AGUIRRE, 2004)

1. Modelagem caixa branca
2. Modelagem (ou identificação) caixa preta
3. Modelagem caixa cinza.

Na modelagem caixa branca é necessário conhecimento pleno das leis físicas relacionadas com o processo e também o conhecimento das grandezas físicas envolvidas, por isso esse tipo de modelagem é também conhecida como *modelagem pela física ou natureza do processo*. Para sistemas muito complexos esta modelagem normalmente torna-se inviável (AGUIRRE, 2004). Por outro lado na modelagem (ou identificação) caixa preta, pouco ou nenhum conhecimento prévio do sistema é necessário. Neste tipo de modelagem pretende-se descrever relações de causa e efeito entre os sinais de entrada e de saída do sistema.

Por fim, a modelagem caixa cinza pode ser colocada entre a modelagem pela física do processo e a modelagem caixa preta. Esse tipo de modelagem caracteriza-se por utilizar informações auxiliares, que não se encontram no conjunto de dados utilizados durante a identificação, ou seja, a estrutura do modelo é escolhida utilizando leis físicas e o modelo é ajustado a partir de dados coletados em um experimento.

A identificação de sistemas pode ser vista como a interface entre o mundo real das aplicações e o mundo matemático dos modelos. Ela é baseada em três elementos (LJUNG, 1999):

1. Um conjunto de dados Z^N , obtido através de dados de entrada e saída do sistema;

$$Z^N = [y(1), u(1), y(2), u(2), \dots, y(N), u(N)], \quad (2.1)$$

onde N é a quantidade de dados amostrados e $u(t)$ e $y(t)$ são, respectivamente, os sinais de entrada e saída do sistema.

2. Um conjunto de modelos candidatos \mathcal{M} capaz de reproduzir, pelo menos de maneira aproximada, as características do sistema original;

3. Um critério que relacione o conjunto de modelos candidatos com o conjunto de dados Z^N .

Os métodos desenvolvidos para a identificação de sistemas podem ser divididos em dois grupos (LJUNG, 2006):

- (a) métodos paramétricos: partem de estruturas matemáticas parametrizadas para descrever o comportamento dinâmico do sistema original. Os parâmetros destas estruturas são estimados usando algoritmos de estimação de parâmetros a partir dos dados medidos (Ljung, 1987).
- (b) métodos não-paramétricos: o comportamento dos sistemas dinâmicos é determinado através de representações gráficas que não têm nenhum parâmetro explícito, como por exemplo, a resposta em frequência do sistema original.

Neste trabalho iremos utilizar apenas estruturas parametrizadas, deste modo utilizaremos métodos paramétricos para a identificação. Além disso, os parâmetros serão invariantes no tempo, ou seja, seu valor permanece constante com o passar do tempo.

2.2 Classe de Modelos

Considere um processo SISO (do inglês *Single-Input, Single-Output*) de tempo discreto, linear e invariante no tempo

$$y(t) = G_0(q)u(t) + \nu(t), \quad (2.2)$$

onde q é o operador de avanço temporal definido como

$$\begin{aligned} qx(t) &= x(t+1), \\ q^{-1}x(t) &= x(t-1), \end{aligned}$$

$G_0(q)$ é a função de transferência do processo, $u(t)$ é o sinal de entrada do sistema, $y(t)$ é o sinal de saída do sistema e $\nu(t)$ é o ruído do processo.

O ruído é um processo quasi-estacionário e pode ser escrito como $\nu(t) = H_0(q)e(t)$, onde $H_0(q)$ é uma função de transferência do ruído, $e(t)$ é ruído branco com variância σ_e^2 e um processo quasi-estacionário é definido a seguir.

Definição 2.2.1. (LJUNG, 1999) *Um sinal $s(t)$ é dito quasi-estacionário se*

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

onde $\bar{E}[\cdot]$ é o operador definido por:

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

com $E[\cdot]$ denotando esperança.

◇

Deste modo podemos escrever o processo (2.2) como

$$y(t) = G_0(q)u(t) + H_0(q)e(t), \quad (2.3)$$

e o sistema \mathcal{S} será definido como

$$\mathcal{S} \triangleq [G_0(q) \ H_0(q)]. \quad (2.4)$$

O processo (2.3) será identificado utilizando uma classe de modelos parametrizada pelo vetor $\theta \in \mathbb{R}^p$:

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t), \quad (2.5)$$

onde $G(q, \theta)$ e $H(q, \theta)$ são funções de transferência da classe de modelo, correspondente ao parâmetro θ .

Para um dado $\theta \in \mathbb{R}^p$, $M(\theta) \triangleq [G(q, \theta) \ H(q, \theta)]$ é chamado de modelo, e o conjunto de modelos \mathcal{M} é definido através de um mapa diferenciável de um conjunto aberto conexo (BAZANELLA; BOMBOIS; GEVERS, 2012).

$$\mathcal{M} : \theta \in D_\theta \rightarrow M(\theta) = [G(q, \theta) \ H(q, \theta)], \quad (2.6)$$

onde $D_\theta \subseteq \mathbb{R}^p$. Neste trabalho assumiremos que a seguinte hipótese é válida:

Hipótese 2.2.1. *O sistema real \mathcal{S} pertence ao conjunto de modelos \mathcal{M} (ou simplesmente $\mathcal{S} \in \mathcal{M}$). Em outras palavras, existe $\theta_0 \in D_\theta$, tal que $M(\theta_0) = \mathcal{S}$.*

◇

Assim dada a Hipótese 2.2.1 podemos definir o valor real de um parâmetro como:

Definição 2.2.2. *O valor de θ_0 tal que $M(\theta_0) = \mathcal{S}$ é chamado de valor real para o parâmetro θ .*

◇

Como veremos a seguir, a estrutura de modelo também pode ser tratada como um mapa parametrizado das entradas e saídas passadas Z^{t-1} no espaço de modelos:

$$\hat{y}(t|t-1, \theta) = g(\theta, Z^{t-1}), \quad (2.7)$$

ou ainda

$$\mathcal{M} : \theta \in D_\theta \rightarrow g(\theta, Z^{t-1}) \quad (2.8)$$

onde, $\hat{y}(t|t-1, \theta)$ é o preditor um passo à frente, θ é o vetor de parâmetros e

$$Z^{t-1} = [y(1), u(1), \dots, y(t-1), u(t-1)].$$

Tabela 2.1: Estrutura de (2.9) comumente utilizadas.

Polinômios de (2.9)	Nome da estrutura
A	AR
B	FIR
AB	ARX
ABC	ARMAX
AC	ARMA
ABD	ARARX
ABCD	ARARMAX
BF	OE
BFGD	BJ

2.2.1 Estruturas de Modelos Padrão

A tabela 2.1 apresenta estruturas de modelos padrão comumente utilizadas, onde assume-se que $G(q, \theta)$ e $H(q, \theta)$ são racionais e os parâmetros são os coeficientes dos polinômios do numerador e denominador. Deste modo podemos detalhar (2.5) como

$$A(q, \theta)y(t) = \frac{B(q, \theta)}{F(q, \theta)}u(t) + \frac{C(q, \theta)}{D(q, \theta)}e(t) \quad (2.9)$$

onde

$$\begin{aligned} A(q, \theta) &= 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}, \\ B(q, \theta) &= b_1q^{-1} + \dots + b_{n_b}q^{-n_b}, \\ F(q, \theta) &= 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}, \\ C(q, \theta) &= 1 + k_1q^{-1} + \dots + k_{n_k}q^{-n_k}, \\ D(q, \theta) &= 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}, \end{aligned}$$

e

$$\theta = [a_1 \dots a_{n_a} \ b_1 \dots b_{n_b} \ f_1 \dots f_{n_f} \ c_1 \dots c_{n_k} \ d_1 \dots d_{n_d}]^T \quad (2.10)$$

é o vetor de parâmetros.

A estrutura (2.9) é uma estrutura bastante genérica e consegue representar uma grande variedade de modelos. Na maioria das aplicações um ou mais dos polinômios são fixados em 1. A Tabela 2.1 (LJUNG, 1999) apresenta os casos mais comuns para os polinômios de (2.9), onde na primeira coluna estão os polinômios que não serão fixados em 1.

2.3 Identificação por minimização do erro de predição

Além de se escolher uma classe de modelos é necessário encontrar um critério capaz de avaliar a capacidade que diferentes modelos tem em descrever os dados observados. Precisamos escolher qual será o método que relacione essa classe de modelos com o conjunto de dados obtidos em um experimento, ou ainda, como será utilizada a informação contida em Z^N a fim de estimar um valor adequado ao vetor de parâmetros estimados $\hat{\theta}_N$, e consequentemente, um elemento de $M(\hat{\theta}_N)$ do

conjunto \mathcal{M} . Em outras palavras, deve-se determinar um mapeamento dos dados Z^N no conjunto D_θ :

$$Z^N \rightarrow \hat{\theta}_N \in D_\theta.$$

Tal mapeamento é um método de estimação de parâmetros. Existem diferentes métodos que realizam este mapeamento. Neste trabalho usaremos o método de *Minimização do Erro de Predição*.

O erro de predição será definido como

$$\epsilon(t, \theta) \triangleq y(t) - \hat{y}(t|t-1, \theta), \quad (2.11)$$

onde $y(t)$ é a saída do processo e $\hat{y}(t|t-1, \theta)$ é o preditor um passo à frente ótimo. O preditor ótimo é o preditor que tem a menor variância para o erro de predição (2.11). Na teoria de identificação de sistemas lineares é utilizado o preditor um passo à frente ótimo dado por (LJUNG, 1999):

$$\begin{aligned} \hat{y}(t|t-1, \theta) &= H^{-1}(q, \theta)G(q, \theta)u(t) + (1 + H^{-1}(q, \theta))y(t) \\ &= W(q, \theta)z(t). \end{aligned} \quad (2.12)$$

com

$$\begin{aligned} W(q, \theta) &= [W_u(q, \theta) \ W_y(q, \theta)] \triangleq [H^{-1}(q, \theta)G(q, \theta) \ (1 + H^{-1}(q, \theta))] \\ z(t) &\triangleq [u(t) \ y(t)]^T. \end{aligned} \quad (2.13)$$

No caso de sistemas não lineares o preditor ótimo não possui uma forma fechada e é de difícil obtenção. Um preditor que pode ser utilizado para sistemas não lineares é o preditor *modelo de simulação* (LJUNG, 1999) definido por:

Definição 2.3.1. (LJUNG, 1999) *Considere o processo SISO de tempo discreto invariante no tempo*

$$y(t) = F(\psi(t), \nu(t)), \quad (2.14)$$

onde $\psi = [y(t-1), \dots, y(t-n), \dots, u(t), \dots, u(t-m)]$, $\nu(t) = [e(t), \dots, e(t-l)]$ e $F(\cdot, \cdot) : \mathbb{R}^{n+m+1} \times \mathbb{R}^{l+1} \rightarrow \mathbb{R}$, que será identificado através da classe de modelos parametrizada:

$$y(t) = f(\psi(t), \nu(t), \theta). \quad (2.15)$$

O preditor modelo de simulação é definido como:

$$\hat{y}(t, \theta) = f(\psi(t), \mathbf{0}, \theta), \quad (2.16)$$

onde se desconsidera o ruído do processo.

◇

O vetor erro de predição (2.11) pode ser visto como um vetor no \mathbb{R}^N . O “tamanho” deste vetor pode ser medido usando qualquer norma no \mathbb{R}^N , quadrática ou não quadrática. Neste trabalho utilizaremos uma norma quadrática definida como:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} \epsilon^2(t, \theta), \quad (2.17)$$

onde a função $V_N(\theta, Z^N)$, para um conjunto de dados Z^N , é uma função escalar bem definida nos parâmetros θ . Esta função será utilizada para avaliar se o modelo $M(\theta)$ é capaz de “descrever o conjunto de dados”. O estimador $\hat{\theta}_N$ é então definido através da minimização de (2.17)

$$\hat{\theta}_N \triangleq \arg \min_{\theta \in D_\theta} V_N(\theta, Z^N), \quad (2.18)$$

onde “arg min” significa o argumento que resulta no valor mínimo da função.

Note que existe uma equivalência entre $[G(q, \theta) H(q, \theta)]$ e $[W_u(q, \theta) W_y(q, \theta)]$. Devido a esta equivalência o modelo $M(\theta)$ pode, e será, tratado indistintamente a qualquer uma das descrições, (2.5) ou (2.12).

2.3.1 Estimador de Mínimos Quadrados

Dada a equação de diferenças linear;

$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t), \quad (2.19)$$

onde, $\theta = [a_1 \ a_2 \ \cdots \ a_{n_a} \ b_1 \ \cdots \ b_{n_b}]^T$. Nota-se que a relação (2.19) é equivalente a (2.5) com

$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad H(q, \theta) = \frac{1}{A(q)}, \quad (2.20)$$

onde

$$\begin{aligned} A(q, \theta) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \\ B(q, \theta) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}. \end{aligned} \quad (2.21)$$

Observe que o modelo (2.20) é o modelo ARX da Tabela 2.1. Escrevendo o preditor (2.12) para o modelo ARX temos

$$\hat{y}(t|t-1, \theta) = B(q)u(t) + [1 - A(q)]y(t) \quad (2.22)$$

Substituindo (2.21) em (2.22) temos que o preditor pode ser reescrito como

$$\hat{y}(t|t-1, \theta) = \varphi^T(t)\theta, \quad (2.23)$$

onde introduzimos o vetor regressor

$$\varphi(t) \triangleq [-y(t-1) \ \cdots \ -y(t-n_a) \ u(t-1) \ \cdots \ u(t-n_b)]^T. \quad (2.24)$$

O preditor (2.23) é o produto escalar entre o vetor regressor e o vetor de parâmetros, e por isso esse tipo de modelo é chamado de *Regressão Linear*. O estimador de mínimos quadrados utiliza a relação (2.23) para compor o vetor erro de predição. Assim, a partir da relação (2.23) temos que o vetor erro de predição será

$$\epsilon(t, \theta) = y(t) - \varphi^T(t)\theta. \quad (2.25)$$

Substituindo (2.25) na função custo (2.17) obtemos:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} [y(t) - \varphi^T(t)\theta]^2. \quad (2.26)$$

A expressão (2.26) é chamada de *Crítério de Mínimos Quadrados*. A função (2.26) é quadrática nos parâmetros θ e portanto poderá ser minimizada de forma analítica (LJUNG, 1999). Desta forma o estimador de mínimos quadrados $\hat{\theta}_N^{MQ}$ será

$$\hat{\theta}_N^{MQ} = \arg \min_{\theta \in D_\theta} V_N(\theta, Z^N) = \left(\sum_{t=1}^N \varphi(t) \varphi^T(t) \right)^{-1} \sum_{t=1}^N \varphi(t) y(t), \quad (2.27)$$

desde que $\left(\sum_{t=1}^N \varphi(t) \varphi^T(t) \right)$ tenha posto completo.

Podemos definir uma forma matricial para a técnica de identificação por mínimos quadrados. Definindo o vetor coluna $N \times 1$

$$Y_N \triangleq \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix} \quad (2.28)$$

e a matriz $N \times d$

$$\Phi_N \triangleq \begin{bmatrix} \varphi^T(1) \\ \vdots \\ \varphi^T(N) \end{bmatrix} \quad (2.29)$$

com $\varphi(t)$ definido em (2.24), podemos reescrever o critério (2.26) como

$$V_N(\theta, Z^N) = \frac{1}{N} |Y_N - \Phi\theta|^2 = \frac{1}{N} (Y_N - \Phi\theta)^T (Y_N - \Phi\theta). \quad (2.30)$$

Assim o estimador de mínimos quadrados será dado por (AGUIRRE, 2004)

$$\hat{\theta}_N^{MQ} = [\Phi_N^T \Phi_N]^{-1} \Phi_N^T Y_N, \quad (2.31)$$

onde a matriz $[\Phi_N^T \Phi_N]^{-1} \Phi_N^T$ é conhecida como pseudoinversa de Φ_N .

2.4 Incertezas nas estimativas dos parâmetros

A qualidade das estimativas é comumente medida através de suas propriedades estatísticas. Para isto, é necessário que o tamanho N do conjunto de dados seja suficientemente grande de forma que possam ser utilizadas condições assintóticas para as propriedades estatísticas da estimativa.

Definição 2.4.1. (SÖDERSTRÖM; STOICA, 2001)

Sejam θ_0 o valor real do parâmetro a ser estimado e $\hat{\theta}_N$ uma estimativa para θ_0 . A estimativa $\hat{\theta}_N$ é dita *tendenciosa* ou *polarizada* se

$$E[\hat{\theta}_N] \neq \theta_0. \quad (2.32)$$

◇

Definição 2.4.2. (SÖDERSTRÖM; STOICA, 2001)

Seja θ_0 o valor real do parâmetro a ser estimado. A estimativa $\hat{\theta}_N$ é dita *consistente* se

$$\lim_{N \rightarrow \infty} \hat{\theta}_N = \theta_0. \quad (2.33)$$

◇

É possível mostrar, sob certas condições no conjunto de dados Z_N (LJUNG, 1999), que a estimativa $\hat{\theta}_N$ possui uma distribuição assintótica a qual respeita

$$\lim_{N \rightarrow \infty} \hat{\theta}_N = \arg \min_{\theta \in D_\theta} E [\epsilon^2(t, \theta)], \quad (2.34)$$

$$\sqrt{N} (\hat{\theta}_N - \theta_0) \xrightarrow{N \rightarrow \infty} \mathcal{N}(0, P_\theta), \quad (2.35)$$

$$\lim_{N \rightarrow \infty} N(\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T = P_\theta, \quad (2.36)$$

onde $\mathcal{N}(0, P_\theta)$ é uma distribuição normal com média zero e covariância P_θ . A matriz P_θ é a matriz de covariância da distribuição assintótica dada por

$$P_\theta = [I(\theta)]^{-1} \Big|_{\theta=\theta_0}, \quad (2.37)$$

com

$$I(\theta) = \frac{1}{\sigma_e^2} E [\psi(t, \theta) \psi^T(t, \theta)], \quad (2.38)$$

e

$$\psi^T(t, \theta) = \frac{\delta \hat{y}(t|t-1, \theta)}{\delta \theta} \Big|_{\theta=\theta_0}. \quad (2.39)$$

A matriz $I(\theta)$ é chamada de *matriz de informação*. Esta matriz contém informações sobre identificabilidade dos parâmetros θ , assim como da riqueza do conjunto de dados Z^N (BAZANELLA; BOMBOIS; GEVERS, 2012). Esse tema será abordado na sessão 2.6.

Assim, se a estimativa $\hat{\theta}_N$ for uma estimativa assintótica que respeite (2.34-2.36) teremos que a distribuição deste estimador tenderá a

$$(\hat{\theta}_N - \theta_0) P_\theta^{-1} (\hat{\theta}_N - \theta_0) \rightarrow \chi_\alpha^2 \quad \text{quando } N \rightarrow \infty$$

e portanto, o valor real do vetor de parâmetros θ_0 pertence, com probabilidade α , ao elipsoide (GEVERS, 2005):

$$U_\theta = \left\{ \theta \mid (\hat{\theta}_N - \theta_0) P_\theta^{-1} (\hat{\theta}_N - \theta_0) \leq \chi_\alpha^2 \right\}, \quad (2.40)$$

onde χ_α^2 é a distribuição Qui-quadrado com probabilidade α . A Tabela 2.2 apresenta alguns valores de χ^2 para diferentes valores de α e diferentes graus de liberdade, p , onde p equivale ao número de parâmetros que serão identificados.

2.4.1 Propriedades estatísticas do estimador de MQ

Agora analisaremos as propriedades estatísticas do estimador de mínimos quadrados (2.27). Suponha que o conjunto de dados observados foi gerado por

$$y(t) = \varphi^T(t) \theta_0 + \nu(t),$$

ou seja, $\mathcal{S} \in \mathcal{M}$, com $\nu(t)$ uma sequência de variáveis aleatórias com média zero e variância σ_ν^2 e θ_0 o valor real do vetor de parâmetros. Utilizando (2.27) pode-se mostrar que (LJUNG, 1999)

Tabela 2.2: Distribuição χ^2 para alguns valores de probabilidade α e diferentes graus de liberdade n .

p	α					
	90%	95%	97,5%	99%	99,5%	
1	2,7055	3,8415	5,0239	6,6349	7,8794	
2	4,6052	5,9915	7,3778	9,2104	10,5965	
3	6,2514	7,8147	9,3484	11,3449	12,8381	
4	7,7794	9,4877	11,1433	13,2767	14,8602	
5	9,2363	11,0705	12,8325	15,0863	16,7496	
6	10,6446	12,5916	14,4494	16,8119	18,5475	
7	12,0170	14,0671	16,0128	18,4753	20,2777	

$$\begin{aligned} \lim_{N \rightarrow \infty} \hat{\theta}_N^{MQ} - \theta_0 &= \lim_{N \rightarrow \infty} \left(\sum_{t=1}^N \varphi(t) \varphi^T(t) \right)^{-1} \sum_{t=1}^N \varphi(t) \nu(t) \\ &= \bar{E} [\varphi(t) \varphi^T(t)] \bar{E} [\varphi(t) \nu(t)], \end{aligned} \quad (2.41)$$

dados que ν e φ são quasi-estacionários (Definição 2.2.1). Assim para que o estimador de mínimos quadrados seja consistente é necessário que:

1. $\bar{E} [\varphi(t) \varphi^T(t)]$ seja não singular;
2. $\bar{E} [\varphi(t) \nu(t)] = 0$. Este será o caso se

Caso 1: $\{\nu(t)\}$ é ruído branco. Então, $\nu(t)$ não irá depender de tempos passados e então $\bar{E} [\varphi(t) \nu(t)] = 0$; ou

Caso 2: O sinal de entrada $\{u(t)\}$ é independente de $\{\nu(t)\}$ e $n_a = 0$ em (2.24). Assim $\varphi(t)$ contém apenas dados de entrada e então $\bar{E} [\varphi(t) \nu(t)] = 0$.

Como mostrado acima, o método de mínimos quadrados retorna uma estimativa consistente sob certas condições. Em alguns casos a falta de consistência pode ser tolerada, por exemplo, se a razão sinal-ruído for grande a estimativa terá um pequeno desvio (tendência) do valor real (SÖDERSTRÖM; STOICA, 2001). Em outras situações, no entanto, pode ser importante ter uma estimativa consistente. Uma maneira de se obter tal estimativa é através do método das variáveis instrumentais.

2.5 Variáveis instrumentais

Como visto na sessão anterior o estimador de mínimos quadrados possui uma forma fechada o que o torna de fácil implementação. Porém pode acontecer da estimativa ser tendenciosa (polarizada). A ideia do método das variáveis instrumentais é modificar o estimador de mínimos quadrados de modo a eliminar a polarização da estimativa, e ainda mantendo uma forma analítica para o estimador.

Considere novamente que o sistema pode ser descrito pela regressão linear (2.23)

$$\hat{y}(t|t-1, \theta) = \varphi^T(t)\theta, \quad (2.42)$$

e que o conjunto de dados possa ser descrito por

$$y(t) = \varphi^T(t)\theta_0 + \nu(t). \quad (2.43)$$

Considere, também, $\zeta(t)$ um vetor cujos elementos são sinais descorrelacionados do ruído $\nu(t)$. O estimador por variáveis instrumentais é obtido através da solução do seguinte sistema de equações

$$\hat{\theta}_N^{VI} = \underset{\theta \in D_\theta}{\text{sol}} \left\{ \frac{1}{N} \sum_{t=1}^N \zeta(t) [y(t) - \varphi^T(t)\theta] = 0 \right\} \quad (2.44)$$

ou

$$\hat{\theta}_N^{VI} = \left[\sum_{t=1}^N \zeta(t)\varphi^T(t) \right]^{-1} \sum_{t=1}^N \zeta(t)y(t). \quad (2.45)$$

Os elementos de $\zeta(t)$ são chamados de *instrumentos* ou *variáveis instrumentais* e

$$\frac{1}{N} \sum_{t=1}^N \zeta(t) [y(t) - \varphi^T(t)\theta], \quad (2.46)$$

é uma estimativa da correlação entre a variável instrumental e o vetor erro de predição. Assim para que $\hat{\theta}_N^{VI}$ tenda a θ_0 quando $N \rightarrow \infty$ é necessário que

- $\bar{E} [\zeta(t)\varphi^T(t)]$ seja não singular, ou seja a matriz $\zeta(t)$ deve ser correlacionada com o vetor regressor $\varphi(t)$.
- $\bar{E} [\zeta(t)\nu(t)] = \mathbf{0}$, ou seja a matriz $\zeta(t)$ deve ser descorrelacionada com o ruído $\nu(t)$.

Existem diversos instrumentos que podem ser utilizados, provavelmente o mais fácil de se implementar é o instrumento composto por uma outra amostragem do sistema. Ou seja, o conjunto de dados que compõe o vetor regressor $\varphi(t)$ deve ser obtido através de uma amostragem do sistema, já o conjunto de dados que compõe o instrumento de uma outra amostragem. Deste modo o vetor regressor e a variável instrumental se diferem apenas da realização do ruído, fazendo assim $\bar{E} [\zeta(t)\nu_0(t)] = \mathbf{0}$ e $\bar{E} [\zeta(t)\varphi^T(t)] \neq \mathbf{0}$.

2.6 Identificabilidade x Informatividade

Muitos conceitos para o termo “identificável” foram propostos na literatura científica, e essas definições foram evoluindo com o passar dos anos. Estes diferentes conceitos não são equivalentes, e nem mesmo tentam expressar a mesma ideia.

O termo “identificável” foi primeiramente utilizado para descrever a propriedade de um problema de identificação ter um único valor de parâmetro como solução (BAZANELLA; BOMBOIS; GEVERS, 2012). Essa definição prevaleceu até o final dos anos 70, início dos anos 80, onde um grande número de trabalhos foram apresentados para tratar da identificabilidade de sistemas operando em laço fechado.

Na identificação por erro de predição, esse conceito é equivalente à função custo $V_N(\theta, Z^N)$ ter um único mínimo global. Porém, essa propriedade depende de dois elementos diferentes do processo de identificação: a estrutura do modelo proposta e as condições do experimento (o quão informativo é o sinal aplicado ao sistema). Por isso, em meados dos anos 80, foi proposto analisar a identificabilidade de sistemas analisando esses dois elementos separadamente, onde a *identificabilidade* de uma estrutura de modelo é uma propriedade somente da estrutura do modelo e não depende do sinal aplicado, enquanto *informatividade* de um conjunto de dados se refere à capacidade que o experimento e o conjunto de dados têm para discriminar entre dois modelos diferentes. Estes são os conceitos adotados neste trabalho.

2.6.1 Identificabilidade

O termo “identificável” indica a propriedade, a qual se refere exclusivamente à estrutura do modelo, de que, dois valores de parâmetros não podem resultar no mesmo modelo. Identificabilidade (global ou local) é uma propriedade de como o modelo $M(\theta)$ é parametrizado.

Definição 2.6.1. (BAZANELLA; BOMBOIS; GEVERS, 2012) *Uma estrutura de modelos \mathcal{M} é localmente identificável no valor θ_1 se $\exists \delta > 0$ tal que, para todo $\theta \in \|\theta - \theta_1\| \leq \delta$:*

$$W(e^{j\omega}, \theta) = W(e^{j\omega}, \theta_1) \quad \text{para quase todo } \omega \Rightarrow \theta = \theta_1.$$

A estrutura de modelo é globalmente identificável em θ_1 se o mesmo é válido para $\delta \rightarrow \infty$. Por fim, a estrutura de modelo é globalmente identificável se ela é globalmente identificável para quase todos θ_1 .

◇

2.6.2 Identificabilidade das estruturas padrões

Como dito anteriormente, o termo “identificável” indica uma propriedade que se refere exclusivamente à estrutura do modelo. O teorema a seguir apresenta, para as estruturas padrão da Tabela 2.1, quando cada uma será globalmente identificável.

Teorema 1. (LJUNG, 1999) *Considere a estrutura de modelos \mathcal{M} correspondente a (2.9), onde θ é o vetor de parâmetros dado por (2.10) e n_a, n_b, n_d, n_f e n_K são os graus dos polinômios A, B, D, F e K respectivamente. Essa estrutura de modelos é globalmente identificável em θ^* se, e somente se, as premissas (i) a (vi) são satisfeitas.*

- i Não existe fator comum entre $q^{n_a}A(q, \theta^*)$, $q^{n_b}B(q, \theta^*)$ e $q^{n_K}K(q, \theta^*)$.*
- ii Não existe fator comum entre $q^{n_b}B(q, \theta^*)$ e $q^{n_f}F(q, \theta^*)$.*
- iii Não existe fator comum entre $q^{n_K}K(q, \theta^*)$ e $q^{n_d}D(q, \theta^*)$.*
- iv Se $n_a \geq 1$, então não deve existir fator comum entre $q^{n_f}F(q, \theta^*)$ e $q^{n_d}D(q, \theta^*)$.*
- v Se $n_d \geq 1$, então não deve existir fator comum entre $q^{n_a}A(q, \theta^*)$ e $q^{n_b}B(q, \theta^*)$.*
- vi Se $n_f \geq 1$, então não deve existir fator comum entre $q^{n_a}A(q, \theta^*)$ e $q^{n_K}K(q, \theta^*)$.*

2.6.3 Informatividade

O conjunto de dados Z^N é a fonte de informação sobre o sistema “real”, este conjunto de dados deve satisfazer a estrutura de modelos \mathcal{M} escolhida. Como mencionado anteriormente, o conceito de identificabilidade refere-se à injetividade do mapa do espaço de parâmetros para o espaço de funções de transferência. Uma questão similar é se o conjunto de dados Z^N permite distinguir entre dois modelos distintos dentro de um conjunto.

Definição 2.6.2. (LJUNG, 1999) Dizemos que dois modelos $M(\theta_1) = W(q, \theta_1)$ e $M(\theta_2) = W(q, \theta_2)$ são iguais se

$$W(e^{j\omega}, \theta_1) = W(e^{j\omega}, \theta_2), \quad \text{para quase todo } \omega$$

◇

A definição 2.6.1 diz que se a estrutura do modelo é globalmente identificável em um valor θ_1 , então não existe nenhum outro valor para o parâmetro $\theta \neq \theta_1$ que resulte no mesmo preditor $W(q, \theta)$. No entanto, não se pode garantir que dois modelos diferentes pertencentes a um conjunto de modelos \mathcal{M} não produzam o mesmo preditor quando utilizado o mesmo conjunto de dados. Esta exigência adicional faz com que o conjunto de dados carregue informação suficiente para distinguir entre diferentes modelos. A definição para a informatividade de um conjunto de dados com respeito a uma estrutura de modelo é apresentada a seguir

Definição 2.6.3. (BAZANELLA; BOMBOIS; GEVERS, 2012) Um conjunto de dados quasi-estacionário $z(t)$ é dito localmente informativo em $\theta_1 \in D_\theta$ com respeito ao conjunto de modelos $\mathcal{M} = \{M(\theta), \theta \in D_\theta\}$ se $\exists \delta > 0$ tal que, para todo θ em $\|\theta - \theta_1\| \leq \delta$, temos

$$\bar{E} \{ [W(q, \theta) - W(q, \theta_1)] z(t) \}^2 = 0 \quad \Rightarrow \quad W(q, \theta) = W(q, \theta_1) \quad (2.47)$$

O conjunto de dados é dito globalmente informativo em θ_1 se o mesmo é válido para $\delta \rightarrow \infty$. Por fim, o conjunto de dados é dito totalmente informativo se ele for globalmente informativo para todo θ .

◇

Observe que a definição de informatividade de uma estrutura paramétrica deve valer para quaisquer par do vetor de parâmetros $(\theta_1, \theta_2) \in D_\theta$. Porém, note que, se a estrutura de modelo for globalmente identificável em θ_1 , por exemplo, então teremos que $\theta_2 = \theta_1$.

Sob a hipótese de que o sistema real pertence ao conjunto de modelo, $\mathcal{S} \in \mathcal{M}$, os conceitos de identificabilidade e informatividade juntos garantem a unicidade do mínimo do critério de identificação. Como mencionado na Seção 2.4 a matriz de informação (2.38) contém informações sobre identificabilidade dos parâmetros θ , assim como da riqueza do conjunto de dados Z^N . A matriz de informação ser positiva definida implica o sistema ser identificável e informativo (GEVERS et al., 2009). Essa propriedade da matriz de informação será utilizada no capítulo 3 como uma maneira alternativa de se testar a identificabilidade de estruturas de modelos predeterminadas.

O teorema a seguir relaciona a função custo, a matriz de informação e o valor real para o parâmetro θ .

Teorema 2. (BAZANELLA; BOMBOIS; GEVERS, 2012) Dado que as funções de transferência $G(q, \theta)$ e $H(q, \theta)$ da classe de modelos (2.5) sejam da forma

$$\begin{aligned} G(q, \theta) &= \frac{ng_0(q^{-1}) + \theta^T ng_1(q^{-1})}{dg_0(q^{-1}) + \theta^T dg_1(q^{-1})}, \\ H(q, \theta) &= \frac{nh_0(q^{-1}) + \theta^T nh_1(q^{-1})}{dh_0(q^{-1}) + \theta^T dh_1(q^{-1})}, \end{aligned} \quad (2.48)$$

onde $\theta \in D_\theta \subseteq \mathbb{R}^p$, ng_0, dg_0, nh_0 e dh_0 representam polinômios em q^{-1} , enquanto ng_1, dg_1, nh_1 e dh_1 são vetores pertencentes a \mathbb{R}^p formados por polinômios em q^{-1} . Esses polinômios satisfazem as restrições $nh_1(\infty) = 0$, $dh_1(\infty) = 0$, $nh_0(\infty) = 1$ e $dh_0(\infty) = 1$. Deste modo $H(\infty, \theta) = 1$.

Dada também a hipótese 2.2.1, então as três condições a seguir são equivalentes:

1. $V_N(\theta, Z^N)$ possui um único mínimo global em θ_0 ;
2. $I(\theta_0) > 0$;
3. Em θ_0 a estrutura de modelo é localmente identificável e o conjunto de dados é localmente informativo.

◇

2.7 Considerações Finais

Neste capítulo foram apresentados os principais aspectos para a identificação de sistemas lineares. Mesmo este trabalho tendo como um dos objetivos a identificação de sistemas não lineares, como o algoritmo de Ritt nos fornece uma forma de transformar uma estrutura não linear parametrizada em regressões lineares nos parâmetros, não é necessário como pré-requisito o conhecimento em identificação de sistemas não lineares. Uma análise de diferentes métodos de identificação de sistemas não lineares, assim como de diferentes representações matemáticas de estruturas não lineares pode ser encontrado em (AGUIRRE, 2004), (LJUNG, 2006).

3 ÁLGEBRA DIFERENCIAL

3.1 Introdução

Neste capítulo introduziremos alguns conceitos sobre álgebra diferencial necessários para compreender o funcionamento do algoritmo de Ritt. Apresentaremos as definições de anéis e ideais diferenciais, polinômios diferenciais e de conjuntos característicos. A partir destas definições mostraremos como é possível descrever um modelo algébrico através de ideais diferenciais, reduzi-los para um conjunto equivalente de polinômios diferenciais, onde então, de forma simples, será possível investigar se o modelo é identificável ou não.

3.2 Definições

Um campo diferencial, ou anel diferencial, é uma coleção de elementos sobre o qual pode-se realizar operações de adição, subtração, multiplicação, divisão e diferenciação. Adição e multiplicação são comutativas e associativas, multiplicação é comutativa com respeito à adição. Para a operação de diferenciação temos:

$$z = (x + y) \rightarrow \dot{z} = \dot{x} + \dot{y} \quad (3.1)$$

e

$$z = xy \rightarrow \dot{z} = y\dot{x} + x\dot{y}, \quad (3.2)$$

onde $(\dot{\cdot})$ é o operador diferenciação, x e y são indeterminados e um indeterminado é um símbolo que não representa nada além de si mesmo. Em particular, não designa uma constante, ou um parâmetro do problema. Indeterminados são usados para construir objetos formais, como polinômios, séries, ou elementos de um grupo livre. Ou seja, é uma variável que não é conhecida (RITT, 1950), (KOLCHIN, 1973).

Um exemplo para um indeterminado é, seja x um indeterminado tal que $2 + 3x = ax + b$, onde $a, b \in \mathbb{Q}$. Assim, x não é conhecido. Porém, se $a = 2$ e $b = 3$, x não será mais um indeterminado, já que seu valor pode ser determinado. Um indeterminado cuja derivada é zero é chamado de *constante*.

Neste trabalho trataremos apenas de sinais temporais, assim todas as derivadas serão com respeito ao tempo, ou seja, o operador $(\dot{\cdot})$ representa $\frac{d(\cdot)}{dt}$. Quando for apresentado o algoritmo de Ritt para sistemas de tempo discreto o operador $(\dot{\cdot})$ será o atraso temporal do sinal, isto é $\dot{x} = q^{-1}x$ (seção 3.4).

Definição 3.2.1. (Ritt1950)

Um anel diferencial $\mathbb{F}\{x_1, x_2, \dots, x_n\}$ é o conjunto de todos polinômios formados pelo conjunto de indeterminados x_1, x_2, \dots, x_n , e suas derivadas

$$\dot{x}_i, \ddot{x}_i, \dots, x_i^{(\nu)}, \dots$$

com coeficientes no campo \mathbb{F} , onde ν é a ordem da derivada de x_i e $i = 1, 2, \dots, n$. Um polinômio diferencial f é um elemento de $\mathbb{F}\{x_1, x_2, \dots, x_n\}$.

◇

Um exemplo de campo diferencial, ou anel diferencial, é $\mathbb{R}\{x_1, x_2\}$ formado pelos indeterminados x_1, x_2 e suas derivadas com coeficientes reais. Já um polinômio diferencial f , pertencente a $\mathbb{R}\{x_1, x_2\}$, é, por exemplo,

$$f = x_1x_2^2 + x_1^2\dot{x}_2. \quad (3.3)$$

Definição 3.2.2. (Ritt1950)

A ordem de um polinômio diferencial f com respeito a um indeterminado x_i é a derivada de maior ordem de x_i em f . Se f não depende de x_i , f é dito de ordem zero em x_i .

◇

Definição 3.2.3. (Ritt1950)

O grau de um polinômio diferencial f com respeito a um indeterminado x_i é o maior expoente de x_i em f .

◇

Como mencionado anteriormente estamos tratando sinais temporais, assim as variáveis tratadas são funções do tempo, ou seja, $x_i = x_i(t)$. Deste modo, uma solução $f(x_1, x_2) = 0$ de um polinômio diferencial nas variáveis $x_1(t)$ e $x_2(t)$ é definida como os sinais x_1 e x_2 tais que o polinômio diferencial f é identicamente nulo.

Dados f e g polinômios diferenciais, um ideal diferencial é um subconjunto I de $\mathbb{F}\{x_1, x_2, \dots, x_m\}$ tal que

1. $f + g \in I$ para todo $f, g \in I$,
2. $fg \in I$ para todo $f \in I$ e $g \in \mathbb{F}\{x_1, x_2, \dots, x_m\}$,
3. $f^{(n)} \in I$ para todo $f \in I$.

Ou seja, um ideal diferencial consiste de todos os polinômios diferenciais formados por elementos de $\mathbb{F}\{x_1, x_2, \dots, x_m\}$ através de multiplicação, adição e diferenciação de polinômios.

Um exemplo de ideal diferencial em $\mathbb{R}\{x_1, x_2\}$ é

$$\left\{ \sum_{i \geq 1} \alpha_i x_1^i + \beta_i x_1^{(i)} + \gamma_i x_2^i + \mu_i x_2^{(i)} : \alpha_i, \beta_i, \gamma_i, \mu_i \in \mathbb{R} \right\}$$

Como mencionado acima, estamos tratando polinômios diferenciais compostos por indeterminados e suas derivadas. Com o objetivo de ordenar essas variáveis define-se um *ranking*.

Definição 3.2.4. (MARGARIA et al., 2001)

Um ranking para os indeterminados (x_1, \dots, x_n) é um ordenamento de todos os indeterminados do conjunto e suas derivadas $x_i^{(m)}$ tal que

1. $x_i^{(m)} \prec x_i^{(m+k)}$, e
2. $x_i^{(m)} \prec x_j^{(l)} \rightarrow x_i^{(m+k)} \prec x_j^{(l+k)}$

para todos $i, j = 1, \dots, n$ e para $\forall m, l, k \in \mathbb{N}$, onde \prec significa ordenado abaixo de.

◇

Um exemplo de ranking para o conjunto de polinômios pertencentes ao anel $\mathbb{R}\{x_1, x_2\}$ é

$$x_1 \prec x_2 \prec \dot{x}_1 \prec \dot{x}_2 \prec \ddot{x}_1 \prec \ddot{x}_2 \prec \dots, \quad (3.4)$$

outro exemplo é

$$x_1 \prec \dot{x}_1 \prec \ddot{x}_1 \prec \dots \prec x_2 \prec \dot{x}_2 \prec \ddot{x}_2 \prec \dots.$$

O líder l_f de um polinômio diferencial f é o elemento, pertencente a f , de mais alto rank. O Separante S_f é a derivada parcial de f com respeito ao líder l_f . O inicial I_f é o coeficiente de maior grau do líder de f . O grau de um indeterminado x em um polinômio diferencial f será denotado por $\deg_x f$. Para o exemplo (3.3) segundo o ranking (3.4) temos que: $l_f = \dot{x}_2$, $S_f = x_1^2$, $I_f = x_1^2$ e $\deg_{l_f} f = 1$.

As definições abaixo relacionam dois polinômios diferenciais entre si, assim como um polinômio diferencial com um conjunto de polinômios.

Definição 3.2.5. (MARGARIA et al., 2001)

Sejam f e g dois polinômios diferenciais com líderes l_f e l_g , respectivamente.

1. f é dito ordenado abaixo de g , denotado por $f \prec g$, se $l_f \prec l_g$ ou se $l_f = l_g$ e $\deg_{l_f} f < \deg_{l_f} g$. Se $l_f = l_g$ e $\deg_{l_f} f = \deg_{l_f} g$ dizemos que f e g têm o mesmo rank, que é anotado como $f \sim g$.
2. g é dito reduzido com respeito a f se não existir derivada de l_f em g e se $\deg_{l_f} g < \deg_{l_f} f$.
3. g é dito reduzido com respeito ao conjunto de polinômios diferenciais $\mathcal{F} = \{f_1, \dots, f_n\}$ se g for reduzido com respeito a f_i para todo $i = 1, \dots, n$.
4. Um conjunto de polinômios diferenciais $\mathcal{F} = \{f_1, \dots, f_n\}$ é chamado autoreduzido se cada elemento de \mathcal{F} for reduzido com respeito a todos os outros, ou seja, f_i é reduzido com respeito a f_j para todo $i \neq j \forall i, j = 1, \dots, n$. Se os polinômios em um conjunto autoreduzido \mathcal{F} estiverem ordenados de menor rank para maior rank, o conjunto é dito "ordenado".

◇

A partir das definições 3.2.5 pode-se mostrar que:

Lema 3.2.1. (LYZELL et al., 2011)

Sejam f e g dois polinômios diferenciais. Se $f \prec g$ com respeito a algum ranking, então f é reduzido com respeito a g , segundo o mesmo ranking.

Prova 1. Se $f \prec g$, temos que

$$l_f \prec l_g \quad (3.5)$$

ou

$$l_f = l_g \text{ e } \deg_{l_f} f < \deg_{l_g} g \quad (3.6)$$

Se f e g satisfazem (3.5) não existirá derivada de l_g em f , já que, pela Definição 3.2.4, $l_f \prec l_g \prec l_g^{(m)}$. Assim f será reduzido com respeito a g . Por outro lado, se f e g satisfazem (3.6), como $f \prec g$ temos que $\deg_{l_f=l_g} f < \deg_{l_f=l_g} g$, assim f será reduzido com respeito a g .

◇

Vale notar que $f \prec g$ é uma condição suficiente para que f seja reduzido com respeito a g , porém não necessária. É fácil de se observar, dado o ranking (3.4), e os polinômios $f = x_2$ e $g = x_1$, que $f \not\prec g$, já que $l_g \prec l_f$. Porém, como f não depende de g , ou seja, os indeterminados que formam f não estão presentes em g , temos que f será reduzido com respeito a g , e mais, g também será reduzido com respeito a f , ou seja, g e f serão autoreduzidos.

Se um polinômio diferencial f não for reduzido com respeito a um segundo polinômio g , pode-se executar uma série de *pseudo-reduções* ou *pseudo-divisões* sobre este polinômio com o objetivo de gerar um novo polinômio, equivalente ao primeiro, que seja reduzido com respeito a g . Essa *pseudo-divisão* ocorre através dos seguintes lemas.

Lema 3.2.2. (LYZELL et al., 2009)

Sejam f e g polinômios na variável x com graus m e n , respectivamente, tais que:

$$f = a_m x^m + \dots + a_0$$

e

$$g = b_n x^n + \dots + b_0,$$

onde $m \geq n$. Então existem polinômios $Q \neq 0$, \bar{Q} e R que satisfazem

$$Qf = \bar{Q}g + R, \quad (3.7)$$

onde $\deg_x R < n$.

◇

Lema 3.2.3. (LYZELL et al., 2009)

Seja g um polinômio com líder l_g e seja f um polinômio que contém $l_g^{(\sigma)}$ para algum $\sigma \in \mathbb{N}$. Então existem polinômios $Q \neq 0$, \bar{Q} e R que satisfazem

$$Qf = \bar{Q}g^{(\sigma)} + R, \quad (3.8)$$

tal que:

1. R não contém $l_g^{(\sigma)}$.
2. Toda solução de $f \equiv 0$, $g \equiv 0$ também é solução de $R \equiv 0$, $g \equiv 0$.

3. Toda solução de $R \equiv 0, g \equiv 0$ com $Q \neq 0$ também é solução de $f \equiv 0, g \equiv 0$.

◇

Lema 3.2.4. (LYZELL et al., 2009)

Sejam f e g polinômios tais que f não é reduzido com respeito a g . Então existem polinômios R e Q , tal que:

1. R é reduzido com respeito a g .
2. Toda solução de $f \equiv 0, g \equiv 0$ também é solução de $R \equiv 0, g \equiv 0$.
3. Toda solução de $R \equiv 0, g \equiv 0$ com $Q \neq 0$ também é solução de $f \equiv 0, g \equiv 0$.

◇

As provas dos Lemas 3.2.2, 3.2.3 e 3.2.4 podem ser encontradas em (LYZELL et al., 2011).

Seja $\mathcal{F} = \{f_1, \dots, f_n\}$ um conjunto autoreduzido, ordenado, de polinômios diferenciais, e g um polinômio diferencial não reduzido com respeito a \mathcal{F} . Aplicando os Lemas 3.2.2, 3.2.3 e 3.2.4 com o objetivo de reduzir g com respeito a todos $f_j, j = 1, \dots, n$, o polinômio resultante será chamado de pseudo-resíduo de g com respeito ao conjunto autoreduzido \mathcal{F} .

Definição 3.2.6. (MARGARIA et al., 2001)

Sejam dois conjuntos de polinômios diferenciais, autoreduzidos, $\mathcal{A} = \{A_1, \dots, A_m\}$ e $\mathcal{B} = \{B_1, \dots, B_n\}$ ordenados de forma crescente segundo um certo ranking, ou seja, $A_1 \prec A_2 \prec \dots \prec A_m$ e $B_1 \prec B_2 \prec \dots \prec B_n$.

\mathcal{A} é de menor ordem que \mathcal{B} se existe um inteiro k tal que $1 \leq k \leq \min(m, n)$ que satisfaz

$$\begin{aligned} A_j &\sim B_j, & j = 1, \dots, k-1 \\ A_k &\sim B_k, \end{aligned}$$

ou, se $m > n$ e $A_j \sim B_j$ para todo $j = 1, \dots, n$.

◇

Seja \mathcal{F} um conjunto ordenado, finito ou infinito, de polinômios diferenciais. É possível, através dos Lemas 3.2.2, 3.2.3 e 3.2.4 formar um subconjunto \mathcal{C} de menor ordem que \mathcal{F} .

Definição 3.2.7. *Conjunto Característico* (MARGARIA et al., 2001)

Conjunto característico \mathcal{C} é um subconjunto autoreduzido de \mathcal{F} , finito, tal que nenhum outro subconjunto de \mathcal{F} tenha ordem menor que \mathcal{C} .

◇

Observe que, para dado *ranking*, o conjunto característico de um ideal não é único. Porém, qualquer conjunto característico de um certo ideal diferencial, para o mesmo *ranking*, deve ser de mesma ordem (SACCOMANI; AUDOLY; D'ANGIÒ, 2003).

Os conceitos de ideais diferenciais e conjunto característico são uma extensão natural de conceitos análogos dentro da álgebra não diferencial (linear) (RITT, 1950). Aqui, o que são ideais diferenciais dentro de campos diferenciais, podem ser considerados como subespaços dentro de espaços vetoriais. Já os conjuntos característicos de um ideal são as bases de um espaço vetorial. A particularidade do conjunto característico é que ele resume toda informação contida em um ideal diferencial em um número finito de polinômios, ou ainda, a partir de combinações dos polinômios pertencentes ao conjunto característico é possível formar todos os polinômios diferenciais pertencentes ao ideal. Analogamente, a partir de combinações lineares dos vetores que formam a base de um espaço vetorial é possível gerar todo aquele espaço vetorial.

Assim como na álgebra não diferencial, qualquer solução de um determinado sistema também é solução da combinação linear com qualquer outro sistema dentro daquele espaço vetorial. Na álgebra diferencial, qualquer solução que satisfaça $f = 0$, com f um polinômio diferencial pertencente a um ideal, também irá satisfazer a equação obtida através da diferenciação, soma, multiplicação por elementos do ideal.

3.3 Algoritmo de Ritt

O algoritmo de Ritt consiste em uma eliminação sistemática de variáveis em equações diferenciais polinomiais. A partir de um conjunto inicial de equações diferenciais pertencentes a um ideal diferencial e dado um ranking para as variáveis, a cada iteração do algoritmo uma nova equação é obtida cuja solução é exatamente a mesma da anterior. Cada nova equação obtida terá seu grau e ordem reduzidos com relação à anterior. Ao término do algoritmo teremos um conjunto característico para o sistema. O algoritmo de Ritt utiliza os lemas 3.2.2, 3.2.3 e 3.2.4 para reduzir os polinômios diferenciais pertencentes a um ideal, preservando a solução do polinômio original.

A ideia do algoritmo de Ritt é associar ideais diferenciais a modelos no espaço de estados. Por exemplo, considere o modelo dado por

$$\begin{aligned} \dot{x}_1 &= \theta_1 x_1^2 + \theta_2 x_2, \\ \dot{x}_2 &= \theta_3 x_1^2 + \theta_4 x_1 x_2, \\ y &= x_1 \end{aligned} \tag{3.9}$$

O modelo descrito por (3.9) corresponde ao ideal em $\mathbb{R}\{x_1, x_2, y\}$ gerado pelos polinômios diferenciais

$$\begin{aligned} f_1 &= \dot{x}_1 - \theta_1 x_1^2 - \theta_2 x_2, \\ f_2 &= \dot{x}_2 - \theta_3 x_1^2 - \theta_4 x_1 x_2, \\ f_3 &= y - x_1 \end{aligned} \tag{3.10}$$

onde $\theta_i \in \mathbb{R}$.

De forma geral,

$$\begin{aligned}\dot{x} &= f(x(t), u(t), \theta), \\ y &= h(x, \theta),\end{aligned}\tag{3.11}$$

onde $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$ representa o estado do sistema, $u(t) \in \mathbb{R}^m, y(t) \in \mathbb{R}^d$ são, respectivamente, os sinais de entrada e saída medidos, $\theta = [\theta_1, \dots, \theta_p]^T \in \mathbb{R}^p$ é o vetor de parâmetros usados para descrever o sistema, $f = [f_1, \dots, f_n]^T, f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ e $h = [h_1, \dots, h_d]^T, h : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$.

Os parâmetros θ serão considerados invariantes no tempo, ou seja

$$\dot{\theta}_i = 0,\tag{3.12}$$

onde a restrição (3.12) é o que distingue o parâmetro θ dos estados x .

O modelo do sistema será dado por um ideal diferencial. O ideal diferencial considerado será dependente do anel diferencial escolhido, ou seja, de quais elementos serão considerados como coeficientes e quais serão indeterminados. O algoritmo de Ritt utiliza o ideal

$$I_1 = \left\{ \dot{x}_1 - f_1, \dots, \dot{x}_n - f_n, y_1 - h_1, \dots, y_d - h_d, \dot{\theta}_1, \dots, \dot{\theta}_p \right\},\tag{3.13}$$

formado pelos polinômios diferenciais pertencentes ao anel $\mathbb{R}\{x, y, u, \theta\}$, onde todos os estados, sinais de entrada e saída e parâmetros são considerados como indeterminados. Deste modo, dado, por exemplo, o ranking

$$\begin{aligned}u &\prec y \prec \theta \prec x : \\ i < j &\Rightarrow u_i^{(\sigma)} \prec u_j^{(\mu)}, y_i^{(\sigma)} \prec y_j^{(\mu)}, \theta_i^{(\sigma)} \prec \theta_j^{(\mu)}, x_i^{(\sigma)} \prec x_j^{(\mu)} \quad \forall \sigma, \mu \in \mathbb{N},\end{aligned}\tag{3.14}$$

o conjunto característico terá a seguinte forma (LJUNG; GLAD, 1994):

$$\begin{aligned}A_1(u, y), \dots, A_d(u, y), \\ B_1(u, y, \theta_1), B_2(u, y, \theta_1, \theta_2), \dots, B_p(u, y, \theta_1, \theta_2, \dots, \theta_p) \\ C_1(u, y, \theta, x_1), \dots, C_n(u, y, \theta, x_1, \dots, x_m),\end{aligned}\tag{3.15}$$

onde, segundo (LJUNG; GLAD, 1994) os polinômios B_i enquadrar-se-ão em um dos três casos descritos abaixo.

Caso 1. Algum $B_i = \dot{\theta}_i$

Caso 2. Para todos $B_i, \deg_{\theta_i} B_i = 1$ e ordem 0 em θ_i

Caso 3. Todos B_i têm ordem 0 em θ_i , e pora algum $B_j, \deg_{\theta_j} B_j > 1$

Note que se os polinômios B_i forem da forma do Caso 2 a matriz formada por estes polinômios será diagonal em θ , já que (3.15) é um conjunto autoreduzido.

Como veremos a seguir esses três casos correspondem a não identificabilidade, identificabilidade global e identificabilidade local da classe de modelos, respectivamente. Os polinômios A_i não dependem dos parâmetros nem do estado do modelo. Assim através destes polinômios podemos testar se a estrutura do modelo utilizada é capaz de descrever os sinais conhecidos, (u, y) , sem precisar estimar qualquer parâmetro do modelo (LJUNG; GLAD, 1994).

Para que possamos escrever o ideal da forma (3.13) é necessário que as funções f_i e h_i sejam funções racionais em x, y, u, θ e em suas derivadas. Essa será uma exigência para as classes de modelos tratadas, mas é importante notar que equações bastante genéricas podem ser levadas a este formato, como podemos observar no exemplo a seguir.

Exemplo 3.3.1. *Dado*

$$x = \sin y \tag{3.16}$$

tomando a derivada de (3.16) e elevando ao quadrado temos

$$\dot{x}^2 = \dot{y}^2(1 - x^2),$$

que é uma expressão polinomial em x e y .

•

Como mencionado anteriormente, o algoritmo de Ritt tem como entrada o anel (3.13) gerado pelo conjunto de polinômios diferenciais

$$\mathcal{I} = \left\{ \dot{x}_1 - f_1, \dots, \dot{x}_n - f_n, y_1 - h_1, \dots, y_d - h_d, \dot{\theta}_1, \dots, \dot{\theta}_p \right\},$$

ou, reescrevendo de forma a simplificar a notação,

$$\mathcal{I} = \{i_1, i_2, \dots, i_m\}$$

com $m = n + d + p$.

Já como saída, o algoritmo de Ritt terá um novo conjunto \mathcal{I} característico, da forma de (3.15) e mais um conjunto $\mathcal{G} = \{g_1, \dots, g_q\}$ que contém informações sobre *separantes* e *iniciais* assim como os polinômios resultantes das *Pseudo-Divisões*. Introduzimos, então, o conceito de solução não degenerada como a solução dos polinômios pertencentes a \mathcal{G} tal que

$$g_i \neq 0 \quad i = 1, \dots, q.$$

Esta condição pode ser interpretada como uma condição de persistência de excitação dos sinais de entrada.

A seguir apresentamos o algoritmo de Ritt de forma sistemática.

Algoritmo 1: Algoritmo de Ritt

```

RANK; % Qualquer ranking seguindo Definição 3.2.4;
 $\mathcal{I} = \{i_1, i_2, \dots, i_m\}$  %  $\mathcal{I}$  ordenado segundo RANK  $i_1 \prec i_2 \prec \dots \prec i_m$ ;
 $\mathcal{G} = \{\emptyset\}$ ; % Inicialização  $\mathcal{G}$ ;
 $\mathcal{A} = \{i_1\}$ ; % Inicialização  $\mathcal{A}$ ;
while  $\mathcal{I} \setminus \mathcal{A} \neq \{\emptyset\}$  do
  % onde  $\setminus$  é a diferença entre os conjuntos;
  for  $j=2, m$  do
    if  $\mathcal{A} \cup \{i_j\} = \text{autoreduzido}$  then
      |  $\mathcal{A} \leftarrow \mathcal{A} \cup \{i_j\}$ ; %  $\mathcal{A}$  é um conjunto autoreduzido de  $\mathcal{I}$ 
    end
  end

   $Q^{i_k} = \sum_{j=0}^{\sigma} \bar{Q}_j A_k^{(\sigma)}$  +  $R$ ; %  $\sigma$  é a ordem do líder de  $i_k$ ;
  %  $i_k$  é o polinômio de maior rank  $\in \mathcal{I} \setminus \mathcal{A}$ ;
  %  $A_k$  é o polinômio de maior rank  $\in \mathcal{A}$  não;
  % reduzido com respeito a  $i_k$ ;

   $\mathcal{G} \leftarrow \mathcal{G} \cup \{Q\}$  % Atualiza  $\mathcal{G}$ ;
  if  $R = 0$  then
    |  $\mathcal{I} \leftarrow \mathcal{I} \setminus \{i_k\}$  % Atualiza  $\mathcal{I}$  retirando o polinômio  $i_k$ ;
  else
    |  $\mathcal{I} \leftarrow (\mathcal{I} \setminus \{i_k\}) \cup \{R\}$  % Atualiza  $\mathcal{I}$  substituindo  $i_k$  por  $R$ ;
  end
end

 $\mathcal{G} \leftarrow \mathcal{G} \cup \{S_{A_j}\} \forall A_j \in \mathcal{A}$ ; % Separante ;
 $\mathcal{G} \leftarrow \mathcal{G} \cup \{I_{A_j}\} \forall A_j \in \mathcal{A}$ ; % Inicial ;
Return( $\mathcal{G}$ ) % Conjunto  $\mathcal{G}$  ;
Return( $\mathcal{I}$ ) % Novo Conjunto  $\mathcal{I}$  característico ;

```

Teorema 3. (LYZELL et al., 2011)

O Algoritmo 1 chega a um conjunto característico após um número finito de iterações.

◇

Teorema 4. (LYZELL et al., 2011)

Toda solução do conjunto inicial de polinômios diferenciais \mathcal{I} no Algoritmo 1 é também solução do conjunto final. Toda solução para o conjunto final na qual os polinômios em \mathcal{G} são não nulos é também solução do conjunto inicial de polinômios.

◇

Como mencionado anteriormente, o conjunto característico formado ao final do algoritmo será triangular nos parâmetros θ e terá a forma (3.15). Através dos Teoremas 5, 6 e 7, abaixo, poderemos determinar se a estrutura do modelo é identificável, e ainda, com os polinômios pertencentes ao conjunto \mathcal{G} fazer um estudo sobre a persistência de excitação.

Teorema 5. (LJUNG; GLAD, 1994)

Dado que os polinômios em \mathcal{G} são não nulos, se o conjunto característico (3.15) tem a forma do Caso 1, então o modelo é não identificável.

◇

Teorema 6. (LJUNG; GLAD, 1994)

Dado que os polinômios em \mathcal{G} são não nulos, e seja θ_* tal que exista alguma solução não-degenerada de (y, u, θ_*, x) . Então se o conjunto característico (3.15) tem a forma do Caso 2, o modelo é globalmente identificável em θ_* .

◇

Teorema 7. (LJUNG; GLAD, 1994)

Dado que os polinômios em \mathcal{G} são não nulos, e seja θ_* tal que exista alguma solução não-degenerada de (y, u, θ_*, x) . Então se o conjunto característico (3.15) tem a forma do Caso 3, o modelo é localmente identificável em θ_* .

◇

Os Teoremas 5, 6 e 7 mostram que a estrutura do modelo é globalmente identificável se e somente se podemos reescrever o conjunto inicial de equações do modelo como regressões lineares, parâmetro por parâmetro, nas variáveis conhecidas:

$$Z_i(u, y) + P_i(u, y)\theta_i = 0 \quad i = 1, \dots, p, \quad (3.17)$$

onde $Z_i(u, y)$ e $P_i(u, y)$ são polinômios diferenciais em u e y . Assim, como visto no capítulo 2, podemos facilmente identificar estruturas desse tipo utilizando os métodos dos mínimos quadrados ou variáveis instrumentais. A utilização dessas expressões reduzidas para a identificação de modelos de tempo discreto utilizando os métodos de mínimos quadrados e variáveis instrumentais será apresentada no capítulo 4.

Agora apresentaremos um exemplo da utilização do algoritmo 1 para o estudo da identificabilidade de uma classe de modelos.

Exemplo 3.3.1. “Goodwin’s Napkin”, (LJUNG; GLAD, 1994)

Considere a equação diferencial homogênea

$$f_1 = \ddot{y} + 2\theta\dot{y} + \theta^2y = 0. \quad (3.18)$$

A equação (3.18) apresenta um único parâmetro desconhecido θ . Com apenas uma simples manipulação algébrica é possível eliminar a não linearidade em θ . Se multiplicarmos (3.18) por \dot{y} e subtrairmos por

$$y \frac{df_1}{dt} = y(y^{(3)} + 2\theta\ddot{y} + \theta^2\dot{y}) \quad (3.19)$$

obtemos

$$(\dot{y}\ddot{y} - yy^{(3)}) + 2\theta(\dot{y}^2 - y\ddot{y}) = 0. \quad (3.20)$$

A equação (3.20) é linear em θ , assim, temos que a estrutura (3.18) é globalmente identificável.

Agora aplicaremos o algoritmo de Ritt para reduzir o sistema. Primeiramente é necessário sinalizar o parâmetro θ como sendo invariante no tempo, deste modo teremos o seguinte conjunto de equações diferenciais:

$$\begin{aligned} f_1 &= \ddot{y} + 2\theta\dot{y} + \theta^2 y = 0, \\ f_2 &= \dot{\theta} = 0 \end{aligned} \quad (3.21)$$

Em seguida, define-se o conjunto $\mathcal{F} = \{f_1, f_2\}$ e o ranking

$$y \prec \dot{y} \prec \ddot{y} \prec \dots \prec \theta \prec \dot{\theta}$$

Temos, então, que $l_{f_1} = \theta$ e $l_{f_2} = \dot{\theta}$, portanto f_1 é reduzido com respeito a f_2 mas f_2 não é reduzido com respeito a f_1 . (Definição 3.2.5). Note que o conjunto \mathcal{F} não forma um conjunto autoreduzido, então devemos executar o algoritmo de Ritt para encontrar um conjunto autoreduzido.

1. O conjunto autoreduzido \mathcal{A} de \mathcal{F} será $\mathcal{A} = \{f_1\}$.
2. $\mathcal{F} \setminus \mathcal{A} = \{f_2\} \neq \{\emptyset\}$
3. Utilizando Lema 3.2.3 reduzimos f_2 com respeito a f_1 ;

$$\begin{aligned} Q_1 f_2 &= \bar{Q}_1 \dot{f}_1 + R_1 \\ Q_1 \dot{\theta} &= \bar{Q}_1 (y^{(3)} + 2\theta\ddot{y} + 2\dot{\theta}(\dot{y} + \theta y) + \theta^2 \dot{y}) + R_1 \end{aligned}$$

tomando $Q_1 = 2(\dot{y} + \theta y)$ e $\bar{Q}_1 = 1$ temos

$$R_1 = -y^{(3)} - 2\theta\ddot{y} - \theta^2 \dot{y}$$

Como o pseudo-resíduo R_1 não é reduzido com respeito a f_1 , mais uma pseudo-redução é necessária

$$\begin{aligned} Q_2 R_1 &= \bar{Q}_2 f_1 + R_2 \\ Q_2 (-y^{(3)} - 2\theta\ddot{y} - \theta^2 \dot{y}) &= \bar{Q}_2 (\ddot{y} + 2\theta\dot{y} + \theta^2 y) + R_2 \end{aligned}$$

tomando $Q_2 = -y$ e $\bar{Q}_2 = \dot{y}$ temos

$$R_2 = yy^{(3)} - \dot{y}\ddot{y} + 2\theta(\dot{y}y - \dot{y}^2), \quad (3.22)$$

que é reduzido com respeito a f_1 .

4. $\mathcal{G} \rightarrow \mathcal{G} \cup \{Q_1, Q_2\}$.
5. $\mathcal{F} \rightarrow (\mathcal{F} \setminus \{f_2\}) \cup \{R_2\}$. Agora o novo conjunto \mathcal{F} será $\mathcal{F} = \{R_2, f_1\}$.
6. \mathcal{F} não é autoreduzido, $\mathcal{F} \setminus \mathcal{A} = \{f_1\} \neq \{\emptyset\}$, já que f_1 não é reduzido com respeito a R_2 , assim $\mathcal{A} = \{R_2\}$.

7. Reduzimos, então, f_1 com respeito a R_2 (Lema 3.2.2)

$$\begin{aligned} Q_3 f_1 &= \bar{Q}_3 R_2 + R_3 \\ Q_3(\ddot{y} + 2\theta\dot{y} + \theta^2 y) &= \bar{Q}_3(y y^{(3)} - \dot{y}\dot{y} + 2\theta(\dot{y}y - \dot{y}^2)) + R_3 \end{aligned}$$

tomando $Q_3 = 4(\dot{y}y - \dot{y}^2)^2$ e $\bar{Q}_3 = (y\theta - y(y y^{(3)} - \dot{y}\dot{y}) + 4\dot{y}(\dot{y}y - \dot{y}^2))$ temos

$$R_3 = y^{(3)2} y^3 - 6y^{(3)} \dot{y}\dot{y}y^2 + 4y^{(3)} \dot{y}^3 y + 4\dot{y}^3 y^2 - 3\dot{y}^2 \dot{y}^2 y \quad (3.23)$$

que é reduzido com respeito a R_2 .

8. $\mathcal{G} \rightarrow \mathcal{G} \cup \{Q_3\}$.

9. $\mathcal{F} \rightarrow (\mathcal{F} \setminus \{f_1\}) \cup \{R_3\}$. Agora o novo conjunto \mathcal{F} será $\mathcal{F} = \{R_3, R_2\}$.

10. R_3 não depende de θ , logo é reduzido com respeito a R_2 e como líder de R_3 , $l_{R_3} = y^{(3)}$ e $\deg_{y^{(3)}} R_2 < \deg_{y^{(3)}} R_3$, R_2 também é reduzido com respeito a R_3 . Portanto, o novo conjunto \mathcal{F} é autoreduzido.

11. $\mathcal{F} \setminus \mathcal{A} = \{\emptyset\}$, Fim do algoritmo.

A análise mais detalhada (identificabilidade e informatividade) para esse sistema será apresentada no Capítulo 4.

•

3.3.1 Ritt Forma Alternativa

A escolha do ideal é crucial para a construção do conjunto característico e é de grande impacto no número e na complexidade das pseudo-reduções necessárias para sua construção. O algoritmo de Ritt utiliza o ideal (3.13), porém poderíamos utilizar o ideal

$$I_2 = \{\dot{x}_1 - f_1, \dots, \dot{x}_n - f_n, y_1 - h_1, \dots, y_d - h_d\}, \quad (3.24)$$

formado pelos polinômios diferenciais pertencentes ao anel $\mathbb{R}\{x, y, u\}$, onde os parâmetros do sistema são os coeficientes dos polinômios diferenciais formados pelos estados, sinal de entrada e saída do modelo e suas derivadas. Com este ideal o conjunto característico, dado, por exemplo, o ranking,

$$\begin{aligned} u &\prec y \prec x : \\ i < j &\Rightarrow u_i^{(\sigma)} \prec u_j^{(\mu)}, y_i^{(\sigma)} \prec y_j^{(\mu)}, x_i^{(\sigma)} \prec x_j^{(\mu)} \quad \forall \sigma, \mu \in \mathbb{N} \end{aligned} \quad (3.25)$$

terá a seguinte forma: (AUDOLY et al., 2001)

$$\begin{aligned} A_1(u, y, \theta), \dots, A_d(u, y, \theta), \\ C_1(u, y, \theta, x_1), \dots, C_n(u, y, \theta, x_1, x_2, \dots, x_n), \end{aligned} \quad (3.26)$$

formado por polinômios diferenciais em u e y e suas derivadas, com coeficientes polinômios algébricos em θ .

Tanto em (3.15) quanto em (3.26) os líderes dos polinômios A_i são derivadas de y_j , onde cada y_j é líder de exatamente um A_i . Em (3.15) os polinômios B_i

são polinômios diferenciais, ou contendo $\dot{\theta}_i$ (Caso 1) ou polinômios sem nenhuma derivada dos parâmetros θ (Caso 2 e Caso 3)(LJUNG; GLAD, 1994).

A utilização do ideal I_2 reduz o número e a complexidade das pseudo-reduções necessárias para se chegar ao conjunto característico do sistema. Essa redução na complexidade das pseudo-reduções se deve ao fato que, ao utilizar um anel $\mathbb{R}\{x, y, u\}$, o ideal será dado por (3.24) que possui p polinômios diferenciais a menos que o ideal (3.13), formado pelo anel $\mathbb{R}\{x, y, u, \theta\}$, onde $p = \dim \theta$. Esses p polinômios a mais são os responsáveis por dar a forma triangular nos parâmetros θ aos polinômios B_i do conjunto característico (3.15) formado ao término do algoritmo de Ritt. Ou seja, após um certo número de iterações dentro do algoritmo de Ritt chegaremos a um conjunto da seguinte forma:

$$A_1(u, y, \theta), \dots, A_d(u, y, \theta),$$

$$\dot{\theta}_1, \dot{\theta}_2, \dots, \dot{\theta}_p$$

$$C_1(u, y, \theta, x_1), \dots, C_n(u, y, \theta, x_1, x_2, \dots, x_n), \quad (3.27)$$

onde A_i são polinômios diferenciais em u e y e suas derivadas com coeficientes polinômios algébricos em θ , ou seja apenas os estados foram eliminados. Note que, até esse ponto, apenas os estados foram eliminados devido ao *ranking* utilizado, já que o *ranking* é quem determina a ordem de eliminação das variáveis no algoritmo de Ritt.

A partir desse ponto o algoritmo de Ritt começa a reduzir os polinômios $\dot{\theta}_i$ com respeito aos A_j de forma a eliminar um por um os parâmetros, até chegar na forma triangular. Serão necessárias mais, no máximo, dp iterações para chegar à forma triangular.

Em modelos mais complexos com vários estados e/ou muitos parâmetros, devido à complexidade das equações e ao custo computacional para lidar com essas equações, a implementação do algoritmo em computadores pessoais se torna inviável. Assim, tentando contornar esse problema, é proposto parar o algoritmo de Ritt quando esse alcança um conjunto com a forma (3.27), e, a partir desse ponto fazer o estudo da identificabilidade do modelo através dos polinômios A_i . Vale notar que parar o algoritmo nesse ponto é o mesmo que tratar o modelo como pertencente ao ideal (3.24).

Os polinômios A_i podem ser escritos como

$$A_i(y, u, \theta) = \xi_0(u, y) + \sum_{i=1}^m \xi_i(u, y) \psi_i(\theta) \quad (3.28)$$

onde $\xi_i(y, u)$ são polinômios diferenciais nas variáveis $y(t)$ e $u(t)$; $\psi_i(\theta)$ são polinômios em θ . Reescrevendo a relação (3.28) temos

$$A_i(y, u, \theta) \triangleq \xi_0(y, u) + \Psi^T(\theta) \Xi(y, u), \quad (3.29)$$

com

$$\Xi(y, u) = [\xi_1 \ \xi_2 \ \dots \ \xi_m]^T,$$

$$\Psi(\theta) = [\psi_1 \ \psi_2 \ \dots \ \psi_m]^T.$$

Assim podemos estudar a identificabilidade do modelo que gera (3.28) através da matriz de informação (2.38), onde neste caso

$$\psi(t, \theta) = \frac{\partial}{\partial \theta} A_i(y, u, \theta),$$

deste modo,

$$I(\theta) = \bar{E} \left[\frac{\partial}{\partial \theta} (\Psi^T(\theta) \Xi(y, u)) \frac{\partial}{\partial \theta} (\Xi^T(y, u) \Psi(\theta)) \right]. \quad (3.30)$$

Como o vetor Ξ não depende dos parâmetros θ , pode-se reescrever (3.30) como

$$I(\theta) = \bar{E} \left[\frac{\partial \Psi^T(\theta)}{\partial \theta} \Xi(y, u) \Xi^T(y, u) \frac{\partial \Psi(\theta)}{\partial \theta} \right], \quad (3.31)$$

utilizando o teorema de Parseval podemos reescrever a matriz de informação como (GEVERS et al., 2009):

$$I(\theta) \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \Psi^T(\theta)}{\partial \theta} \Xi(e^{j\omega}) \Xi^T(e^{j\omega}) \frac{\partial \Psi(\theta)}{\partial \theta} d\omega, \quad (3.32)$$

ou ainda, como $\frac{\partial \Psi^T(\theta)}{\partial \theta}$ são polinômios em θ , ou seja não dependem de ω

$$I(\theta) = \frac{1}{2\pi} \frac{\partial \Psi^T(\theta)}{\partial \theta} \left[\int_{-\pi}^{\pi} \Xi(e^{j\omega}) \Xi^T(e^{j\omega}) d\omega \right] \frac{\partial \Psi(\theta)}{\partial \theta}. \quad (3.33)$$

Utilizando a desigualdade de Sylvester,

$$\rho(AB) \leq \min(\rho(A), \rho(B)), \quad (3.34)$$

onde $\rho(\cdot)$ denota o posto da matriz, pode-se mostrar o seguinte (BAZANELLA; BOMBOIS; GEVERS, 2012)

Proposição 1. *Uma estrutura de modelo parametrizada $\mathcal{M}(\theta)$ é localmente identificável em θ_1 se*

$$\rho \left(\frac{\partial \Psi^T(\theta)}{\partial \theta} \right) \Big|_{\theta=\theta_1} = p. \quad (3.35)$$

◇

Do mesmo modo temos que

$$\rho \left(\int_{-\pi}^{\pi} \Xi(e^{j\omega}) \Xi^T(e^{j\omega}) d\omega \right) \geq p, \quad (3.36)$$

para que seja possível identificar o sistema (BAZANELLA; BOMBOIS; GEVERS, 2012)

Note que a expressão (3.33) mostra como a matriz de informação combina as informações sobre a identificabilidade da estrutura do modelo e a informatividade do conjunto de dados, através das matrizes $\frac{\partial \Psi^T(\theta)}{\partial \theta}$ e $\Xi(e^{j\omega}) \Xi^T(e^{j\omega})$, respectivamente.

A relação (3.35) é uma condição sobre o vetor de parâmetros do modelo, ou seja, nos traz informações sobre a identificabilidade da estrutura do modelo. Por outro

lado, (3.36) é uma condição sob os sinais de entrada e saída do sistema, assim como das condições iniciais, ou seja diz respeito à riqueza do conjunto de dados.

Quando o algoritmo de Ritt encontra um conjunto característico equivalente para a classe de modelo estudada, os teoremas 5, 6 e 7 fornecem condições necessárias e suficientes que o modelo deve satisfazer para que seja possível determinar se a classe de modelos é localmente ou globalmente identificável, e mais, é possível através das relações obtidas fazer a identificação do sistema. Por outro lado, parando o algoritmo quando este encontra um conjunto com a forma de (3.27) só será possível fazer a análise da identificabilidade da estrutura do modelo, determinar se a estrutura do modelo é localmente identificável através da Proposição 1. Como a estrutura encontrada não será linear nos parâmetros a identificação através da estrutura reduzida não poderá ser feita através dos métodos de mínimos quadrados e/ou variáveis instrumentais.

3.4 Algoritmo de Ritt Tempo Discreto

No Algoritmo 1 apresentamos o algoritmo de Ritt para sistemas em tempo contínuo. Nesta seção formalizaremos os conceitos necessários para poder usar o algoritmo em sistemas de tempo discreto.

Poucas alterações são necessárias, e o algoritmo é facilmente adaptado. Agora ao invés de sistemas dinâmicos descritos por (3.11), teremos sistemas de equações de diferença da forma

$$\begin{aligned} f_i(z(t), \dots, z(t-N), \theta) &= 0 \quad i = 1, \dots, n, \\ \theta(t) - \theta(t-1) &= 0 \end{aligned} \quad (3.37)$$

onde f_i são polinômios, $\theta \in \mathbb{R}^p$ é o vetor de parâmetros do sistema, e a variável z será dada por

$$z(t)^T = (y(t)^T, u(t)^T, e(t)^T, x(t)^T),$$

onde $y(t)$ e $u(t)$ são, respectivamente, os sinais de saída e entrada do sistema, $e(t)$ é o ruído do sistema e $x(t)$ são as variáveis de estado.

Assim como nos sistemas de tempo contínuo, onde as soluções dos polinômios diferenciais são dadas através de manipulações algébricas envolvendo o operador de diferenciação, nos sistemas de equações de diferenças utilizaremos o operador de atraso temporal, q^{-1} , onde

$$q^{-1}f(x(t), u(t)) = f(x(t-1), u(t-1)).$$

De forma a generalizar a notação, definimos

$$\begin{aligned} \dot{u} &\triangleq q^{-1}u(t) = u(t-1), \\ \ddot{u} &\triangleq q^{-2}u(t) = u(t-2), \\ u^{(k)} &\triangleq q^{-k}u(t) = u(t-k) \quad k \in \mathbb{N} > 2. \end{aligned} \quad (3.38)$$

Um conceito fundamental na álgebra diferencial é o de *rank* das variáveis. Para as variáveis de tempo discreto, ao invés de ordenar as variáveis segundo a ordem das

derivadas dentro de um polinômio usaremos o atraso temporal da variável:

$$x(t) \prec x(t-1) \prec x(t-2) \prec \dots \prec x(t-N).$$

Como na álgebra diferencial qualquer *ranking* é válido desde que esteja de acordo com a Definição 3.2.4, os conceitos de líder, separante, inicial, polinômios reduzidos, conjunto de polinômios, conjunto característico, etc, são análogos aos de tempo contínuo (LYZELL et al., 2011). Também são análogos os Lemas 3.2.2, 3.2.3 e 3.2.2 usados para reduzir polinômios diferenciais.

A principal diferença entre o algoritmo de Ritt em tempo contínuo e sua versão em tempo discreto é o número de pseudo-divisões necessárias para se chegar ao conjunto característico. Podemos observar facilmente essa diferença no seguinte exemplo.

Exemplo 3.4.1. *Sejam*

$$\begin{aligned} f_1 &= x^2 + \theta, \\ f_2 &= \dot{x} + \theta, \end{aligned} \quad (3.39)$$

dois polinômios diferenciais. Dado ranking

$$\theta \prec \dot{\theta} \prec x \prec \dot{x} \prec \dots, \quad (3.40)$$

temos que os líderes de f_1 e f_2 são, respectivamente, $l_{f_1} = x$ e $l_{f_2} = \dot{x}$. Temos também que, pela definição 3.2.5 item 3, o polinômio f_2 não é reduzido com respeito a f_1 , deste modo podemos utilizar o Lema 3.2.3 para reduzir f_2 através da seguinte pseudo-divisão

$$\begin{aligned} Q_C f_2 &= \bar{Q}_C \dot{f}_1 + R \\ Q_C(\dot{x} + \theta) &= \bar{Q}_C(2x\dot{x} + \dot{\theta}) + R \end{aligned} \quad (3.41)$$

tomando $Q_C = 2x$ e $\bar{Q}_C = 1$ temos

$$R = 2x\theta - \dot{\theta}, \quad (3.42)$$

que tem como líder $l_R = l_{f_1} = x$, Assim, como $\deg_x R < \deg_x f_1$, R será reduzido com respeito a f_1 e não será necessária outra iteração (Lema 3.2.4).

Agora se considerarmos os polinômios (3.39) como sendo equações de diferenças (tempo discreto), dado ranking (3.40), teremos que, assim como no caso contínuo, os líderes de f_1 e f_2 são, respectivamente, $l_{f_1} = x$ e $l_{f_2} = \dot{x}$ e que o polinômio f_2 não será reduzido com respeito a f_1 . Assim podemos reduzir f_2 com respeito a f_1 da seguinte forma:

$$Q_1 f_2 = \bar{Q}_1 \dot{f}_1 + R_1. \quad (3.43)$$

Diferentemente do caso contínuo onde $\dot{f}_1 = 2x\dot{x} + \dot{\theta}$, temos que $\dot{f}_1 = \dot{x}^2 + \dot{\theta}$, onde agora $\dot{f}_1(t) = f_1(t-1)$ (utilizamos o operador atraso temporal). Deste modo

$$Q_1(\dot{x} + \theta) = \bar{Q}_1(\dot{x}^2 + \dot{\theta}) + R_1, \quad (3.44)$$

tomando $Q_1 = \dot{x}$ e $\bar{Q}_1 = 1$ temos

$$R_1 = \dot{x}\theta - \dot{\theta}. \quad (3.45)$$

Como R contém \dot{x} mais uma redução será necessária,

$$\begin{aligned} Q_2 f_2 &= \bar{Q}_2 R_1 + R_2, \\ Q_2(\dot{x} + \theta) &= \bar{Q}_2(\dot{x}\theta - \dot{\theta}) + R_2, \end{aligned} \quad (3.46)$$

tomando $Q_2 = \theta$ e $\bar{Q}_2 = 1$, temos

$$R_2 = \theta^2 + \dot{\theta}, \quad (3.47)$$

cujo líder é $l_{R_1} = \dot{\theta}$. Assim, R_2 será reduzido com respeito a f_1 . Note que foram necessárias duas pseudo-divisões para chegar a um polinômio reduzido a f_1 enquanto para o caso contínuo apenas uma.

•

Diferentemente do caso contínuo, no caso discreto não se obtém condições necessárias e suficientes para que a estrutura do modelo seja globalmente identificável. Porém duas situações ainda podem ser encontradas:

Teorema 8. (LYZELL et al., 2009) *Se a saída do algoritmo 1 contém uma expressão da forma*

$$P_i(u, y)\theta_i - Z_i(u, y) = 0 \quad i = 1, \dots, p,$$

onde, $Z_i(u, y)$ e $P_i(u, y)$ são equações de diferenças em u e y , então a estrutura de modelo é globalmente identificável.

◇

Teorema 9. (LYZELL et al., 2009)

Dado o ranking

$$\left\{ u^{(\mu)} \prec y_j^{(\nu)} \prec \theta_1 \prec \dot{\theta}_1 \prec \dots \prec \theta_d \prec \dot{\theta}_d \prec \dots \prec x_k^{(\sigma)} \right\},$$

onde x_k contém qualquer variável imensurável. Suponha que a saída do algoritmo é um conjunto auto-reduzido com a seguinte forma

$$\begin{aligned} A_1(u, y), \dots, A_d(u, y), \\ \dot{\theta}_1 - \theta_1, B_2(u, y, \theta_1, \theta_2), \dots, B_p(u, y, \theta_1, \theta_2, \dots, \theta_p) \\ C_1(u, y, \theta, x_1), \dots, C_n(u, y, \theta, x_1, \dots, x_m). \end{aligned} \quad (3.48)$$

Além disso, assume-se que existe uma solução para este conjunto tal que todos os polinômios em \mathcal{G} são diferentes de zero. Então, existem infinitos valores θ que satisfazem as equações (3.48), isto é, o sistema é não identificável.

◇

O Teorema 9 implica que, se o algoritmo de Ritt não consegue eliminar a equação referente ao atraso temporal de algum dos parâmetros, $(\dot{\theta}_1 - \theta_1)$ não necessariamente θ_1 , então a estrutura do modelo é não identificável. E ainda, do Teorema 8, não se pode afirmar que se não for possível escrever a estrutura inicial como regressões lineares nos parâmetros a estrutura será não identificável.

Segundo (LYZELL et al., 2011) a prova (para sistemas de tempo contínuo) de que o modelo será identificável se somente se for possível escrevê-lo como uma regressão linear é baseado na propriedade de que se duas funções $f(t)$ e $g(t)$ são analíticas e satisfazem $f(t)g(t) \equiv 0$, então ou $f(t) \equiv 0$ ou $g(t) \equiv 0$, já o mesmo não é válido para o caso discreto. Um exemplo simples de sinal que não satisfaz essa relação é a delta de Kronecker:

$$\delta(t - k) = \begin{cases} 1 & t = k \\ 0 & t \neq k. \end{cases}$$

Ou seja, $\delta(t)\delta(t - 1) \equiv 0$ mesmo $\delta(t)$ e $\delta(t - 1)$ não sendo identicamente nulas.

4 ESTUDO DE CASOS

Neste capítulo apresentaremos uma série de exemplos onde foi utilizado o algoritmo de Ritt para estudar a identificabilidade de diversas classes de modelos. Dividiremos esse capítulo em duas sessões, uma que trata os exemplos de tempo contínuo e outra para tempo discreto. Para os sistemas de tempo contínuo trataremos de analisar a identificabilidade das estruturas propostas sem considerar o efeito de ruído. Já para os sistemas de tempo discreto, além de fazer o estudo da identificabilidade da estrutura, faremos a identificação do sistema considerando ruído no sistema, e então apresentaremos um estudo da consistência e precisão das estimativas.

O algoritmo de Ritt, tanto para o caso contínuo quanto para o caso discreto, foi implementado através da plataforma Mupad do MATLAB. Foram criadas rotinas capazes de executar o Algoritmo 1, onde os dados de entrada são um *ranking* e um ideal da forma de (3.13). O algoritmo retorna um conjunto característico para esse ideal e o conjunto \mathcal{G} de polinômios resultantes das pseudo-reduções. As rotinas criadas são apresentadas no Apêndice A.

Utilizou-se um computador pessoal para executar o algoritmo de Ritt implementado. O computador utilizado é um com processador *Intel i7 2.0GHz*, com 8GB de memória RAM rodando *Windows 7 x64 professional*. A versão do MATLAB utilizada é a R2010a. Como será apresentado no decorrer deste capítulo, devido a limitação do *hardware* utilizado, a implementação do algoritmo de Ritt torna-se muito demorada para estruturas de modelos com muitos parâmetros. Por isso, para estes casos foi utilizado o algoritmo na forma alternativa.

4.1 Tempo Contínuo

Nesta sessão será discutida a utilização do algoritmo de Ritt no estudo da identificabilidade de estruturas de modelos em sistemas de tempo contínuo encontradas na literatura. Os resultados obtidos serão comparados com os encontrados na literatura. As três primeiras estruturas foram retiradas de (LJUNG; GLAD, 1994). As demais estruturas utilizadas foram retiradas de outros trabalhos, onde foram utilizados diferentes métodos para determinar se a estrutura do modelo é identificável.

4.1.1 Caso 1 – “Goodwin’s Napkin”

Como mencionado acima, esta estrutura de modelo foi retirada de (LJUNG; GLAD, 1994), e é a mesma apresentada no Exemplo 3.3.1. Executando o algoritmo de Ritt implementado no conjunto (3.21) com o *ranking* (3.14) obtemos os seguintes

polinômios diferenciais que formam um conjunto característico:

$$y^{(3)^2}y^2 - 6y^{(3)}\ddot{y}\dot{y}y + 4y^{(3)}\dot{y}^3 + 4\dot{y}^3y - 3\dot{y}^2\dot{y}^2 = 0 \quad (4.1)$$

$$y^{(3)}y - \ddot{y}\dot{y} + 2\theta\ddot{y}y - 2\theta\dot{y}^2 = 0 \quad (4.2)$$

Fazendo a relação entre os polinômios do conjunto característico com os polinômios $A_i(u, y)$, $B_i(u, y, \theta)$ e $C_i(u, y, \theta, x)$ de (3.13) temos que, como o sistema não possui estados não existem polinômios $C_i(u, y, \theta, x)$, o polinômio $A(u, y)$ é o dado por (4.1) e o polinômio $B(u, y, \theta)$ é dado por (4.2). Deste modo, de acordo com o Teorema 6 a partir da relação (4.2) podemos afirmar que o conjunto de modelos descrito por (3.18) é *Globalmente Identificável*, já que conseguiu-se encontrar uma forma reduzida do tipo regressão linear nos parâmetros (Caso 2 de (3.13)). O algoritmo ainda retorna uma expressão que não depende do parâmetro θ , (4.1), esta é uma equação diferencial que qualquer $y(t)$ que obedeça (3.18) deve satisfazer independentemente do valor do parâmetro θ .

Também obtemos o conjunto \mathcal{G} de polinômios

$$y(2\dot{y} + 2\theta y) \neq 0 \quad (4.3)$$

$$-4(\dot{y}^2 - y\ddot{y})^2(2\dot{y}^2 - 2y\ddot{y}) \neq 0. \quad (4.4)$$

4.1.1.1 Informatividade

O conjunto \mathcal{G} é formado pelas seguintes condições

$$y \neq 0, \quad (4.5)$$

$$(\dot{y} + \theta y) \neq 0, \quad (4.6)$$

$$(\dot{y}^2 - y\ddot{y}) \neq 0. \quad (4.7)$$

A relação (4.5) é a condição trivial, se (4.7) não for satisfeita, ou seja,

$$(\dot{y}^2 - y\ddot{y}) = 0$$

teremos que, reescrevendo a equação (4.2),

$$\underbrace{y^{(3)}y - \ddot{y}\dot{y}}_{Z(y)} + 2\theta \underbrace{(\ddot{y}y - \dot{y}^2)}_{\varphi(y)} = 0 \quad (4.8)$$

o vetor regressor $\varphi(y)$ de (4.8) será identicamente nulo. Neste caso, a estrutura do modelo não poderá ser identificada, já que o sinal $y(t)$ não será informativo o suficiente.

Outra condição que deve ser satisfeita é (4.6). Facilmente se consegue determinar quais são as condições iniciais que satisfazem (4.6). Sabe-se que a solução geral do modelo 3.18 é

$$y(t) = Ae^{-\theta t} + tBe^{-\theta t}, \quad (4.9)$$

com A e B , constantes, dependentes das condições iniciais. Escrevendo a solução geral em termos das condições iniciais obtemos:

$$y(t) = y(0)e^{-\theta t} + t(\dot{y}(0) + \theta y(0))e^{-\theta t}. \quad (4.10)$$

Derivando e substituindo (4.10) em (4.6) obtemos

$$\dot{y}(0) + \theta y(0) \neq 0. \quad (4.11)$$

Ou seja, qualquer condição inicial que não satisfaça (4.11) fará com que não seja possível identificar o sistema. Para esse caso é simples de observar isso. Se derivarmos e substituirmos (4.10) no vetor regressor $\varphi(y)$ de (4.8) obtemos

$$\varphi(y) = - [\dot{y}(0) + \theta y(0)]^2 e^{-2\theta t} \quad (4.12)$$

Assim, se (4.11) não for satisfeita o vetor regressor (4.12) será identicamente nulo, e portanto, não será possível identificar o sistema, mesmo a estrutura sendo globalmente identificável. Note que esta condição é redundante, visto que (4.7) também faz com que o vetor regressor seja identicamente nulo.

4.1.2 Caso 2 – Modelo não linear

A estrutura de modelo estudada será dada por (LJUNG; GLAD, 1994)

$$\begin{aligned} \dot{x}_1 &= \theta x_2^2 \\ \dot{x}_2 &= u \\ y &= x_1. \end{aligned} \quad (4.13)$$

Executando o algoritmo de Ritt em (4.77) com o *ranking*

$$u \prec \dot{u} \prec \dots \prec y \prec \dot{y} \prec \dots \prec \theta \prec \dot{\theta} \prec x_1 \prec \dot{x}_1 \prec \dots \prec x_2 \prec \dot{x}_2 \prec \dots \quad (4.14)$$

obtemos os seguintes polinômios diferenciais que compõem o conjunto característico da classe de modelos (4.13):

$$2\dot{u}\dot{y}\dot{y} - 2uy^{(3)}\dot{y} + u\dot{y}^2 = 0, \quad (4.15)$$

$$(4u^2\dot{y})\theta - \dot{y}^2 = 0, \quad (4.16)$$

$$y - x_1 = 0, \quad (4.17)$$

$$2u\dot{y} - x_2\dot{y} = 0. \quad (4.18)$$

Observe que (4.15) é o polinômio $A(u, y)$ de (3.15), (4.16) é o polinômio $B(u, y, \theta)$ de (3.15) e (4.17) e (4.18) são os polinômios $C_i(u, y, \theta, x_1, x_2)$. Assim, como os polinômios $B(u, y, \theta)$ são como os do Caso 2 pelo Teorema 6, a estrutura do modelo (4.77) será globalmente identificável.

4.1.2.1 Informatividade

O algoritmo ainda retorna o conjunto \mathcal{G} de polinômios

$$2\theta^2 x_2 \neq 0, \quad (4.19)$$

$$8\theta^3 u^3 \neq 0, \quad (4.20)$$

$$4u^2 \dot{y} \neq 0, \quad (4.21)$$

$$16u^4 \dot{y}^2 \neq 0, \quad (4.22)$$

As relações (4.19-4.22) são condições de persistência de excitação, onde nota-se que (4.21) e (4.22) são redundantes, visto que (4.21) é a raiz quadrada de (4.22). E ainda, (4.21) corresponde ao vetor regressor de (4.16).

4.1.3 Caso 3 – Modelo compartimental

Este modelo foi retirado de (LJUNG; GLAD, 1994). Para este caso serão utilizados o algoritmo de Ritt na sua forma original e na forma alternativa apresentado na Sessão 3.3.1.

A estrutura compartimental é dada por

$$\begin{aligned}\dot{x} &= -\frac{V_m x}{k_m + x} - k_{01}x, \\ y &= cx,\end{aligned}\tag{4.23}$$

onde $\theta = [V_m \ c \ k_m \ k_{01}]^T$ são os parâmetros a se identificar.

Utilizando o *ranking*

$$\left\{ u \prec \dot{u} \prec \dots \prec y \prec \dot{y} \prec \dots \prec V_m \prec \dot{V}_m \prec c \prec \dot{c} \prec k_m \prec \dot{k}_m \prec k_{01} \prec \dot{k}_{01} \prec x \prec \dot{x} \prec \dots \right\},\tag{4.24}$$

o algoritmo de Ritt retorna as seguintes expressões que formam o conjunto característico

$$\begin{aligned}\sigma_1(2y^{(4)}\dot{y}^4y - 2y^{(4)}\ddot{y}^2y^2 \\ + 3y^{(3)^2}\dot{y}^2y^2 + 2y^{(3)}\ddot{y}^2\dot{y}y^2 - 14y^{(3)}\ddot{y}\dot{y}^3y + 6y^{(3)}\dot{y}^5 \\ - 3\dot{y}^4y^2 + 12\dot{y}^3\dot{y}^2y - 6\dot{y}^2\dot{y}^4) = 0,\end{aligned}\tag{4.25}$$

onde σ_1 é dado por

$$\sigma_1 = (\dot{y}^2 - y\ddot{y})^8(y^{(3)}\dot{y}y^2 - \ddot{y}^2y^2 - 2\ddot{y}\dot{y}^2y + 2\dot{y}^4),$$

$$\dot{V}_m = 0,\tag{4.26}$$

$$\begin{aligned}- (\dot{y}^2 - y\ddot{y})^3(V_mcy^{(3)^2}\dot{y}^2y^4 - 2V_mcy^{(3)}\ddot{y}^2\dot{y}y^4 \\ - 4V_mcy^{(3)}\ddot{y}\dot{y}^3y^3 + 4V_mcy^{(3)}\dot{y}^5y^2 + V_mcy^{(4)}y^4 \\ - 4\dot{y}^3\dot{y}^3y^3 + 4V_mcy^{(3)}\dot{y}^2y^3 + 12\ddot{y}^2\dot{y}^5y^2 \\ - 12\ddot{y}\dot{y}^7y - 8V_mcy^{(3)}\dot{y}^6y + 4\dot{y}^9 + 4V_mcy^{(8)}) = 0,\end{aligned}\tag{4.27}$$

$$\begin{aligned}- (\dot{y}^2 - y\ddot{y})^5(V_my^{(3)^2}\dot{y}^2y^5 - 2V_my^{(3)}\ddot{y}^2\dot{y}y^5 \\ - 2V_my^{(3)}\ddot{y}\dot{y}^3y^4 + 2V_my^{(3)}\dot{y}^5y^3 + V_my^{(4)}y^5 + 4k_m\dot{y}^3\dot{y}^3y^3 + 2V_m\ddot{y}^3\dot{y}^2y^4 \\ - 12k_m\ddot{y}^2\dot{y}^5y^2 - 2V_m\ddot{y}^2\dot{y}^4y^3 + 12k_m\ddot{y}\dot{y}^7y - 4k_m\dot{y}^9) = 0,\end{aligned}\tag{4.28}$$

$$y^{(3)}\dot{y}^2y + k_{01}y^{(3)}\dot{y}y^2 - 3\ddot{y}^2\dot{y}y - k_{01}\ddot{y}^2y^2 + 2\ddot{y}\dot{y}^3 - 2k_{01}\ddot{y}\dot{y}^2y + 2k_{01}\dot{y}^4 = 0,\tag{4.29}$$

$$\begin{aligned}- (\dot{y}^2 - y\ddot{y})^3(V_my^{(3)^2}\dot{y}^2y^5 - 2V_my^{(3)}\ddot{y}^2\dot{y}y^5 \\ - 4V_my^{(3)}\ddot{y}\dot{y}^3y^4 + 4V_my^{(3)}\dot{y}^5y^3 + V_my^{(4)}y^5 \\ - 4x\dot{y}^3\dot{y}^3y^3 + 4V_m\dot{y}^3\dot{y}^2y^4 + 12x\ddot{y}^2\dot{y}^5y^2 \\ - 12x\ddot{y}\dot{y}^7y - 8V_m\ddot{y}\dot{y}^6y^2 + 4xy^9 + 4V_my^{(8)}y) = 0,\end{aligned}\tag{4.30}$$

Novamente fazendo a relação entre os polinômios do conjunto característico com os polinômios $A_i(u, y)$, $B_i(u, y, \theta)$ e $C_i(u, y, \theta, x)$ de (3.13) temos que: (4.25) é o polinômio $A(u, y)$, (4.26 - 4.29) são os polinômios $B_i(u, y, \theta)$ e (4.30) o polinômio $C(u, y, \theta, x)$. A partir do Teorema 5 observamos que a estrutura (4.23) não pode ser identificável devido à expressão (4.26), Caso 2 dos polinômios $B_i(u, y, \theta)$, porém o parâmetro k_{01} pode ser identificado através da relação (4.29). Note que, se V_m fosse conhecido, seria possível identificar todos os demais parâmetros, ou seja, a estrutura seria globalmente identificável.

Encontramos também o conjunto de polinômios \mathcal{G}

$$-c^5 \neq 0, \quad (4.31)$$

$$y^2(y + ck_m)^2 \neq 0, \quad (4.32)$$

$$2c^3(y + ck_m)(\dot{y}^2 - y\ddot{y})^2 \neq 0, \quad (4.33)$$

$$-2c\dot{y}(\dot{y}^2 - y\ddot{y})^2 \neq 0, \quad (4.34)$$

$$-8c^3\dot{y}^3(\dot{y}^2 - y\ddot{y})^6 \neq 0, \quad (4.35)$$

$$-2c\dot{y}(\dot{y}^2 - y\ddot{y})^2 \neq 0, \quad (4.36)$$

$$-V_m(\dot{y}^2 - y\ddot{y})^3(y^{(3)}\dot{y}y^2 - \ddot{y}^2y^2 - 2\ddot{y}\dot{y}^2y + 2\dot{y}^4)^2 \neq 0, \quad (4.37)$$

$$-V_m(\dot{y}^2 - y\ddot{y})^3(y^{(3)}\dot{y}y^2 - \ddot{y}^2y^2 - 2\ddot{y}\dot{y}^2y + 2\dot{y}^4)^2 \neq 0, \quad (4.38)$$

$$V_m^2(\dot{y}^2 - y\ddot{y})^6(y^{(3)}\dot{y}y^2 - \ddot{y}^2y^2 - 2\ddot{y}\dot{y}^2y + 2\dot{y}^4)^4 \neq 0. \quad (4.39)$$

Novamente, as relações (4.31 - 4.39) são condições de persistência de excitação. Note que as relações (4.35 - 4.36) e (4.37 - 4.39) são redundantes, já que são combinações uma das outras.

Agora utilizaremos o algoritmo na forma alternativa, ou seja, executaremos o algoritmo até encontrarmos uma expressão da forma de (3.29), com o objetivo de encontrar a matriz de informação do sistema e então utilizar a Proposição 1 para a análise da identificabilidade. Executando o algoritmo de Ritt na forma alternativa obtemos a seguinte expressão

$$V(y, u, \theta) = \underbrace{y\dot{y}}_{\xi_0} + \underbrace{[k_{01} \quad V_m c + ck_{01}k_m \quad ck_m]}_{\Psi(\theta)} \underbrace{[y^2 \quad y \quad \dot{y}]^T}_{\Xi(u, y)}$$

calculando o gradiente da função $\Psi(\theta)$ temos

$$\frac{\partial \Psi^T(\theta)}{\partial \theta} = \begin{pmatrix} 0 & c & 0 \\ 0 & V_m + k_{01}k_m & k_m \\ 0 & ck_{01} & c \\ 1 & ck_m & 0 \end{pmatrix}. \quad (4.40)$$

Segundo a Proposição 1 o posto da matriz (4.40) deve ser igual ao tamanho do vetor de parâmetros, que nesse caso é quatro, porém $\rho\left(\frac{\partial \Psi^T}{\partial \theta}\right) = 3 < 4$. Assim sendo, a estrutura do modelo (4.23) não é identificável, que é o mesmo resultado obtido utilizando o algoritmo de Ritt.

4.1.4 Caso 4 – Três modelos Compartimentais

A estrutura do modelo para o caso 4 foi retirado de (SACCOMANI; AUDOLY; D'ANGIÒ, 2003) e descreve a dinâmica de uma determinada substância em um

tecido. A estrutura do modelo é dada por:

$$\begin{aligned}\dot{x}_1 &= p_{13}x_3 + p_{12}x_2 - p_{21}x_1 + u, \\ \dot{x}_2 &= -p_{12}x_2 + p_{21}x_1, \\ \dot{x}_3 &= -p_{13}x_3, \\ y &= x_2,\end{aligned}\tag{4.41}$$

onde $\theta = [p_{12} \ p_{13} \ p_{21}]^T$ são os parâmetros do modelo.

Apesar de ser um sistema relativamente pequeno, com três parâmetros, por esta estrutura apresentar três estados, x_1, x_2 e x_3 , não se conseguiu chegar a um conjunto característico, deste modo utilizando o algoritmo de Ritt na sua forma alternativa encontramos com o *ranking*

$$\{u \prec \dot{u} \prec \dots \prec y \prec \dot{y} \prec \dots \prec p_{12} \prec \dot{p}_{12} \prec p_{13} \prec \dot{p}_{13} \prec p_{21} \prec \dot{p}_{21} \prec x_1 \prec \dot{x}_1 \prec \dots \prec x_2 \prec \dot{x}_2 \prec \dots \prec x_3 \prec \dot{x}_3 \prec \dots\},\tag{4.42}$$

a seguinte equação

$$V(y, u, \theta) = \underbrace{y^{(3)}}_{\xi_0} + \underbrace{[(p_{12}p_{13} + p_{13}p_{21}) \ (p_{12} + p_{13} + p_{21}) \ -p_{13}p_{21} \ -p_{21}]}_{\Psi(\theta)} \underbrace{[\dot{y} \ \ddot{y} \ u \ \dot{u}]^T}_{\Xi(y,u)}.\tag{4.43}$$

Assim podemos utilizar a matriz de informação formada pela estrutura (4.43) para fazer a análise da identificabilidade da estrutura original. Temos que

$$\frac{\partial \Psi^T(\theta)}{\partial \theta} = \begin{pmatrix} p_{13} & 1 & 0 & 0 \\ p_{12} + p_{13} & 1 & p_{21} & 0 \\ p_{13} & 1 & p_{13} & 1 \end{pmatrix}.\tag{4.44}$$

Como o posto $\rho\left(\frac{\partial \Psi^T(\theta)}{\partial \theta}\right) = 3$, pela Proposição 1, a estrutura (4.41) é localmente identificável. Vale ressaltar que para esse caso não foi possível a utilização do algoritmo de Ritt na forma original, porém o resultado obtido, identificabilidade, é o mesmo encontrado por (SACCOMANI; AUDOLY; D'ANGIÒ, 2003).

4.1.5 Caso 5

Estrutura retirada de (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001) dada por:

$$\begin{aligned}\dot{x}_1 &= \theta_1 x_1^2 + \theta_2 x_1 x_2 \\ \dot{x}_2 &= \theta_3 x_1^2 + x_1 x_2 \\ y &= x_1\end{aligned}\tag{4.45}$$

Aplicando o algoritmo sobre a estrutura (4.45) com o *ranking*

$$\left\{u \prec \dot{u} \prec \dots \prec y \prec \dot{y} \prec \dots \prec \theta_1 \prec \dot{\theta}_1 \prec \theta_2 \prec \dot{\theta}_2 \prec \theta_3 \prec \dot{\theta}_3 \prec x_1 \prec \dot{x}_1 \prec x_2 \prec \dot{x}_2 \prec \dots\right\},$$

encontramos o seguinte conjunto de equações reduzidas

$$\begin{aligned}
& -y^4(y^{(3)})^2 + y^4\ddot{y}y^{(4)} - 2y^3\dot{y}^2y^{(4)} + 3y^3\dot{y}y^{(3)}\ddot{y} \\
& \quad - 5y^3\ddot{y}^3 + 6y^2\dot{y}^3y^{(3)} + 12y^2\dot{y}^2\ddot{y}^2 - 30y\dot{y}^4\ddot{y} + 16\dot{y}^6 = 0, \quad (4.46)
\end{aligned}$$

$$(\theta_1 + 1)y^3\ddot{y} + (-2\theta_1 - 2)y^2\dot{y}^2 - y^2y^{(3)} + 5y\dot{y}\ddot{y} - 4\dot{y}^3 = 0, \quad (4.47)$$

$$\begin{aligned}
& (\theta_2\theta_3 + 1)y^5\ddot{y} + (-2\theta_2\theta_3 - 2)y^4\dot{y}^2 - y^4y^{(3)} \\
& \quad + 5y^3\dot{y}\ddot{y} - 4y^2\dot{y}^3 + y^2\dot{y}y^{(3)} - y^2\dot{y}^2 - 2y\dot{y}^2\ddot{y} + 2\dot{y}^4 = 0, \quad (4.48)
\end{aligned}$$

$$\dot{\theta}_2 = 0, \quad (4.49)$$

$$y - x_1 = 0, \quad (4.50)$$

$$-y^3\ddot{y} + 2y^2\dot{y}^2 + y^2y^{(3)} + (\theta_2y^2\ddot{y} - 2\theta_2y\dot{y}^2)x_2 - 6y\dot{y}\ddot{y} + 6\dot{y}^3 = 0 \quad (4.51)$$

Fazendo a relação entre os polinômios do conjunto característico com os polinômios $A_i(u, y)$, $B_i(u, y, \theta)$ e $C_i(u, y, \theta, x)$ de (3.13) temos que; (4.46) é o polinômio $A(u, y)$, (4.47 - 4.49) são os polinômios $B_i(u, y, \theta)$ e (4.50) e (4.51) os polinômios $C_i(u, y, \theta, x)$. Os polinômios $B_i(u, y, \theta)$ correspondem ao Caso 1.

Assim, de acordo com o Teorema 5 a estrutura do modelo (4.45) não é identificável, já que existe uma expressão do tipo $\dot{\theta} = 0$, equação (4.49). Porém, podemos identificar o parâmetro θ_1 através da equação (4.47), assim como o produto $\theta_2\theta_3$ pela relação (4.48). Note que se o estado x_2 fosse conhecido poderíamos utilizar a relação (4.51) para identificar o modelo. Este resultado é coerente com aqueles descritos em (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001).

Para o conjunto \mathcal{G} de polinômios encontramos

$$\theta_2^3y^3 \neq 0, \quad (4.52)$$

$$(y^4\theta_1 - y^2\dot{y})^2 \neq 0, \quad (4.53)$$

$$-\theta_2y^4(2\dot{y}^2 - y\ddot{y})(y^5\ddot{y}\theta_2 - 2y^4\dot{y}^2\theta_2) \neq 0, \quad (4.54)$$

$$\theta_2y^4(2\dot{y}^2 - y\ddot{y}) \neq 0, \quad (4.55)$$

Utilizando o algoritmo de Ritt na forma alternativa encontramos a seguinte expressão reduzida:

$$V(y, u, \theta) = \underbrace{y\ddot{y} - \dot{y}^2 - y^2\dot{y}}_{\xi_0} + \underbrace{[(\theta_1 - \theta_2\theta_3) - \theta_1]}_{\Psi(\theta)} \underbrace{[y^4 \quad y^2\dot{y}]^T}_{\Xi(u, y)},$$

deste modo,

$$\frac{\partial \Psi^T(\theta)}{\partial \theta} = \begin{pmatrix} 1 & 1 \\ -\theta_3 & 0 \\ -\theta_2 & 0 \end{pmatrix}. \quad (4.56)$$

Como $\rho\left(\frac{\partial \Psi^T(\theta)}{\partial \theta}\right) = 2 < 3$, segundo a Proposição 1 a estrutura do modelo não é identificável.

4.1.6 Caso 6 – Modelo Farmacocinético

Esta estrutura de modelos foi retirada de (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001). A estrutura deste modelo é bastante similar ao Caso 4.1.3, porém se acrescenta um estado ao modelo (4.23), obtendo a seguinte estrutura:

$$\begin{aligned}\dot{x}_1 &= \alpha_1(x_2 - x_1) - \frac{V_m x_1}{k_c + x_1}, \\ \dot{x}_2 &= \alpha_2(x_1 - x_2), \\ y &= x_1.\end{aligned}\quad (4.57)$$

O vetor de parâmetros é dado por $\theta = [\alpha_1 \ \alpha_2 \ V_m \ k_c]^T$.

Apesar de ter o mesmo número de parâmetros que o Caso 4.1.3, o acréscimo de um estado aumenta o número de pseudo-reduções e a complexidade das pseudo-reduções necessárias para encontrar o conjunto característico para o modelo. Por isso não foi possível chegar a um conjunto autoreduzido utilizando o algoritmo de Ritt. Porém, dado o *ranking*

$$\{u \prec \dot{u} \prec \dots \prec y \prec \dot{y} \prec \dots \prec V_m \prec \dot{V}_m \prec \alpha_2 \prec \dot{\alpha}_2 \prec k_c \prec \dot{k}_c \prec \alpha_1 \prec \dot{\alpha}_1 \prec x_1 \prec \dot{x}_1 \prec \dots \prec x_2 \prec \dot{x}_2 \prec \dots\}, \quad (4.58)$$

podemos utilizar o algoritmo de Ritt na forma alternativa. Deste modo conseguimos chegar a uma expressão da forma (3.29), ou seja, uma forma que só depende dos sinais de entrada, saída e parâmetros:

$$\begin{aligned}V(y, u, \theta) &= y^2 \ddot{y} + (\alpha_1 + \alpha_2) y^2 \dot{y} + (2\alpha_1 k_c + 2\alpha_2 k_c) y \dot{y} + V_m \alpha_2 y^2 \\ &\quad + 2k_c y \ddot{y} + V_m \alpha_2 k_c y + (V_m k_c + \alpha_1 k_c^2 + \alpha_2 k_c^2) \dot{y} + k_c^2 \ddot{y}\end{aligned}\quad (4.59)$$

calculando

$$\frac{\partial \Psi^T(\theta)}{\partial \theta} \Xi(y, u) = \begin{pmatrix} 1 & 0 & k_c & 0 & 0 & k_c^2 & 0 \\ 1 & V_m & k_c & 0 & V_m k_c & k_c^2 & 0 \\ 0 & \alpha_2 & 0 & 0 & \alpha_2 k_c & k_c & 0 \\ 0 & 0 & \alpha_1 + \alpha_2 & 1 & V_m \alpha_2 & V_m + \alpha_1 k_c + \alpha_2 k_c + k_c(\alpha_1 + \alpha_2) & 2k_c \end{pmatrix} \begin{pmatrix} y^2 \ddot{y} \\ y^2 \\ 2y \dot{y} \\ 2y \ddot{y} \\ y \\ \dot{y} \\ \ddot{y} \end{pmatrix}. \quad (4.60)$$

Como $\rho\left(\frac{\partial \Psi^T(\theta)}{\partial \theta}\right) = 4$, segundo a Proposição 1, a estrutura do modelo (4.57) é localmente identificável. Mesmo resultado encontrado em (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001).

4.1.7 Caso 7 – Modelo Compartimental

Essa estrutura de modelo descreve a cinética de um medicamento no corpo humano. (AUDOLY et al., 2001).

$$\begin{aligned}\dot{x}_1 &= -\left(k_{21} + \frac{V_m}{K_m + x_1}\right) x_1 + k_{12} x_2 + b_1 u \\ \dot{x}_2 &= k_{21} x_1 - (k_{02} + k_{12}) x_2 \\ y &= c_1 x_1,\end{aligned}\quad (4.61)$$

onde $\theta = [k_{21} \ V_m \ K_m \ k_{12} \ b_1 \ k_{02} \ c_1]^T$.

Devido ao número de parâmetros, não foi possível encontrar um conjunto característico para a estrutura (4.61). Assim faremos a análise da identificabilidade desta estrutura através da matriz de informação. Deste modo, dado que

$$\begin{aligned}
V(y, u, \theta) = & (k_{02}k_{21} + 2k_{12}k_{21})y^3 + (k_{02} + k_{12} - k_{21})y^2\dot{y} \\
& - y^2\ddot{y} + b_1c_1y^2\dot{u} - (b_1c_1k_{02} + b_1c_1k_{12})y^2u + b_1c_1^3k_m^2\dot{u} \\
& + (V_m c_1 k_{02} + V_m c_1 k_{12} + 2c_1 k_{02} k_{21} k_m + 4c_1 k_{12} k_{21} k_m)y^2 \\
& + (c_1^2 k_{02} k_m^2 + c_1^2 k_{12} k_m^2 - c_1^2 k_{21} k_m^2 - V_m c_1^2 k_m)\dot{y} \\
& - (2b_1c_1^2k_{02}k_m + 2b_1c_1^2k_{12}k_m)yu + (2c_1k_{02}k_m + 2c_1k_{12}k_m - 2c_1k_{21}k_m)y\dot{y} \\
& - 2c_1k_my\ddot{y} + 2b_1c_1^2k_my\dot{u} \\
& + (V_m c_1^2 k_{02} k_m + V_m c_1^2 k_{12} k_m + c_1^2 k_{02} k_{21} k_m^2 + 2c_1^2 k_{12} k_{21} k_m^2)y \\
& - c_1^2 k_m^2 \ddot{y} - (b_1 c_1^3 k_{02} k_m^2 + b_1 c_1^3 k_{12} k_m^2)u, \quad (4.62)
\end{aligned}$$

temos que

$$\frac{\partial \Psi^T(\theta)}{\partial \theta} \Xi(y, u) = [J \ K \ L] \Xi(y, u),$$

com J , K , L e $\Phi(y, u)$ dados por (4.63), (4.64), (4.65) e (4.66), respectivamente. Como $\rho\left(\frac{\partial \Psi^T(\theta)}{\partial \theta}\right) = 6 < 7$, pela Proposição 1, a estrutura do modelo (4.61) não é identificável. O resultado obtido é o mesmo encontrado pelos autores em (DENIS-VIDAL; JOLY-BLANCHARD; NOIRET, 2001), porém neste trabalho foi utilizada *bases de Gröbner* para fazer a análise de identificabilidade da estrutura (CALAN-DRINI et al., 1997), (BOULIER et al., 2009).

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 & c_1(2k_{02}k_{21} + 4k_{12}k_{21}) & c_1(k_{02} + k_{12} - k_{21}) & c_1 \\ k_{02} + 2k_{12} & -1 & 0 & 0 & c_1(2k_{02}k_m + 4k_{12}k_m) & -c_1k_m & 0 \\ 2k_{21} & 1 & b_1c_1 & 0 & c_1(Vm + 4k_{21}k_m) & c_1k_m & 0 \\ 0 & 0 & c_1(k_{02} + k_{12}) & c_1 & 0 & 0 & 0 \\ k_{21} & 1 & b_1c_1 & 0 & c_1(Vm + 2k_{21}k_m) & c_1k_m & 0 \\ 0 & 0 & b_1(k_{02} + k_{12}) & b_1 & Vmk_{02} + Vmk_{12} + 2k_{02}k_{21}k_m + 4k_{12}k_{21}k_m & k_m(k_{02} + k_{12} - k_{21}) & k_m \\ 0 & 0 & 0 & 0 & c_1(k_{02} + k_{12}) & 0 & 0 \end{pmatrix} \quad (4.63)$$

$$K = \begin{pmatrix} b_1c_1^2(k_{02} + k_{12}) & b_1c_1^2 & c_1^2(Vmk_{02} + Vmk_{12} + k_{02}k_{21}k_m + 2k_{12}k_{21}k_m) + c_1^2k_m(k_{02}k_{21} + 2k_{12}k_{21}) & \\ 0 & 0 & c_1^2k_m(k_{02}k_m + 2k_{12}k_m) & \\ b_1c_1^2k_m & 0 & c_1^2k_m(Vm + 2k_{21}k_m) & \\ c_1^2k_m(k_{02} + k_{12}) & c_1^2k_m & 0 & \\ b_1c_1^2k_m & 0 & c_1^2k_m(Vm + k_{21}k_m) & \\ 2b_1c_1k_m(k_{02} + k_{12}) & 2b_1c_1k_m & 2c_1k_m(Vmk_{02} + Vmk_{12} + k_{02}k_{21}k_m + 2k_{12}k_{21}k_m) & \\ 0 & 0 & c_1^2k_m(k_{02} + k_{12}) & \end{pmatrix} \quad (4.64)$$

$$L = \begin{pmatrix} c_1^2(Vm - k_{02}k_m - k_{12}k_m + k_{21}k_m) - c_1^2k_m(k_{02} + k_{12} - k_{21}) & 2c_1^2k_m & 2b_1c_1^3k_m(k_{02} + k_{12}) & 2b_1c_1^3k_m \\ c_1^2k_m^2 & 0 & 0 & 0 \\ -c_1^2k_m^2 & 0 & b_1c_1^3k_m^2 & 0 \\ 0 & 0 & c_1^3k_m^2(k_{02} + k_{12}) & c_1^3k_m^2 \\ -c_1^2k_m^2 & 0 & b_1c_1^3k_m^2 & 0 \\ 2c_1k_m(Vm - k_{02}k_m - k_{12}k_m + k_{21}k_m) & 2c_1k_m^2 & 3b_1c_1^2k_m^2(k_{02} + k_{12}) & 3b_1c_1^2k_m^2 \\ 0 & c_1^2k_m & 0 & 0 \end{pmatrix} \quad (4.65)$$

$$\Phi(y, u) = \begin{pmatrix} y^3 \\ y^2\dot{y} \\ -uy^2 \\ y^2\dot{u} \\ y^2 \\ 2y\dot{y} \\ -2y\dot{y} \\ -2uy \\ 2y\dot{u} \\ y \\ -\dot{y} \\ -\ddot{y} \\ -u \\ \dot{u} \end{pmatrix} \quad (4.66)$$

4.2 Tempo Discreto

Nesta sessão apresentaremos os resultados obtidos para sistemas de tempo discreto. Assim como para sistemas de tempo contínuo, o algoritmo de Ritt foi implementado na plataforma MuPad do MATLAB. Trataremos sistemas mais simples do que os estudados na sessão anterior, porém agora inseriremos ruído ao sistema e utilizaremos as expressões reduzidas encontradas pelo algoritmo de Ritt para realizar a identificação dos parâmetros.

Como visto no capítulo 3 se encontrarmos ao término do algoritmo de Ritt expressões reduzidas da forma

$$Z_i(y(t), u(t)) = \varphi_i(y(t), u(t))\theta_i + \nu_i(e(t)), \quad (4.67)$$

o parâmetro θ_i pode ser estimado diretamente utilizando o estimador de mínimos quadrados (2.27) ou o estimador por variáveis instrumentais (2.44), onde cada parâmetro θ_i será identificado separadamente. Para isto precisamos definir um preditor. O preditor escolhido é conhecido como *simulation model predictor* (LJUNG, 1999), que nada mais é que o preditor obtido desconsiderando o ruído do processo:

$$\hat{Z}_i(t, \theta_i) = \varphi_i(y(t), u(t))\theta_i. \quad (4.68)$$

Assim a função custo a ser minimizada será dada por:

$$J_N(\theta_i) = \sum_{t=1}^N [Z_i(y(t), u(t)) - \hat{Z}_i(t, \theta_i)]^2. \quad (4.69)$$

Exemplo 4.2.1. (BAZANELLA; RUI, 2012)

Considere o seguinte conjunto de modelos lineares com parametrização não linear:

$$\mathcal{M}_1 : \quad y(t) = \theta_1 y(t-1) + \theta_1^2 u(t-1) + e(t), \quad (4.70)$$

Definindo o ranking

$$e^{(\cdot)} \prec u^{(\cdot)} \prec y^{(\cdot)} \prec \theta_1^{(\cdot)}. \quad (4.71)$$

Aplicando o algoritmo de Ritt com este ranking encontramos o seguinte modelo equivalente a (4.70),

$$\underbrace{y\ddot{u} - \dot{y}\dot{u}}_{Z(y,u)} = \underbrace{(\dot{y}\ddot{u} - \ddot{y}\dot{u})}_{\varphi(y,u)} \theta_1 + \underbrace{(\ddot{u}e - \dot{u}\dot{e})}_{\nu(e)}, \quad (4.72)$$

onde omitimos a dependência temporal dos sinais $y(t)$, $u(t)$ e $e(t)$ e também estamos utilizando a notação (3.38).

Como visto no Capítulo 3, os sinais $y(t)$, $u(t)$ e o parâmetro θ_1 satisfazem (4.72) se e somente se eles também satisfizerem a equação original (4.70). Deste modo podemos calcular a estimativa por mínimos quadrados através da relação (2.27)

$$\hat{\theta}_N = \left[\sum_{t=1}^N \varphi(y(t), u(t))^2 \right]^{-1} \sum_{t=1}^N \varphi(y(t), u(t)) Z(y(t), u(t)).$$

A propriedade fundamental de um estimador é a consistência. Como visto no capítulo 2, para o estimador de MQ ser consistente é necessário que o vetor regressor seja ortogonal (descorrelacionado) com o ruído, ou seja

$$\bar{E}[\varphi(y(t), u(t))\nu(e(t))] = \mathbf{0}. \quad (4.73)$$

Substituindo (4.72) em (4.73) temos

$$\begin{aligned}
\bar{E}[\varphi(y(t), u(t))\nu(e(t))] &= \bar{E}[(\dot{y}\ddot{u} - \dot{y}\dot{u}) \times (\ddot{u}e - \dot{u}\dot{e})] \\
&= \bar{E}[(\dot{y}\ddot{u} \times \ddot{u}e)] - \bar{E}[(\dot{y}\dot{u} \times \ddot{u}e)] - \bar{E}[(\dot{y}\ddot{u} \times \dot{u}\dot{e})] + \bar{E}[(\dot{y}\dot{u} \times \dot{u}\dot{e})] \\
&= \bar{E}[(\dot{y}\ddot{u}) \times (-\dot{u}\dot{e})] \\
&\neq 0.
\end{aligned} \tag{4.74}$$

É fácil de ver que $\bar{E}[(\dot{y}\ddot{u}) \times (-\dot{u}\dot{e})] \neq 0$; como

$$y = \theta_1 \dot{y} + \theta_1^2 \dot{u} + e, \tag{4.75}$$

temos que

$$\dot{y} = \theta_1 \ddot{y} + \theta_1^2 \ddot{u} + \dot{e}, \tag{4.76}$$

assim y é função apenas de e , do mesmo modo que \dot{y} é função apenas de \dot{e} . Deste modo como (4.73) não é satisfeita o estimador de mínimos quadrados será polarizado.

•

Exemplo 4.2.2. (BAZANELLA; RUI, 2012)

Considerando o seguinte modelo FIR não linear.

$$\mathcal{M}_2 : \quad y(t) = \theta_1 u^2(t) + \theta_1^2 u(t-1) + e(t), \tag{4.77}$$

Novamente, aplicando o algoritmo de Ritt em (4.77), com o ranking (4.71), encontramos:

$$\underbrace{y\ddot{u} - \dot{y}\dot{u}}_{Z(y,u)} = \underbrace{(u^2\ddot{u} - \dot{u}^3)}_{\varphi(y,u)} \theta_1 + \underbrace{(\ddot{u}e - \dot{u}\dot{e})}_{\nu(e)} \tag{4.78}$$

e então, do mesmo modo que em (4.74)

$$\begin{aligned}
\bar{E}[\varphi(y(t), u(t))\nu(e(t))] &= \bar{E}[(u^2\ddot{u} - \dot{u}^3) \times (\ddot{u}e - \dot{u}\dot{e})] \\
&= 0,
\end{aligned} \tag{4.79}$$

já que por hipótese a entrada $u(t)$ é descorrelacionada com o ruído. Logo, para a classe de modelo (4.2.2) o estimador de MQ será não tendencioso.

•

O algoritmo de Ritt encontrou uma forma reduzida linear em θ para as classes de modelos \mathcal{M}_1 e \mathcal{M}_2 , assim, segundo o Teorema 8, o sistema é *globalmente identificável*. Apresentaremos os estudos numéricos das propriedades de consistência do estimador de mínimos quadrados na próxima seção, onde quando o estimador de mínimos quadrados for tendencioso será necessária a utilização do estimador das variáveis instrumentais (LJUNG, 1999),(AGUIRRE, 2004) para que a estimativa tenda ao valor real do parâmetro a ser identificado.

4.3 Estudos numéricos

Nesta sessão apresentaremos estudos numéricos de três exemplos propostos. Os dois primeiros são sistemas cujas estruturas de modelo foram apresentadas na seção anterior, estruturas (4.70) e (4.77), onde os valores reais para os parâmetros são $\theta_1 = 0,5$ para \mathcal{S}_1 e $\theta_1 = 1$ para \mathcal{S}_2 :

$$\mathcal{S}_1 : \quad y(t) = 0,5y(t-1) + 0,25u(t-1) + e(t), \quad (4.80)$$

$$\mathcal{S}_2 : \quad y(t) = u^2(t) + u(t-1) + e(t). \quad (4.81)$$

Em cada um dos casos a identificação foi feita para dois sinais de entradas distintos chamados $u_1(t)$ e $u_2(t)$, sendo $u_1(t)$ um sinal aleatório com distribuição uniforme entre $[0 : 1]$

$$u_1(t) \stackrel{\Delta}{=} \text{rand}[0 : 1] \quad (4.82)$$

e $u_2(t)$ uma combinação de duas senoides dada por:

$$u_2(t) \stackrel{\Delta}{=} \sin(t/3) + \sin(t). \quad (4.83)$$

Em todas as simulações o ruído $e(t)$ é ruído branco Gaussiano com média zero e desvio padrão $\sigma_e = 0.1$ - chamado $e_1(t)$ - ou $\sigma_e = 0.01$ - chamado $e_2(t)$. Em cada caso, para estudar as propriedades estatísticas das estimativas, simulou-se o sistema 500 vezes (500 rodadas de Monte Carlo) para um conjunto de dados $N = 10.000$. Nos dois casos o sistema real pertence ao conjunto de modelos: $\mathcal{S}_1 \in \mathcal{M}_1$ e $\mathcal{S}_2 \in \mathcal{M}_2$. O outro exemplo é:

$$\mathcal{S}_3 : \quad y(t) = -0,5y(t-1) + u(t-1) + u(t) + e(t) + 0,5e(t-1), \quad (4.84)$$

que será identificado através da estrutura

$$\mathcal{M}_3 : \quad y(t) = -\theta_1 y(t-1) + \theta_2 u(t-1) + u(t) + e(t) + \theta_1 e(t-1), \quad (4.85)$$

Novamente, para o terceiro exemplo o sistema real pertence ao conjunto de modelo: $\mathcal{S}_3 \in \mathcal{M}_3$.

4.3.1 Resultados Sistema \mathcal{S}_1

As equações resultantes do algoritmo de Ritt para o modelo \mathcal{M}_1 foram apresentadas em (4.72). Como visto anteriormente a classe de modelos \mathcal{M}_1 é linear com parametrização não linear. Apesar de ser um modelo simples com apenas um parâmetro, a não linearidade na parametrização torna a função custo não convexa.

Como visto no capítulo 2 a identificação pode ser feita através da minimização da função custo

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t, t-1|\theta))^2, \quad (4.86)$$

Ao invés disto, propomos utilizar o algoritmo o Ritt seguido da minimização da função custo (4.69). Ou seja, utilizamos o algoritmo de Ritt para encontrar as expressões reduzidas da forma de (4.67), e então, minimizando a função custo (4.69) encontramos uma estimativa para os parâmetros θ .

As duas funções custo resultantes de um experimento utilizando a entrada $u_1(t)$, definida em (4.82), são apresentadas na Figura 4.1. Podemos observar que a função custo $V_N(\theta, Z^N)$ apresenta um mínimo local próximo a $\theta = -0,8$ e um mínimo global em $\theta = 0,5$. Esse mínimo local poderia ser um incômodo na hora da identificação, enquanto a função custo (4.69) é convexa, já que esta é uma função quadrática por construção assim apresenta apenas um mínimo em $\theta = 0,5$.

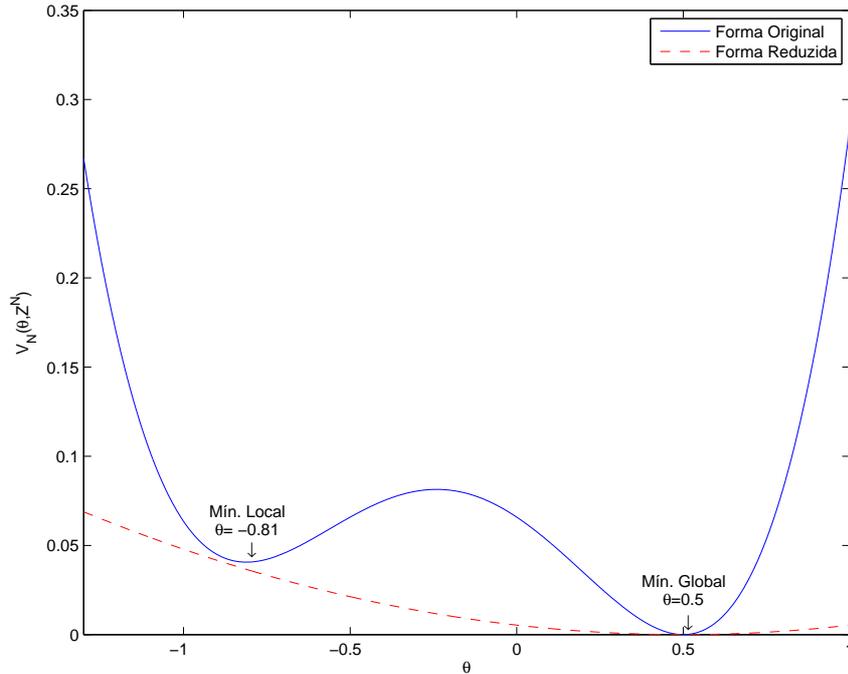
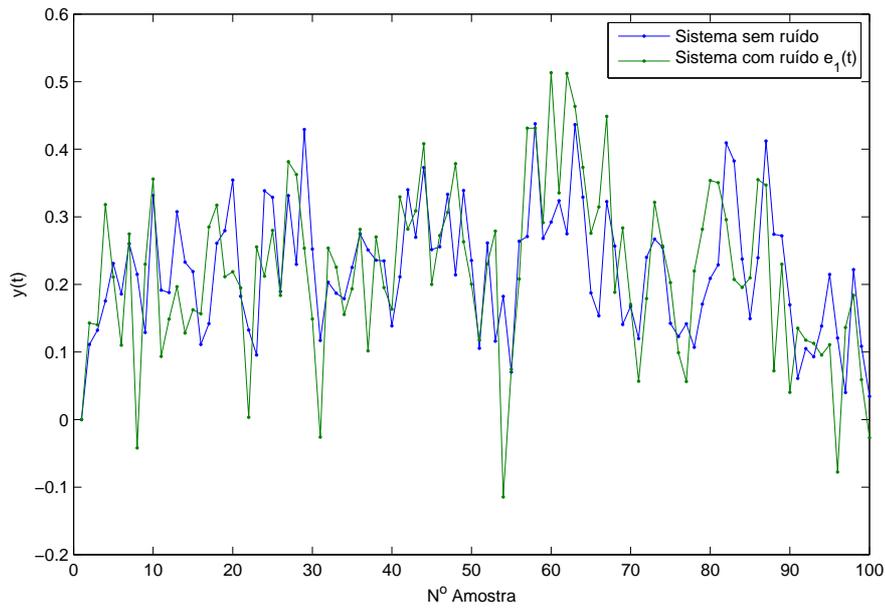
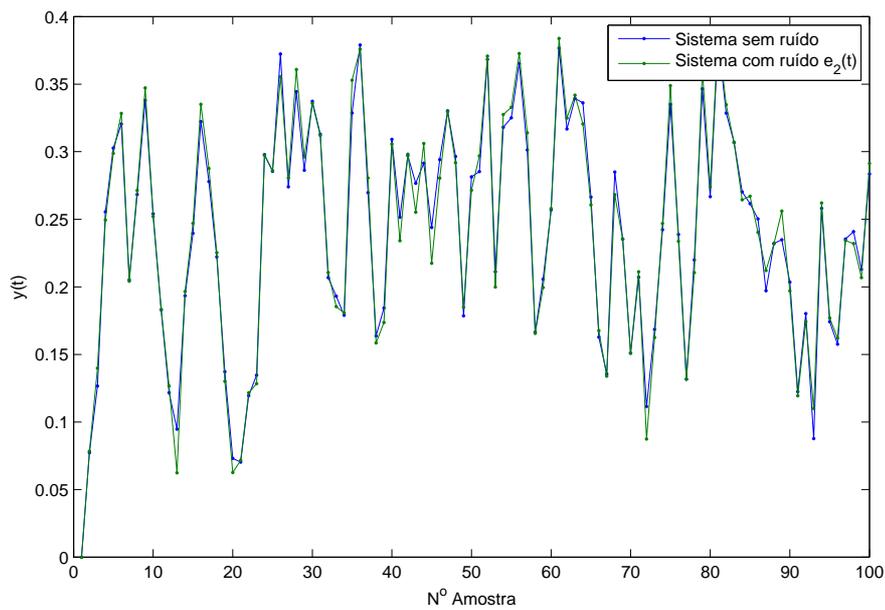


Figura 4.1: Funções custo $V_N(\theta, Z^N)$ linha cheia e $J_N(\theta, Z^N)$ linha tracejada para o sistema \mathcal{S}_1

A Tabela 4.1 apresenta os valores para as estimativas encontradas para o sistema \mathcal{S}_1 através dos métodos de mínimos quadrados e variáveis instrumentais. Como mostrado em (4.74) o estimador de mínimos quadrados para a classe de modelos (4.70) seria polarizado devido à correlação entre o vetor regressor e o ruído; essa polarização pode ser observada nas colunas dois e quatro da Tabela 4.1, onde nota-se que a média dos estimadores é significativamente diferente do valor real θ_1 . Por outro lado, para o estimador das variáveis instrumentais, o resultado obtido através da expressão (2.44) corrige essa polarização na estimativa (colunas 6 e 8). Como mencionado no início do capítulo, o sistema foi simulado 500 vezes, porém como o instrumento utilizado para calcular o estimador das variáveis instrumentais foi uma realização independente das demais, precisou-se simular uma vez mais o sistema para compor o instrumento $\zeta(t)$.

Observe que, para o ruído $e_1(t)$, o desvio do estimador de mínimos quadrados com relação ao valor real é de aproximadamente 20%, enquanto, para o ruído $e_2(t)$, o desvio é inferior a 1%; essa diferença é devida à variância do ruído. Podemos observar pela Figura 4.2a que o sinal de saída do sistema com o ruído $e_1(t)$ varia muito quando comparado ao sinal sem ruído, por outro lado para o caso da saída com o ruído $e_2(t)$ (Figura 4.2b) a variação é muito pequena.

(a) Sinal com ruído $e_1(t)$.(b) Sinal com ruído $e_2(t)$.Figura 4.2: Sinal $y(t)$ para o sistema \mathcal{S}_1 .Tabela 4.1: Média e desvio padrão para os estimadores, Sistema \mathcal{S}_1

	Mínimos Quadrados				Variável Instrumental			
	Ruído $\sigma = 0,1$		Ruído $\sigma = 0,01$		Ruído $\sigma = 0,1$		Ruído $\sigma = 0,01$	
	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$
u_1	0,40231	0,07135	0,49873	0,00784	0,50018	0,09268	0,49991	0,00790
u_2	0,46676	0,03492	0,49984	0,0035	0,49679	0,03859	0,49990	0,00335

4.3.2 Resultados Sistema \mathcal{S}_2

Como para o sistema \mathcal{S}_1 , começaremos apresentando as funções custo $J_N(\theta, Z^N)$ e $V_N(\theta, Z^N)$ (Figura 4.3), Podemos observar que a função custo da forma original possui dois mínimos; um mínimo local em $\theta = -1,45$ e o mínimo global em $\theta = 1$, enquanto a função da forma reduzida é convexa com um mínimo $\theta = 1$.

Para este exemplo, mostrou-se em (4.79) que o estimador de mínimos quadrados não seria polarizado. Podemos confirmar esse resultado observando a tabela 4.2, onde apresentamos os valores dos estimadores encontrados para o sistema \mathcal{S}_2 .

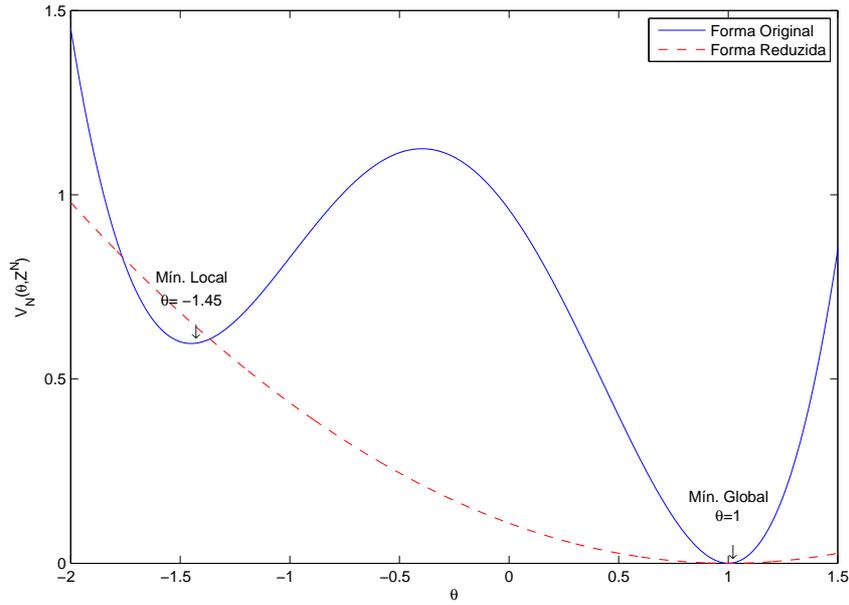


Figura 4.3: Funções custo $V_N(\theta, Z^N)$ linha cheia e $J_N(\theta, Z^N)$ linha tracejada para o sistema \mathcal{S}_2

Tabela 4.2: Média e desvio padrão para os estimadores. Sistema \mathcal{S}_2

	Mínimos Quadrados				Variável Instrumental			
	Ruído $\sigma = 0.1$		Ruído $\sigma = 0.01$		Ruído $\sigma = 0.1$		Ruído $\sigma = 0.01$	
	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$	$\mu_{\hat{\theta}_1}$	$\sigma_{\hat{\theta}_1}$
u_1	0.99991	0.00321	0.99998	0.00321	0.99969	0.03541	1.00013	0.00360
u_2	0.99927	0.01253	0.99990	0.00111	1.00076	0.01078	1.00001	0.00119

4.3.3 Resultados Sistema \mathcal{S}_3

Aplicando o algoritmo de Ritt na expressão (4.85) com o ranking

$$\left\{ u^{(\cdot)} \prec y^{(\cdot)} \prec \theta_1 \prec \dot{\theta}_1 \prec \theta_2 \prec \dot{\theta}_2 \right\} \quad (4.87)$$

encontramos as seguintes expressões reduzidas para θ_1 e θ_2 , respectivamente,

$$\underbrace{y\ddot{u} + \dot{u}^2 - \dot{y}\dot{u} - \ddot{u}u}_{Z_1(y,u)} = \underbrace{(\dot{y}\dot{u} - \dot{y}\ddot{u} - \dot{u}\ddot{e} + \ddot{u}\dot{e})}_{\varphi_1(y,u)} \theta_1 \underbrace{-\dot{u}\dot{e} + \ddot{u}e}_{\nu_1(y,u)} \quad (4.88)$$

e

$$\underbrace{\dot{y}^2 - \dot{y}\ddot{u} - \dot{y}\ddot{y} + \dot{y}\ddot{u}}_{Z_2(y,u)} = - \underbrace{(\ddot{y}\ddot{u} - \ddot{y}\ddot{u} - \ddot{u}\ddot{e} + \ddot{u}\ddot{e})}_{\varphi_2(y,u)} \theta_2 + \underbrace{2\dot{y}\dot{e} - \dot{e}^2 - \dot{u}\dot{e} - \dot{y}\dot{e} - \dot{y}\dot{e} + \dot{e}\dot{e} + \dot{u}\dot{e}}_{\nu_2(y,u)} \quad (4.89)$$

Para o sistema \mathcal{S}_3 a função custo será uma superfície (Figura 4.4). Utilizando o algoritmo de Ritt podemos dividir a classe de modelos original em duas regressões lineares nos parâmetros que por construção são quadráticas e convexas. As Figuras 4.5a e 4.5b apresentam as funções custo das formas reduzidas para θ_1 e θ_2 respectivamente. Assim, como o algoritmo chegou a duas expressões lineares nos parâmetros, segundo o Teorema 8, o sistema é *globalmente identificável*.

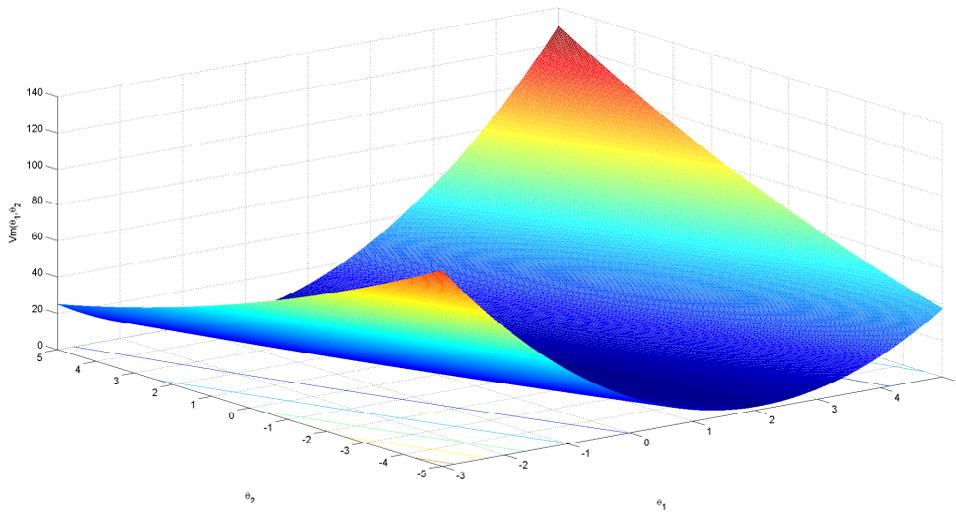
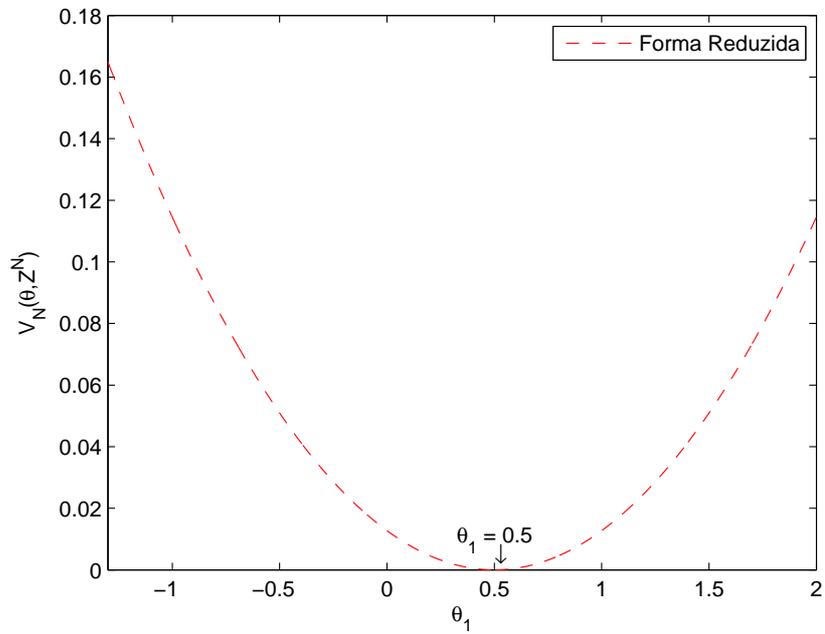
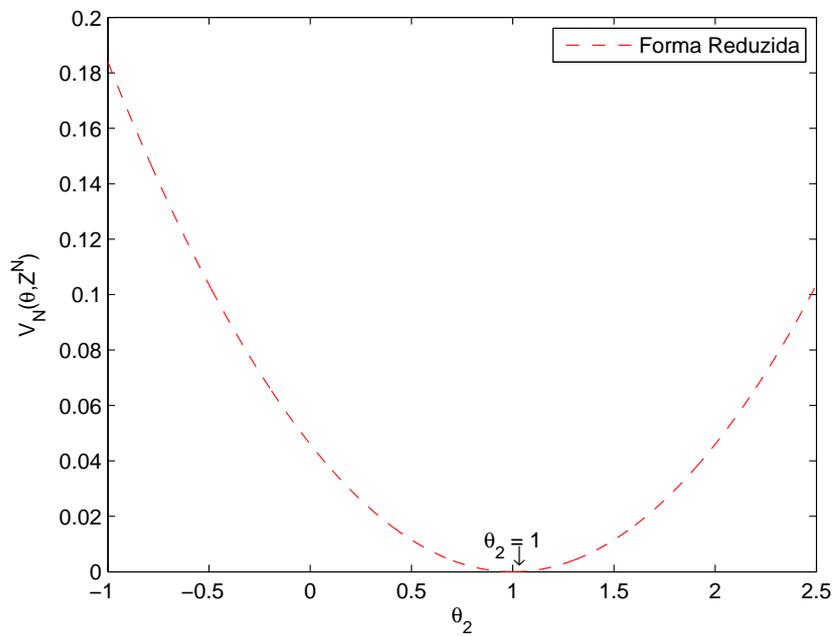


Figura 4.4: Função custo sistema \mathcal{S}_3

Observe que o regressor $\varphi_1(y, u) = -\varphi_2(y, u)$ e que ambos são correlacionados com os ruídos $\nu_1(y, u)$ e $\nu_2(y, u)$ respectivamente, já que na estrutura de modelo (4.85) y é correlacionado com e e \dot{e} ; do mesmo modo \dot{y} com \dot{e} e \ddot{e} , deste modo os estimadores seriam polarizados. Essa correlação fica evidente nas Figuras 4.6a e 4.7a, onde nota-se uma visível polarização nos estimadores. O mesmo ocorre para o ruído $e_2(t)$, figuras 4.6b e 4.7b. Note também que dentro dos vetores regressores $\varphi_1(y, u)$ e $\varphi_2(y, u)$ aparecem termos dependentes do ruído - $(-\ddot{u}\ddot{e} + \ddot{u}\ddot{e})$ e $(\ddot{u}\ddot{e} - \ddot{u}\ddot{e})$, respectivamente. Por esse motivo utilizaremos o preditor *modelo de simulação*, definido na Definição 2.3.1, para realizar a identificação.

As figuras 4.8a, 4.9a, 4.8b e 4.9b apresentam a dispersão dos estimadores utilizando o estimador das variáveis instrumentais. Nota-se que para o ruído $e_1(t)$, figuras 4.8a e 4.9a, parece ainda existir uma polarização no estimador. Essa polarização aparece devido à correlação entre o instrumento utilizado e o ruído da forma reduzida. Como podemos observar, a expressão do ruído, principalmente para (4.89), torna-se bastante complexa e ainda depende dos parâmetros, assim uma análise mais detalhada deve ser feita para conseguir encontrar um instrumento que não seja correlacionado com o ruído da forma reduzida.

Essa correlação fica ainda mais evidente quando aumentamos o nível do ruído. As figuras 4.10a e 4.10b apresentam a dispersão dos parâmetros para os estimadores de

(a) Função custo $J_N(\theta_1, Z^N)$ (b) Função custo $J_N(\theta_2, Z^N)$ Figura 4.5: Função custo $J_N(\theta, Z^N)$ para sistema \mathcal{S}_3

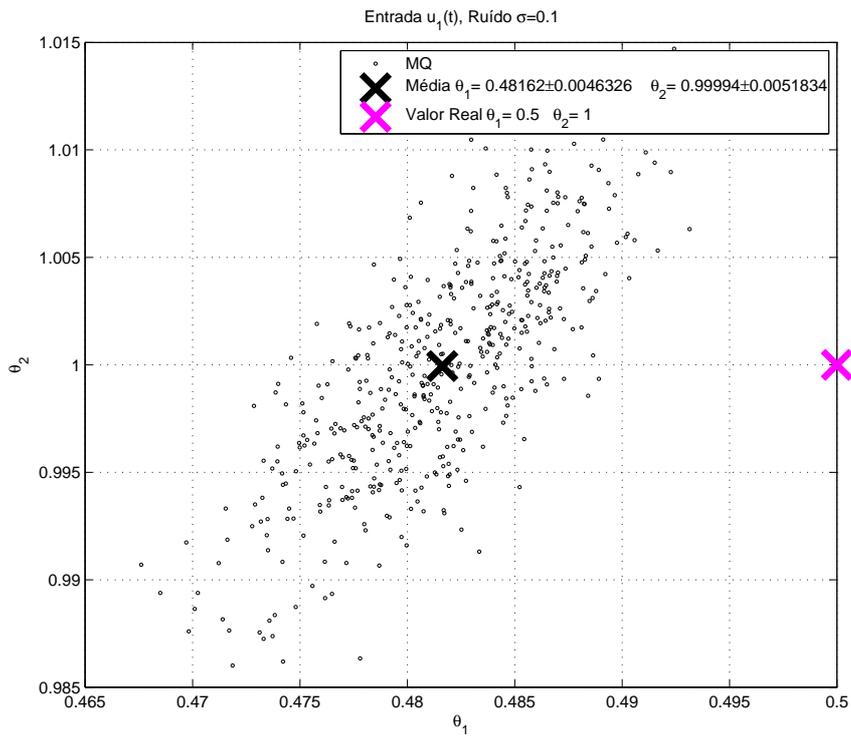
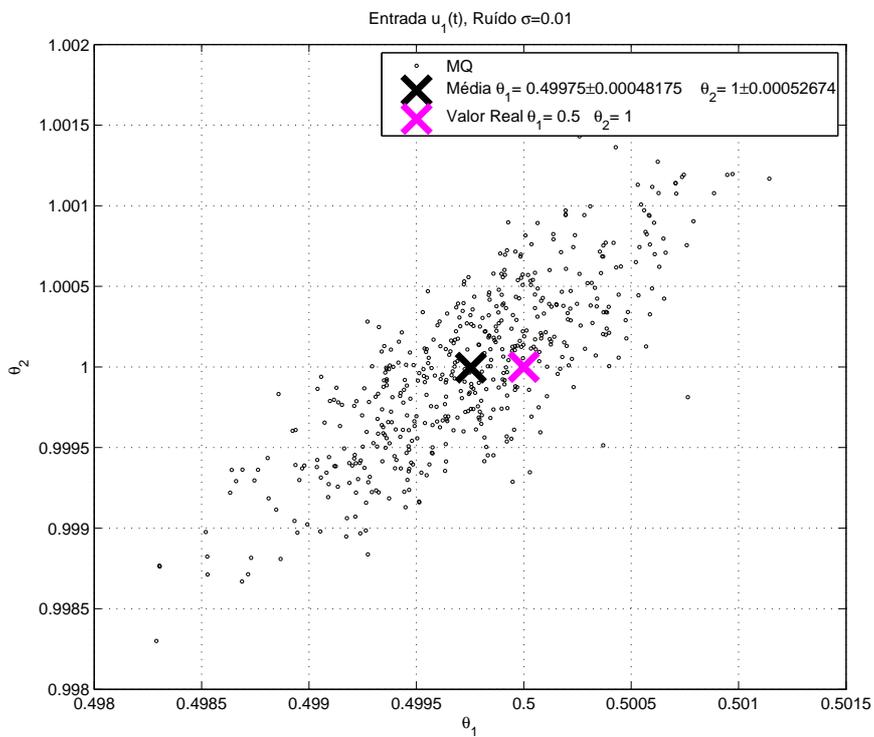
(a) Ruído $e_1(t)$ (b) Ruído $e_2(t)$

Figura 4.6: Dispersão dos parâmetros utilizando o estimador de MQ para entrada $u_1(t)$

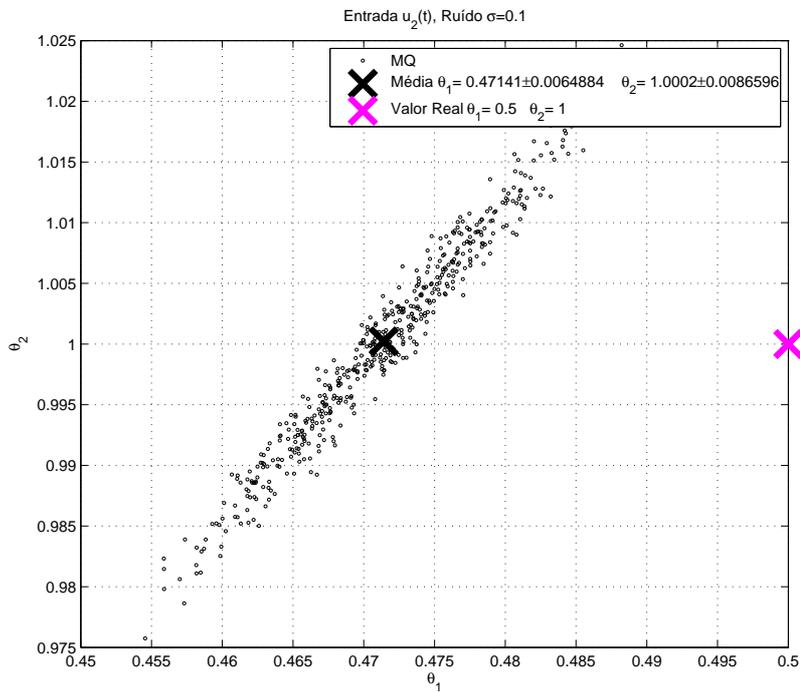
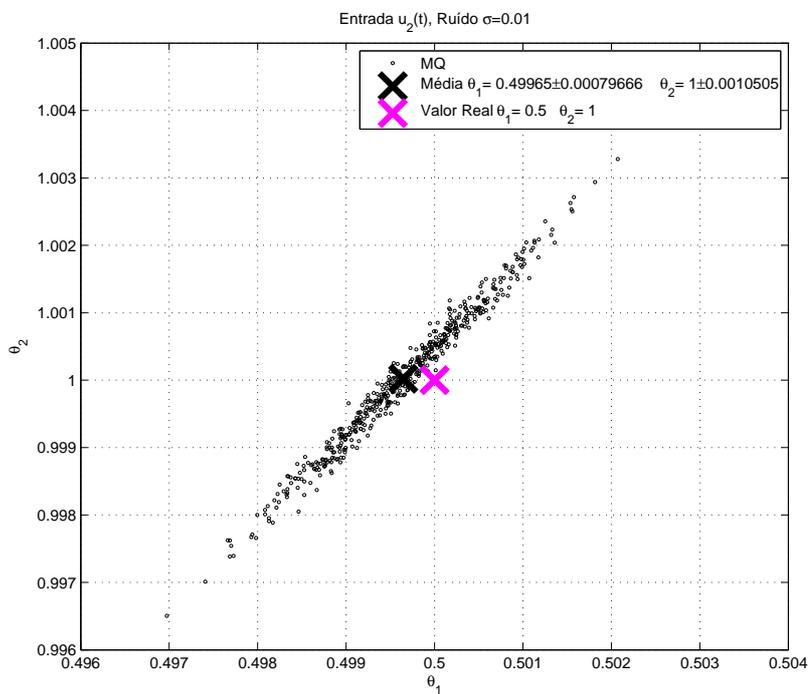
(a) Ruído $e_1(t)$ (b) Ruído $e_2(t)$

Figura 4.7: Dispersão dos parâmetros utilizando o estimador de MQ para entrada $u_2(t)$

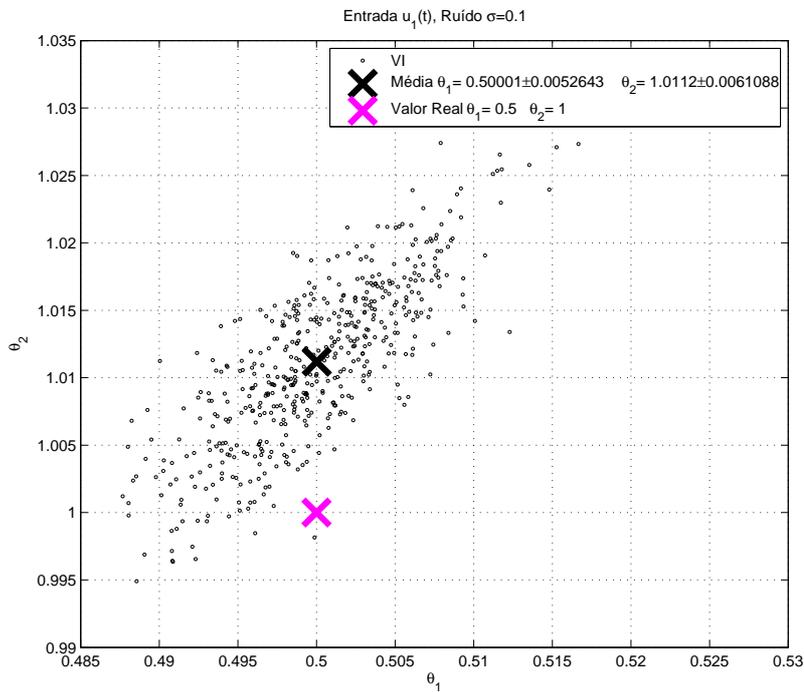
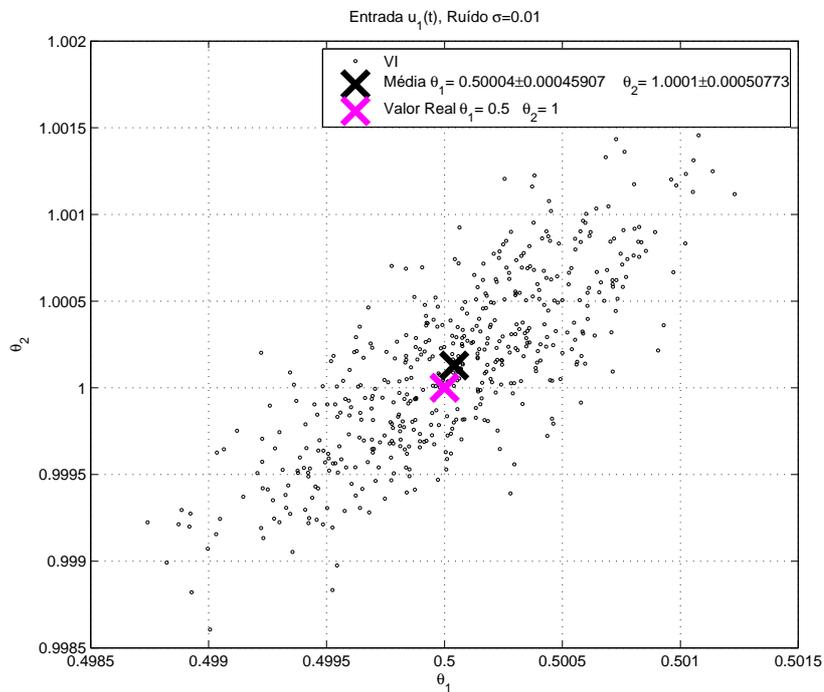
(a) Ruído $e_1(t)$ (b) Ruído $e_2(t)$

Figura 4.8: Dispersão dos parâmetros utilizando o estimador de VI para entrada $u_1(t)$

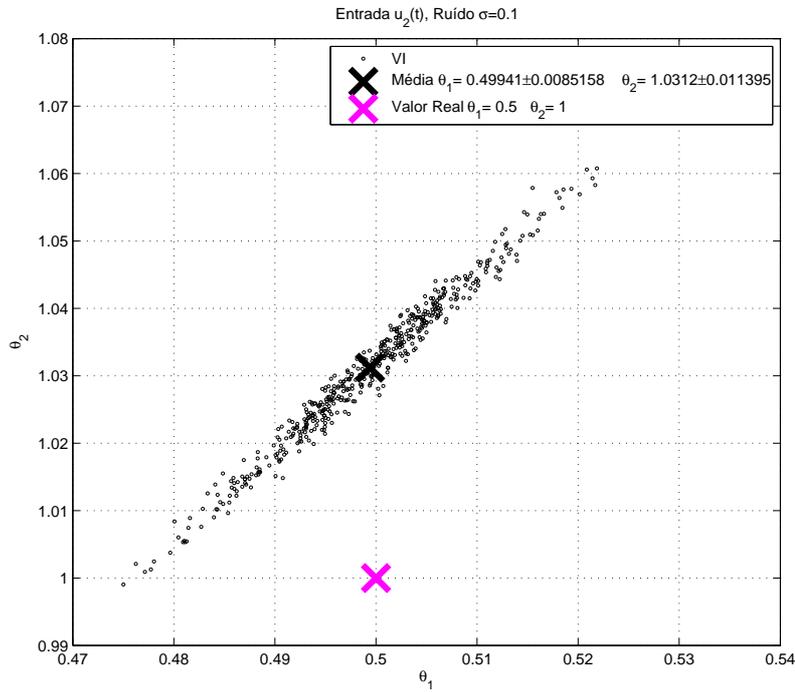
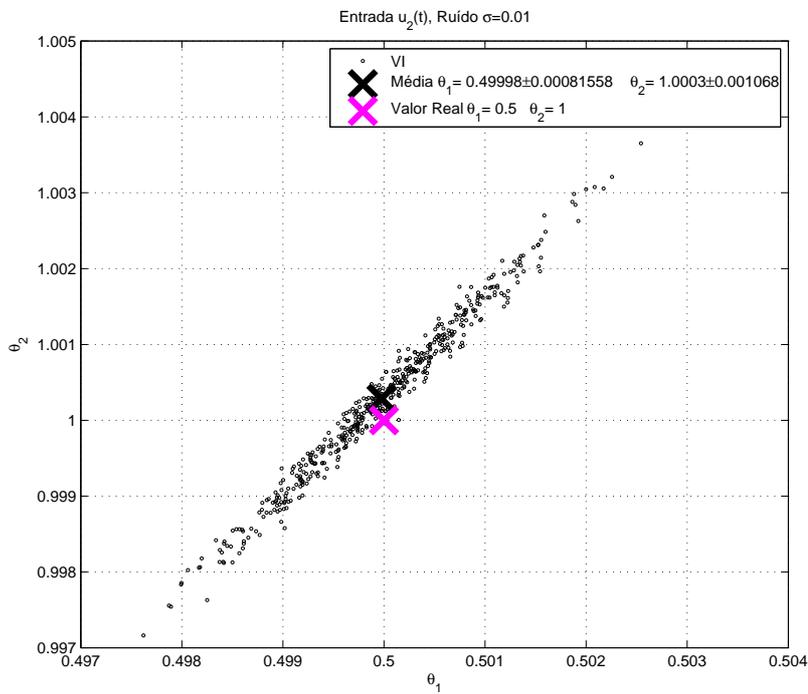
(a) Ruído $e_1(t)$ (b) Ruído $e_2(t)$

Figura 4.9: Dispersão dos parâmetros utilizando o estimador de VI para entrada $u_2(t)$

mínimos quadrados e variáveis instrumentais para as entradas u_1 e u_2 , com $e(t)$ ruído branco gaussiano com $\sigma_e = 0,6$. Observe que o estimador de mínimos quadrados apresenta polarização na direção de θ_1 , enquanto para o parâmetro θ_2 não parece existir polarização ($E[\hat{\theta}_1^{MQ}] = 0,3956$ e $E[\hat{\theta}_2^{MQ}] = 0,99964$ para a entrada u_1 e $E[\hat{\theta}_1^{MQ}] = 0,38211$ e $E[\hat{\theta}_2^{MQ}] = 0,99973$ para a entrada u_2). O mesmo acontece para o estimador por variáveis instrumentais, porém observa-se que não parece existir polarização em θ_1 , onde o erro entre o valor real e o valor médio do estimador é menor que 0,1% para ambos os sinais de entrada. Por outro lado, para θ_2 o erro entre o valor real e o valor médio do estimador é maior que 20%.

Por fim, mesmo existindo uma polarização no estimador por variáveis instrumentais se observa que para sinais de ruído de pequena variância, Figuras (4.8b e 4.9b) o erro entre o valor real do parâmetro e o valor médio do estimador é inferior a 0.1%.

4.4 Considerações Finais

Neste capítulo apresentou-se os resultados da utilização do algoritmo de Ritt no estudo da identificabilidade de estruturas, assim como na identificação de sistemas. Referente ao estudo da identificabilidade de estruturas de modelos, observou-se que para estruturas com mais de quatro parâmetros a implementação do algoritmo de Ritt em computadores pessoais se torna inviável visto que a complexidade das expressões analíticas resultantes tende a crescer. Porém, utilizando o algoritmo de Ritt na forma alternativa conseguimos determinar se estruturas com até sete parâmetros são identificáveis.

Quanto à utilização do algoritmo de Ritt para a identificação de sistemas discretos, foram apresentados sistemas mais simples com até dois parâmetros, porém se acrescentou ruído aos sistemas. A identificação feita a partir das expressões reduzidas resultantes do algoritmo de Ritt mostrou-se bastante simples de ser implementada e através do método de mínimos quadrados ou variáveis instrumentais foi possível identificar os parâmetros do modelo. Porém, constatou-se que o vetor regressor obtido através da forma reduzida encontrada pelo algoritmo de Ritt, em alguns casos, era correlacionado com o ruído tornando a estimativa tendenciosa. Por este motivo utilizou-se o estimador das variáveis instrumentais para eliminar a polarização. Para esse caso observou-se que mesmo com o estimador das variáveis instrumentais a estimativa para o parâmetro, em alguns casos, era polarizada. Entretanto, como observado no sistema \mathcal{S}_3 , nem todas as estimativas eram polarizadas. Como o algoritmo de Ritt transforma uma estrutura predeterminada em regressões lineares parâmetro por parâmetro o instrumento utilizado para realizar a identificação de um dos parâmetros era descorrelacionado com o ruído referente a este mesmo parâmetro, enquanto, para o outro parâmetro existia uma correlação entre o ruído e o instrumento.

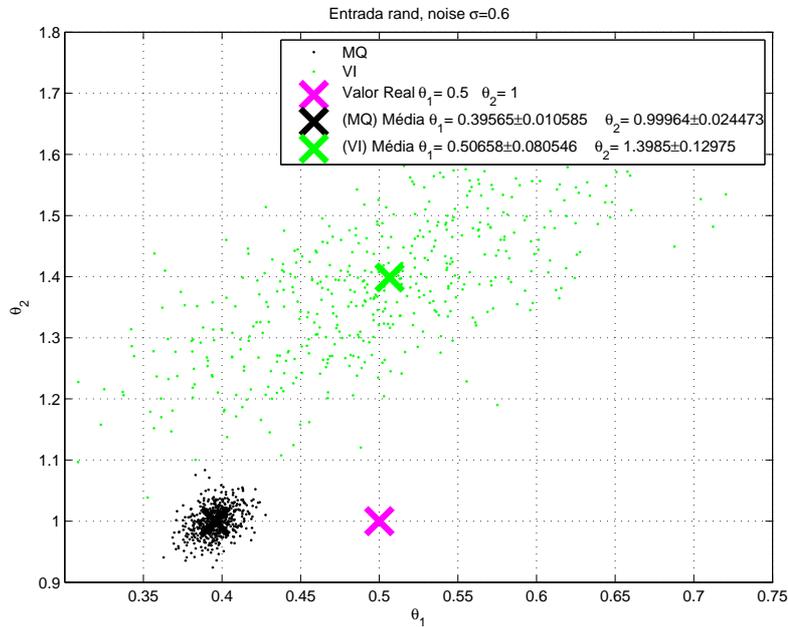
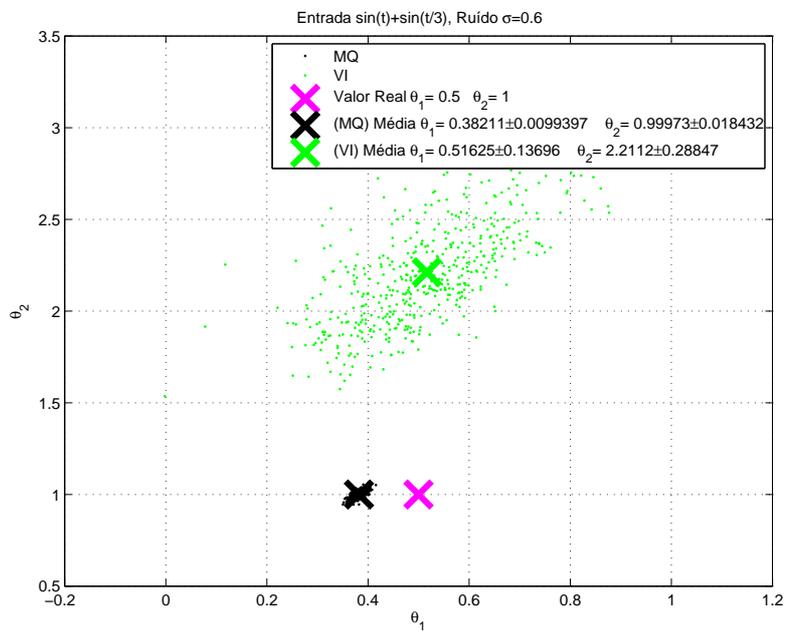
(a) Entrada u_1 (b) Entrada u_2

Figura 4.10: Dispersão dos parâmetros utilizando o estimador de MQ e VI com ruído $\sigma_e = 0,6$

5 CONCLUSÃO

Neste trabalho foi apresentado o algoritmo de Ritt para o estudo da identificabilidade e identificação de estruturas de modelos polinomiais de tempo contínuo e discreto. O algoritmo foi implementado na plataforma MuPad do MATLAB e executado em um computador pessoal cujas especificações de hardware e software foram apresentadas no capítulo 4. O algoritmo foi implementado com sucesso, tanto para sistemas de tempo contínuo, quanto sistemas de tempo discreto.

Durante a realização do trabalho verificou-se que para estruturas com mais de quatro parâmetros o custo computacional necessário para se obter um conjunto característico da estrutura inicial tornava a execução do algoritmo inviável na plataforma utilizada. Para estes modelos após um certo número de iterações do algoritmo a memória RAM do computador utilizado se esgotava fazendo com que o algoritmo utilizasse a memória do HD, tornando assim a execução do algoritmo bastante demorada. Por esse motivo propôs-se uma forma alternativa para a utilização do algoritmo de Ritt no teste de identificabilidade de estruturas de modelo polinomiais. A forma proposta utiliza o algoritmo de Ritt para encontrar uma forma equivalente da estrutura original que só contenha variáveis conhecidas do sistema – entrada, saída e parâmetros. Assim, a partir da matriz de informação da estrutura equivalente consegue-se determinar se a estrutura original é identificável.

Mesmo com a limitação computacional pode-se estudar maneiras de tornar o algoritmo mais eficiente do ponto de vista computacional. Uma alternativa é a utilização de supercomputadores ou clusters para execução do algoritmo. Para isso a implementação do algoritmo deve ser feita em uma linguagem como C ou Fortran. Já existem pacotes capazes de tratar problemas de álgebra diferencial implementados em C (FRANÇOIS; FRANÇOIS, 2009) que poderiam ser utilizados. Porém para se obter o máximo desempenho nesses tipos de computadores é necessário, além de estender esses pacotes para utilização de processamento paralelo ou até mesmo GPU (do inglês *Graphics processing unit*), estudar como paralelizar o algoritmo de Ritt. Para o caso de sistemas de tempo discreto, a adaptação do algoritmo de Ritt é bastante recente (LYZELL et al., 2009), por isso ainda não se encontram rotinas em C ou Fortran para tratar sistemas desse tipo. Outra possibilidade bastante interessante é analisar detalhadamente o algoritmo de modo a tentar estendê-lo para outros tipos de classes de modelo, ou ainda, tentar encontrar outras classes que possam ser escritas na forma polinomial.

De um ponto de vista teórico, os conceitos de identificabilidade e informatividade para sistemas não lineares, assim como nos conceitos de persistência de excitação e riqueza do sinal ainda não são tão claramente estabelecidos. Esses conceitos parecem estar ainda bastante atrelados e a separação destes se faz necessária para uma melhor

compreensão desse tipo de sistema, de forma similar à teoria de identificabilidade de sistemas lineares (LJUNG; GLAD, 1994). Neste trabalho foi dada uma definição precisa de identificabilidade para parametrizações não-lineares.

A aplicação do algoritmo em sistemas discretos é bastante incipiente, os únicos trabalhos encontrados são (LYZELL et al., 2009) e (LYZELL et al., 2011). Nestes trabalhos são apresentados o algoritmo de Ritt para sistemas de tempo discreto e a sua utilização na identificação de modelos. Naqueles trabalhos, o uso do algoritmo para a identificação foi feita somente para sistemas FIR, enquanto que no presente trabalho foi realizada a identificação de um sistema OE (Sistema \mathcal{S}_3). Para este caso observou-se que a regressão linear resultante do algoritmo de Ritt pode ser usada para estimar os parâmetros, porém um estudo mais aprofundado deve ser feito com o objetivo de encontrar um instrumento que resulte em uma estimativa não polarizada.

Por fim, para sistemas de tempo discreto, ainda existem diversas questões teóricas em aberto, como por exemplo, não se sabe para quais estruturas de modelos podemos garantir que o algoritmo de Ritt resulte em uma regressão linear, muito embora se obtenha ótimos resultados para classes de modelos não lineares com não linearidades polinomiais (LYZELL et al., 2011)(BAZANELLA; RUI, 2012). Como apresentado na seção 3.4, a presente teoria não fornece condições necessárias e suficientes para garantir a identificabilidade de estruturas como acontece para o algoritmo de Ritt em sistemas de tempo contínuo (LJUNG; GLAD, 1994), onde os Teoremas 5, 6 e 7 fornecem as condições necessárias e suficientes para que uma estrutura de modelo seja identificável.

REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas**. Belo Horizonte: UFMG, 2004.
- AUDOLY, S. et al. Global identifiability of nonlinear models of biological systems. **IEEE Transactions on Biomedical Engineering**, New York, v.48, n.1, p.55–65, Jan. 2001.
- BAZANELLA, A.; BOMBOIS, X.; GEVERS, M. Necessary and sufficient conditions for uniqueness of the minimum in Prediction Error Identification. **Automatica**, Tarrytown, v.48, n.8, p.1621 – 1630, 2012.
- BAZANELLA, A.; RUI, R. Identification of arbitrary model parameterizations: the uses of ritt’s algorithm. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 19., 2012, Campina Grande. **Anais. . .** São Paulo: SBA, 2012. p.3577–3584.
- BEN-ZVI, A. A Computationally Efficient Algorithm for Testing the Identifiability of Polynomial Systems with Applications to Biological Systems. **Industrial & Engineering Chemistry Research**, Washington, v.49, n.13, p.6125–6134, 2010.
- BOULIER, F. et al. Computing representations for radicals of finitely generated differential ideals. **Applicable Algebra in Engineering, Communication and Computing**, New York, v.20, p.73–121, 2009. 10.1007/s00200-009-0091-7.
- CALANDRINI, G. et al. Ritt’s algorithm, Gröbner basis and discretization. **Mathematical and Computer Modelling**, Oxford, v.26, n.5, p.67 – 77, 1997.
- CHIS, O.-T.; BANGA, J. R.; BALSACANTO, E. Structural Identifiability of Systems Biology Models: a critical comparison of methods. **PLoS ONE**, London, v.6, n.11, p.e27755, 11 2011.
- DENIS-VIDAL, L.; JOLY-BLANCHARD, G.; NOIRET, C. Some effective approaches to check the identifiability of uncontrolled nonlinear systems. **Mathematics and Computers in Simulation**, New York, v.57, p.35 – 44, 2001.
- FRANÇOIS, B.; FRANÇOIS, L. Differential algebra and QSSA methods in biochemistry. In: IFAC SYMPOSIUM ON SYSTEM IDENTIFICATION, 15., 2009, St. Malo. **Anais. . .** Laxemburg: IFAC, 2009. v.15, p.33–38.
- GEVERS, M. Identification for Control: from the early achievements to the revival of experiment design. In: IEEE CONFERENCE ON DECISION AND

CONTROL, 2005 AND 2005 EUROPEAN CONTROL CONFERENCE., 44., 2005, Sevilha. **Anais...** New York: IEEE, 2005. v.11, p.1–18.

GEVERS, M. et al. Identification and the Information Matrix: how to get just sufficiently rich? **Automatic Control, IEEE Transactions on**, New York, v.54, n.12, p.2828 –2840, Dec. 2009.

HARDLE, W. **Applied nonparametric regression**. Cambridge: ELT Press, 1990. v.26.

KARLSSON, J.; ANGUELOVA, M.; JIRSTRAND, M. An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems. In: IFAC SYMPOSIUM ON SYSTEM IDENTIFICATION, 16., 2012, Brussels. **Anais...** Laxemburg: IFAC, 2012. v.16, p.941–946.

KOLCHIN, E. **Differential Algebra and Algebraic Groups**. New York: Academic Press, 1973.

LJUNG, L. **System Identification: theory for the user**. 2nd. ed. Englewood Cliffs: Prentice-Hall, 1999.

LJUNG, L. Some aspects on nonlinear system identification. In: IFAC SYMPOSIUM ON SYSTEM IDENTIFICATION, 14., 2006, Newcastle. **Anais...** Laxemburg: IFAC, 2006. p.110–121.

LJUNG, L.; GLAD, T. On global identifiability for arbitrary model parametrizations. **Automatica**, Tarrytown, v.30, n.2, p.265 – 276, 1994.

LYZELL, C. et al. Identification Aspects of Ritt’s Algorithm for Discrete-Time Systems. In: IFAC SYMPOSIUM ON SYSTEM IDENTIFICATION, 15., 2009, Saint-Malo. **Anais...** Laxemburg: IFAC, 2009. p.681–686.

LYZELL, C. et al. Difference algebra and system identification. **Automatica**, Tarrytown, v.47, n.9, p.1896 – 1904, 2011.

MARGARIA, G. et al. Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. **Mathematical Biosciences**, New York, v.174, n.1, p.1 – 26, 2001.

MESHKAT, N.; EISENBERG, M.; III, J. J. D. An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases. **Mathematical Biosciences**, New York, v.222, n.2, p.61 – 72, 2009.

NEMCOVÁ, J. Structural identifiability of polynomial and rational systems. **Mathematical Biosciences**, New York, v.223, n.2, p.83 – 96, 2010.

POHJANPALO, H. System identifiability based on the power series expansion of the solution. **Mathematical Biosciences**, New York, v.41, p.21 – 33, 1978.

RITT, J. **Differential Algebra**. Providence: American Mathematical Society, 1950.

SACCOMANI, M. P.; AUDOLY, S.; D'ANGIÒ, L. Parameter identifiability of nonlinear systems: the role of initial conditions. **Automatica**, Tarrytown, v.39, n.4, p.619 – 632, 2003.

SÖDERSTRÖM, T.; STOICA, P. **System Identification**. Cambridge: Prentice Hall international, 2001.

APÊNDICE A ROTINAS MUPAD

A.1 Rotina principal

Rotina principal, retorna o conjunto característico e o conjunto \mathcal{G}

```

RITT:=proc(F,RANK,thetas)
  local T,A,reduzirA,reduzirFA,a,f,R1,AUX, cont,G;

  begin
    cont:=0;
    T:=FIND_LIDER(F,RANK);
    A:=CONJUNTO(T);
    while(A[2]<>NIL) do
      red_aux:=QUAL_REDUZIR(A[1],A[2]):
      reduzirA:=red_aux[1]:
      reduzirFA:=red_aux[2]:
      a:=linalg::row(A[1], reduzirA);
      f:=linalg::row(A[2], reduzirFA);
      R1:=LEMMA9(a,f,RANK,thetas);
      G[cont]:=R1[2];
      R1[1]:=simplific(R1[1],G,RANK);
      AUX:=linalg::setRow(A[2], reduzirFA, R1[1]);
      T:=linalg::stackMatrix(A[1],AUX);
      T:=ORDEM(T);
      A:=CONJUNTO(T);
    end_while;
    G:=matrix(cont,1,G);
    return(A,G);
  end_proc

```

A.2 Encontra lider

Essa rotina tem como objetivo encontrar os lideres de um conjunto de polinômios.

```

FIND_LIDER:= proc(F,RANK)

  local pol, t2, aux, Pos, cont;
  local Dleader, local leader, local i, n, FF;
  local grau, ordem, iPos, iMin;
  begin
    cont:=0;
    t2 := table():
    n:=nops(RANK):
    for pol in F do
      for i from n downto 1 step 1 do

```

```

    aux:=degree(pol,RANK[i]):
    if is(aux<>0) then
        leader:=RANK[i]:
        Dleader:=degree(pol,RANK[i]):
        Pos:=i:
        cont:=cont+1:
        t2[cont,1] := pol:
        t2[cont,2] := Pos:
        t2[cont,3] := Dleader:
        t2[cont,4] := leader:
        break:
    end_if;
end_for;
end_for;
FF:=matrix(cont,4,t2):

grau:=linalg::col(FF, 3):
ordem:=linalg::col(FF, 2):
cont:=nops(ordem);
for iPos from 1 to cont-1 step 1 do
    iMin:=iPos:
    for i from iPos+1 to cont step 1 do
        if is(ordem[i]<ordem[iMin]) then
            iMin:=i:
        else
            if is(ordem[i]=ordem[iMin] and grau[i]<grau[iMin]) then
                iMin:=i:
            end_if
        end_if:
    end_for:
    if is(iMin<>iPos) then
        FF:=linalg::swapRow(FF,iPos,iMin);
        ordem:=linalg::swapRow(ordem,iPos,iMin);
        grau:=linalg::swapRow(grau,iPos,iMin)
    end_if
end_for:

//Retorna Matriz FF em Ordem segundo RANK onde:  1o Coluna: polinomio
//                                                2o Coluna: posicao relativa rank
//                                                3o Coluna: grau do lider
//                                                4o Coluna: lider
return(FF);
end_proc;

```

A.3 Conjunto

Tem como entrada um conjunto de polinômios \mathcal{F} e retorna um conjunto característico \mathcal{A} de \mathcal{F} e um conjunto $\mathcal{F} \setminus \mathcal{A}$

```

CONJUNTO:= proc(FF)
    local A, aux, nlist, COLremovida;
    local lista, lista2, iPos, i, j, diff_aux;
    local lider_teste, removida, FA, AA, FFA;
    local grau, ordem, iMin;
begin
    A := table():
    removida:=table():

```

```

aux:=0: nlista:=1:
COLremovida:=1:
A[1,1]:=FF[1,1]:
iPos:=0:
ordem:=linalg::col(FF,2):
cont:=nops(ordem):
lista:=matrix([[j $ j = 1..cont]]):

while tamanho(lista)>1 do

  aux:=aux+1:
  A[aux,1]:=FF[lista[1],1]:
  A[aux,2]:=FF[lista[1],2]:
  A[aux,3]:=FF[lista[1],3]:
  A[aux,4]:=FF[lista[1],4]:
  iPos:=lista[1]:
  lista:=linalg::delCol(lista,1):
  lista2:=lista:
  COLremovida:=1:
  for i in lista do
    if is(degree(FF[iPos,1],FF[i,4])=0 and degree(FF[i,1],FF[iPos,4])=0) then
      lider_teste:=17:
      diff_aux:=FF[iPos,4]:
      for j from FF[iPos,2] to FF[i,2] step 1 do
        diff_aux:=D(diff_aux):
        if is(degree(FF[i,1],diff_aux)<>0) then
          lider_teste:=7:
          removida[nlista]:=lista2[COLremovida]:
          nlista:=nlista+1:
          lista2:=linalg::delCol(lista2, COLremovida):
          break
        end_if:
      end_for:
      if is(lider_teste=17) then
        lista2:=lista2:
        COLremovida:=COLremovida+1:
        if nops(lista2)<=1 then
          aux:=aux+1:
          A[aux,1]:=FF[lista2[1],1]:
          A[aux,2]:=FF[lista2[1],2]:
          A[aux,3]:=FF[lista2[1],3]:
          A[aux,4]:=FF[lista2[1],4]:
        end_if
      else;
        if tamanho(lista2)=1 and COLremovida<>1 then
          aux:=aux+1:
          A[aux,1]:=FF[lista2[1],1]:
          A[aux,2]:=FF[lista2[1],2]:
          A[aux,3]:=FF[lista2[1],3]:
          A[aux,4]:=FF[lista2[1],4]:
        end_if;
      end_if:
    else;
      if is(degree(FF[iPos,1],FF[i,4])=0 and degree(FF[i,1],FF[iPos,4])<degree(FF[iPos,1],FF[iPos,4])) then

        lider_teste:=17:
        diff_aux:=FF[iPos,4]:

```

```

for j from FF[iPos,2] to FF[i,2] step 1 do
  diff_aux:=D(diff_aux):
  if is(degree(FF[i,1],diff_aux)<>0) then
    lider_teste:=7:
    removeida[nlista]:=lista2[COLremoveida];
    nlista:=nlista+1:
    lista2:=linalg::delCol(lista2, COLremoveida):
    break
  end_if:
end_for:
if is(lider_teste=17) then
  lista2:=lista2:
  COLremoveida:=COLremoveida+1;
  if nops(lista2)<=1 then
    aux:=aux+1:
    A[aux,1]:=FF[lista2[1],1]:
    A[aux,2]:=FF[lista2[1],2]:
    A[aux,3]:=FF[lista2[1],3]:
    A[aux,4]:=FF[lista2[1],4]:
  end_if
else;
  if tamanho(lista2)=1 and COLremoveida<>1 then
    aux:=aux+1:
    A[aux,1]:=FF[lista2[1],1]:
    A[aux,2]:=FF[lista2[1],2]:
    A[aux,3]:=FF[lista2[1],3]:
    A[aux,4]:=FF[lista2[1],4]:
  end_if;
end_if:
else;
  removeida[nlista]:=lista2[COLremoveida];
  nlista:=nlista+1:
  lista2:=linalg::delCol(lista2, COLremoveida);
  if tamanho(lista2)=1 and COLremoveida<>1 then
    aux:=aux+1:
    A[aux,1]:=FF[lista2[1],1]:
    A[aux,2]:=FF[lista2[1],2]:
    A[aux,3]:=FF[lista2[1],3]:
    A[aux,4]:=FF[lista2[1],4]:
  end_if;
end_if:
end_if:

end_for:
lista:=lista2:
end_while:
AA:=matrix(aux,4,A):
if (nlista=1) then
  FFA:=NIL;
  return(AA,FFA)
else;

removeida:=matrix(1,nlista-1,removeida):
aux:=0:
for pol in removeida do
  aux:=aux+1:
  FA[aux,1]:=FF[pol,1]:
  FA[aux,2]:=FF[pol,2]:

```

```

    FA[aux,3]:=FF[pol,3]:
    FA[aux,4]:=FF[pol,4]:
end_for:
FA:=matrix(aux,4,FA):
grau:=linalg::col(FA,3):
ordem:=linalg::col(FA,2):
cont:=aux;
//ordena matriz FA
for iPos from 1 to cont-1 step 1 do
    iMin:=iPos:
    for i from iPos+1 to cont step 1 do
        if is(ordem[i]<ordem[iMin]) then
            iMin:=i:
        else
            if is(ordem[i]=ordem[iMin] and grau[i]<grau[iMin]) then
                iMin:=i:
            end_if:
        end_if:
    end_for:
    if is(iMin<>iPos) then
        FA:=linalg::swapRow(FA,iPos,iMin);
        ordem:=linalg::swapRow(ordem,iPos,iMin);
        grau:=linalg::swapRow(grau,iPos,iMin);
    end_if:
end_for:
// Retorna Conjunto Caracteristico AA e  $F \setminus A = FFA$ 
FFA:=FA:
return(AA,FFA):
end_if;
end_proc

```

A.4 Qual reduzir

A partir dos conjuntos \mathcal{A} e $\mathcal{F} \setminus \mathcal{A}$ resultantes da rotina “Conjunto”Vretorna quais polinômios devem ser reduzidos.

```

QUAL_REDUIR:= proc(A,F)
    local nA, nF, grauA, grauF, ordemA, ordemF, i, teste, Ai,f,a;
begin
    grauA:=linalg::col(A, 3):
    ordemA:=linalg::col(A, 2):
    grauF:=linalg::col(F, 3):
    ordemF:=linalg::col(F, 2):
    nA:=nops(grauA):
    nF:=nops(grauF):
    Ai:=0;
    f:=linalg::row(F,nF):
    reduzir:=0;
    for i from nA downto 1 step 1 do
        a:=linalg::row(A,i):
        teste:=reduzido(f,a):
        if is(teste=0) then
            reduzir:=i:
            break;
        end_if:
    end_for;
    if (reduzir=0) then
        error("Algo errado no F");
    end_if;
end;

```

```

else:
  return(reduzir,nF);
  // Retorna em qual Linha de A esta o polinomio cujo f nao eh reduzido. e a linha de F
end_if:
end_proc;

```

A.5 Reduções

Reduz um polinômio com respeito a outro.

```

LEMMA9:= proc(a,f,RANK,thetas)
  local m,n,Q,QQ,AUX,R,ff,na,G,aa, fazer,var, teste,i, teste_der,j;
  //Retorna Vetor R em Ordem segundo RANK onde:  1o Coluna: polinomio
  //                                               2o Coluna: posicao relativa rank
  //                                               3o Coluna: grau do lider
  //                                               4o Coluna: lider
begin

  fazer:=1:
  G:=1:
  while fazer=1 do
    teste:=17;
    if (degree(a[1],f[4])>0 and degree(f[1],f[4])>=degree(a[1],f[4])) then
      R:=f[4]^f[3];
      na:=degree(a[1],f[4]):
      ff:=f:
      while degree(R,f[4])>=na do //LEMMA8
        teste:=7;
        m:=degree(ff[1],ff[4]):
        n:=degree(a[1],ff[4]):
        Q:=coeff(a[1],ff[4],n)^(m-n+1);
        QQ:=((ff[4]^(m-n))*(coeff(ff[1],ff[4],m))*(coeff(a[1],ff[4],n)^(m-n)));
        G:=G*Q;
        G:=simplify(G);
        // print("q do zero",Q);
        AUX:=expand(Q*ff[1]-QQ*a[1]);
        ff[1]:=AUX;
        R:=simplify(AUX);
        // print("zero?",R);
      end_while;
      if(teste=17) then
        error("error LEMA -- R nao mudo");
      end_if;

      R:=FIND_LIDER({R},RANK):
    else
      teste_der:=ORDEM_DERIVADA(a[4],f[1]);
      if (degree(a[1],f[4])=0 and degree(f[1],a[4])<>0 and teste_der=0) then
        na:=degree(a[1],a[4]):
        R:=a[4]^na;
        ff:=f:
        while degree(R,a[4])>=na do
          teste:=7;
          m:=degree(ff[1],a[4]):
          n:=degree(a[1],a[4]):
          Q:=coeff(a[1],a[4],n)^(m-n+1);
          QQ:=((a[4]^(m-n))*(coeff(ff[1],a[4],m))*(coeff(a[1],a[4],n)^(m-n)));
          G:=G*Q;

```

```

G:=simplify(G);
AUX:=expand(Q*ff[1]-QQ*a[1]);
ff[1]:=AUX;
R:=simplify(AUX);
end_while;
if(teste=17) then
  error("error LEMA -- R nao mudo");
end_if;
R:=FIND_LIDER({R},RANK):
else;
//print("AQUI");
if (teste_der>0) then
  var:=indets(f[4]);
  teste:=7:
  aa:=a;
  for i from 1 to teste_der step 1 do
    aa[1]:=D(aa[1]);
    for j in thetas do
      aa[1]:=subs(aa[1],D(D(j))=0,EvalChanges);
    end_for
  end_for;
ff:=f;
m:=degree(ff[1],ff[4]):
n:=degree(aa[1],ff[4]):
if(n<>0) then
  if (m>=n) then
    R:=f[4]^f[3];
    na:=degree(aa[1],f[4]):
    while degree(R,f[4])<>0 do
      teste:=7;
      Q:=coeff(aa[1],ff[4],n)^((m-n)+1);
      QQ:=((ff[4]^(m-n))*(coeff(ff[1],ff[4],m))*(coeff(aa[1],ff[4],n))^(m-n));
      G:=G*Q;
      G:=simplify(G);
      AUX:=expand(Q*ff[1]-QQ*aa[1]);
      ff[1]:=AUX;
      R:=simplify(AUX);
    end_while
  else
    R:=f[4]^f[3];
    na:=degree(aa[1],f[4]):
    while degree(R,f[4])<>0 do
      teste:=7;
      m:=degree(ff[1],ff[4]):
      n:=degree(aa[1],ff[4]):
      Q:=coeff(ff[1],ff[4],m);
      QQ:=((ff[4]^(n-m))*(coeff(aa[1],ff[4],n)));
      G:=G*Q;
      G:=simplify(G);
      AUX:=expand(Q*aa[1]-QQ*ff[1]);
      aa[1]:=AUX;
      R:=simplify(AUX);
    end_while;
  end_if;
if(teste=17) then
  error("error LEMA -- R nao mudo");
end_if;
R:=FIND_LIDER({R},RANK):

```

```

else
  aa:=FIND_LIDER({aa[1]},RANK):
  m:=degree(ff[1],aa[4]):
  n:=degree(aa[1],aa[4]):
  if (m>=n) then
    R:=aa[4]^m;
    while degree(R,aa[4])<>0 do
      teste:=7;
      // print(m,n,ff[1]);
      Q:=coeff(aa[1],aa[4],n)^((m-n)+1);
      QQ:=((aa[4]^(m-n))*(coeff(ff[1],aa[4],m))*(coeff(aa[1],aa[4],n))^(m-n));
      G:=G*Q;
      G:=simplify(G);
      AUX:=expand(Q*ff[1]-QQ*aa[1]);
      ff[1]:=AUX;
      R:=simplify(AUX);
    end_while;
  else;
    R:=aa[4]^aa[3];
    while degree(R,aa[4])<>0 do
      teste:=7;
      m:=degree(ff[1],aa[4]):
      n:=degree(aa[1],aa[4]):
      Q:=coeff(ff[1],aa[4],m);
      QQ:=((aa[4]^(n-m))*(coeff(aa[1],aa[4],n)));
      G:=G*Q;
      G:=simplify(G);
      AUX:=expand(Q*aa[1]-QQ*ff[1]);
      aa[1]:=AUX;
      R:=simplify(AUX);
    end_while;
  end_if;
  if(teste=17) then
    error("error LEMA -- R nao mudo");
  end_if;
  R:=FIND_LIDER({R},RANK):

end_if;

else;
  if (teste_der<0) then
    error("error LEMA -- derivada negativa");
  end_if;

end_if;

end_if;
end_if;

R:=simplifica(R,RANK);

if (reduzido(R,a)=0) then
  fazer:=1;
  f:=R;
else
  fazer:=0;
end_if;
end_while;

```

```

return(R,G);
end_proc

```

A.6 Simplifica 1

Simplifica termos constantes de polinômios

```

simplifica:= proc(F,RANK)
//Retorna Vetor F simplificando termos em comun em Ordem segundo RANK onde:
// 1o Coluna: polinomio
// 2o Coluna: posicao relativa rank
// 3o Coluna: grau do lider
// 4o Coluna: lider local R, fator, pol, n;

begin
R:=F[1];
fator:=factor(R);
n:=tamanho(fator);
if (fator[2]=FAIL) then
R:=F;
return(R);
else;
for pol from 1 to n-1 do
if is(has(fator[pol],_plus))=FALSE then
R:=divide(R,fator[pol]):
if (R[2]<>0) then
error("SimplificAA erro... resto nao eh zero");
end_if;
R:=simplify(R[1]);
end_if;
end_for;
end_if;
R:=FIND_LIDER({R},RANK);
return(R);
end_proc;

tamanho:=proc(F)
local falso, aux;
begin
if F=NIL then
return(0);
else
aux:=0:
falso:=nops(F):
for i from 1 to falso do
if F[i]=FAIL then
return(aux);
else
aux:=aux+1:
end_if;
end_for;
end_if;
end_proc

```

A.7 Simplifica 2

A partir dos polinômios de \mathcal{G} simplifica os polinômios de \mathcal{F} .

```
simplific:= proc(F,listaQ,RANK)

//Retorna Vetor F simplificando termos em comun em Ordem segundo RANK onde:
// 1o Coluna: polinomio
// 2o Coluna: posicao relativa rank
// 3o Coluna: grau do lider
// 4o Coluna: lider

local R, fatorR, Q, pol, diviv;
begin
R:=F[1];
fatorR:=factor(R):
Q:=matrix(nops(listaQ),1,listaQ);
if (fatorR[2]=FAIL) then
R:=F;
return(R);
else;
for pol in Q do
diviv:=gcd(R,pol);
R:=divide(R,diviv);
if (R[2]<>0) then
error("Simplific_v4 erro... resto nao eh zero");
end_if;
R:=simplify(R[1]);
end_for;
end_if;
R:=FIND_LIDER({R},RANK);
return(R);
end_proc;
```

A.8 Ordem

Ordena os polinômios segundo ranking

```
//ORDENA POLINOMIO SEGUNDO RANK
// FF matriz onde 1 coluna = Polinomios
//                2 coluna = RANK
//                3 Grau do lider
//                4 Lider
// cont= numero de polinomios em FF
ORDEM:=proc(FF)
local grau, ordem, iPos, iMin, i;
begin
grau:=linalg::col(FF, 3):
ordem:=linalg::col(FF, 2):
cont:=nops(ordem);
for iPos from 1 to cont-1 step 1 do
iMin:=iPos;
for i from iPos+1 to cont step 1 do
if is(ordem[i]<ordem[iMin]) then
iMin:=i;
else
if is(ordem[i]=ordem[iMin] and grau[i]<grau[iMin]) then
iMin:=i;
end_if;
end_for;
end_for;
```

```

        end_if
    end_if:
end_for:
if is(iMin<>iPos) then
    FF:=linalg::swapRow(FF,iPos,iMin);
    ordem:=linalg::swapRow(ordem,iPos,iMin);
    grau:=linalg::swapRow(grau,iPos,iMin)
end_if
end_for:
// Retorna Matriz FF em ordem segundo RANK
return(FF);
end_proc:

```

A.9 Ordem da Derivada

Retorna a ordem da maior derivada de um indeterminado dentro de um polinômio diferencial.

```

//Determina a maior derivada de v em A
// retornando um o respectivo inteiro... ex. ORDEM_DE(y,A)=7 existe a setima dericada de y em A
ORDEM_DERIVADA:= proc(v,A)
    local operando, pol, var, ind, parar, cont, teste, dif, grau, ind2, maior, inicial;

begin
    operando:=op(expand(A));
    var:=args(1);
    var:=indets(var);
    var:=var[1];
    cont:=0;
    parar:=7;
    maior:=0;
    inicial:=var;
    if (has(A,_plus))=FALSE then
        pol:=A;
        ind:=indets(pol);
        if (has(var,ind))=TRUE then
            parar:=7;
            cont:=0;
            dif:=var;
            grau:=degree(pol,dif);
            teste:=simplify(pol/dif^grau);
            while(parar=7) do
                ind2:=indets(teste);
                if (has(var,ind2))=TRUE then
                    cont:=cont+1;
                    dif:=D(dif);
                    grau:=degree(teste,dif);
                    teste:=simplify(teste/dif^grau);
                else
                    parar:=17;
                end_if;
            end_while;
            if maior<=cont then
                maior:=cont;
            end_if;
        end_if;
    else
        for pol in operando do
            ind:=indets(pol);

```

```

if (has(var,ind))=TRUE then
  parar:=7;
  cont:=0;
  dif:=var;
  grau:=degree(pol,dif);
  teste:=simplify(pol/dif^grau);
  while(parar=7) do
    ind2:=indets(teste);
    if (has(var,ind2))=TRUE then
      cont:=cont+1;
      dif:=D(dif);
      grau:=degree(teste,dif);
      teste:=simplify(teste/dif^grau);
    else
      parar:=17;
    end_if;
  end_while;
  if maior<=cont then
    maior:=cont;
  end_if;
end_if;
end_for;
cont:=0;
if(inicial=args(1)) then
  return(maior);
else;
  while(inicial<>args(1)) do
    inicial:=D(inicial);
    cont:=cont+1;
  end_while;
  return(maior-cont);
end_if;
end_proc

```

A.10 Rankink

Cria automaticamente um ranking

```

// theta    vetor com os parametros
// var      vetor variavel x    ex: [x1,x2,x3...xN]
// diff_i   até qual derivada em i vc quer ir. onde: i= u,y,x
// tipo     tipo de rank: tipo 1
// u<y<theta_i<DIFF(theta_i)<x
//                                     tipo 2
// u<y<theta1<theta2<...<thetaN<DIFF(theta1)<DIFF(theta2)...<DIFF(thetaN)<x

RANKEAR:=proc(theta,var,diff_u, diff_y, diff_x, tipo)
  local aux,j,i,pol,rank,n,RANK;
begin
  aux:=u:
  for i from 1 to diff_u do
    RANK[1,i]:=aux:
    aux:=D(aux):
  end_for:
  aux:=y:
  for i from 1 to diff_y do
    RANK[1,i+diff_u]:=aux:

```

```

    aux:=D(aux):
end_for:
if tipo=1 then
  j:=1:
  for pol in theta do
    RANK[1,diff_u+diff_y+j]:=pol:
    j:=j+1:
    RANK[1,diff_u+diff_y+j]:=D(pol):
    j:=j+1:
  end_for:
  for pol in var do
    aux:=pol:
    for i from 1 to diff_x do
      RANK[1,diff_u+diff_y+j]:=aux:
      j:=j+1:
      aux:=D(aux):
    end_for:
  end_for:
else;
  if tipo=2 then
    j:=1:
    for i from 1 to nops(theta) do
      RANK[1,diff_u+diff_y+j]:=theta[i]:
      j:=j+1:
    end_for:
    for i from 1 to nops(theta) do
      RANK[1,diff_u+diff_y+j]:=D(theta[i]):
      j:=j+1:
    end_for:
    for pol in var do
      aux:=pol:
      for i from 1 to diff_x do
        RANK[1,diff_u+diff_y+j]:=aux:
        j:=j+1:
        aux:=D(aux):
      end_for:
    end_for:
  end_if;
end_if;
n:=nops(RANK):
rank:=matrix(1,n,RANK);
return(rank);
end_proc

```

A.11 Reduzido

Verifica se dois polinômios são reduzidos um com respeito ao outro

```

// g reduzido com respeito a f?
// se sim return(1)
// se nao return(0)
// f e g vetores com:
//
//          1o Coluna: polinomio
//          2o Coluna: posicao relativa rank
//          3o Coluna: grau do lider
//          4o Coluna: lider

reduzido:=proc(g,f) //Aparentemente OK

```

```

local lider_teste, diff_aux, j;
begin
  if (g[2]<f[2])then
    if (degree(g[1],g[4])<degree(f[1],g[4])) then
      return(1);
    else
      lider_teste:=17;
      if is(degree(g[1],f[4])<degree(f[1],f[4])) then
        diff_aux:=f[4];
        for j from g[2] to f[2] step 1 do
          diff_aux:=D(diff_aux);
          // print("AUQUI",diff_aux,f[4]);
          if is(degree(g[1],diff_aux)<>0) then
            lider_teste:=7;
            return(0);
            break;
          end_if;
        end_for;
        if is(lider_teste=17) then
          return(1);
        end_if;
      else;
        return(0);
      end_if;
    end_if;
  end_if;
  if (g[2]=f[2]) then //LEMMA 6 em (Difference algebra and system identification) caso g=f
    if(g[3]<f[3]) then
      return(1);
    end_if;
    if(g[3]>=f[3]) then
      return(0);
    end_if;
  end_if;
  lider_teste:=17;
  if is(degree(g[1],f[4])<degree(f[1],f[4])) then
    diff_aux:=f[4];
    for j from f[2] to g[2] step 1 do
      diff_aux:=D(diff_aux);
      if is(degree(g[1],diff_aux)<>0) then
        lider_teste:=7;
        return(0);
        break;
      end_if;
    end_for;
    if is(lider_teste=17) then
      return(1);
    end_if;
  else;
    return(0);
  end_if;
end_proc

```

A.12 Identificável

Verifica se o conjunto característico resultante da rotina “algoritmo de Ritt” é identificável.

//Entradas: Matriz A característica

```

//          parametros do sistema
//          variaveis desconhecidas do sistema
IDENTIFICAVEL:= proc(A,parametros,variaveis)
  local j,i,nA,AA,nP,remover,pol,aux,lista,syms,iPos,teste;
begin
  aux:=0;
  AA:=A;
  nA:=tamanho(linalg::col(AA,1));
  nP:=tamanho(parametros);
  for pol in variaveis do
    aux:=0;
    lista:=0;
    for i from nA downto 1 step 1 do
      syms:=indets(AA[i,1]);
      for iPos in syms do
        if (iPos=pol) then
          aux:=aux+1;
          lista[aux]:=i;
        end_if;
      end_for;
    end_for;
    remover:=[lista[j] $ j=1..aux];
    AA:=linalg::delRow(AA, remover);
    nA:=tamanho(linalg::col(AA,1));
  end_for;
  aux:=0;
  for j from nA downto 1 step 1 do
    for iPos from 1 to nP step 1 do
      if (degree(AA[j,1],parametros[iPos])=1 and degree(AA[j,1],D(parametros[iPos]))=0) then
        teste:=17;
        if iPos<>nP then
          for i from iPos+1 to nP step 1 do
            if is(degree(AA[j,1],parametros[i])>0) or (degree(AA[j,1],D(parametros[i]))>0) then
              teste:=7;
              aux:=aux+1;
              lista[aux]:=j;
              break;
            end_if;
          end_for;
        end_if;
        if teste=7 then
          break;
        end_if;
      else;
        if (degree(AA[j,1],D(parametros[iPos]))>0) then
          aux:=aux+1;
          lista[aux]:=j;
        end_if;
      end_if;
    end_for;
  end_for;
  remover:=[lista[j] $ j=1..aux];
  AA:=linalg::delRow(AA, remover);
  nA:=tamanho(linalg::col(AA,1));
  aux:=0;
  for j from nA downto 1 step 1 do
    syms:=indets(AA[j,4]);
    if syms[1]=u or syms[1]=y then

```

```
        aux:=aux+1;
        lista[aux]:=j;
    end_if;
end_for;
remover:=[lista[j] $ j=1..aux];
AA:=linalg::delRow(AA, remover);
if AA=NIL then
    return("NÃO IDENTIFICAVEL");
else
    return(AA);
end_if;
end_proc
```