

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

**Identificação de sistemas contínuos  
utilizando o método dos momentos de  
Poisson**

DISSERTAÇÃO DE MESTRADO

Leandro Dariva Pinto

**Porto Alegre**

**2003**

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

**Identificação de sistemas contínuos  
utilizando o método dos momentos de  
Poisson**

Leandro Dariva Pinto

Dissertação de Mestrado apresentada como  
requisito parcial para obtenção do título de  
Mestre em Engenharia

**Orientador:**  
**Prof. Dr. Jorge Otávio Trierweiler**

**Porto Alegre**

**2003**

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

A Comissão Examinadora, abaixo assinada, aprova a Dissertação *Identificação de sistemas contínuos utilizando o método dos momentos de Poisson*, elaborada por Leandro Dariva Pinto, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:

---

Prof. Dr. Alexandre Zanfelice Bazanella

---

Prof. Dr. Argimiro Resende Secchi

---

Prof. Dr. Luís Gustavo Soares Longhi

*“...Só padece da solidão  
aqueles que se abstêm  
das lutas de seu tempo...”*

*Dom Quixote  
Miguel de Cervantes*

# Agradecimentos

Ao finalizar esse trabalho de dissertação gostaria de deixar registrado aqui o meu agradecimento a algumas pessoas que me acompanharam durante esse período da minha vida e que de alguma forma ou de outra ajudaram na síntese desse trabalho. Se fosse listar todas essas pessoas provavelmente esses agradecimentos ocupariam mais espaço que a dissertação em si. Portanto, se algum nome foi deixado de lado, não o foi por esquecimento mas sim pela falta de espaço.

Gostaria de começar agradecendo a minha família que durante toda a minha vida me deu o suporte em todos os sentidos para que eu atingisse o ponto onde hoje chego. Sem eles a jornada teria sido, com certeza, muito mais difícil.

Gostaria de agradecer também a empresa onde trabalho que, apesar de ter trocado de nome duas vezes desde que comecei a trabalhar lá (hoje chama-se Braskem) nunca mudou sua crença de que as pessoas são os seus ativos mais importantes e, portanto, incentiva que seus membros busquem uma educação continuada. Gostaria de agradecer especialmente a meu líder, Esdras P. Demoro, que logo após minha formatura me “mandou” ir fazer o mestrado. Agradeço também as chances que tive de discutir o meu trabalho com ele pois com sua experiência me ensinou a sempre enxergar as coisas por uma perspectiva diferente.

Agradeço também ao meu orientador, professor Jorge O. Trierweiler, que com seu dinamismo, inteligência e, principalmente, bom humor me incentivou nos momentos em que minha crença no meu trabalho diminuía.

Aos meus colegas de trabalho, Artur, Gustavo, Alexsandro, Rossano, Tiago, Nina e Paula pela amizade, pelas trocas de idéias e pelo ambiente bem humorado que desfrutei nos últimos quatro anos. Um abraço aos meus inúmeros amigos e colegas de mestrado pela amizade e incentivo. Gostaria de agradecer a uma amiga em especial que durante esses anos sempre foi uma confidente e exemplo de determinação e superação. Simone! Do fundo do meu coração, muito obrigado.

Aos demais amigos muito obrigado pela compreensão pela minha ausência e pelo incentivo.

Finalmente gostaria de agradecer a minha amada Dione que nesse último ano compreendeu minha ausência para me dedicar à redação desse trabalho.

A todos os meus mais sinceros agradecimentos!

Leandro Dariva Pinto

# Resumo

A maioria dos métodos de síntese e sintonia de controladores, bem como métodos de otimização e análise de processos necessitam de um modelo do processo em estudo. A identificação de processos é portanto uma área de grande importância para a engenharia em geral pois permite a obtenção de modelos empíricos dos processos com que nos deparamos de uma forma simples e rápida. Mesmo não utilizando leis da natureza, os modelos empíricos são úteis pois descrevem o comportamento específico de determinado processo.

Com o rápido desenvolvimento dos computadores digitais e sua larga aplicação nos sistemas de controle em geral, a identificação de modelos discretos foi amplamente desenvolvida e empregada, entretanto, modelos discretos não são de fácil interpretação como os modelos contínuos pois a maioria dos sistemas com que lidamos são de representação contínua. A identificação de modelos contínuos é portanto útil na medida que gera modelos de compreensão mais simples.

A presente dissertação estuda a identificação de modelos lineares contínuos a partir de dados amostrados discretamente. O método estudado é o chamado método dos momentos de Poisson. Este método se baseia em uma transformação linear que quando aplicada a uma equação diferencial ordinária linear a transforma em uma equação algébrica evitando com isso a necessidade do cálculo das derivadas dos sinais de entrada e saída.

Além da análise detalhada desse método, onde demonstramos o efeito de cada parâmetro do método de Poisson sobre o desempenho desse, foi realizado também um estudo dos problemas decorrentes da discretização de sinais contínuos, como por exemplo o efeito *aliasing* decorrente da utilização de tempos de amostragem muito grandes e de problemas numéricos da identificação de modelos discretos utilizando dados com tempos de amostragem muito pequenos de forma a destacar as vantagens da identificação contínua sobre a identificação discreta. Também foi estudado um método para compensar a presença de *offsets* nos sinais de entrada e saída, método esse inédito quando se trata do método dos momentos de Poisson.

Esse trabalho também comprova a equivalência entre o método dos momentos de Poisson e uma metodologia apresentada por Rolf Johansson em um artigo de 1994.

Na parte final desse trabalho são apresentados métodos para a compensação de erros de modelagem devido à presença de ruído e distúrbios não medidos nos dados utilizados na identificação. Esses métodos permitem que o método dos momentos de Poisson concorra com os métodos de identificação discretos normalmente empregados como por exemplo ARMAX e Box-Jenkins.

# Abstract

The majority of methods for synthesis and tuning of controllers as well as methods for optimization and analysis of process need some kind of model of the process in study. System identification is, therefore, a very important area to the engineering in general because it allows the obtaining of empiric models of the process which we came across in a simply and fast way.

With the fast development of digital computer and its application to control systems in general, discrete-time system identification was largely developed and employed, however, discrete models are not of easy interpretation as continuous models because most of the systems which we deal with are of continuous nature. System identification of continuous models is, therefore, useful in the way that it generates models which understanding is simpler.

The present dissertation studies system identification of continuous linear models using discretized data. The studied method is the one called Poisson moment function method. This method is based on a linear transformation which, when applied to an ordinary linear differential equation, transform it into an algebraic equation avoiding, in this way, the necessity of differentiation of the input and output signals.

Besides the detailed analysis of this method, where we demonstrate the effect of each parameter on its performance, it was also accomplished an analysis of problems which arose from discretization of continuous signals, for example, the aliasing effect which arose as result of using large sampling times and numerical problems in discrete system identification arisen from using too small sampling intervals, which justify the advantages of continuous identification over discrete identification. It is also studied a method to avoid the effects of offsets present in the input and output signals used in the identification. This method is unpublished considering its application in the Poisson method.

This work also proves the equivalence between the Poisson Moment Functional method and a methodology presented by Rolf Johansson in an article of 1994.

In the last part of this work, methods for compensation of modeling errors due to the presence of non-measurable disturbance in the used data are presented. These methods allow that the Poisson moment function method can compete with discrete identification methods usually employed, for example, ARMAX and Box-Jenkins.

# Sumário

<b>Introdução .....</b>	<b>1</b>
1.1 Motivação.....	1
1.2 Objetivo.....	2
1.3 Estrutura da Dissertação.....	2
<b>Identificação de sistema lineares .....</b>	<b>4</b>
2.1 Estrutura de modelos lineares .....	4
2.1.1 Modelos não paramétricos .....	5
2.1.2 Modelos paramétricos sem distúrbios.....	6
2.1.3 Modelos paramétricos com distúrbios .....	9
2.2 Avaliação do erro de predição.....	15
2.3 Métodos para ajuste de parâmetros.....	17
2.3.1 Método dos mínimos quadrados para regressão linear .....	17
2.3.2 Métodos de busca iterativos.....	18
2.4 Identificação de modelos lineares no espaço de tempo contínuo .....	19
2.4.1 Método das funções de modulação .....	21
2.4.2 Método dos filtros lineares.....	23
2.4.3 Método da caracterização espectral dos sinais (método das funções ortogonais) .....	23
2.5 Considerações sobre a identificação de modelos lineares no domínio de tempo discreto e no espaço de tempo contínuo.....	25
2.6 Referências Bibliográficas .....	31
<b>Método dos Momentos de Poisson.....</b>	<b>32</b>
3.1 Introdução .....	32
3.2 Cálculo dos momentos de Poisson das derivadas de um sinal.....	33
3.3 Implementação do cálculo dos momentos de Poisson.....	36
3.4 Aplicação dos momentos de Poisson na identificação de modelos lineares invariantes no tempo .....	38
3.4.1 Método 1 –Vários estágios da cadeia de filtros de Poisson e um instante de tempo.....	40
3.4.2 Método 2 –Número mínimo de estágios da cadeia de filtros de Poisson e vários instantes de tempo. ....	42
3.5 Método dos Momentos de Poisson Generalizado .....	45
3.6 Estimação recursiva dos parâmetros .....	47
3.7 Influência da constante da cadeia de filtros de Poisson na precisão da estimativa do modelo. ....	54
3.8 Considerações sobre a identificação das condições iniciais. ....	59
3.9 Comparação entre a identificação contínua e discreta.....	66
3.10 Análise de uma forma alternativa de interpretar o método do momentos de Poisson. ....	70
3.11 Referências Bibliográficas .....	76



<b>Compensação de <i>bias</i> dos parâmetros .....</b>	<b>77</b>
4.1 Análise do <i>bias</i> dos parâmetros ajustados pelo método dos mínimos quadrados .....	78
4.2 Análise do valor do <i>bias</i> em função do tempo de amostragem e da constante da cadeia de filtros para um sistema de 1ª ordem .....	80
4.3 Métodos de compensação de <i>bias</i> .....	88
4.3.1 Método da máxima verossimilhança .....	89
4.3.2 Método da variável instrumental .....	94
4.3.3 Método de compensação de <i>bias</i> por pré-filtros .....	96
4.4 Estudo comparativo dos métodos de compensação de <i>bias</i> .....	102
4.5 Referências Bibliográficas .....	108
<b>Conclusões e sugestões de trabalhos futuros .....</b>	<b>110</b>
5.1 Conclusões .....	110
5.2 Sugestões .....	113
<b>Apêndice A - Cálculo do gradiente de <i>y</i> predito para modelo ARMAX .....</b>	<b>I</b>
<b>Apêndice B - Efeito <i>Aliasing</i> .....</b>	<b>III</b>

# Índice de figuras

<b>Figura 2.1:</b> Comparação entre modelo e sistema.....	5
<b>Figura 2.2:</b> Representação em forma de diagrama de blocos de um sistema de equações de estado lineares [Ogata, 1997]. .....	8
<b>Figura 2.3:</b> Estrutura de modelo ARX.....	11
<b>Figura 2.4:</b> Estrutura de modelo ARMAX. ....	12
<b>Figura 2.5:</b> Estrutura da família de modelos tipo <i>equation error</i> . ....	13
<b>Figura 2.6:</b> Estrutura de modelo OE ( <i>output error</i> ).....	14
<b>Figura 2.7:</b> Estrutura de modelo Box-Jenkins (BJ).....	14
<b>Figura 2.8:</b> Diagrama do método geral de identificação de modelos lineares no espaço de tempo contínuo [Unbehauen e Rao, 1987]......	20
<b>Figura 2.9:</b> Classificação das diversas transformações lineares utilizadas na identificação de sistema lineares no espaço de tempo contínuo [Unbehauen e Rao, 1987]. .....	21
<b>Figura 2.10:</b> Exemplo típico de função de modulação. Nesse caso a função utilizadas é um polinômio com a seguinte forma: $\omega(t) = t^2(a_3t^3 + a_2t^2 + a_1t + a_0)$ . .....	22
<b>Figura 2.11:</b> Convergência do parâmetro do sistema e de sua correspondente margem de incerteza a medida que o tempo de amostragem é reduzido [Lee <i>et al.</i> , 2001]. .....	27
<b>Figura 2.12:</b> Erros de modelagem relativos de modelos no espaço de tempo discreto (a) e um modelo no espaço de tempo contínuo (b) em relação ao tempo de amostragem. ( $\Delta t = 0,1$ , linha sólida; $\Delta t = 0,02$ , linha tracejada; $\Delta t = 0,004$ , linha pontilhada) [Lee <i>et al.</i> , 2001]......	30
<b>Figura 3.1:</b> Cadeia de Filtros de Poisson. ....	37
<b>Figura 3.2:</b> Comportamento do condicionamento da matriz $M_k^T M_k$ em função do tamanho da cadeia de filtro de Poisson. ....	46
<b>Figura 3.3:</b> Dados de entrada e saída utilizados na estimativa do parâmetros do modelo definido por (3.49) utilizando o método dos mínimos quadrados recursivo. ....	50
<b>Figura 3.4:</b> Norma do erro quadrática em função do tempo. Primeiro gráfico $\lambda = 1$ , segundo gráfico $\lambda = 2$ , terceiro gráfico $\lambda = 5$ . Linha continua - $P_0=1.10^2I_6$ , linha tracejada - $P_0=1.10^4I_6$ , Linha ponto-traço - $P_0=1.10^6I_6$ , onde $I_6$ representa a matriz identidade de tamanho 6. ....	51
<b>Figura 3.5:</b> Norma do erro quadrática em função do tempo. Primeiro gráfico $P_0=1.10^2I_6$ , segundo gráfico $P_0=1.10^4I_6$ , terceiro gráfico - $P_0=1.10^6I_6$ , onde $I_6$ representa a matriz identidade de tamanho 6. Linha continua $\lambda = 1$ , linha tracejada - $\lambda = 2$ , Linha ponto-traço - $\lambda = 5$ . ....	52
<b>Figura 3.6:</b> Norma do erro quadrática em função do tempo. Comparação entre os métodos ordinários e normalizado para $\lambda = 2$ . Primeiro gráfico $P_0=1.10^2I_6$ , segundo gráfico $P_0=1.10^4I_6$ , terceiro gráfico - $P_0 = 1.10^6I_6$ , onde $I_6$ representa a matriz identidade de tamanho 6. Linha contínua – método ordinário, linha tracejada – método normalizado. ....	53

<b>Figura 3.7:</b> Norma do erro quadrática em função do tempo. Comparação entre os métodos ordinários e normalizado para $\lambda = 5$ . Primeiro gráfico $P_0=1.10^2I_6$ , segundo gráfico $P_0=1.10^4I_6$ , terceiro gráfico - $P_0=1.10^6I_6$ , onde $I_6$ representa a matriz identidade de tamanho 6. Linha contínua – método ordinário, linha tracejada – método normalizado. ....	53
<b>Figura 3.8:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ na ausência de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha traço-ponto – sistema 3 e linha contínua com pontos – sistema 4. ....	55
<b>Figura 3.9:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ na presença de 10% de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha ponto-traço – sistema 3 e linha contínua com pontos – sistema 4. ....	56
<b>Figura 3.10:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ na presença de 20% de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha traço-ponto – sistema 3 e linha contínua com pontos – sistema 4. ....	57
<b>Figura 3.11:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ para o sistema 1. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído. ....	57
<b>Figura 3.12:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ para o sistema 2. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha ponto-traço – 20% de ruído. ....	58
<b>Figura 3.13:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ para o sistema 3. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído. ....	58
<b>Figura 3.14:</b> Comportamento da norma quadrática do erro em função da constante $\lambda$ para o sistema 4. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído. ....	59
<b>Figura 3.15:</b> Funções pulso de Poisson para $\lambda = 1$ e $\lambda = 2$ . Linha contínua $k = 1$ , linha tracejada $k = 2$ , linha ponto-traço $k = 3$ . ....	60
<b>Figura 3.16:</b> Comportamento da norma quadrática do erro em função do número de tempo máximos da função de Poisson de maior índice no caso a eliminação da influência das condições iniciais do sinais de entrada e saída. Linha contínua $\lambda = 1$ e linha pontilhada $\lambda = 2$ . ....	60
<b>Figura 3.17:</b> Momentos de Poisson de um degrau unitário. Linha contínua – $M_0\{1\}$ , linha tracejada $M_1\{1\}$ e linha pontilhada $M_2\{1\}$ . ....	63
<b>Figura 3.18:</b> Derivadas das repostas da cadeia de filtros de Poisson ao degrau unitário. ....	65
<b>Figura 3.19:</b> Comportamento da norma quadrática do erro em função do número de tempo máximos da função de Poisson de maior índice no caso da eliminação de influência de <i>bias</i> nos sinais de entrada e saída. Linha contínua $\lambda = 1$ , linha tracejada $\lambda = 2$ e linha pontilhada $\lambda = 3$ . ....	66
<b>Figura 3.20:</b> Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson ( $\lambda = 5$ ) - linha pontilhada e do modelo ARX221 - linha contínua com pontos. ....	67

<b>Figura 3.21:</b> Sinal de saída com 3% de ruído branco utilizado na identificação do sistema definido na equação (3.49).....	68
<b>Figura 3.22:</b> Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson ( $\lambda = 5$ ) - linha pontilhada e do modelo ARX221 - linha continua com pontos. Modelos identificados com 3% de ruído no sinal de saída. ....	68
<b>Figura 3.23:</b> Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson - linha pontilhada e do modelo ARX15 15 1 - linha continua com pontos. Modelos identificados com 15% de ruído no sinal de saída. ....	69
<b>Figura 3.24:</b> Resposta ao degrau unitário do sistema identificado , linha contínua, do modelo gerado pelo método de Poisson, linha pontilhada e do modelo ARX30 30 1, linha continua com pontos, identificados com 15% de ruído no sinal de saída. ....	70
<b>Figura 4.1:</b> Sistema contínuo perturbado por um ruído genérico $w(t)$ .....	78
<b>Figura 4.2:</b> Esquema de identificação de um sistema contínuo de primeira ordem. ....	82
<b>Figura 4.3:</b> Esquema de identificação de um sistema contínua na forma discreta equivalente. ....	83
<b>Figura 4.4:</b> Norma quadrática do erro em função da constante $\lambda$ para diferentes tempos e amostragem, $T_e$ . ....	87
<b>Figura 4.5:</b> Norma quadrática do erro em função do tempo de amostragem, $T_e$ , para diferentes valores da constante $\lambda$ . ....	88
<b>Figura 4.6:</b> Modelo auxiliar para método da variável instrumental baseado em modelo auxiliar (VIMA). ....	95
<b>Figura 4.7:</b> Diagrama de blocos representativo do método de variável instrumental combinado com o método de Poisson. ....	96
<b>Figura 4.8:</b> Magnitude do sistema - linha pontilhada, do modelo ARMA de geração de ruído - linha tracejada, do pré-filtro $F(p)$ - linha ponto-tracejada e da cadeia de filtros de Poisson - linha contínua. ....	103
<b>Figura 4.9:</b> Variação dos índices de performance em função do grau de ruído ( $\lambda=\beta=2$ ). Linha tracejada, mínimos quadrados, linha ponto-tracejada, variável instrumental, linha contínua, compensação por pré-filtro. Primeira linha – Norma quadrática do erro dos parâmetros, segunda linha – Norma do erro da resposta ao degrau. ....	105
<b>Figura 4.10:</b> Variação dos índices de performance em função da constante da cadeia de filtros e Poisson $\lambda$ , (RRS = 30%). Linha tracejada, mínimos quadrados, linha ponto-tracejada, variável instrumental, linha contínua, compensação por pré-filtro. Primeira linha – Norma quadrática do erro dos parâmetros, segunda linha – Norma do erro da resposta ao degrau. ....	106

## Índice de tabelas

<b>Tabela 2.1:</b> Equação do erro de predição para diferentes estruturas de modelos lineares. ....	15
<b>Tabela 3.1:</b> Resultados da identificação do sistema (3.51) utilizando as duas variações do métodos de Poisson com dados sem e com ruído (7%). ....	43
<b>Tabela 3.2:</b> Resultados do teste de sensibilidade da primeira variação do método de Poisson em relação à quantidade de dados utilizado na identificação sem a presença de ruídos nesses. ....	44
<b>Tabela 3.3:</b> Resultados do teste de sensibilidade da primeira variação do método de Poisson em relação à quantidade de dados utilizado na identificação com a presença de ruídos nesses. ....	44
<b>Tabela 3.4:</b> Resultados do teste de sensibilidade da segunda variação do método de Poisson em relação à quantidade de dados utilizado na identificação com a presença de ruídos nesses. ....	45
<b>Tabela 3.5:</b> Sistemas lineares utilizados na análise da influência da constante $\lambda$ sobre a precisão do método de Poisson. ....	54
<b>Tabela 3.6:</b> Resultados da identificação com diferentes níveis de <i>off-set</i> nos sinais de entrada e saída. ....	61
<b>Tabela 4.1:</b> Resultados da identificação com dados contaminados com ruído branco. .	104
<b>Tabela 4.2:</b> Resultados da identificação com dados contaminados com ruído colorido. ....	105

# Capítulo 1

## Introdução

### 1.1 Motivação

O que define uma ciência é a existência de um objeto de estudo e método de análise definido. O método científico por sua vez consiste na elaboração de hipóteses e a confrontação dessas com os fatos observados. A elaboração de hipótese pode ser compreendida como a elaboração de modelos que tentam explicar o sistema em estudo. O objetivo da ciência é, portanto, a elaboração de modelos para tentar explicar e prever a natureza dos sistemas que nos cercam.

Um modelo pode ser compreendido como um conjunto de regras e princípios que se propõem a explicar o comportamento de um determinado sistema. Nas ciências onde as causas e efeitos são quantidades mensuráveis essas regras e princípios podem ser traduzidas na forma de conjuntos de equações matemáticas.

Os modelos matemáticos por sua vez são compostos por relações e constantes. As relações podem ser divididas basicamente em dois grupos. O primeiro consiste nas relações obtidas a partir da aplicação de conceitos fundamentais, balanço de massa, balanço de energia, balanço de momento, etc, ao sistema em estudo. O segundo grupo consiste em relações empíricas utilizadas para tentar prever o comportamento de sistemas dos quais não temos conhecimento suficiente dos princípios fundamentais que os regem. Constantes são, por sua vez, valores introduzidos nas relações dos modelos de forma a ajustar a trajetória prevista com os dados observados.

A identificação, objeto de estudo dessa dissertação, pode ser definida como o conjunto de técnicas e métodos utilizados para ajustar os valores das constantes dos modelos aos dados

coletados em experimentos e observações dos sistemas em estudo. Mesmo as constantes físicas mais fundamentais são parâmetros de ajuste de algum modelo.

Normalmente a maioria dos sistemas apresenta um comportamento não-linear, isto é, sua resposta a um estímulo externo depende do ponto de operação onde se encontra, caso esteja em estado estacionário, e do tamanho e forma do sinal de entrada, ou seja, é um sistema que não obedece ao princípio da superposição. Apesar disso muitos sistemas não-lineares podem ser descritos de uma forma satisfatória por sistemas lineares. Muitas pesquisas foram feitas nesse campo de tal forma que a identificação de sistemas normalmente é entendida como identificação de modelos lineares.

Apesar da grande quantidade de pesquisa já realizada nesse campo, a identificação de sistemas lineares ainda apresenta alguns desafios a serem superados. Entre esses desafios está a identificação de modelos no domínio de tempo contínuo ao invés do tradicional domínio discreto, a identificação de modelos em malha fechada e a avaliação de desempenho de controladores.

O desenvolvimento da identificação de modelos no espaço contínuo é de grande interesse para a engenharia química uma vez que os sistemas com que essa lida são em sua maioria contínuos no tempo. Além disso os sistemas de controle utilizados na maioria das plantas industriais, apesar de serem digitais, possuem a interface dos controladores, PID em sua maioria, como se as equações implementadas fizessem parte do domínio de tempo contínuo. Como exemplo podemos citar os sistemas de controle PROVOX e Delta-V ambos da Emerson Process.

## **1.2 Objetivo**

O presente trabalho tem por objetivo introduzir os principais modelos de identificação linear no domínio de tempo contínuo bem como as principais técnicas para identificá-los. Dentre os principais métodos conhecidos para identificação de modelos no espaço de tempo contínuo será dada maior atenção ao método dos momentos de Poisson.

## **1.3 Estrutura da Dissertação**

Essa dissertação está organizada da seguinte forma:

O capítulo 1 descreve a motivação para a realização desse trabalho e o seu objetivo bem como a estrutura da dissertação.

O capítulo 2 descreve os principais conceitos de identificação linear e apresenta uma revisão dos principais métodos de identificação. Os métodos de identificação no espaço de tempo contínuo são explorados de uma maneira mais detalhada. Dentre os métodos de identificação no espaço de tempo contínuo é dado destaque ao método dos momentos de Poisson. Além disso é feita uma comparação entre as vantagens e desvantagens das técnicas de identificação no espaço de tempo discreto, atualmente as mais difundidas, com as técnicas de identificação no espaço de tempo contínuo.

No capítulo 3 é apresentado em detalhes o método dos momentos de Poisson. Nesse capítulo é demonstrada a fundamentação teórica do método e algumas variações aplicadas à identificação de modelos variantes no tempo e de modelos distribuídos. Também é apresentado um outro método de identificação no espaço contínuo sugerido por Rolf Johansson em seu artigo “Identification of Continuous-Time Models” de 1994. Nesse trabalho demonstramos que esse método é equivalente ao método dos momentos de Poisson.

No capítulo 4 analisamos três métodos para a compensação de erro de modelagem devido à presença de distúrbios não medidos nos sinais de entrada e saída do sistema a ser identificado. Esses métodos são: Máxima verossimilhança, variável instrumental e compensação de *bias* por pré-filtragem dos sinais de entrada.

Finalmente no capítulo 5 são apresentadas as conclusões desse estudo e sugestões para trabalhos futuros.

A principal contribuição desse trabalho é a análise inédita da questão da presença de *off-sets* nos sinais utilizados na identificação de modelos lineares contínuos utilizando o método dos momentos de Poisson. Além disso esse trabalho apresenta uma compilação dos principais resultados sobre o método dos momentos de Poisson para a identificação de sistemas lineares invariantes no tempo.

Outro aspecto analisado nessa dissertação é a limitação da utilização de dados amostrados para a identificação de modelos dinâmicos tanto quanto o tempo de amostragem empregado e muito grande ou quanto este é muito pequeno, limitação essa existente apenas para modelos discretos.



## Capítulo 2

### Identificação de sistemas lineares

A identificação de sistemas lineares é uma área da ciência muito importante uma vez que a maioria dos sistemas pode ser descrita de forma aproximada por modelos lineares.

A identificação de modelos lineares pode ser basicamente dividida em três etapas:

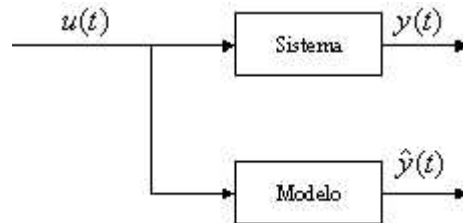
- 1) Escolha da estrutura dos modelos a serem identificados e derivação do sistema de equações para estimação dos parâmetros desconhecidos;
- 2) Planejamento dos experimentos para gerar os dados necessários e pré-tratamento desses.
- 3) Ajuste dos parâmetros dos modelos;
- 4) Validação dos modelos obtidos.

O presente capítulo irá descrever as principais famílias de modelos utilizados na identificação de sistemas lineares, os principais métodos de ajuste dos parâmetros e as principais técnicas de validação dos parâmetros.

#### 2.1 Estrutura de modelos lineares

Um modelo é uma relação capaz de reproduzir com razoável precisão a resposta de um determinado sistema não autônomo a um estímulo externo. Um modelo linear é um modelo que satisfaz o princípio da superposição, isto é, a resposta a uma combinação de entradas é igual a combinação das saídas geradas por essas entradas separadamente.

A diferença entre sistema e modelo é ilustrada na figura 2.1 onde  $u(t)$  representa um estímulo externo, comumente chamado de entrada,  $y(t)$  representa a resposta do sistema ao estímulo, comumente chamado de saída, e  $\hat{y}(t)$  representa a saída do sistema predita pelo modelo.



**Figura 2.1:** Comparação entre modelo e sistema.

Normalmente a saída predita pelos modelos não coincide perfeitamente com a saída do sistema correspondente. Essa diferença é chamada de erro de predição e é dada pela equação (2.1).

$$e(t) = y(t) - \hat{y}(t) \quad (2.1)$$

Basicamente os modelos lineares podem ser classificados em modelos paramétricos e não-paramétricos [Sinha *et* Kustza, 1983].

### 2.1.1 Modelos não paramétricos

Modelos não paramétricos são os modelos que não possuem um vetor finito de parâmetros para representar a resposta do sistema modelado.

Os modelos não-paramétricos se baseiam no fato de poder-se determinar a resposta de um sistema a um distúrbio qualquer baseado no conhecimento da resposta do sistema a um impulso unitário. O modelo do sistema nesse caso é um vetor contendo a resposta do sistema a um impulso unitário em diferentes instantes.

A resposta de um sistema causal a um distúrbio qualquer pode ser calculado através da integral de convolução como demonstrado na equação (2.2)

$$\hat{y}(t) = \int_0^t u(\tau) \delta(t - \tau) d\tau, \quad u(t) = 0 \quad \forall t < 0 \quad (2.2)$$

onde  $\delta(t)$  é a resposta do sistema a um impulso unitário.

### 2.1.2 Modelos paramétricos sem distúrbios

Modelos paramétricos são modelos que possuem um vetor de parâmetros finito para representar a resposta do sistema a um estímulo externo. Existem basicamente dois tipos de modelos paramétricos para sistemas lineares: função de transferência e espaço de estado.

Uma função de transferência é normalmente representada pela razão de dois polinômios como representado pela equação (2.3).

$$G(\xi, \theta) = \frac{B(\xi, \theta_b)}{A(\xi, \theta_a)} = \frac{b_0 + b_1\xi + \dots + b_m\xi^m}{1 + a_1\xi + a_2\xi^2 + \dots + a_n\xi^n} \quad (2.3)$$

onde  $\xi$  representa um operador linear qualquer,  $\theta_a$  e  $\theta_b$  representam os vetores de parâmetros dos polinômios  $A$  e  $B$ , sendo que, para sistemas próprios  $n \geq m$ . O vetor  $\theta$  é formado pela concatenação dos vetores  $\theta_a$  e  $\theta_b$ . A relação entre entrada e saída de um sistema é dada por:

$$\hat{y}(t, \theta) = G(\xi, \theta)u(t) \quad (2.4)$$

A outra forma de representar um sistema dinâmico é através das equações de espaço de estado. O estado de um sistema dinâmico é o menor conjunto de variáveis (chamadas variáveis de estado) tais que o conhecimento dessas variáveis no instante  $t_0$  junto com o conhecimento das variáveis de entrada para  $t \geq t_0$  determinam completamente o comportamento do sistema para qualquer instante  $t \geq t_0$ .

Na análise de espaço de estado nós lidamos com três tipos de variáveis que estão envolvidas na modelagem de sistemas dinâmicos: variáveis de entrada, variáveis de saída e variáveis de estado. O espaço de estado é o espaço  $n$ -dimensional cujos eixos coordenados são formados pelas  $n$  variáveis de estado,  $x_1, x_2, \dots, x_n$  do sistema [Ogata, 1997].

Um sistema dinâmico deve envolver elementos que armazenem o valor das variáveis de entrada para os instantes  $t \geq t_0$ . No caso de sistemas contínuos, os integradores servem como dispositivos de memória de tal forma que as saídas de tais integradores podem ser consideradas como variáveis que definem os estados internos do sistema. O número de variáveis de estado necessárias para definir completamente a dinâmica de um sistema é igual ao número de integradores envolvidos no sistema.

Um sistema dinâmico na forma explícita contendo  $r$  variáveis de entrada,  $m$  variáveis de saída e  $n$  variáveis de estado pode ser descrito pelos seguintes conjuntos de equações:

$$\begin{aligned}
\dot{x}_1(t) &= f_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
\dot{x}_2(t) &= f_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
&\vdots \\
\dot{x}_n(t) &= f_n(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t)
\end{aligned} \tag{2.5}$$

$$\begin{aligned}
y_1(t) &= g_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
y_2(t) &= g_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
&\vdots \\
y_m(t) &= g_r(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t)
\end{aligned} \tag{2.6}$$

Se definirmos:

$$\begin{aligned}
\mathbf{x}(t) &= \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ f_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \end{bmatrix} \\
\mathbf{y}(t) &= \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix}, \quad \mathbf{g}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} g_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ g_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ \vdots \\ g_m(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \end{bmatrix}, \quad \mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_r(t) \end{bmatrix}
\end{aligned} \tag{2.7}$$

então as equações (2.5) e (2.6) podem ser reescritas na forma vetorial:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{2.8}$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}, \mathbf{u}, t) \tag{2.9}$$

As equações (2.8) e (2.9) são a forma explícita geral de representação do espaço de estado de um sistema dinâmico não autônomo qualquer. Para o presente trabalho é de interesse apenas o caso linear que pode ser obtido a partir da linearização das equações (2.8) e (2.9) em torno de um ponto de operação. O resultado dessa operação são as equações de estado linearizado:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \tag{2.10}$$

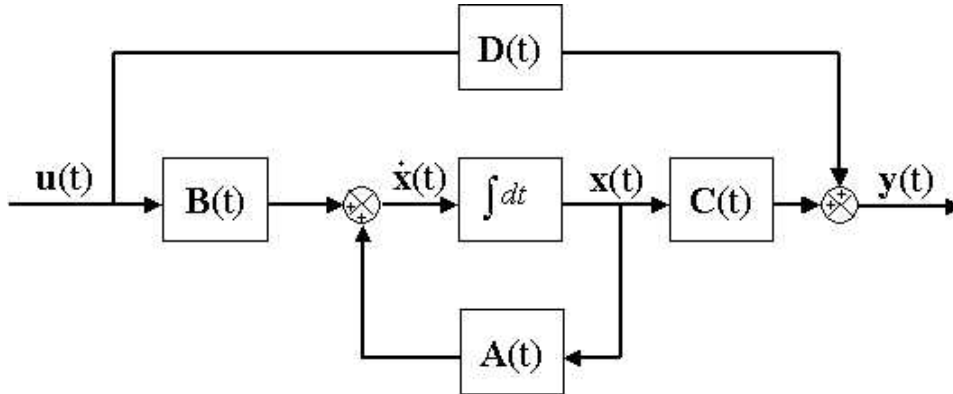
$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) \tag{2.11}$$

onde  $\mathbf{A}(t)$  é chamada de matriz de estados,  $\mathbf{B}(t)$  de matriz de entrada,  $\mathbf{C}(t)$  matriz de saída e  $\mathbf{D}(t)$  matriz de transmissão direta. A representação em forma de diagrama de blocos das equações de estado linearizadas pode ser vista na figura 2.2. Se as funções vetoriais  $\mathbf{f}$  e  $\mathbf{g}$  não envolvem diretamente o tempo as equações (2.10) e (2.11) podem ser simplificadas.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (2.12)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \quad (2.13)$$

A equação (2.12) é a equação de estado de um sistema linear invariante no tempo e a equação (2.13) é a equação de saída para o mesmo sistema.



**Figura 2.2:** Representação em forma de diagrama de blocos de um sistema de equações de estado lineares [Ogata, 1997].

A representação em espaço de estado é muito útil por sua capacidade de representar sistemas multivariáveis de uma forma compacta e elegante, o que é muito importante para o estudo de estratégias de controle avançado uma vez que sistemas de controle multivariável são cada vez mais comuns na indústria de processos em geral.

A representação de um modelo na forma de espaço de estado pode ser facilmente convertida para a forma de função de transferência. Para isso basta aplicar as equações (2.12) e (2.13) alguma transformada linear como, por exemplo, a transformada de Laplace. Como o escopo principal desse trabalho é a identificação de modelos no espaço de tempo contínuo utilizaremos a transformada de Laplace. Assim as equações (2.12) e (2.13) ficam da seguinte forma.

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \quad (2.14)$$

$$\mathbf{Y}(s) = \mathbf{C}\mathbf{X}(s) + \mathbf{D}\mathbf{U}(s) \quad (2.15)$$

Como a função de transferência de um sistema é definida como a razão entre as transformadas de Laplace das entradas e saídas quando as condições iniciais são nulas, assumimos que  $\mathbf{x}(0)$  na equação (2.14) é um vetor nulo. Dessa forma a equação (2.14) pode ser reescrita da seguinte forma.

$$(s\mathbf{I} - \mathbf{A})\mathbf{X}(s) = \mathbf{B}\mathbf{U}(s) \quad (2.16)$$

multiplicando a equação (2.16) por  $(s\mathbf{I} - \mathbf{A})^{-1}$  em ambos os lados temos:

$$X(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}U(s) \quad (2.17)$$

e substituindo a equação (2.17) na equação (2.15) temos:

$$\mathbf{Y}(s) = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]U(s) \quad (2.18)$$

A partir da equação (2.18) podemos concluir que a função de transferência do sistema é dada por:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (2.19)$$

Como a operação de inversão de uma matriz envolve a determinação de seu determinante a equação (2.19) pode ser reescrita da seguinte forma:

$$\mathbf{G}(s) = \frac{\mathbf{Q}(s)}{|s\mathbf{I} - \mathbf{A}|} \quad (2.20)$$

onde  $\mathbf{Q}(s)$  é uma matriz de polinômios em  $s$  e  $|s\mathbf{I} - \mathbf{A}|$  é igual ao polinômio característico de  $G(s)$ . Em outras palavras os auto-valores de  $\mathbf{A}$  são idênticos aos pólos de  $G(s)$ .

### 2.1.3 Modelos paramétricos com distúrbios

As estruturas de modelos apresentadas na seção anterior representam apenas a relação entre entradas medidas e saídas de um sistema linear. Entretanto é sabido que sistema reais estão sujeitos a distúrbios não medidos. Esses distúrbios podem ter origem nos instrumentos de medidas ou em variáveis não medidas que afetam o sistema como por exemplo a temperatura ambiente e a concentração de reagentes e impurezas nas correntes de alimentação.

Para que técnicas de identificação tenham valor prático é necessário, de alguma forma, levar em conta esses distúrbios não medidos. Normalmente a forma mais comum de se levar em conta esses fatores é a introdução no modelo de um termo adicional da seguinte forma:

$$e(t) = H(\xi, \theta)w(t) \quad (2.21)$$

onde  $e(t)$  é o distúrbio no processo e  $w(t)$  é um sinal tipo ruído branco.

Adicionando esse termo à equação (2.4) temos que a predição da saída do modelo é dada por:

$$\hat{y}(t, \theta) = G(\xi, \theta)u(t) + H(\xi, \theta)w(t) \quad (2.22)$$

Considerando o operador  $\xi$  como sendo igual ao operador deslocamento  $q$  cujo efeito é demonstrado na equação (2.23):

$$q^k f(t) = f(t + kT) \quad (2.23)$$

onde  $T$  representa o intervalo de amostragem dos dados, a equação (2.22) fica da seguinte forma:

$$\hat{y}(t, \theta) = G(q, \theta)u(t) + H(q, \theta)w(t) \quad (2.24)$$

que é a forma geral de representação de modelos lineares SISO no espaço de tempo discreto. Nesse trabalho utilizamos a mesma notação adotada por [Ljung, 1987] onde  $t$  representa o instante de tempo atual.

Considerando os principais estudos na área de identificação de sistemas lineares no domínio de tempo discreto podemos descrever algumas das principais estruturas utilizadas para a identificação discreta com compensação de distúrbios externos.

A forma mais simples de representar uma relação entre entradas e saídas no espaço de tempo discretos é através de equação de diferenças lineares como a apresentada na equação (2.25), abaixo.

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_1 u(t-1) + \dots + b_m u(t-m) + w(t) \quad (2.25)$$

Como o termo de ruído branco  $w(t)$  entra diretamente na equação de diferenças, o modelo definido pela equação (2.25) é chamado, em inglês, de *equation error model*. O vetor de parâmetros ajustáveis para essa estrutura de modelos é:

$$\theta = [a_1 \quad a_2 \quad \dots \quad a_n \quad b_1 \quad \dots \quad b_m]^T \quad (2.26)$$

Comparando a equação (2.3) com esse vetor de parâmetros notamos a falta do termo  $b_0$  que, na representação discreta de sistemas lineares, representa um influência instantânea da entrada  $u(t)$  na saída  $y(t)$ . Devido ao processo de amostragem e ao fato de que se lê e escreve no mesmo instante, não é possível se saber o valor da saída  $y(t)$  produzido pela entrada  $u(t)$ .

Definindo os polinômios  $A$  e  $B$  da seguinte forma:

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_n q^{-n} \quad (2.27)$$

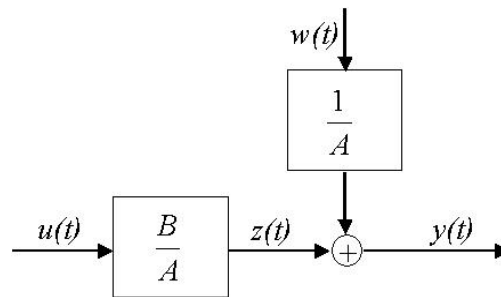
$$B(q) = b_1 q^{-1} + b_2 q^{-2} + \dots + b_m q^{-m} \quad (2.28)$$

podemos verificar que as funções  $G(q, \theta)$  e  $H(q, \theta)$ , para essa estrutura em particular, apresentam a seguinte forma:

$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad H(q, \theta) = \frac{1}{A(q)} \quad (2.29)$$

A estrutura de modelo descrita pela equação (2.25) é também conhecida como ARX sigla de *AutoRegressive with eXogeneous input*, onde AR se refere a parte auto-regressiva  $A(q)y(t)$  e X se refere a entrada externa  $B(q)u(t)$ . No caso especial onde  $m = 0$  essa estrutura é chamada de FIR, *finite impulse response*.

A estrutura ARX pode ser representada em formato de diagrama de blocos como apresentado na figura 2.3.



**Figura 2.3:** Estrutura de modelo ARX.

A principal desvantagem da estrutura ARX é a falta de adequada liberdade para modelar as propriedades do termo de distúrbio. É possível resolver essa falta de flexibilidade através da descrição do erro de modelagem como sendo uma média móvel de um ruído branco. Fazendo isso o temos um modelo descrito pela seguinte equação:

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_1 u(t-1) + \dots + b_m u(t-m) + w(t) + c_1 w(t-1) + \dots + c_l w(t-l) \quad (2.30)$$

Considerando o polinômio  $C(q)$  como sendo dado por:

$$C(q) = 1 + c_1 q^{-1} + \dots + c_l q^{-l} \quad (2.31)$$

a equação (2.30) pode ser reescrita da seguinte forma:

$$A(q)y(t) = B(q)u(t) + C(q)w(t) \quad (2.32)$$

Assim as funções  $G(q, \theta)$  e  $H(q, \theta)$  correspondem a:

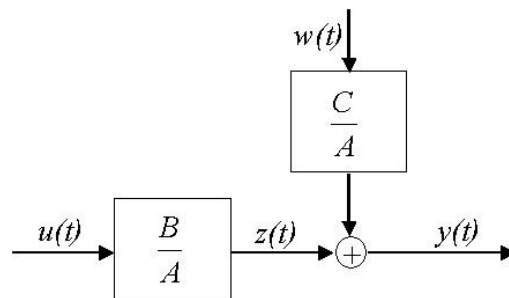
$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad H(q, \theta) = \frac{C(q)}{A(q)} \quad (2.33)$$

e agora o vetor de parâmetros é dado por:



$$\theta = [a_1 \quad \dots \quad a_n \quad b_1 \quad \dots \quad b_m \quad c_1 \quad \dots \quad c_l]^T \quad (2.34)$$

Devido ao termo de média móvel, *moving average*, (MA)  $C(q)w(t)$  a estrutura de modelo dada pela equação (2.32) é chamada de ARMAX. Essa estrutura é uma ferramenta padrão em estudos de teoria de controle e em econometria tanto para descrição de sistemas como para projeto de controladores. A estrutura ARMAX pode ser representado em formato de diagrama de blocos como apresentado na figura 2.4.



**Figura 2.4:** Estrutura de modelo ARMAX.

Ao invés de modelar o erro como uma média móvel de um sinal tipo ruído branco também é possível modelá-lo utilizando um termo de autoregressão. Fazendo isso temos um modelo com a seguinte estrutura:

$$A(q)y(t) = B(q)u(t) + \frac{1}{D(q)}w(t) \quad (2.35)$$

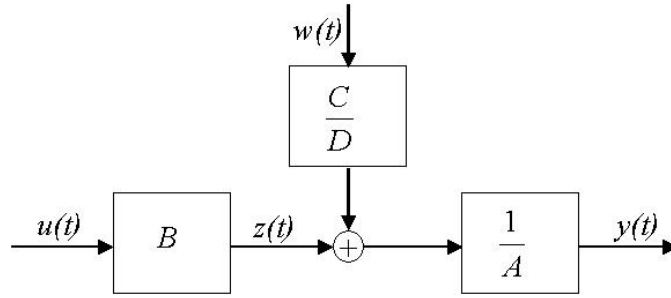
onde

$$D(q) = 1 + d_1q^{-1} + \dots + d_kq^{-k} \quad (2.36)$$

Essa estrutura de modelo pode ser chamada, devido aos dois termos de autoregressão como ARARX. Podemos generalizar a estrutura tipo equação de erro (*equation error*), introduzindo na equação (2.36) um termo de média móvel para descrever o erro de predição de tal forma que a estrutura do modelo fica da seguinte forma:

$$A(q)y(t) = B(q)u(t) + \frac{C(q)}{D(q)}w(t) \quad (2.37)$$

A estrutura de modelo descrito pela equação (2.37) contém as estruturas descritas pelas equações (2.32) e (2.35) como casos especiais. A família de estruturas tipo *equação do erro* pode ser representada em forma de diagrama de blocos como apresentado na figura 2.5.



**Figura 2.5:** Estrutura da família de modelos tipo equação do erro.

A característica principal das estruturas tipo equação do erro é que os termos  $G$  e  $H$  possuem uma correlação através do polinômio  $A$  que aparece no denominador de ambos. De um ponto de vista físico é mais interessante parametrizar esses termos de uma forma independente.

Se supusermos que a relação entre a entrada e a saída sem distúrbios,  $z(t)$ , pode ser descrita por uma equação de diferenças e que o distúrbio consiste apenas de ruído branco temos as seguintes descrições:

$$\begin{aligned} z(t) + f_1 z(t-1) + \dots + f_v z(t-v) &= b_1 u(t-1) + \dots + b_m u(t-m) \\ y(t) &= z(t) + w(t) \end{aligned} \quad (2.38)$$

Definindo o polinômio  $F$  como:

$$F(q) = 1 + f_1 q^{-1} + \dots + f_v q^{-v} \quad (2.39)$$

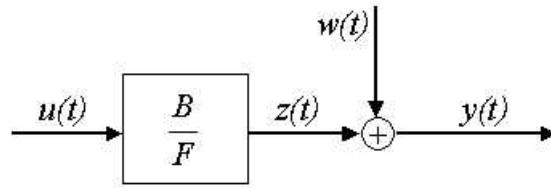
podemos escrever a estrutura do modelo definido pela equação (2.38) da seguinte forma:

$$y(t) = \frac{B(q)}{F(q)} u(t) + w(t) \quad (2.40)$$

A estrutura definida pela equação (2.40) é conhecida como estrutura tipo *output error* (OE). O vetor de parâmetros dessa estrutura fica sendo:

$$\theta = [b_1 \quad b_2 \quad \dots \quad b_m \quad f_1 \quad f_2 \quad \dots \quad f_v] \quad (2.41)$$

e o fluxo de sinais do modelo pode ser representado em forma de diagrama de blocos como apresentado na figura 2.6.

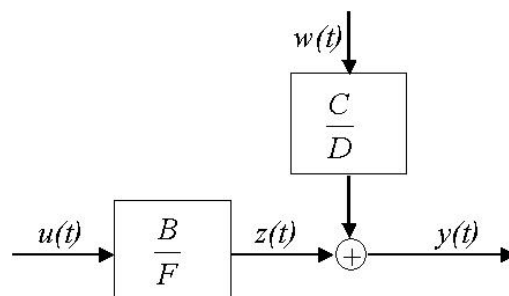


**Figura 2.6:** Estrutura de modelo OE (*output error*).

Um desenvolvimento natural da estrutura OE é a modelagem do erro de predição. Isso pode ser feito introduzindo um termo tipo ARMA, autoregressão e média móvel, para modelar o erro de predição  $e(t)$  em função de um sinal tipo ruído branco  $w(t)$ . Fazendo isso temos a seguinte estrutura de modelo.

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}w(t) \quad (2.42)$$

Essa estrutura é a mais natural do ponto de vista físico, uma vez que os termos  $G$  e  $H$  da equação (2.22) são parametrizados de forma independente. Essa estrutura foi sugerida e tratada por [Box e Jenkins, 1970] e é conhecida por seus nomes, ou seja, Box-Jenkins (BJ). Essa estrutura pode ser representada na forma de diagrama de blocos como apresentado na figura 2.7.



**Figura 2.7:** Estrutura de modelo Box-Jenkins (BJ).

Na representação dos sistemas lineares no formato de espaço de estado também é possível levar em conta os efeitos de distúrbios não medidos. Isso pode ser feito através da introdução de um termo de distúrbios nas equações (2.12) e (2.13).

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \quad (2.43)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{e}(t) \quad (2.44)$$

onde  $\mathbf{e}(t)$  é dado por:

$$\mathbf{e}(t) = \mathbf{H}(\xi, \theta)\mathbf{w}(t) \quad (2.45)$$

O termo  $\mathbf{D}u(t)$  foi removido da equação (2.44) pois para a maioria dos sistemas físicos ele é nulo, ou seja, na maioria dos casos reais não existe uma interferência instantânea das entradas sobre as saídas. Normalmente as entradas devem primeiramente afetar os estados do sistema para que esses alterem as respectivas saídas. Nas equações (2.43) e (2.44) o termo do erro é adicionado na equação de saída do modelo. Entretanto, o erro também pode ser adicionado à equação dos estados representando, dessa forma, ruídos de processo. Essa forma é apresentada mais adiante nesse capítulo.

A parametrização da matriz  $\mathbf{H}$  que é um modelo para distúrbios ou ruídos não medidos e que na figura 2.7 corresponde a razão  $C/D$ , pode ter um certo grau de dependência com as demais matrizes do modelo. O grau de dependência irá depender do conhecimento prévio que se tem sobre as fontes de distúrbios que afetam o sistema. Entende-se por conhecimento das fontes de distúrbios o conhecimento das covariâncias dos sinais de ruído. Essa idéia será discutida na seção seguinte que trata da predição das saídas do modelos.

## 2.2 Avaliação do erro de predição

Definida a estrutura do modelo a próxima etapa do processo de identificação é o ajuste dos parâmetros do modelo. Para isso é necessária a utilização de um método de otimização que ajuste as saídas do modelo com as saídas do sistema a ser modelado. Essa otimização normalmente é feita através da minimização do respectivo erro de predição.

Resolvendo a equação (2.22) para  $w(t)$  e substituindo esse por  $e(t)$  temos a seguinte expressão para o cálculo do erro de predição:

$$w(t) = H^{-1}(\xi, \theta)y(t) - H^{-1}(\xi, \theta)G(\xi, \theta)u(t) \quad (2.46)$$

onde  $e(t)$  é parcela da saída  $y(t)$  que não pode ser predita pelo conhecimento dos dados passados. Por esse motivo essa parcela também é chamada de inovação no tempo  $t$ .

**Tabela 2.1:** Equação do erro de predição para diferentes estruturas de modelos lineares.

Estrutura	Equação do erro de predição
ARX	$w(t) = A(q)y(t) - B(q)u(t)$
ARMAX	$w(t) = \frac{A(q)}{C(q)}y(t) - \frac{B(q)}{C(q)}u(t)$
OE	$w(t) = y(t) - \frac{B(q)}{F(q)}u(t)$
BJ	$w(t) = \frac{D(q)}{C(q)}y(t) - \frac{D(q)B(q)}{C(q)F(q)}u(t)$

Na Tabela 2.1 , acima, vemos a forma da expressão para o cálculo do erro de predição para diferentes estruturas de modelos lineares.

No caso da representação de sistemas lineares na forma de espaço de estados, temos duas opções para incluir a correção da predição da saída pelos erros de predição passados. A primeira é a forma apresentada pelas equações (2.43) a (2.45) onde esse termo é adicionado na equação de saída.

Nessa forma de representação é fácil demonstrar que a equação do erro de predição, removendo-se a parametrização das matrizes para simplificar a notação, é dada por:

$$\mathbf{e}(t) = -\mathbf{H}^{-1}\mathbf{C}(s\mathbf{I} - \mathbf{A})\mathbf{B}u(t) + \mathbf{H}^{-1}\mathbf{y}(t) \quad (2.47)$$

Entretanto a forma mais comum de compensar os erros de predição é através da adição do termo de correção na equação de estados do sistema. Essa abordagem é muito difundida em função da utilização dos observadores de estados para a predição de variáveis não medidas. Esses observadores de estados são também conhecidos como filtros de Kalman [Ljung, 1987]. Nesse caso a estrutura do modelo linear com correção pelo erro de predição passado passa a ser dado pelas seguintes equações:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) + \mathbf{K}\mathbf{e}(t) \quad (2.48)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \quad (2.49)$$

onde  $\mathbf{K}$  é normalmente conhecida como matriz de ganhos de Kalman.

Aplicando-se a transformada de Laplace à equação (2.48) e resolvendo para o vetor de estados temos a seguinte equação:

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}[\mathbf{B}u(s) + \mathbf{K}\mathbf{e}(s)] \quad (2.50)$$

Aplicando-se a transformada de Laplace à equação (2.49) e substituindo-se a equação (2.50) temos a seguinte equação para as saídas dos sistemas:

$$\mathbf{y}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}[\mathbf{B}u(s) + \mathbf{K}\mathbf{e}(s)] \quad (2.51)$$

que resolvendo para  $\mathbf{e}(t)$  fornece a equação do erro de predição para esse caso.

$$\mathbf{e}(s) = -\mathbf{K}^{-1}\mathbf{B}u(s) + \mathbf{K}^{-1}(s\mathbf{I} - \mathbf{A})\mathbf{C}^{-1}\mathbf{y}(s) \quad (2.52)$$

## 2.3 Métodos para ajuste de parâmetros

Sabendo como avaliar o erro de predição de uma estrutura de modelo é possível realizar o ajuste dos respectivos parâmetros de forma a predizer da melhor forma possível as saídas do sistema. Normalmente esse ajuste é feito minimizando alguma função objetivo que leve em conta o erro de predição, de forma que os parâmetros de um modelo sejam dados por:

$$\hat{\theta}_N = \arg \min V_N(\theta, \varphi(t)) \quad (2.53)$$

onde  $V_N$  é uma função objetivo dada,  $\varphi(t)$  é o vetor de regressão que contém os valores das entradas e saídas medidas do sistema,  $N$  é o número de pontos coletados das entradas e saídas e  $\hat{\theta}_N$  é o vetor de parâmetros que minimiza  $V_N$ .

### 2.3.1 Método dos mínimos quadrados para regressão linear

Normalmente a função utilizada como função objetivo no processo de ajuste de parâmetros de um modelo linear é a soma dos erros quadráticos.

$$V_N(\theta, \varphi(t)) = \frac{1}{2N} \sum_{t=1}^N e^2(\theta, \varphi(t)) \quad (2.54)$$

onde  $e(t)$  é dado pela equação (2.1).

No caso da estrutura de modelos lineares ARX é possível, comparando as equações (2.1) e (2.25), mostrar que a predição da saída do sistema é dada por:

$$\hat{y}(t) = -a_1 y(t-1) + \dots - a_n y(t-n) + b_1 u(t-1) + \dots + b_m u(t-m) \quad (2.55)$$

que pode ser reescrita da seguinte forma:

$$\hat{y}(t) = \hat{\theta} \varphi(t)^T \quad (2.56)$$

onde  $\hat{\theta}$ , nesse caso é dado pela equação (2.26) e  $\varphi(t)$  é dado por:

$$\varphi(t) = [-y(t-1) \quad \dots \quad -y(t-n) \quad u(t-1) \quad \dots \quad u(t-m)] \quad (2.57)$$

Assim a equação (2.54) pode ser reescrita da seguinte forma:

$$V_N(\theta, \varphi(t)) = \frac{1}{2N} \sum_{t=1}^N e^2(t) = \frac{1}{2N} [y(t) - \hat{\theta} \varphi(t)^T]^T [y(t) - \hat{\theta} \varphi(t)^T] \quad (2.58)$$

Derivando a equação (2.58) em relação ao vetor de parâmetros e igualando a zero obtemos a seguinte equação:

$$\frac{dV_N(\hat{\theta}, \varphi(t))}{d\hat{\theta}} = -\varphi(t)^T y(t) + \varphi(t)^T \varphi(t) \hat{\theta} = 0 \quad (2.59)$$

que resolvendo para o vetor  $\hat{\theta}$  fornece a expressão para o cálculo dos parâmetros do modelo ARX.

$$\hat{\theta} = [\varphi(t)^T \varphi(t)]^{-1} \varphi(t)^T y(t) \quad (2.60)$$

A grande vantagem do método dos mínimos quadrados aplicado para regressão linear é que sua função é quadrática e portanto possui solução analítica.

### 2.3.2 Métodos de busca iterativos

A solução analítica do problema de identificação linear só é possível quando utilizamos estruturas possíveis de serem transformadas em regressores lineares como a estrutura ARX. No caso de estruturas como ARMAX e Box-Jenkins isso não é possível de ser aplicado visto que o erro de predição passado afeta a predição da saída.

Para esses casos é necessário utilizar um método de ajuste iterativo. Normalmente para o caso da identificação linear utiliza-se algum método de métrica variável dado por uma equação do tipo:

$$\hat{\theta}^{(i+1)} = \hat{\theta}^i + \mu [R(\hat{\theta}, \varphi(t, \hat{\theta}))]^{-1} V'_N(\hat{\theta}, \varphi(t, \hat{\theta})) \quad (2.61)$$

onde  $R$  é uma matriz que modifica a direção de busca. No caso de utilizarmos  $R$  como sendo uma matriz identidade temos o método conhecido como máxima descida. No caso de utilizarmos a matriz hessiana da função objetivo  $V$  temos o método de Gauss-Newton. O parâmetro  $\mu$  é o tamanho do passo de cada iteração que é escolhido de forma que:

$$V'_N(\hat{\theta}^{(i+1)}, \varphi(t, \hat{\theta})) < V'_N(\hat{\theta}^{(i)}, \varphi(t, \hat{\theta})) \quad (2.62)$$

Observa-se que agora o vetor de regressão é função também dos parâmetros do modelo uma vez que o erro de predição está incluído dentro desse vetor.

O gradiente da função objetivo é calculado de acordo com a seguinte expressão.

$$V'_N(\hat{\theta}, \varphi(t, \hat{\theta})) = -\sum_{t=1}^N \psi(\hat{\theta}, \varphi(t, \hat{\theta})) e(\hat{\theta}, \varphi(t, \hat{\theta})) \quad (2.63)$$

onde  $\psi(\hat{\theta}, \varphi(t, \hat{\theta}))$  é o gradiente do erro de predição. O cálculo desse termo depende da estrutura do modelo utilizado na identificação. Para o caso particular da estrutura ARMAX pode-se facilmente demonstrar que o gradiente do erro de predição é dado por:

$$\psi(\hat{\theta}, \varphi(t)) = \frac{1}{C(q)} \varphi(t, \hat{\theta}) \quad (2.64)$$

A demonstração da equação (2.64) é dada no apêndice A.

É sabido que o método de máxima descida é ineficiente perto do mínimo. Métodos do tipo Newton apresentam melhor desempenho nessa situação. Para utilizarmos o método de Gauss-Newton é necessário o cálculo da matriz hessiana da função objetivo. No caso da função objetivo definida pela equação (2.54) a matriz hessiana é dada por:

$$V_N''(\hat{\theta}, \varphi(t, \hat{\theta})) = \sum_{t=1}^N \psi(\hat{\theta}, \varphi(t, \hat{\theta})) \psi(\hat{\theta}, \varphi(t, \hat{\theta}))^T + \sum_{t=1}^N \psi'(\hat{\theta}, \varphi(t, \hat{\theta})) \psi(\hat{\theta}, \varphi(t, \hat{\theta})) \quad (2.65)$$

onde  $\psi'(\hat{\theta}, \varphi(t, \hat{\theta}))$  é a hessiana do erro de predição.

O cálculo do segundo termo da equação (2.65) pode ter um custo computacional elevado devido ao cálculo de todos os termos de  $\psi'(\hat{\theta}, \varphi(t, \hat{\theta}))$ . Entretanto esse termo, próximo do mínimo global, é desprezível. Além disso, quando a função objetivo entre a iteração atual e o mínimo global não pode ser bem aproximada por uma função quadrática, o efeito da correção da direção de busca pela hessiana não é muito importante. Adicionalmente, a omissão do segundo termo da equação (2.65) garante que a aproximação da hessiana será sempre positiva definida, o que garante a convergência do processo de busca iterativa para um ponto de mínimo [Ljung, 1987].

## 2.4 Identificação de modelos lineares no espaço de tempo contínuo

No caso da identificação de modelos lineares no espaço de tempo contínuo a estrutura dos modelos a serem identificados, sem considerar compensação de distúrbios não medidos, é a seguinte:

$$\sum_{i=0}^n a_i \frac{d^i y(t)}{dt^i} = \sum_{j=0}^m b_j \frac{d^j u(t)}{dt^j} \quad n \geq m \quad (2.66)$$

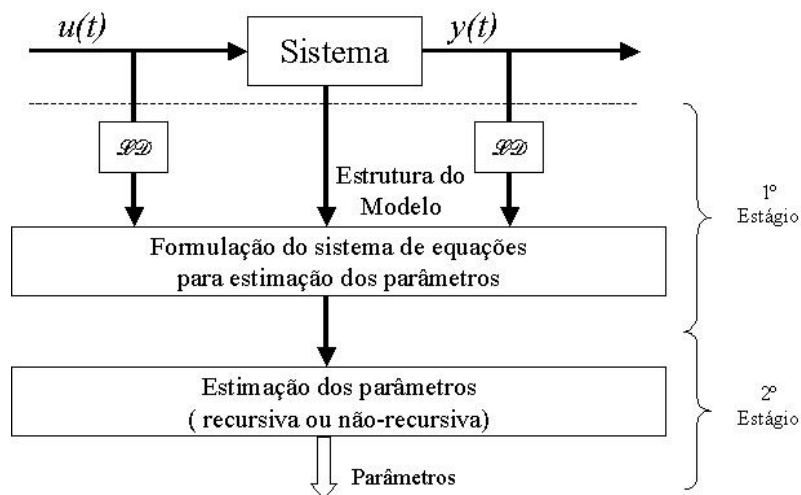
O vetor regressor para a identificação dos parâmetros da estrutura de modelo representado pela equação (2.66) contém, portanto, as derivadas dos sinais de entrada e saída do sistema a ser identificado. Na maioria dos casos práticos as derivadas desses sinais não são medidas e o cálculo dessas não é viável na prática devido a presença natural de ruídos de



medição nos sinais de entradas e saídas. A operação de derivação aplicada sobre esse tipo de sinal provoca a ampliação da magnitude do ruído. Como consequência os métodos de identificação que fazem uso do cálculo direto das derivadas só são aplicáveis nos casos em que a presença de ruído nos sinais é insignificante, o que é praticamente inexistente na indústria de processos.

O problema da identificação de modelos lineares no espaço de tempo contínuo está portanto na primeira etapa do processo geral de identificação, ou seja, na montagem das equações para ajuste dos parâmetros uma vez que o vetor de regressores  $\varphi(t)$  contém as derivadas dos sinais de entrada e saída, as quais são medidas diretamente ou, mais freqüentemente, não podem ser calculadas devido à presença de ruído nos sinais de entrada e saída.

A solução para esse problema é a aplicação de uma operação linear nos sinais de entrada e saída do sistema bem como na estrutura do modelo a ser identificado a fim de transformar a equação diferencial (2.66) em uma equação algébrica. O processo de identificação para modelos lineares no espaço de tempo contínuo pode ser resumido pela figura 2.8.

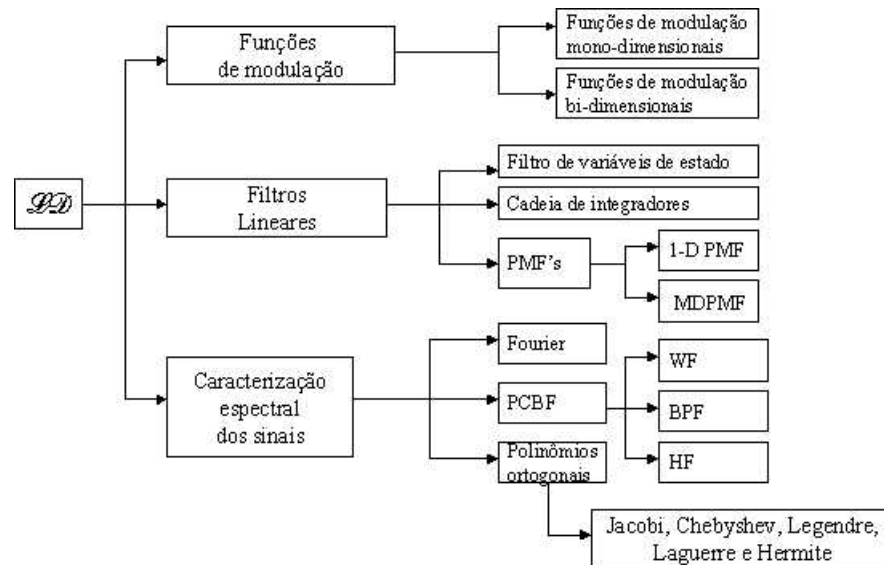


**Figura 2.8:** Diagrama do método geral de identificação de modelos lineares no espaço de tempo contínuo [Unbehauen e Rao, 1987].

Os diferentes métodos para identificação de modelos lineares no espaço de tempo contínuo diferenciam-se basicamente em relação ao tipo de transformação linear utilizada na primeira etapa do processo de identificação.

As principais transformações lineares utilizadas no processo de identificação de modelos lineares no espaço de tempo contínuo podem ser agrupadas em três grandes categorias. A primeira grande categoria agrupa as transformações que fazem uso de funções de modulação, a segunda categoria agrupa as transformações que utilizam cadeias de filtros lineares, categoria onde se encaixa o método dos momentos de Poisson que é o método explorado nesse trabalho e finalmente a terceira categoria que utiliza transformações que

trabalham com a caracterização espectral dos sinais de entrada e saída do sistema em relação a sistemas de funções ortonormais. A figura 2.9 mostra um diagrama onde as diversas transformações lineares estão agrupadas nas categorias acima descritas:



**Figura 2.9:** Classificação das diversas transformações lineares utilizadas na identificação de sistema lineares no espaço de tempo contínuo [Unbehauen e Rao, 1987].

Para exemplificar cada uma das categorias descritas acima utilizaremos como modelo a seguinte função de transferência:

$$G(s) = \frac{b}{as + 1} \quad (2.67)$$

que corresponde à seguinte equação diferencial ordinária:

$$a \frac{dy(t)}{dt} + y(t) = bu(t) \quad (2.68)$$

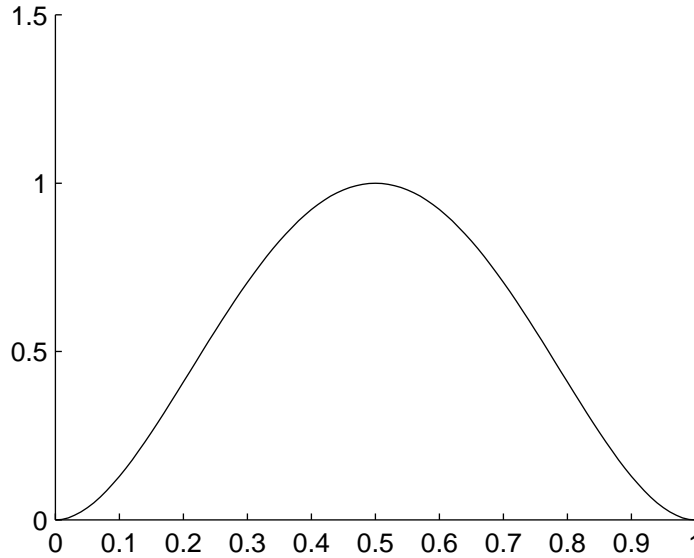
### 2.4.1 Método das funções de modulação

O método das funções de modulação consiste basicamente em multiplicar os sinais de entrada e saída por funções de modulação e integrar esses produtos dentro do intervalo de tempo dos dados observados  $[0, t_0]$ , [Unbehauen e Rao, 1987].

As funções de modulação são funções  $\omega(t)$ , contínuas e infinitamente diferenciáveis, e com suas derivadas conhecidas de forma analítica. Para o exemplo da equação (2.68), um sistema de primeira ordem, as funções de modulação devem possuir as seguintes características.

$$\omega(0) = \omega(t_0) = \left. \frac{d\omega(t)}{dt} \right|_{t=0} = \left. \frac{d\omega(t)}{dt} \right|_{t=t_0} = 0 \quad (2.69)$$

Um exemplo típico dessas funções de modulação pode ser visto na figura 2.10 a seguir.



**Figura 2.10:** Exemplo típico de função de modulação. Nesse caso a função utilizada é um polinômio com a seguinte forma:  $\omega(t) = t^2(a_3t^3 + a_2t^2 + a_1t + a_0)$ .

Considerando um conjunto de funções de modulação  $\omega_n(t)$ , onde  $n = 1, 2, 3, \dots$ . A primeira etapa do método das funções de modulação consiste em multiplicar a equação do modelo a ser ajustada, no caso a equação (2.69), por  $\omega_n(t)$  e integrar o produto no intervalo  $[0, t_0]$ .

$$a \int_0^{t_0} \frac{dy(t)}{dt} \omega_n(t) dt + \int_0^{t_0} y(t) \omega_n(t) dt = b \int_0^{t_0} u(t) \omega_n(t) dt, \quad n = 1, 2, 3, \dots \quad (2.70)$$

Integrando por partes o primeiro termo do lado esquerdo da equação (2.70) temos o seguinte:

$$ay(t)\omega_n(t)\Big|_0^{t_0} - a \int_0^{t_0} y(t) \frac{d\omega_n(t)}{dt} dt + \int_0^{t_0} y(t) \omega_n(t) dt = b \int_0^{t_0} u(t) \omega_n(t) dt, \quad n = 1, 2, 3, \dots \quad (2.71)$$

De acordo com as propriedades das funções de modulação utilizadas, equação (2.69), o primeiro termo do lado esquerdo da equação (2.71) é nulo o que leva finalmente a equação utilizada para a estimação dos parâmetros do modelo dado pelas equação (2.67).

$$a \int_0^{t_0} y(t) \frac{d\omega_n(t)}{dt} dt + b \int_0^{t_0} u(t) \omega_n(t) dt = \int_0^{t_0} y(t) \omega_n(t) dt, \quad n = 1, 2, 3 \dots \quad (2.72)$$

Observando a equação (2.72) podemos notar que o termo da derivada de primeira ordem do sinal de saída foi substituído pela deriva de primeira ordem da função de modulação que não possui ruído de medição.

Como o modelo definido pela equação (2.67) possui dois parâmetros a serem determinados são necessárias duas equações de estimação de parâmetros definidas pela equação (2.72), ou seja, são necessárias duas funções de modulação para identificar esse modelo. O sistema de equações para identificação desse sistema pode então ser escrito da seguinte forma:

$$\begin{bmatrix} \int_0^{t_0} y(t) \frac{d\omega_1(t)}{dt} dt & \int_0^{t_0} u(t) \omega_1(t) dt \\ \int_0^{t_0} y(t) \frac{d\omega_2(t)}{dt} dt & \int_0^{t_0} u(t) \omega_2(t) dt \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \int_0^{t_0} y(t) \omega_1(t) dt \\ \int_0^{t_0} y(t) \omega_2(t) dt \end{bmatrix} \quad (2.73)$$

### 2.4.2 Método dos filtros lineares

O método dos filtros lineares consiste na aplicação de um filtro ou uma cadeia de filtros lineares nos sinais de entrada e saída do sistema a ser modelado. Dentre esse métodos o mais estudado é o método dos momentos de Poisson que é o foco desse trabalho.

O método dos momentos de Poisson faz uso de cadeias de filtros lineares de primeira ordem cujas saídas podem ser combinadas de forma a fornecer uma transformação linear das derivadas dos sinais de entrada e saída do sistema a ser modelado. A aplicação dessa transformação em uma equação diferencial do tipo (2.66) a transforma em uma equação algébrica. Esse método é abordado em detalhes no capítulo 3 dessa dissertação.

### 2.4.3 Método da caracterização espectral dos sinais (método das funções ortonormais)

O método das funções ortonormais baseia-se na caracterização espectral dos sinais de entrada e saída em relação a uma base de funções ortonormais quaisquer  $[\Psi_n]$ . Um sinal qualquer  $f(t)$  pode ser representado através de uma série infinita de funções ortonormais da seguinte forma:

$$f(t) = \sum_{i=1}^{\infty} f_i \Psi_i(t) \quad (2.74)$$

onde os parâmetros  $f_i$  são calculados de acordo com

$$f_i = (f(t), \Psi_i(t)) = \frac{\int_0^{t_0} f(t) dt}{\|\Psi_i(t)\|} \quad (2.75)$$

No caso do modelo linear de primeira ordem definido pela equação (2.67), como temos dois parâmetros a serem ajustados, é necessário expandir os sinais de entrada e saída em pelo menos duas funções ortonormais para termos o número mínimo de equações para a obtermos a solução para esses parâmetros. Assim, temos que os sinais de entrada e saída podem ser aproximados por:

$$\begin{aligned} y(t) &\approx y_1 \Psi_1(t) + y_2 \Psi_2(t) \\ u(t) &\approx u_1 \Psi_1(t) + u_2 \Psi_2(t) \end{aligned} \quad (2.76)$$

onde os parâmetros  $y_1, y_2, u_1$  e  $u_2$  são calculados de acordo com a equação (2.75).

Integrando a equação (2.68) no intervalo  $[0, t_0]$  temos:

$$ay(t_0) - ay(0)s(t) + \int_0^{t_0} y(\tau) d\tau = b \int_0^{t_0} u(\tau) d\tau \quad (2.77)$$

onde  $s(t)$  é a função degrau unitário em  $t=0$ . Essa função, bem como as integrais de  $\Psi_n(t)$ ,  $n=1,2$ , podem também ser representadas em função da base de funções ortonormais utilizadas, de forma que:

$$\begin{aligned} s(t) &\approx s_1 \Psi_1(t) + s_2 \Psi_2(t) \\ \int_0^{t_0} \Psi_1(\tau) d\tau &\approx e_{11} \Psi_1(t) + e_{12} \Psi_2(t) \\ \int_0^{t_0} \Psi_2(\tau) d\tau &\approx e_{21} \Psi_1(t) + e_{22} \Psi_2(t) \end{aligned} \quad (2.78)$$

Inserindo as equações (2.76) e (2.78) na equação (2.77) temos:

$$\begin{aligned} &a[y_1 \Psi_1(t) + y_2 \Psi_2(t) - y(0)s_1 \Psi_1(t) - y(0)s_2 \Psi_2(t)] \\ &+ [y_1 [e_{11} \Psi_1(t) + e_{12} \Psi_2(t)] + y_2 [e_{21} \Psi_1(t) + e_{22} \Psi_2(t)]] \\ &= b[u_1 [e_{11} \Psi_1(t) + e_{12} \Psi_2(t)] + u_2 [e_{21} \Psi_1(t) + e_{22} \Psi_2(t)]] \end{aligned} \quad (2.79)$$

Agrupando os termos da equação (2.79) que contém a mesma função base e igualando as parcelas obtidas nos dois lados da equação obtemos as seguintes equações:

$$\begin{aligned} \Psi_1 : a[y_1 - y(0)s_1] + [y_1 e_{11} + y_2 e_{21}] &= b[u_1 e_{11} + u_2 e_{21}] \\ \Psi_2 : a[y_2 - y(0)s_2] + [y_1 e_{12} + y_2 e_{22}] &= b[u_1 e_{12} + u_2 e_{22}] \end{aligned} \quad (2.80)$$

que podem ser reescritas na forma de um sistema de equações lineares:

$$\begin{bmatrix} (y(0)s_1 - y_1) & (u_1e_{11} + u_2e_{21}) \\ (y(0)s_2 - y_2) & (u_1e_{12} + u_2e_{22}) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} (y_1e_{11} + y_2e_{21}) \\ (y_1e_{12} + y_2e_{22}) \end{bmatrix} \quad (2.81)$$

Como pôde ser visto pelos exemplos acima os sistemas de equações para estimação dos parâmetros do modelo gerados pelos diferentes métodos são sistema algébricos lineares. As transformações lineares no primeiro estágio do processo de identificação de sistemas lineares no espaço de tempo contínuo portanto têm a função de transformar um modelo descrito por uma equação diferencial em uma equação algébrica que pode por sua vez ser transformada em um sistema linear de equações algébricas.

O segundo estágio do processo de identificação de modelos lineares no espaço de tempo contínuo é idêntico ao caso da identificação no espaço de tempo discreto, pois esse estágio é apenas a aplicação de algum método de regressão ou otimização ao sistema de equações algébricas utilizadas para o ajuste dos parâmetros do modelo.

## 2.5 Considerações sobre a identificação de modelos lineares no domínio de tempo discreto e no espaço de tempo contínuo

A identificação de modelos lineares no espaço de tempo discreto teve um grande desenvolvimento após o advento dos sistemas digitais de controle em função da natureza discreta desses e pelo fato dos métodos de identificação discretos não necessitarem da determinação das derivadas dos sinais de entrada e saída do sistema a ser modelado. Entretanto a identificação no espaço de tempo discreto apresenta algumas desvantagens e problemas.

Em primeiro lugar os sistemas reais, principalmente na indústria de processos, são de natureza contínua. Além disso, os sistemas de controle digitais, apesar de sua natureza discreta, utilizam, em sua grande maioria, equações de controladores contínuos. Portanto para a sintonia baseada em identificação é muito útil a obtenção de modelos no espaço de tempo contínuo. Uma solução para esse problema é a transformação dos modelos discretos em modelos contínuos. Essa solução, entretanto apresenta um inconveniente. Apesar de o mapeamento dos pólos entre os espaço contínuo (Laplace) e discreto (Transformada Z) ser único e dado pela seguinte expressão

$$z_i = e^{p_i \Delta t} \quad (2.82)$$

o mesmo não vale para os zeros. O mapeamento do zeros é, segundo [Aström, 1994], dependente do intervalo de amostragem dos dados o que torna a transformação de um modelo discreto em um modelo contínuo uma operação não trivial.

Outro problema enfrentado pela identificação no espaço de tempo discreto é o fato de que todos os métodos dessa categoria devem adotar uma suposição sobre a dinâmica dos dados de processo entre os instantes de amostragem. A suposição mais simples e mais empregada, é considerar que os sinais de entrada são constantes entre os instantes de amostragem (*zero order hold*).

Entretanto as únicas variáveis que realmente se mantêm constantes entre os intervalos de amostragem são aquelas manipuladas diretamente pelo sistema de controle, como por exemplo aberturas de válvulas de controle. As demais variáveis, ao contrário, possuem, normalmente, dinâmicas entre os intervalos de amostragem. Assim muitos sistemas a serem identificados não satisfazem a suposição de valor constante entre os intervalos de amostragem. Nesse caso métodos de identificação discretos com tempos de amostragem típicos (aproximadamente um quarto da constante de tempo ou metade do tempo morto) produzem modelos de pouca qualidade devido a não consideração das dinâmicas entre os tempos de amostragem. Esse efeito, conhecido como efeito *aliasing*, afeta a identificação de qualquer tipo de modelo dinâmico (discreto ou contínuo) e é demonstrado com detalhes no apêndice B.

Por outro lado se reduzirmos o tempo de amostragem o suficiente para garantirmos a qualidade dos modelos enfrentamos outro problema. De acordo com a equação (2.82) os pólos, e por conseqüência os parâmetros dos modelos no espaço de tempo discretos dependem do tempo de amostragem. De acordo com essa correlação a medida que aproximamos o intervalo de amostragem de zero os pólos convergem para o ponto (1,0) no plano z. No caso de modelos no espaço de tempo contínuo essa correlação não existe pois os parâmetros dos modelos, nesse caso, não dependem do tempo de amostragem. Essa diferença resulta em comportamentos do erro de modelagem em função do intervalo de amostragem completamente diferentes [Lee *et al.*, 2001]. Para exemplificar esse fato tomemos como exemplo um sistema de primeira ordem descrito por:

$$\begin{aligned}\tau \frac{dz(t)}{dt} + z(t) &= u(t) \\ y(t) &= z(t) + e(t)\end{aligned}\tag{2.83}$$

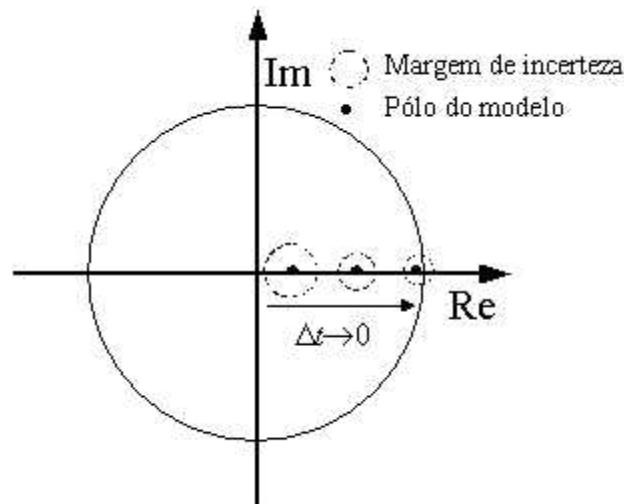
onde  $e(t)$  é um ruído branco discreto (constante entre os tempos de amostragem) com média zero e distribuição gaussiana. Os sinais  $u(t)$  e  $y(t)$  representam a entrada e saída medidas, de forma discreta, do sistema e  $z(t)$  a saída real (não contaminada pelo ruído). O intervalo de amostragem é  $\Delta t$ . A versão discreta do modelo (2.83) é:

$$\begin{aligned}\hat{z}(q) &= \hat{a}\hat{z}(q-1) + (1-\hat{a})u(q-1) \\ \hat{y}(q) &= \hat{z}(q)\end{aligned}\tag{2.84}$$

onde  $\hat{a} = \exp(-\Delta t / \hat{\tau})$  é a estimativa do parâmetros real do sistema  $a = \exp(-\Delta t / \tau)$  e  $\Delta t$  é o intervalo de amostragem. Podemos supor que a estimativa do parâmetro é afetada por um erro de modelagem  $\delta$  de forma que  $\hat{a} = a + \delta$ . Isso significa que a estimativa da constante de tempo do sistema definido pela equação (2.83) é dado por:

$$\tau = -\frac{\Delta t}{\ln(\delta + a)} \quad (2.85)$$

Deve-se observar que, como já mencionado, o parâmetro  $a$  é função do tempo de amostragem. A figura 2.11 mostra graficamente o movimento do parâmetro  $a$  e sua margem de incerteza ( $\delta$ ) à medida que o tempo de amostragem tende para zero.



**Figura 2.11:** Convergência do parâmetro do sistema e de sua correspondente margem de incerteza a medida que o tempo de amostragem é reduzido [Lee *et al.*, 2001].

À medida que o tempo de amostragem é reduzido o parâmetro do modelo discretizado converge para 1. Assim, para garantir uma acuracia constante na identificação da constante de tempo em relação ao tempo de amostragem é necessário que a margem de incerteza convirja para zero com mesma taxa de convergência que o parâmetro  $a$  converge para 1. De outra forma diferentes níveis de acuracia para a constante de tempo estimada serão obtidos em função do tempo de amostragem.

O problema dos métodos de identificação no espaço de tempo discreto com tempos de amostragem muito pequenos provém da tentativa de estimar parâmetros muito próximos de 1, que é o limite de estabilidade no plano  $z$ . Nesse caso, mesmo um pequeno erro  $\delta$  pode levar a estimativas inaceitáveis ou mesmo instáveis (fora do círculo unitário).

O parâmetro do sistema  $a$  converge para 1 na taxa de convergência  $1 - a = 1 - \exp(-\Delta t / \tau) \leq \Delta t / \tau$ . Se assumirmos que a covariância do ruído  $e(t)$  não se altera em função do tempo de amostragem e que o tempo total do teste para obtenção de dados é



fixo, o erro esperado para o parâmetro do modelo discreto  $a$  é aproximadamente inversamente proporcional à raiz quadrada do número total de pontos amostrados ( $\propto 1/\sqrt{N}$ ), ou, de forma equivalente, diretamente proporcional à raiz quadrada do tempo de amostragem ( $\propto \sqrt{\Delta t}$ ), [Ljung, 1987]. Assim o erro esperado para o parâmetro do modelo discreto  $a$  pode ser representado aproximadamente por:

$$\delta = \delta_a \sqrt{\Delta t} \quad (2.86)$$

onde  $\delta_a$  é uma constante.

É possível se demonstrar que sempre existe um tempo de amostragem pequeno o suficiente que faz com que  $\delta$ , definido pela equação (2.86), seja maior que  $\Delta t / \tau$ . Em tais casos a identificação no espaço de tempo discreto com intervalo de amostragem pequeno pode resultar em modelos instáveis. É importante, portanto, reconhecer que a taxa de convergência do parâmetro do modelo discreto  $a$  para 1,  $(\Delta t / \tau)$  e a taxa de convergência para zero da margem de erro ( $\delta_a \sqrt{\Delta t}$ ) são diferentes. Essa diferença resulta em diferentes graus de acuracia em função do tempo de amostragem podendo mesmo gerar modelos instáveis para tempos de amostragem muito pequenos.

Por outro lado, a constante de tempo do sistema definido pela equação (2.83) pode ser estimada com os mesmos dados do caso discreto, utilizando para isso um método de identificação no espaço de tempo contínuo, o que dá uma taxa de convergência da margem de erro para zero proporcional a  $1/\sqrt{N}$ .

$$\begin{aligned} \hat{\tau} \frac{d\hat{z}(t)}{dt} + \hat{z}(t) &= u(t) \\ \hat{y}(t) &= \hat{z}(t) \end{aligned} \quad (2.87)$$

Nos métodos de identificação no espaço de tempo contínuo, portanto, não existe o problema do intervalo de amostragem, uma vez que o parâmetro do sistema  $\tau$  não depende do tempo de amostragem. No caso da identificação no espaço de tempo contínuo, portanto, quanto menor o tempo de amostragem, melhor é a qualidade do modelo.

Podemos confirmar essas características de uma forma mais sistemática comparando as respostas no domínio da frequência do erros de modelagem relativos dos dois métodos. Para analisar a influência do tempo de amostragem assumimos que a função de covariância do ruído e que a entrada do processo não variam com o tempo de amostragem e que o tempo total é fixo. Então podemos assumir o seguinte erro esperado dos parâmetros de um modelo discreto em função do número de pontos amostrados.

$$\hat{a} = a + \delta_a / \sqrt{N} \quad (2.88)$$

$$\hat{G}(z) = \frac{1 - \hat{a}}{z - \hat{a}} = \frac{1 - a - \delta_a / \sqrt{N}}{z - a - \delta_a / \sqrt{N}} \quad (2.89)$$

$$G(z) = \frac{1-a}{z-a} \quad (2.90)$$

onde  $N$  é o número total de pontos amostrados ( $N = t_f/\Delta t$ , onde  $t_f$  é o tempo de duração do experimento) e  $\delta_d$  é uma constante. A equação (2.90) é a função de transferência exata do processo e a equação (2.89) é uma função de transferência estimada determinada usando um método de identificação discreto qualquer. Então, a resposta no domínio da frequência do erro de modelagem relativo do modelos discreto identificado é dada por:

$$E_d(z) = \left| \frac{G(z) - \hat{G}(z)}{G(z)} \right|_{z=\exp(j\omega\Delta t)} \quad (2.91)$$

Da mesma maneira, podemos obter o seguinte para um modelo identificado no espaço de tempo contínuo.

$$\hat{\tau} = \tau + \delta_c / \sqrt{N} \quad (2.92)$$

$$\hat{G}(s) = \frac{1}{(\tau + \delta_c / \sqrt{N})s + 1} \quad (2.93)$$

$$G(s) = \frac{1}{\tau s + 1} \quad (2.94)$$

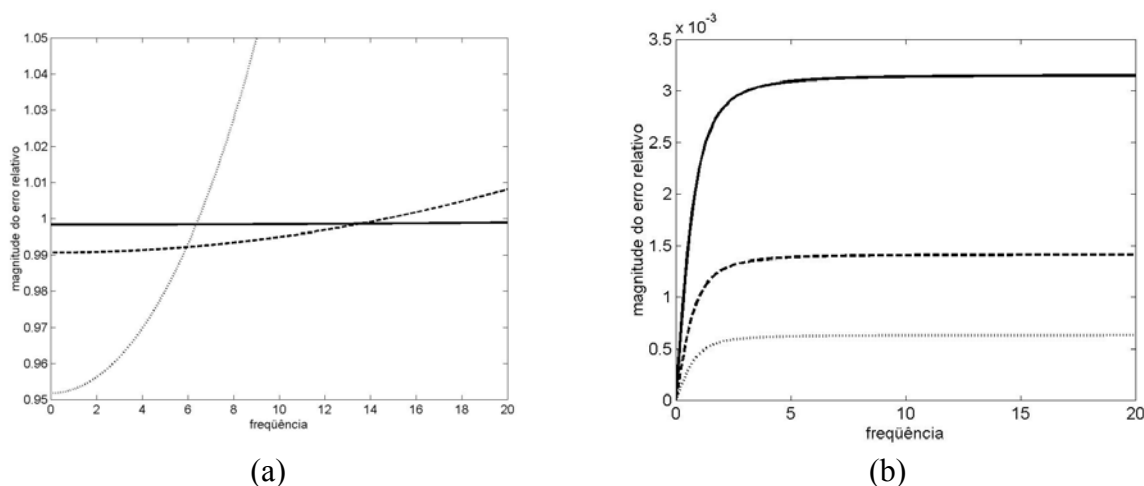
onde  $\delta_c$  é uma constante. Então, a resposta no domínio da frequência do erro de modelagem relativo do modelo contínuo identificado é dada por:

$$E_c(s) = \left| \frac{G(s) - \hat{G}(s)}{G(s)} \right|_{s=j\omega} \quad (2.95)$$

A figura 2.12 mostra a resposta no domínio da frequência dos erros de modelagem relativos quando  $t_f = 1$  e  $\delta_c = \delta_d = 0,01$  por exemplo. Podemos ver marcante diferença entre o comportamento dos erros de ajuste com relação ao tempo de amostragem entre os dois métodos. A magnitude do erro de ajuste do modelo contínuo, equação (2.95), diminui, mas a magnitude do erro de ajuste do modelo discreto, equação (2.91), aumenta com a redução do tempo de amostragem. Isso ocorre porque o parâmetro real e o erro de ajuste convergem com taxa diferentes no caso do modelo discreto. Em contraste, no caso do modelo contínuo o erro de ajuste diminui a medida que o tempo de amostragem é reduzido, ao passo que o parâmetro real permanece constante.

De um ponto de vista mais geral, podemos explicar o problema do tempo de amostragem no caso da identificação do domínio de tempo discreto da seguinte forma: Nos métodos de identificação discretos, todos os pólos localizados no lado negativo do plano da

transformada de Laplace, pólos estáveis, convergem para o ponto (1,0) do plano da transformada-z, e os correspondentes erros de ajuste também convergem para zeros à medida que o tempo de amostragem é reduzido. Não há tal problema de tempo de amostragem caso as taxas de convergência dos pólos do modelo discretos e os correspondentes erros de ajuste sejam a mesma. Entretanto, esse pode não ser o caso em situações reais. A magnitude do erro de ajuste pode aumentar à medida que o tempo de amostragem é reduzido para certos ruídos. Em contraste, no caso dos métodos de identificação no domínio de tempo contínuo, todos os pólos são constantes e o erro de ajuste converge para zero. Então, o erro de ajuste se torna menor à medida que o tempo de amostragem é reduzido.



**Figura 2.12:** Erros de modelagem relativos de modelos no domínio de tempo discreto (a) e um modelo no domínio de tempo contínuo (b) em relação ao tempo de amostragem. ( $\Delta t = 0,1$ , linha sólida;  $\Delta t = 0,02$ , linha tracejada;  $\Delta t = 0,004$ , linha pontilhada) [Lee *et al.*, 2001].

Na verdade, a variância do ruído varia com o tempo de amostragem em vez de se manter constante. Entretanto isso não muda a relativa vantagem do método de identificação contínuo comparado com o discreto uma vez que a variância do ruído afeta os dois métodos da mesma forma.

Outro problema decorrente da redução do tempo de amostragem nos métodos de identificação discretos é que, no caso da existência de mais de um pólo, a convergência desses para o ponto (1,0) no plano da transformada-z torna necessário o uso de uma precisão numérica cada vez mais elevada pois o menor erro de truncamento pode alterar, de maneira significativa, a dinâmica do modelo identificado.

Com base no acima exposto podemos justificar a necessidade de métodos de identificação contínuos uma vez que os métodos de identificação discretos são afetados diretamente pelo tempo de amostragem dos dados utilizados na identificação.

---

## 2.6 Referências Bibliográficas

- Ljung, L., “System Identification: Theory for the User”, Prentice Hall: Englewood Cliffs, NJ (1987).
- Sinha, N. K. E Kuszta, B., “Modeling and Identification of Dynamic Systems”, Van Nostrand Reinhold Eletrical/Computer Science and Engineering series: Van Nostrand Reinhold company (1983).
- Ogata, K., “Modern Control Engineering”, Prentice Hall: Englewood Cliffs, NJ, (1997) pp 710-718.
- Unbehauen, H. e Rao, G. P., “Identification of Continuous Systems”, North-Holland Systems and Control Series, V.10, pp 167-199 (1987).
- Aström, K. J., Hagander, P. e Sternby, J., “Zeros os Sampled Systems”, *Automática*, v. 10, N° 1, pp 31-38, (1994).
- Lee, S. Y., Kwak, H. J. e Lee, I., “Continuous-Time Subspace System Identification Method”, *Ind. Eng. Chem. Res.* V. 40, pp 2886-2896, (2001).
- Box, G. E. O. e Jenkins, G. M., “Time series Analysis, Forecasting and Control”, Ed. Holden Day, San Francisco, (1970).

## Capítulo 3

# Método dos Momentos de Poisson

### 3.1 Introdução

Como dito no capítulo anterior o principal problema da identificação de modelos contínuos é a necessidade de se trabalhar com as derivadas dos sinais de entrada e saída do processo a ser identificado. Vimos que para contornar esse problema é necessária a aplicação de uma transformação linear nos dados de entrada e saída. Um dos métodos mais interessantes de transformação linear é o uso dos momentos de Poisson. Nesse capítulo iremos abordar esse método em detalhes.

As primeiras referências à utilização dos momentos de Poisson na identificação de sistema lineares são do final da década de 1960 [Piovozo e Bolgiano, **1968, 1969**]. Nesses trabalhos os momentos de Poisson são apresentados como uma alternativa às expansões em bases de funções trigonométricas (Fourier) para análise espectral de sinais. Em um artigo anterior [Green e Messel, **1953**] demonstram a utilização de derivadas de alta ordem da função delta de Dirac,  $\delta(t)$ , em uma expansão em série de uma função  $f(t)$  qualquer. Nesse artigo é demonstrado que, como a função  $\delta(t)$  não pode ser tratada como uma função ordinária e sim como uma distribuição, a função  $f(t)$  também o deve ser.

Logo após a publicação dos artigos de Piovozo e Bolgiano surge a primeira referência à utilização dos momentos de Poisson na identificação de modelos lineares no domínio de tempo contínuo. [Fairman e Shen, **1970**] conjugam as idéias da expansão de uma função através de uma série de derivadas da função delta com a utilização dos momentos de Poisson dessa mesma função e demonstram que uma função pode ser representada por uma série do tipo:

$$f(t) = f_k = \sum_{k=0}^{\infty} M_k \{f(t)\} e^{-\lambda(t-t_0)} \delta^{(k)}(t-t_0) \quad (3.1)$$

onde:

$$M_k \{f(t)\} \equiv f_k^{t_0} = \int_0^{t_0} f(t) p_k(t_0 - t) dt \quad (3.2)$$

e

$$p_k(t) \equiv p_k^{t_0} = \frac{t_0^k}{k!} e^{-\lambda t} \quad (3.3)$$

A expressão (3.2) é normalmente referida como o momento de Poisson da função  $f(t)$  ao passo que a expressão (3.3) é referida como função pulso de Poisson. Nas equações acima  $t_0$  representa o instante final de um determinado sinal e  $\lambda$  é uma constante real positiva.

Nesse artigo os autores não fornecem maiores detalhes sobre o desenvolvimento e aplicação dessa expansão focando-se apenas na utilização dos momentos de Poisson, definidos pela equação (3.2) na identificação de modelos lineares contínuos.

## 3.2 Cálculo dos momentos de Poisson das derivadas de um sinal

A característica mais interessante do cálculo do momento de Poisson de uma função  $f(t)$  é quando esse é aplicado às derivadas dessa. Substituindo-se a função  $f(t)$  na equação (3.2) por sua derivada de primeira ordem temos:

$$M_k \left\{ \frac{df(t)}{dt} \right\} = \int_0^{t_0} \frac{df(t)}{dt} p_k(t_0 - t) dt \quad (3.4)$$

Inserindo-se a equação (3.3) na equação acima temos:

$$M_k \left\{ \frac{df(t)}{dt} \right\} = \int_0^{t_0} \frac{df(t)}{dt} \frac{(t_0 - t)^k}{k!} e^{-\lambda(t_0 - t)} dt \quad (3.5)$$

que pode ser integrada por partes de forma que:

$$M_k \left\{ \frac{df(t)}{dt} \right\} = \left[ \frac{(t_0 - t)^k}{k!} e^{-\lambda(t_0 - t)} f(t) \right]_0^{t_0} - \int_0^{t_0} \left[ -\frac{(t_0 - t)^{(k-1)}}{(k-1)!} e^{-\lambda(t_0 - t)} + \lambda \frac{(t_0 - t)^k}{k!} e^{-\lambda(t_0 - t)} \right] f(t) dt \quad (3.6)$$

O primeiro termo da equação (3.6) anula-se no limite  $t_0$ . Dessa forma podemos reescrever a equação (3.6) da seguinte forma:

$$M_k \left\{ \frac{df(t)}{dt} \right\} = -\frac{t_0}{k!} e^{-\lambda t_0} f(0) + \int_0^{t_0} \frac{(t_0 - t)^{(k-1)}}{(k-1)!} e^{-\lambda(t_0 - t)} f(t) dt - \lambda \int_0^{t_0} \frac{(t_0 - t)^k}{k!} e^{-\lambda(t_0 - t)} f(t) dt \quad (3.7)$$

Comparando-se a equação (3.7) com as equações (3.2) e (3.3) podemos afirmar que:

$$M_k \left\{ \frac{df(t)}{dt} \right\} = f_{k-1}^{t_0} - \lambda f_k^{t_0} - p_k^{t_0} f(0) \quad (3.8)$$

ou seja, o momento de Poisson da derivada de primeira ordem da função  $f(t)$  é função dos momentos de Poisson da própria função e de sua condição inicial.

Aplicando o mesmo procedimento para a derivada de segunda ordem obtemos o seguinte resultado:

$$M_k \left\{ \frac{d^2 f(t)}{dt^2} \right\} = M_{k-1} \left\{ \frac{df(t)}{dt} \right\} - \lambda M_k \left\{ \frac{df(t)}{dt} \right\} - p_k^{t_0} \frac{df(0)}{dt} \quad (3.9)$$

Observando essa equação notamos que ela tem a mesma estrutura que a equação (3.8) com a diferença que em vez de o momento da derivada de segunda ordem ser função dos momentos da própria função  $f(t)$  ela é função dos momentos das derivadas de primeira ordem. A mesma expressão pode ser desenvolvida para derivadas de ordens mais elevadas de forma que:

$$M_k \left\{ \frac{d^n f(t)}{dt^n} \right\} = M_{k-1} \left\{ \frac{d^{(n-1)} f(t)}{dt^{(n-1)}} \right\} - \lambda M_k \left\{ \frac{d^{(n-1)} f(t)}{dt^{(n-1)}} \right\} - p_k^{t_0} \frac{d^{(n-1)} f(0)}{dt^{(n-1)}} \quad (3.10)$$

Agora se substituirmos a equação (3.8) na equação (3.9), ajustando os índices dos momentos de forma apropriada, obteremos a expressão que nos dá a relação entre o momento de Poisson da derivada de segunda ordem da função  $f(t)$  com os momentos de Poisson da própria função  $f(t)$ , ou seja,

$$M_k \left\{ \frac{d^2 f(t)}{dt^2} \right\} = f_{(k-2)}^{t_0} - 2\lambda f_{(k-1)}^{t_0} + \lambda^2 f_k^{t_0} - (p_{(k-1)}^{t_0} - \lambda p_k^{t_0}) f(0) - p_k^{t_0} \frac{df(0)}{dt}. \quad (3.11)$$

Extrapolando a equação (3.11) podemos chegar à expressão para o cálculo do momento de Poisson da  $n$ -ésima derivada da função  $f(t)$ , dada por:

$$M_k \left\{ \frac{d^n f(t)}{dt^n} \right\} = \sum_{i=0}^n (-\lambda)^i \binom{n}{i} f_{(k-n+i)}^{t_0} - \sum_{j=0}^{n-1} \sum_{m=0}^j (-\lambda)^m \binom{j}{m} p_{(k-j+m)}^{t_0} \frac{d^{(n-j-1)} f(0)}{dt^{(n-j-1)}} \quad (3.12)$$

onde

$$\binom{n}{i} \equiv C_{n,i} = \frac{n!}{i!(n-i)!} \quad (3.13)$$

Portanto é possível calcular os momento de Poisson das derivadas da função  $f(t)$  sem a necessidade de se trabalhar diretamente com as derivadas dessa. Essa é a característica da transformação de Poisson que a torna uma transformação linear interessante de ser aplicada na identificação de modelos lineares contínuos.

Outra forma mais fácil de lidar com o cálculo dos momentos de Poisson das derivadas de um sinal é expressando a equação (3.12) em forma matricial considerando todas as condições iniciais como sendo nulas. Isso é feito definindo-se a seguinte matriz  $\gamma_n$ .

$$\gamma_n = \begin{matrix} j & & & & & & & & i \\ & 0 & 1 & 2 & & n & & & \\ \left[ \begin{array}{cccccccc} (-1)^0 C_{n,0} \lambda^0 & (-1)^1 C_{n,1} \lambda^1 & (-1)^2 C_{n,2} \lambda^2 & \cdots & (-1)^n C_{n,n} \lambda^n & & & & 0 \\ 0 & (-1)^0 C_{(n-1),0} \lambda^0 & (-1)^1 C_{(n-1),1} \lambda^1 & \cdots & (-1)^{(n-1)} C_{(n-1),(n-1)} \lambda^{(n-1)} & & & & 1 \\ 0 & & & & & & & & \\ \vdots & & & & & & & & \\ 0 & 0 & \cdots & \cdots & & 1 & & & n \end{array} \right] & \end{matrix} \quad (3.14)$$

cuja regra de formação é a seguinte:

$$\begin{aligned} \gamma(i, j) &= (-1)^{(j-i)} C_{(n-i),(j-i)} \lambda^{(j-i)}, \text{ se } i \leq j \\ \gamma(i, j) &= 0, \text{ se } i > j \end{aligned} \quad (3.15)$$

Definindo o vetor  $F_n^{t_0}$  como sendo o vetor dos momentos de Poisson do sinal  $f$  no instante  $t_0$ ,  $F_n(0)$  como o vetor das condições iniciais e  $P_n^{t_0}$  como o vetor das funções pulso de Poisson no instante  $t_0$ , isto é,

$$\begin{aligned} F_n^{t_0} &= [f_0(t_0) \quad f_1(t_0) \quad \cdots \quad f_n(t_0)]^T \\ F_n(0) &= [f^{(0)}(0) \quad f^{(1)}(0) \quad \cdots \quad f^{(n)}(0)]^T \\ P_n^{t_0} &= [p_1(t_0) \quad p_2(t_0) \quad \cdots \quad p_n(t_0)]^T \end{aligned} \quad (3.16)$$



e o vetor  $\mathbf{M}_f$  como o vetor dos momentos de Poisson das  $n$  primeiras derivadas de  $f$  no instante  $t_0$ :

$$\mathbf{M}_f = \left[ M_n \left\{ \frac{d^n f(t_0)}{dt^n} \right\} \quad M_n \left\{ \frac{d^{(n-1)} f(t_0)}{dt^{(n-1)}} \right\} \quad M_n \left\{ \frac{d^{(n-2)} f(t_0)}{dt^{(n-2)}} \right\} \quad \dots \quad M_n \{f(t_0)\} \right] \quad (3.17)$$

Podemos, então, definir o vetor  $\mathbf{M}_f$  em função de  $F_n^{t_0}$ ,  $F_n(0)$  e  $P_n^{t_0}$  da seguinte forma:

$$\mathbf{M}_f = \gamma_n F_n^{t_0} - \begin{bmatrix} Q_n & 0_{n \times 1} \\ 0_{1 \times n} & 0 \end{bmatrix} \begin{bmatrix} F_{n-1}(0) \\ 0 \end{bmatrix} \left( \begin{bmatrix} \gamma_{n-1} & 0_{n \times 1} \\ 0_{1 \times n} & 0 \end{bmatrix} \begin{bmatrix} P_n^T \\ 0 \end{bmatrix} \right)^T \quad (3.18)$$

onde  $Q_n$  é uma matriz quadrada definida pela seguinte regra de formação:

$$\begin{aligned} q(i, j) &= 1, & i \geq j \\ q(i, j) &= 0, & i < j \end{aligned} \quad (3.19)$$

### 3.3 Implementação do cálculo dos momentos de Poisson

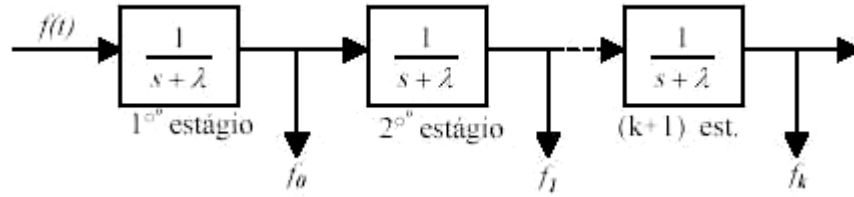
Outra característica interessante do método dos momentos de Poisson é a forma como os momentos da função  $f(t)$  são calculados. Introduzindo-se a equação (3.3) na equação (3.2) temos o seguinte:

$$M_k \{f(t)\} = \int_0^{t_0} f(t) \frac{(t_0 - t)^k}{k!} e^{-\lambda(t_0 - t)} dt \quad (3.20)$$

A equação (3.20) representa a convolução da função  $f(t)$  com a função pulso de Poisson,  $p_k^0(t)$ . Agora analisando-se a função pulso de Poisson vemos que ela corresponde à transformada de Laplace inversa de um filtro linear com pólos múltiplos de ordem  $k$ :

$$\mathcal{L}^{-1} \left\{ \frac{1}{(s + \lambda)^k} \right\} = \frac{t^{(k-1)} e^{-\lambda t}}{(k-1)!}, \quad k \in I \geq 0 \quad (3.21)$$

Isso significa que os momentos de Poisson da função  $f(t)$  podem ser obtidos como as saídas intermediárias de uma cadeia de filtros lineares onde cada estágio da cadeia é composto por um filtro linear de primeira ordem,  $1/(s + \lambda)$ , conforme o mostrado pela figura 3.1. Essa cadeia de filtros é chamada de Cadeia de Filtros de Poisson.



**Figura 3.1:** Cadeia de Filtros de Poisson.

O tamanho da cadeia depende de quantas derivadas queremos utilizar. A cadeia de filtros de Poisson pode ser representada na forma de espaço de estados. Para isso podemos reescrever a cadeia de filtros de Poisson da seguinte forma:

$$\begin{aligned}
 f_0(s) &= \frac{1}{s+\lambda} f(s) \quad \Rightarrow \quad sf_0(s) + \lambda f_0(s) = f(s) \\
 f_1(s) &= \frac{1}{s+\lambda} f_0(s) \quad \Rightarrow \quad sf_1(s) + \lambda f_1(s) = f_0(s) \\
 &\vdots \\
 f_n(s) &= \frac{1}{s+\lambda} f_{(n-1)}(s) \quad \Rightarrow \quad sf_n(s) + \lambda f_n(s) = f_{(n-1)}(s)
 \end{aligned} \tag{3.22}$$

Aplicando-se a transformada inversa de Laplace ao sistema (3.22) obtemos:

$$\begin{aligned}
 \dot{f}_0(t) &= -\lambda f_0(t) + f(t) \\
 \dot{f}_1(t) &= -\lambda f_1(t) + f_0(t) \\
 &\vdots \\
 \dot{f}_n(t) &= -\lambda f_n(t) + f_{(n-1)}(t)
 \end{aligned} \tag{3.23}$$

onde  $\dot{f}$  representa a derivada de primeira ordem de  $f$ . Esse sistema, por sua vez, pode ser reescrito da seguinte forma:

$$\begin{bmatrix} \dot{f}_0(t) \\ \dot{f}_1(t) \\ \dot{f}_2(t) \\ \vdots \\ \dot{f}_n(t) \end{bmatrix} = \begin{bmatrix} -\lambda & 0 & 0 & \cdots & 0 \\ 1 & -\lambda & 0 & \cdots & 0 \\ 0 & 1 & -\lambda & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & 1 & -\lambda \end{bmatrix} \begin{bmatrix} f_0(t) \\ f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} f(t) \tag{3.24}$$

Assim podemos dizer que os momentos de Poisson da função  $f(t)$  podem ser calculados pela seguinte representação no espaço de estados linear:

$$\dot{F} = AF + Bf(t) \tag{3.25}$$

### 3.4 Aplicação dos momentos de Poisson na identificação de modelos lineares invariantes no tempo

Um sistema linear SISO invariante no tempo pode ser representado pela seguinte expressão genérica:

$$\sum_{i=0}^n a_i \frac{d^i y(t)}{dt^i} = \sum_{j=0}^m b_j \frac{d^j u(t)}{dt^j}, \quad n > m \quad (3.26)$$

Pode-se expressar a equação (3.26) de forma vetorial. Como convenção define-se que  $a_n = 1$ .

$$d\mathbf{y}(t)\underline{a}^T = d\mathbf{u}(t)\underline{b}^T \quad (3.27)$$

onde:

$$\begin{aligned} \underline{a} &= [1 \quad a_{n-1} \quad a_{n-2} \quad \cdots \quad a_0]^T \\ \underline{b} &= [b_m \quad b_{m-1} \quad b_{m-2} \quad \cdots \quad b_0]^T \\ d\mathbf{y}(t) &= \left[ \begin{array}{cccc} \frac{d^n y(t)}{dt^n} & \frac{d^{n-1} y(t)}{dt^{n-1}} & \frac{d^{n-2} y(t)}{dt^{n-2}} & \cdots \quad y(t) \end{array} \right] \\ d\mathbf{u}(t) &= \left[ \begin{array}{cccc} \frac{d^m u(t)}{dt^m} & \frac{d^{m-1} u(t)}{dt^{m-1}} & \frac{d^{m-2} u(t)}{dt^{m-2}} & \cdots \quad u(t) \end{array} \right] \end{aligned} \quad (3.28)$$

Aplicando-se a transformação de Poisson sobre a equação (3.27) essa pode ser reescrita da seguinte forma:

$$\mathbf{M}_y \underline{a} = \mathbf{M}_u \underline{b} \quad (3.29)$$

onde os vetores  $\mathbf{M}_u$  e  $\mathbf{M}_y$  são definidos conforme (3.17).

Como os momentos de Poisson das derivadas de um sinal são funções também das condições iniciais do sinal e de suas derivadas é possível estimar o valor dessas expandindo-se os vetores de parâmetros  $\underline{a}$  e  $\underline{b}$  do seguinte modo:

$$\begin{aligned} \underline{a}' &= [1 \quad \cdots \quad a_0 \quad | \quad \chi_n \quad \cdots \quad \chi_1] \\ \underline{b}' &= [b_m \quad \cdots \quad b_0 \quad | \quad \eta_m \quad \cdots \quad \eta_1] \end{aligned} \quad (3.30)$$

onde:

$$\begin{aligned}
 \chi_n &= y(0) \\
 \chi_{(n-1)} &= y^{(1)}(0) + a_{(n-1)}y(0) \\
 &\vdots \\
 \chi_1 &= y^{(n-1)}(0) + a_{(n-1)}y^{(n-2)}(0) + \dots + a_1y(0)
 \end{aligned} \tag{3.31}$$

e

$$\begin{aligned}
 \eta_m &= b_mu(0) \\
 \eta_{(m-1)} &= b_mu^{(1)}(0) + b_{(m-1)}u(0) \\
 &\vdots \\
 \eta_1 &= b_mu^{(m-1)}(0) + b_{(m-1)}u^{(m-2)}(0) + \dots + b_1u(0)
 \end{aligned} \tag{3.32}$$

Os regressores  $\mathbf{M}_y$  e  $\mathbf{M}_u$  também devem ser expandidos da seguinte forma:

$$\begin{aligned}
 \mathbf{M}_y^* &= \left[ \gamma_n Y_n^{t_0} \mid \gamma_{(n-1)} P_n^{t_0} \right] \\
 \mathbf{M}_u^* &= \left[ \gamma_m U_m^{t_0} \mid \gamma_{(m-1)} P_m^{t_0} \right]
 \end{aligned} \tag{3.33}$$

onde os vetores  $P_n^{t_0}$  e  $P_m^{t_0}$  são vetores definidos em (3.16) que contêm os valores das funções pulso de Poisson no instante  $t_0$ .

Como definimos que o parâmetro  $a_n$  é igual a 1 podemos reescrever a equação (3.29) com os vetores expandidos da seguinte forma:

$$\left[ \gamma_n(2:n,:)Y_n^{t_0} \mid \gamma_{(n-1)}P_n^{t_0} \right] \underline{a}' + \gamma_n(1,:)Y_n^{t_0} = \left[ \gamma_m U_m^{t_0} \mid \gamma_{(m-1)} P_m^{t_0} \right] \underline{b}' \tag{3.34}$$

onde  $\gamma_n(2:n,:)$  representa a matriz  $\gamma_n$  sem a primeira linha e a vetor  $\gamma_n(1,:)$  representa a primeira linha de  $\gamma_n$ .

Como os regressores das condições iniciais das entradas e saída são idênticos (funções pulso de Poisson) não podemos estimá-las independentemente sendo necessário estimar uma combinação dessas.

Agrupando-se os regressores referentes às entradas, à saída e às condições iniciais podemos reescrever a equação (3.34) da seguinte forma:

$$\left[ \gamma_n(2:n,:)Y_n^{t_0} \mid -\gamma_m U_m^{t_0} \mid -\gamma_{(n-1)}P_n^{t_0} \right] \left[ \underline{a} \mid \underline{b} \mid \underline{c} \right]^T = -\gamma_n(1,:)Y_n^{t_0} \tag{3.35}$$

onde o vetor  $\underline{c}$  é a combinação das condições iniciais dos sinais de entrada e saída ponderados pelos coeficientes do modelo da seguinte forma:

$$\begin{bmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ \vdots \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ a_{n-1} & 1 & 0 & \cdots & 0 & -b_m & 0 & 0 \\ a_{n-2} & a_{n-1} & 1 & \cdots & 0 & -b_{m-1} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & 1 & -b_1 & -b_2 & -b_3 \end{bmatrix} \begin{bmatrix} y^{(0)}(0) \\ y^{(1)}(0) \\ y^{(2)}(0) \\ \vdots \\ y^{(n)}(0) \\ \hline u^{(0)}(0) \\ \vdots \\ u^{(m)}(0) \end{bmatrix} \quad (3.36)$$

A equação (3.35) é um regressor linear que pode ser resumido da seguinte forma:

$$\mathbf{M}_k \theta = \mathbf{y}_k^{t_0} \quad (3.37)$$

onde:

$$\begin{aligned} \mathbf{M}_k &= [\gamma_n(2:n, :) Y_n^{t_0} \mid -\gamma_m U_m^{t_0} \mid -\gamma_{(n-1)} P_n^{t_0}] \\ \mathbf{y}_k^{t_0} &= -\gamma(1, :) Y_k^{t_0} \\ \theta &= [\underline{a} \mid \underline{b} \mid \underline{c}] \end{aligned} \quad (3.38)$$

Como a equação (3.37) é um regressor linear podemos achar o valor da estimativa dos parâmetros do modelo utilizando o método dos mínimos quadrados, da mesma forma como é feito para o ajuste dos parâmetros da estrutura ARX na identificação de modelos discretos. Assim a estimativa do vetor de parâmetros  $\theta$  é dada por:

$$\hat{\theta} = (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \mathbf{M}_k^T \mathbf{y}_k^{t_0} \quad (3.39)$$

Tendo o vetor  $\hat{\theta}$  podemos achar as condições iniciais dos sinais de entrada e saída através da resolução do sistema dado por (3.36).

Existem duas variações do método de Poisson que podem ser aplicadas na identificação de modelos lineares. Essas variações diferem entre si com relação ao método utilizado na montagem da matriz  $\mathbf{M}_k$  da equação (3.37).

### 3.4.1 Método 1 – Vários estágios da cadeia de filtros de Poisson e um instante de tempo.

Nesse primeiro método utilizamos apenas um instante de tempo dos dados amostrados, normalmente o último, e o número necessário de estágios da cadeia de filtros de Poisson para montar o sistema de equações para estimação dos parâmetros.

Para entender melhor a diferença entre as duas variações iremos usar como termo independente na equação (3.37) o termo  $\gamma_n(1,:)Y_n^{t_0}$ , isto é, igual a saída do  $(n+1)$ -ésimo estágio da cadeia de filtros de Poisson ao invés de  $\gamma_n(n,:)Y_n^{t_0}$ .

Como estamos utilizando apenas um instante de tempo, por exemplo o tempo final  $t_f$ , devemos utilizar uma cadeia de filtros maior. Para exemplificar essa idéia vamos aplicá-la a um sistema linear de primeira ordem dado por:

$$a \frac{dy(t)}{dt} + y(t) = bu(t) \quad (3.40)$$

Aplicando-se a transformação de Poisson sobre a equação (3.40) temos a seguinte equação algébrica:

$$a[y_{k-1}^{t_i} - \lambda y_k^{t_i} - p_k^{t_i} y(0)] + y_k^{t_i} = bu_k^{t_i} \quad (3.41)$$

que pode ser reescrita da seguinte forma agrupando os parâmetros a serem estimados em um vetor e os regressores em outro:

$$\left[ y_{k-1}^{t_i} - \lambda y_k^{t_i} - p_k^{t_i} y(0) \quad -u_k^{t_i} \right] \begin{bmatrix} a \\ b \end{bmatrix}^T = -y_k^{t_i} \quad (3.42)$$

Desejamos montar um sistema de equações do tipo:

$$\mathbf{M}_k \boldsymbol{\theta} = y_k^{t_f} \quad (3.43)$$

onde a matriz  $\mathbf{M}_k$  é dada por:

$$\mathbf{M}_k = \begin{bmatrix} m_1 & m_2 & \cdots & m_k \end{bmatrix}^T \quad (3.44)$$

e onde  $m_k$  é um vetor dado por,

$$m_k = \begin{bmatrix} y_{k-1}^{t_f} - \lambda y_k^{t_f} - p_k^{t_f} y(0) & -u_k^{t_f} \end{bmatrix} \quad (3.45)$$

Assim, para esse primeiro método o sistema de equações lineares originárias de (3.45) é:

$$\begin{bmatrix} y_0^{t_f} - \lambda y_1^{t_f} - p_1^{t_f} y(0) & -u_1^{t_f} \\ y_1^{t_f} - \lambda y_2^{t_f} - p_2^{t_f} y(0) & -u_2^{t_f} \\ \vdots & \vdots \\ y_{k-1}^{t_f} - \lambda y_k^{t_f} - p_k^{t_f} y(0) & -u_k^{t_f} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -y_1^{t_f} \\ -y_2^{t_f} \\ \vdots \\ -y_k^{t_f} \end{bmatrix} \quad (3.46)$$

que pode ser resolvido usando o método dos mínimos quadrados conforme a equação (3.39), caso  $\mathbf{M}_k$  não seja uma matriz quadrada. No caso de  $\mathbf{M}_k$  ser uma matriz quadrada podemos obter a solução para o vetor de parâmetros de uma forma mais direta. Para isso basta inverter a matriz  $\mathbf{M}_k$  e multiplicá-la pelo termo independente conforme:

$$\theta = \mathbf{M}_k^{-1} y_k^{t_f} \quad (3.47)$$

Podemos fazer uma analogia entre esse método e o ajuste de uma curva polinomial de ordem  $n$ . Para isso podemos fazer uso da informação do valor de  $n+1$  pontos ou então da informação do valor de um ponto e de suas  $n$  primeiras derivadas. Esse método é equivalente a segunda alternativa, isto é, utilizamos a informação dos momentos de Poisson de várias ordens de um mesmo instante de tempo para o ajuste dos parâmetros do modelo.

### **3.4.2 Método 2 –Número mínimo de estágios da cadeia de filtros de Poisson e vários instantes de tempo.**

Nesse segundo método utilizamos um número mínimo de estágios da cadeia de filtros de Poisson e vários instantes de tempo. De acordo com a analogia do ajuste de uma curva polinomial utilizada para o primeiro método esse segundo método seria equivalente a primeira opção, ou seja, utilizar a informação de vários pontos para o ajuste da curva.

Para exemplificar vamos aplicar esse método ao exemplo anterior. Nesse método o sistema de equações dado pela equação (3.43) é dado por:

$$\begin{bmatrix} y_0^{t_1} - \lambda y_1^{t_1} - p_1^{t_1} y(0) & -u_1^{t_1} \\ y_0^{t_2} - \lambda y_1^{t_2} - p_1^{t_2} y(0) & -u_1^{t_2} \\ \vdots & \\ y_0^{t_f} - \lambda y_1^{t_f} - p_1^{t_f} y(0) & -u_1^{t_f} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -y_1^{t_1} \\ -y_1^{t_2} \\ \vdots \\ -y_1^{t_f} \end{bmatrix} \quad (3.48)$$

que pode ser resolvido utilizando-se o método dos mínimos quadrados conforme dado pela equação (3.39).

Teoricamente as duas versões do método de Poisson devem apresentar os mesmos resultado, entretanto, na prática isso não se comprova. Na presença de ruído nos dados do processo o segundo método apresenta melhor desempenho do que o primeiro método. Além disso, o primeiro método apresenta uma sensibilidade quanto ao valor final dos dados de entrada e saída do processo.

Para testar o desempenho dos dois métodos foi simulado um processo linear com a seguinte função de transferência.

$$G(s) = \frac{2s + 2}{s^2 + 5s + 6} \quad (3.49)$$

Como sinal de entrada foi utilizado um sinal tipo PRBS (*Pseudo Random Binary Signal*). Foi utilizado um tempo de amostragem de 0.02 e um tempo total de simulação de 100. À saída do processo foi adicionado um ruído branco cuja razão ruído sobre sinal de entrada é igual a aproximadamente 7%. O primeiro método foi utilizado com uma cadeia de filtros com dez estágios.

No primeiro teste de identificação comparou-se os dois métodos utilizando todos os dados (5000 amostragens) com e sem ruído nos dados de saída. Os resultados desse teste podem ser observados na tabela 3.1.

**Tabela 3.1:** Resultados da identificação do sistema (3.51) utilizando as duas variações do métodos de Poisson com dados sem e com ruído (7%).

	MÉTODO 1					MÉTODO 2				
Ruído/Sinal	$a_1$	$a_0$	$b_1$	$b_0$	$\sum(\theta - \hat{\theta})^2$	$a_1$	$a_2$	$b_1$	$b_0$	$\sum(\theta - \hat{\theta})^2$
0	4,979	5,986	1,990	1,995	0,0273	4,986	5,985	1,993	1,995	0,0220
7%	1,050	8,557	0,335	2,872	5,0665	4,501	5,435	0,335	2,872	0,8007
Real	5	6	2	2	-	5	6	2	2	-

Para avaliar o desempenho de ambos os métodos utilizou-se a norma quadrática da diferença entre o vetor de parâmetros do modelo simulado e o vetor de parâmetros estimados. Pode-se observar que quando os sinais do sistema não estão contaminados por ruído ambos os métodos apresentam resultados praticamente idênticos. No entanto na presença de ruído os desempenhos mudam drasticamente. Nesse caso o segundo método apresenta-se muito superior com uma norma de erro quadrado aproximadamente 6 vezes menor do que o primeiro método.

Outro teste realizado com o primeiro método é da sua sensibilidade em relação ao valor final dos dados de entrada e saída do processo. Para isso realizou-se a identificação do modelo (3.49) com quantidades crescente de dados. Primeiramente realizou-se esse teste com os dados livres de ruído. O resultados desse teste estão resumido na tabela 3.2.

Como pode ser visto a partir desses resultados, na ausência de ruído, a primeira variante do método de Poisson apresenta resultados muito parecidos independentemente da quantidade de dados utilizados, desde que, os dados utilizados contenham informação suficiente sobre o sistema.



O segundo teste foi feito com os dados contaminados por ruído branco nos dados de saída utilizando-se as mesmas quantidades de dados do teste anterior. Os resultados desse teste podem ser vistos na tabela 3.3.

**Tabela 3.2:** Resultados do teste de sensibilidade da primeira variação do método de Poisson em relação à quantidade de dados utilizado na identificação sem a presença de ruídos nesses.

# pontos	$a_1$	$a_0$	$b_1$	$b_0$	norm
300	4,9890	5,9849	1,9939	1,9950	0,0203
500	4,9843	5,9834	1,9922	1,9944	0,0248
700	4,9866	5,9882	1,9932	1,9961	0,0195
900	4,9856	5,9832	1,9925	1,9944	0,0241
1100	4,9908	5,9836	1,9946	1,9945	0,0241
1300	4,9832	5,9824	1,9922	1,9942	0,0263
1500	4,9885	5,9836	1,9937	1,9945	0,0217
1700	4,9905	5,9863	1,9946	1,9954	0,0181
1900	4,9891	5,9848	1,9940	1,9949	0,0203
2100	4,9827	5,9853	1,9916	1,9951	0,0247
2300	4,9908	5,9848	1,9947	1,9950	0,0192
2500	4,9789	5,9862	1,9901	1,9954	0,0275
				<i>Média</i>	0,0225
				<i>Des. Pad.</i>	0,0031
				<i>Des.Pad/Média</i>	0,1360

**Tabela 3.3:** Resultados do teste de sensibilidade da primeira variação do método de Poisson em relação à quantidade de dados utilizado na identificação com a presença de ruídos nesses.

# pontos	$a_1$	$a_0$	$b_1$	$b_0$	norm
300	2,2388	5,8951	1,0401	1,9901	2,9252
500	10,9759	14,6813	4,1247	5,1396	11,2003
700	4,4926	6,0578	1,8613	2,0130	0,5293
*900	-3,4300	1,4517	-1,5546	0,4050	10,3407
1100	6,3397	4,4921	2,4997	1,4522	2,1490
1300	5,1053	4,8741	2,0467	1,6057	1,1985
1500	6,4357	5,4416	2,4538	1,8095	1,6172
1700	3,6973	3,8014	1,1609	1,2240	2,7995
*1900	-1,2733	2,9068	-0,3924	0,9430	7,4675
2100	4,8521	8,7903	2,1019	2,9426	2,9506
2300	4,8521	8,7903	2,1019	2,9426	2,9506
♣2500	1,9747	8,4833	0,6930	2,8430	4,2117
				<i>Média</i>	4,1950
				<i>Des. Pad.</i>	3,5350
				<i>Des.Pad/Média</i>	0,8427

\* modelos instáveis, ♣ modelo oscilatório amortecido.

Como os resultados da tabela 3.3 indicam, na presença de ruído, o desempenho dessa primeira versão do método de Poisson apresenta um comportamento aleatório em relação à

quantidade de dados utilizados, podendo, em alguns casos, gerar modelos instáveis. Esse comportamento pode ser entendido pelo fato de estarmos utilizando a informação de um único ponto para ajustar o modelo. Na presença de ruído aleatório esse ponto pode estar mais ou menos distante do valor real. Na segunda versão do método de Poisson não observamos esse comportamento pelo fato de estarmos utilizando informação de vários pontos para identificar o modelos, assim, as variações causadas pelos ruído acabam sendo compensadas.

Para demonstrar a eficiência da segunda variação do método de Poisson, realizamos a identificação do sistema com os dados contaminados por ruído branco com as mesmas amostragem utilizadas no teste anterior. O resultados desse teste podem ser vistos na tabela 3.4, abaixo.

**Tabela 3.4:** Resultados do teste de sensibilidade da segunda variação do método de Poisson em relação à quantidade de dados utilizado na identificação com a presença de ruídos nesses.

# pontos	$a_1$	$a_0$	$b_1$	$b_0$	norm
300	4,4391	5,8583	1,8234	1,9542	0,6066
500	4,1148	5,9747	1,6929	1,9997	0,9373
700	4,2085	5,8170	1,7385	1,9328	0,8560
900	4,5062	5,5090	1,8395	1,8251	0,7357
1100	4,5436	5,3354	1,8375	1,7637	0,8557
1300	4,5473	5,3198	1,8426	1,7576	0,8667
1500	4,4981	5,1878	1,8092	1,7169	1,0140
1700	4,4455	5,3309	1,7828	1,7673	0,9254
1900	4,4593	5,3350	1,7876	1,7684	0,9128
2100	4,4878	5,4287	1,8116	1,7989	0,8153
2300	4,4238	5,4326	1,7952	1,8009	0,8576
2500	4,5013	5,4340	1,8174	1,8005	0,8014
				<i>Média</i>	0,8487
				<i>Des. Pad.</i>	0,1046
				<i>Des.Pad/Média</i>	0,1360

Como pode ser visto a partir dos dados da tabela acima a segunda variante do método de Poisson é muito mais estável que a primeira quando da presença de ruído nos sinais de entrada e saída do sistema a ser identificado.

### 3.5 Método dos Momentos de Poisson Generalizado

O método de Poisson apresentado até o momento é conhecido como método de Poisson ordinário e apresenta como característica o numerador de cada estágio da cadeia de filtros de Poisson igual a 1. Isso implica que o ganho de cada estágio é igual a  $(1/\lambda)$  ou seja no caso de utilizarmos um parâmetro  $\lambda$  maior do que 1 o sinal filtrado pela cadeia de filtros de

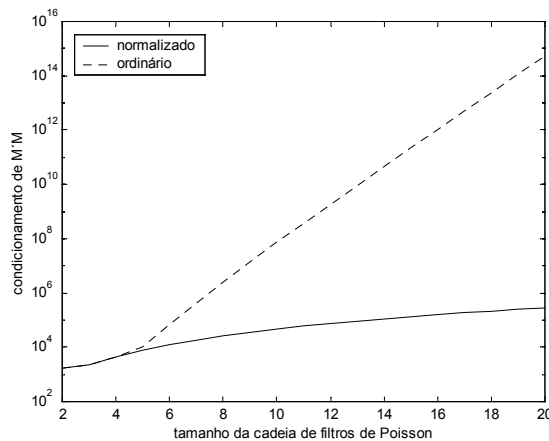
Poisson irá ter sua magnitude reduzida após a passagem por cada estágio da cadeia o que afeta o condicionamento do sistema de equações de estimação dos parâmetros do modelo.

Para evitar esse problema deve-se alterar o ganho dos estágios da cadeia de filtro através da introdução de um fator  $\beta$  [Unbehauen e Rao, 1987]. A introdução desse fator provoca pequenas alterações nas equações desenvolvidas para o método de Poisson ordinário. A equação (3.12) que dá o momento de Poisson da  $n$ -ésima deriva do sinal  $f(t)$  em função dos momentos de Poisson desse fica da seguinte forma:

$$M_k \left\{ \frac{d^n f(t)}{dt^n} \right\} = \sum_{i=0}^n (-\lambda)^i \beta^{(n-i)} \binom{n}{i} f_{(k-n+i)}^{t_0} - \sum_{j=0}^{n-1} \sum_{m=0}^j (-\lambda)^m \beta^{(j-m)} \binom{j}{m} p_{(k-j+m)}^{t_0} \frac{d^{(n-j-1)} f(0)}{dt^{(n-j-1)}} \quad (3.50)$$

Na equação (3.24) usada para a geração dos momentos de Poisson do sinal  $f(t)$  deve-se substituir todos os elementos das matrizes da representação em espaço de estado da cadeia de filtros de Poisson iguais a um pelo parâmetro  $\beta$ .

Essa variação do método de Poisson é conhecida como método dos momentos de Poisson generalizado. No caso particular em que o parâmetro  $\beta$  é igual ao parâmetro  $\lambda$  o método é conhecido como método dos momentos de Poisson normalizado. No método normalizado o ganho de cada estágio da cadeia de filtros é igual a um. Dessa forma a magnitude dos momentos de Poisson dos sinais filtrados pela cadeia não é alterada.



**Figura 3.2:** Comportamento do condicionamento da matriz  $M_k^T M_k$  em função do tamanho da cadeia de filtro de Poisson.

Para exemplificar o efeito da normalização do ganho dos estágios da cadeia de filtros de Poisson sobre o condicionamento do sistema de equações para identificação dos parâmetros do modelo vamos tomar como exemplo o sistema definido pela equação (3.49). Como sinal de entrada foi utilizado um sinal tipo PRBS. Foi utilizado um tempo de amostragem de 0.02 e um tempo total de simulação de 100. Não foi adicionado ruído aos

dados de entrada e saída do sistema. Nesse teste variou-se o tamanho da cadeia de filtros de Poisson de 2 (tamanho mínimo para a identificação de um sistema de 2ª ordem) até 20 e para cada tamanho da cadeia calculou-se o condicionamento da matriz  $M_k^T M_k$  que deve ser invertida no cálculo dos mínimos quadrados como definido em (3.39). A definição de condicionamento utilizada aqui foi a razão entre o maior valor singular e o menor. Utilizou-se um parâmetro  $\lambda$  igual a 2. Os resultados do teste podem ser vistos na figura 3.2, acima.

Como pode ser observado pelo gráfico acima o método normalizado apresenta um comportamento muito mais robusto do que o método ordinário. Quanto maior for a ordem do modelo a ser identificado e quanto maior for o valor da constante  $\lambda$  menor será o condicionamento do sistema. Para esses casos o método normalizado deve ser empregado.

### 3.6 Estimação recursiva dos parâmetros

O algoritmo dos mínimos quadrados discutido anteriormente para a estimação do vetor de parâmetros  $\theta$  utilizando a equação (3.39) possui a desvantagem de necessitar de um número fixo e grande de observações dos sinais de entrada e saída. Quando uma nova observação desses é disponibilizada é necessário recalculá-la toda a matriz  $\mathbf{M}_k$  e o vetor  $\mathbf{y}_k^{t_0}$  da equação (3.37) com as novas medidas e a inversão da matriz  $\mathbf{M}_k^T \mathbf{M}_k$  deve ser feita novamente.

Esse procedimento não se mostra apropriado para o cálculo iterativo de parâmetros pois é necessário o armazenamento do histórico dos dados de entrada e saída para o cálculo da matriz  $\mathbf{M}_k$  e do vetor  $\mathbf{y}_k^{t_0}$ . Além disso a operação de cálculo dos momentos de Poisson e a inversão da matriz  $\mathbf{M}_k^T \mathbf{M}_k$  são operações que consomem a capacidade computacional do computador de processo utilizado.

Entretanto o método dos mínimos quadrados pode ser alterado para uma forma recursiva que não tem a necessidade nem do armazenamento dos dados históricos nem da inversão da matriz  $\mathbf{M}_k^T \mathbf{M}_k$ .

Quando uma nova observação dos sinais de entrada e saída é coletada a matriz  $\mathbf{M}_k$  e o vetor  $\mathbf{y}_k^{t_0}$  podem ser expandidos da seguinte forma:

$$\begin{aligned} \mathbf{M}_{(k+1)} &= \begin{bmatrix} \mathbf{M}_k \\ m_{(k+1)}^T \end{bmatrix} \\ \mathbf{y}_{(k+1)}^{t_0} &= \begin{bmatrix} \mathbf{y}_k^{t_0} & y_{(k+1)} \end{bmatrix}^T \end{aligned} \quad (3.51)$$

onde  $m_{(k+1)}^T$  e  $y_{(k+1)}^{t_0}$  são definidos por (3.38) com a única diferença que os vetores  $Y_n^{t_0}$ ,  $U_n^{t_0}$  e  $P_n^{t_0}$  possuem apenas uma coluna referente a nova observação.

Assim a nova estimativa do vetor de parâmetros é dada por:

$$\hat{\theta}_{(k+1)} = \left( \mathbf{M}_{(k+1)}^T \mathbf{M}_{(k+1)} \right)^{-1} \mathbf{M}_{(k+1)}^T \mathbf{y}_{(k+1)}^{t_0} \quad (3.52)$$

Para simplificar a notação vamos a definir a matriz  $\tilde{\mathbf{M}}_k$  de forma que:

$$\tilde{\mathbf{M}}_k = \mathbf{M}_k^T \mathbf{M}_k \quad (3.53)$$

Assim podemos reescrever a equação (3.52) da seguinte forma:

$$\hat{\theta}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1} \begin{bmatrix} \mathbf{M}_k \\ \mathbf{m}_{(k+1)}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{y}_{(k+1)}^{t_0} \\ \mathbf{y}_{(k+1)}^{t_0} \end{bmatrix} \quad (3.54)$$

Multiplicando as matrizes do lado direito da equação acima temos:

$$\hat{\theta}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1} \left[ \mathbf{M}_k^T \mathbf{y}_k^{t_0} + \mathbf{m}_{(k+1)} \mathbf{y}_{(k+1)} \right] \quad (3.55)$$

Como

$$\tilde{\mathbf{M}}_k \hat{\theta}_k = \mathbf{M}_k^T \mathbf{M}_k \hat{\theta}_k = \mathbf{M}_k^T \mathbf{y}_k^{t_0} \quad (3.56)$$

que pode ser substituída em (3.55) de forma que:

$$\hat{\theta}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1} \left[ \tilde{\mathbf{M}}_k \hat{\theta}_k + \mathbf{m}_{(k+1)} \mathbf{y}_{(k+1)} \right] \quad (3.57)$$

Além disso combinando as equações (3.51) e (3.53) temos que:

$$\tilde{\mathbf{M}}_{(k+1)} = \tilde{\mathbf{M}}_k + \mathbf{m}_{(k+1)} \mathbf{m}_{(k+1)}^T \quad (3.58)$$

Se isolarmos  $\tilde{\mathbf{M}}_k$  na equação acima e substituirmos na equação acima temos:

$$\hat{\theta}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1} \left[ \tilde{\mathbf{M}}_{(k+1)} \hat{\theta}_k - \mathbf{m}_{(k+1)} \mathbf{m}_{(k+1)}^T \hat{\theta}_k + \mathbf{m}_{(k+1)} \mathbf{y}_{(k+1)}^{t_0} \right] \quad (3.59)$$

que pode ser simplificada da seguinte forma:

$$\hat{\theta}_{(k+1)} = \hat{\theta}_k + \tilde{\mathbf{M}}_{(k+1)}^{-1} \mathbf{m}_{(k+1)} \left[ \mathbf{y}_{(k+1)}^{t_0} - \mathbf{m}_{(k+1)}^T \hat{\theta}_k \right] \quad (3.60)$$

Definindo o vetor de ganhos de Kalman como:

$$\mathbf{h}_{(k+1)} = \mathbf{Q}_{(k+1)} \mathbf{m}_{(k+1)} \quad (3.61)$$

onde  $\mathcal{Q}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1}$  para simplificar a notação e observando que a expressão entre colchetes na equação (3.60) pode ser vista como uma predição do erro de modelagem no estágio  $(k+1)$  baseado na estimativa dos parâmetros no estágio anterior. Essa predição é denotada como  $\hat{v}_{(k+1)}^{f_0}$  de forma que:

$$\hat{v}_{(k+1)}^{f_0} = y_{(k+1)}^{f_0} - m_{(k+1)}^T \hat{\theta}_k \quad (3.62)$$

Assim da equação (3.60) obtemos a equação para o cálculo na nova estimativa do vetor de parâmetros como sendo:

$$\hat{\theta}_{(k+1)} = \hat{\theta}_k + h_{(k+1)} \hat{v}_{(k+1)}^{f_0} \quad (3.63)$$

Falta agora definir o vetor de ganhos de Kalman em função de quantidades disponíveis. Para isso consideremos em primeiro lugar que, de acordo com (3.58):

$$\mathcal{Q}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}^{-1} = [\tilde{\mathbf{M}}_k + m_{(k+1)} m_{(k+1)}^T]^{-1} \quad (3.64)$$

Com um pouco de álgebra podemos demonstrar a seguinte propriedade:

$$[\mathbf{A} + \mathbf{BC}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} [\mathbf{I} + \mathbf{CA}^{-1} \mathbf{B}]^{-1} \mathbf{CA}^{-1} \quad (3.65)$$

que pode ser aplicada na equação (3.64) uma vez que a matriz  $\tilde{\mathbf{M}}_k$  é inversível. Fazendo isso podemos reescrever essa equação da seguinte forma:

$$\mathcal{Q}_{(k+1)} = \mathcal{Q}_k - \mathcal{Q}_k m_{(k+1)} [1 + m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}]^{-1} m_{(k+1)}^T \mathcal{Q}_k \quad (3.66)$$

onde a expressão entre colchetes é uma quantidade escalar. Inserindo-se essa equação na equação (3.61) obtemos:

$$h_{(k+1)} = \left\{ \mathcal{Q}_k - \mathcal{Q}_k m_{(k+1)} [1 + m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}]^{-1} m_{(k+1)}^T \mathcal{Q}_k \right\} m_{(k+1)} \quad (3.67)$$

Como a quantidade entre colchetes é um escalar podemos representá-la como um denominador comum de forma que a equação acima pode ser reescrita da seguinte forma:

$$h_{(k+1)} = \frac{\mathcal{Q}_k m_{(k+1)} [1 + m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}] - \mathcal{Q}_k m_{(k+1)} m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}}{1 + m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}} \quad (3.68)$$

que simplificada nos conduz a expressão final para o cálculo do vetor de ganhos de Kalman:

$$h_{(k+1)} = \frac{\mathcal{Q}_k m_{(k+1)}}{1 + m_{(k+1)}^T \mathcal{Q}_k m_{(k+1)}} \quad (3.69)$$

Introduzindo-se essa equação na equação (3.66) temos:

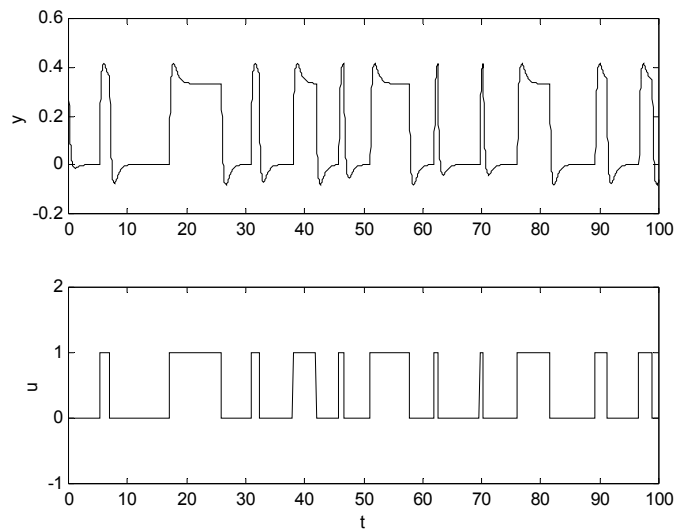
$$\mathcal{Q}_{(k+1)} = \mathcal{Q}_k - h_{(k+1)} m_{(k+1)}^T \mathcal{Q}_k \quad (3.70)$$

As equações (3.62), (3.63), (3.69) e (3.70) formam o núcleo básico do algoritmo dos mínimos quadrados recursivo.

A matriz  $\mathcal{Q}_k$  é chamada de matriz de covariância do vetor de parâmetros  $\hat{\theta}$ . Esse algoritmo recursivo pode ser relacionado à teoria dos filtros de Kalman.

De uma forma geral o algoritmo recursivo pode ser resumido da seguinte forma através dos seguintes passos:

- 1) Estabelecimento de uma estimativa inicial para a matriz de covariância;
- 2) Cálculo do erro de predição da saída do modelo em função da estimativa do vetor de parâmetros estimado na iteração anterior através da equação (3.62);
- 3) Cálculo do vetor de ganhos de Kalman através da equação (3.69);
- 4) Cálculo da nova estimativa do vetor de parâmetros através da equação (3.63);
- 5) Cálculo da nova matriz de covariância através da equação (3.70), volta ao passo 2.



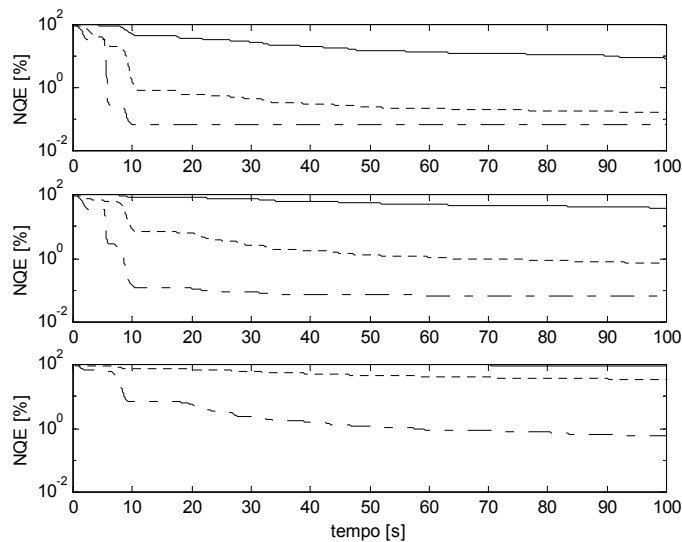
**Figura 3.3:** Dados de entrada e saída utilizados na estimativa do parâmetros do modelo definido por (3.49) utilizando o método dos mínimos quadrados recursivo.

A estimativa inicial da matriz de covariância nos diz o grau de confiança que temos na estimativa inicial do vetor de parâmetros. Quanto maior for o valor do elementos da diagonal da matriz de covariância menor é nossa confiança no vetor de parâmetros. Normalmente tomamos como estimativa inicial da matriz de covariância a matriz identidade multiplicada por um escalar muito maior do que a unidade. O valor dos elementos da diagonal dessa matriz também influenciam na taxa de convergência do método recursivo. Quanto maior forem esses valores maior é a taxa de convergência. O valor da constante  $\lambda$  da cadeia de filtros de Poisson também influencia na taxa de convergência da estimativa do vetor de parâmetros.

Para testar estas características utilizou-se o modelo definido em (3.49) com condições iniciais  $y(0)=0,4$  e  $y'(0)=-1,6$ . O conjunto de dados utilizados na identificação do sistema pode ser visto na figura 3.3. O tempo de amostragem dos dados é 0,02. Foram testados três diferentes valores de  $\lambda$ ,  $\lambda=1$ ,  $\lambda=2$  e  $\lambda=5$ . Também foram testados três diferentes valores para os elementos da diagonal da estimativa inicial da matriz de covariância,  $1.110^2$ ,  $1.10^4$  e  $1.10^6$ . Também foi testado o desempenho quando da utilização do método de Poisson ordinário e o método normalizado. Para testar a taxa de convergência do método recursivo para cada uma das situações acima descritas utilizou-se a norma quadrática normalizada do erro, definida por:

$$NQE[\%] = 100 \frac{\|\theta - \hat{\theta}\|_2}{\|\theta\|_2} \quad (3.71)$$

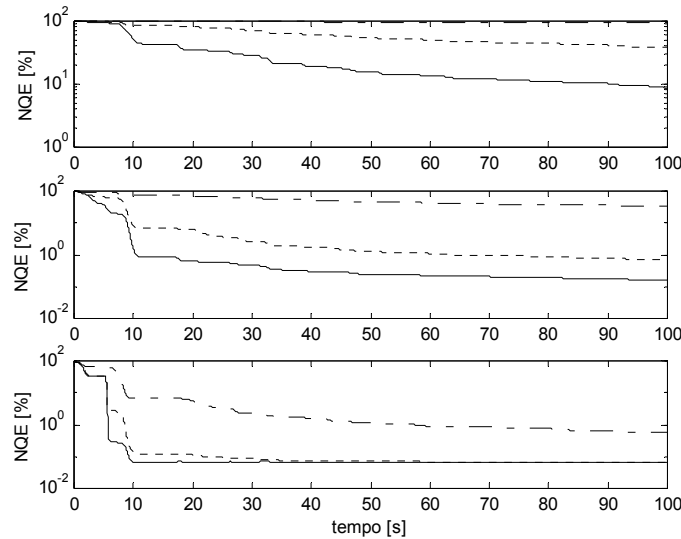
Na figura 3.4 vemos a análise desses dados agrupados em função da constante  $\lambda$  da cadeia de filtros de Poisson e na figura 3.5 vemos os mesmos resultados só que agrupados em função do valor dos elementos da diagonal da estimativa inicial da matriz de covariância.



**Figura 3.4:** Norma do erro quadrática em função do tempo. Primeiro gráfico  $\lambda = 1$ , segundo gráfico  $\lambda = 2$ , terceiro gráfico  $\lambda = 5$ . Linha contínua -  $P_0=1.10^2 I_6$ , linha tracejada -  $P_0=1.10^4 I_6$ , Linha ponto-traço -  $P_0=1.10^6 I_6$ , onde  $I_6$  representa a matriz identidade de tamanho 6.

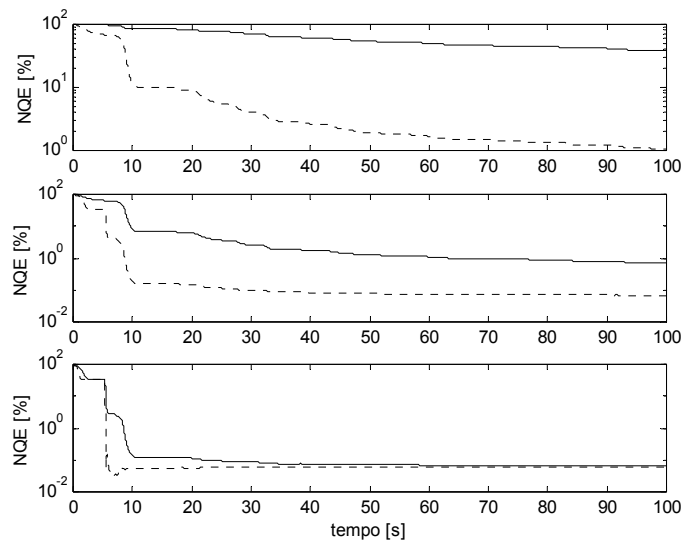


Como pode ser observado a partir dessas duas figuras quanto maior for o valor dos elementos da diagonal da estimativa inicial da matriz de covariância,  $P_0$ , para um mesmo valor de  $\lambda$ , mais rapidamente o método iterativo converge para o valor ótimo da estimativa do vetor de parâmetros  $\hat{\theta}$ . Também pode-se observar que no método iterativo o valor da constante da cadeia de filtros de Poisson também influencia na taxa de convergência do método. Nesse caso específico quando  $\lambda = 1$  a convergência é maior e quando  $\lambda = 5$  a convergência é mais lenta.

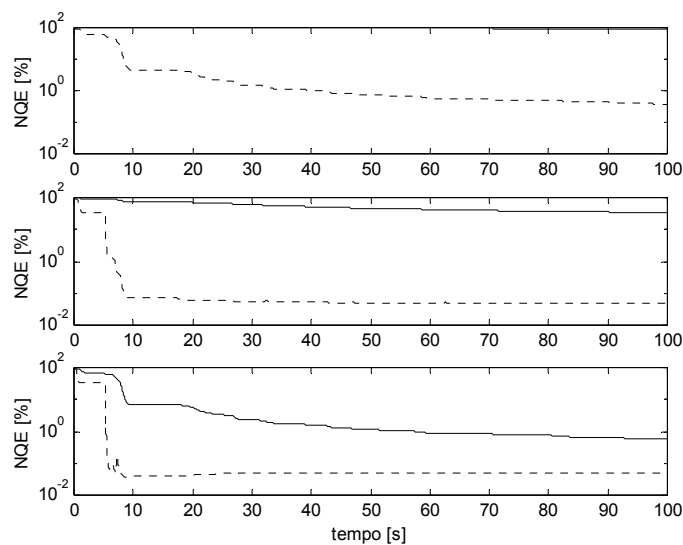


**Figura 3.5:** Norma do erro quadrática em função do tempo. Primeiro gráfico  $P_0=1.10^2 I_6$ , segundo gráfico  $P_0=1.10^4 I_6$ , terceiro gráfico -  $P_0=1.10^6 I_6$ , onde  $I_6$  representa a matriz identidade de tamanho 6. Linha contínua  $\lambda = 1$ , linha tracejada -  $\lambda = 2$ , Linha ponto-traço -  $\lambda = 5$ .

Como dito o fato de utilizarmos o método de Poisson ordinário ou normalizado também influencia o comportamento da convergência do método recursivo. Nas figuras 3.6 e 3.7 podemos ver a diferença na taxa de convergência quando da utilização do método ordinário e quando da utilização do método normalizado.



**Figura 3.6:** Norma do erro quadrática em função do tempo. Comparação entre os métodos ordinários e normalizado para  $\lambda = 2$ . Primeiro gráfico  $P_0=1.10^2 I_6$ , segundo gráfico  $P_0=1.10^4 I_6$ , terceiro gráfico -  $P_0=1.10^6 I_6$ , onde  $I_6$  representa a matriz identidade de tamanho 6. Linha contínua – método ordinário, linha tracejada – método normalizado.



**Figura 3.7:** Norma do erro quadrática em função do tempo. Comparação entre os métodos ordinários e normalizado para  $\lambda = 5$ . Primeiro gráfico  $P_0=1.10^2 I_6$ , segundo gráfico  $P_0=1.10^4 I_6$ , terceiro gráfico -  $P_0=1.10^6 I_6$ , onde  $I_6$  representa a matriz identidade de tamanho 6. Linha contínua – método ordinário, linha tracejada – método normalizado.

Podemos verificar, portanto, a partir desses resultados que o método normalizado apresenta uma taxa de convergência mais rápida do que o método normalizado para o mesmo valor de  $\lambda$ .

O método recursivo aqui apresentado foi implementado no MATLAB 5.3 a fim de complementar o *Continuous-time System Identification Toolbox* (CONTSID) versão 5.0 [Garnier e Mensler, 1999] que foi utilizado nesse trabalho.

### 3.7 Influência da constante da cadeia de filtros de Poisson na precisão da estimativa do modelo.

Como visto na análise do método dos mínimos quadrados recursivo o valor da constante  $\lambda$  dos estágios da cadeia de filtros de Poisson influencia na taxa de convergência. Além de influencia na taxa de convergência, a constante  $\lambda$  influencia também na acurácia da estimativa dos parâmetros do modelo. O valor ótimo da constante  $\lambda$  depende das características dinâmicas do processo e da quantidade de ruído presente nos dados. Para analisar essa dependência foram utilizados os quatro sistemas dinâmicos descritos na tabela 3.5.

Utilizou-se para cada sistema três diferentes níveis de ruído no sinal de saída, 0, 10% e 20%, definindo a potência do ruído através da razão entre o desvio padrão do ruído adicionado ao sistema sobre o desvio padrão do sinal de saída, conforme:

$$rrs[\%] = 100 \frac{\text{desvio\_padrão}(w)}{\text{desvio\_padrão}(y)} \quad (3.72)$$

Na análise feita foi utilizado apenas o método de Poisson normalizado ajustado pelo método dos mínimos quadrados não recursivo. O desempenho do método foi medido através da norma quadrática do erro, conforme definido em (3.71), analisado o valor da constante  $\lambda$  entre 0,1 e 10. Para identificar o sistema foi utilizado o mesmo sinal tipo PRBS como sinal de entrada. O tempo de amostragem utilizado para todos os sistemas foi 0,02.

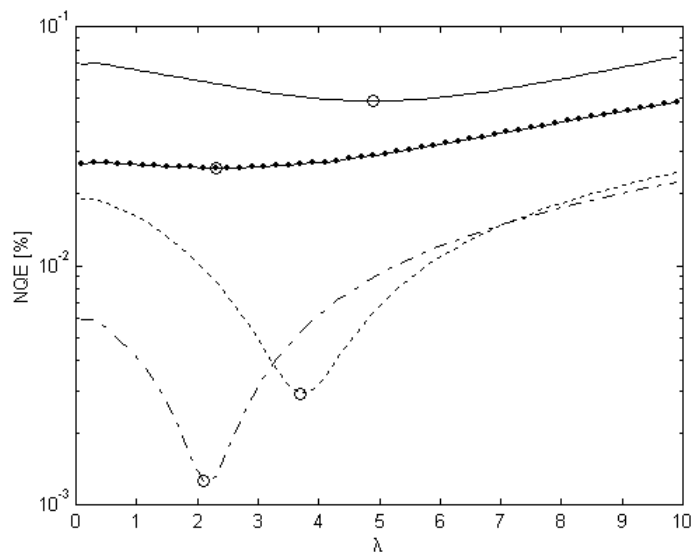
**Tabela 3.5:** Sistemas lineares utilizados na análise da influência da constante  $\lambda$  sobre a precisão do método de Poisson.

Sistema	Pólos	Zeros	Pólo dominante	Função de Transferência
1	-2 e -3	-1	-3	$G(s) = \frac{2s + 2}{s^2 + 5s + 6}$
2	-2 e -3	-	-3	$G(s) = \frac{2}{s^2 + 5s + 6}$
3	-1 e -2	-	-2	$G(s) = \frac{2}{s^2 + 3s + 2}$
4	-1 e -2	-0,5	-2	$G(s) = \frac{2s + 1}{s^2 + 3s + 2}$

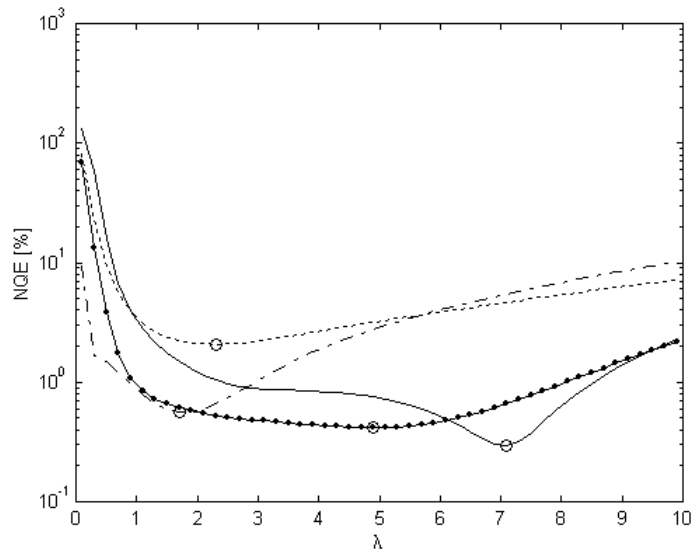
Nas figuras 3.8, 3.9 e 3.10 é apresentado o comportamento da norma quadrática do erro em função da constante  $\lambda$  para os quatro sistemas com o mesmo nível de ruído no sinal de saída em cada gráfico.

Pela análise das figuras abaixo podemos observar que o valor ótimo da constante  $\lambda$  varia para cada sistema. Também podemos observar que a presença de ruído no sinal de saída altera o comportamento relativo da norma do erro quadrático dos sistemas analisados.

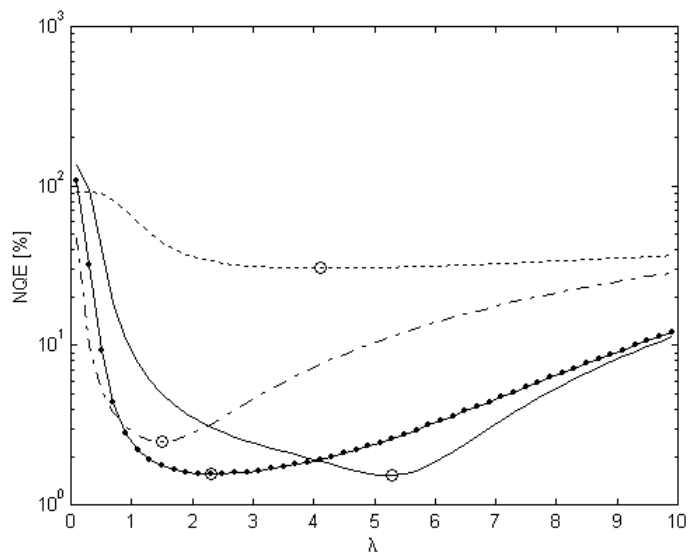
Nas figuras 3.11, 3.12, 3.13 e 3.14 é apresentado o comportamento da norma do erro quadrático para os quatro sistemas analisados em função da constante  $\lambda$  comparando o efeito da adição de ruído no sinal de saída. Cada gráfico corresponde a um sistema diferente considerando os três casos de ruídos analisados.



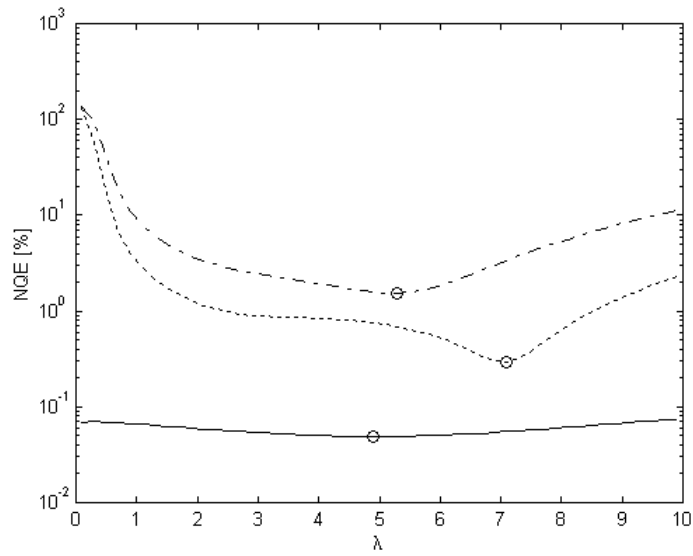
**Figura 3.8:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  na ausência de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha traço-ponto – sistema 3 e linha contínua com pontos – sistema 4.



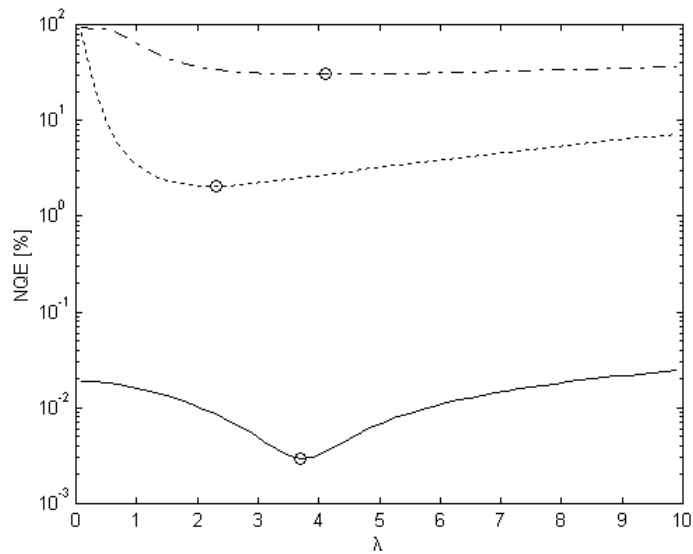
**Figura 3.9:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  na presença de 10% de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha ponto-traço – sistema 3 e linha contínua com pontos – sistema 4.



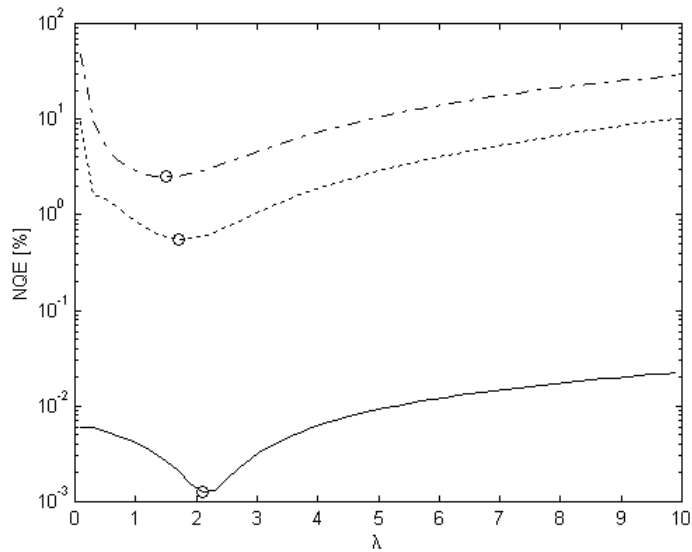
**Figura 3.10:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  na presença de 20% de ruído no sinal de saída. Linha contínua – sistema 1, linha tracejada – sistema 2, linha traço-ponto – sistema 3 e linha contínua com pontos – sistema 4.



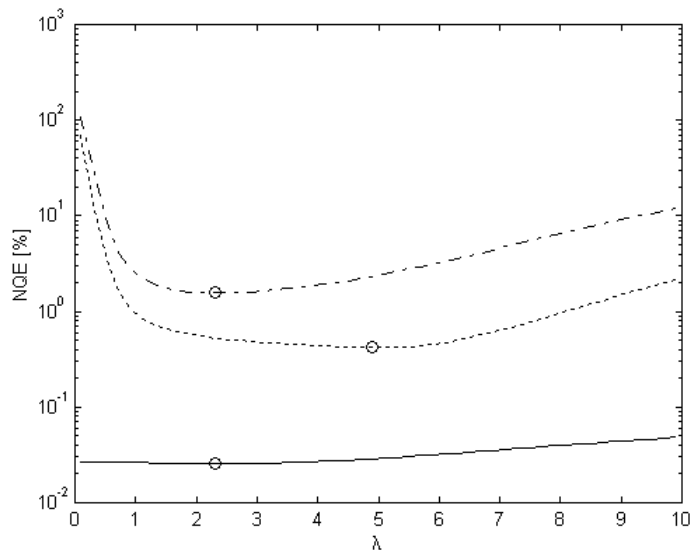
**Figura 3.11:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  para o sistema 1. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído.



**Figura 3.12:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  para o sistema 2. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha ponto-traço – 20% de ruído.



**Figura 3.13:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  para o sistema 3. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído.



**Figura 3.14:** Comportamento da norma quadrática do erro em função da constante  $\lambda$  para o sistema 4. Linha contínua – sem ruído, linha tracejada – 10% de ruído e linha traço-ponto – 20% de ruído.

Podemos observar pelos resultados apresentados nas figuras acima que a presença de ruído no sinal de saída não só afeta a precisão da identificação do sistema como também altera o valor ótimo da constante  $\lambda$  no qual o valor da norma quadrática o erro é mínima.

Pode-se concluir que a escolha do valor da constante  $\lambda$  é de grande importância para o sucesso da identificação de um sistema linear pelo método de Poisson sendo que seu valor ótimo depende tanto da dinâmica do sistema como da quantidade de ruído presente nos dados

utilizados na identificação. Entretanto, apesar dessa importante influência, não se estabeleceu ainda um relação que possa determinar qual o valor ótimo de  $\lambda$  para cada caso que não seja o de tentativa e erro. A única coisa que se pode concluir é que valores muito pequenos, menores do que 1, bem como valores muito grandes, maiores do que 10, tendem a produzir estimativas pouco precisas.

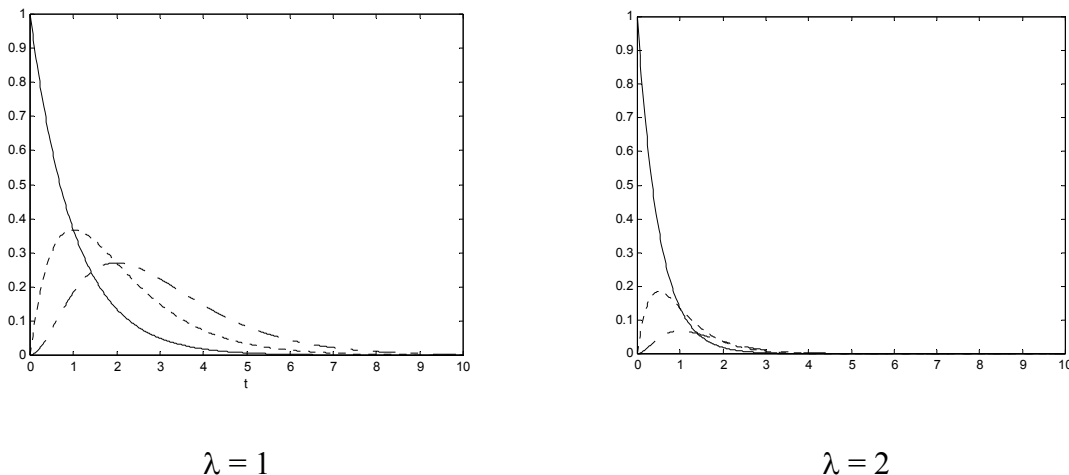
Uma outra coisa interessante de se observar é que há sistemas em que, quando livres de ruído, a constante  $\lambda$  praticamente não afeta o desempenho do método do momentos de Poisson. Dos exemplos analisados os que se enquadram nessa categoria são os sistema 1 e 4.

No capítulo 4 abordaremos novamente esse problema de um outro ponto de vista.

### 3.8 Considerações sobre a identificação das condições iniciais.

Como visto, o método de Poisson permite que as condições iniciais dos sinais de entrada e saída sejam estimadas juntamente com os parâmetros do modelo. Isso é útil no sentido que muitas vezes é difícil deixar um sistema, principalmente sistemas químicos, em estado estacionário, no qual podemos considerar que as condições iniciais em termos de variáveis de desvio, principalmente dos sinais de saída sejam nulas. Se a condição de estado estacionário do sinais não é observada é de fundamental importância considerar as condições iniciais na identificação do modelo.

Entretanto como já visto na equação (3.38) os vetores na matriz  $\mathbf{M}_k$  correspondentes as condições iniciais são formados pelas funções pulso de Poisson,  $p_k$ . O gráfico das funções pulso de Poisson normalizadas para  $\lambda = 1$  e  $\lambda = 2$  podem ser vistas nas figuras 3.15.



**Figura 3.15:** Funções pulso de Poisson para  $\lambda = 1$  e  $\lambda = 2$ . Linha contínua  $k = 1$ , linha tracejada  $k = 2$ , linha ponto-traço  $k = 3$ .

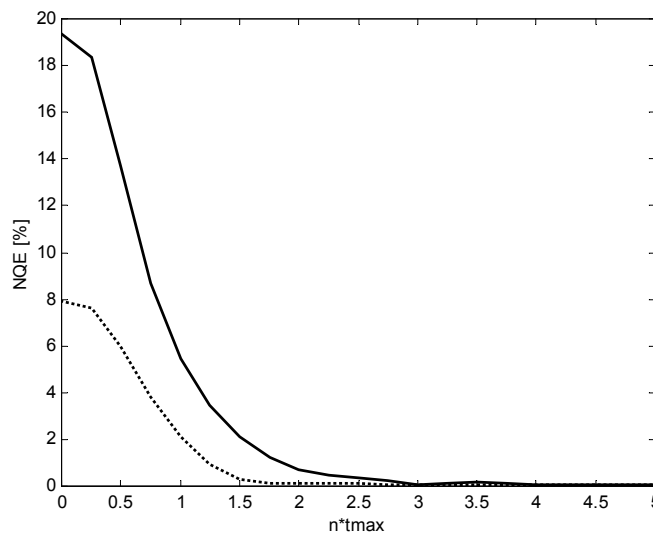


Como pode ser observado a partir das figuras acima as funções pulso decaem rapidamente até zero e quanto maior o valor da constante  $\lambda$  maior é a taxa de decaimento. Dessa forma é possível desconsiderar os dados iniciais, pois a partir de um certo instante os vetores correspondentes às condições iniciais são praticamente nulos, de forma a não necessitarmos estimar as condições iniciais dos dados de entrada e saída. Isso é uma vantagem uma vez que diminui a quantidade de parâmetros a serem estimados

Falta determinar qual a quantidade de dados a serem excluídos. Observando a figura 3.15 parece que um parâmetro adequado para ajustar essa quantidade é o instante de tempo no qual a função pulso de maior índice, a de decaimento mais lento, atinge o seu máximo. Esse valor é facilmente obtido diferenciando-se a equação (3.3) e igualando-se essa derivada a zero. O valor assim obtido é:

$$t_{\max} = \frac{k}{\lambda} \quad (3.73)$$

Para avaliar o efeito da eliminação dos dados iniciais utilizou-se como teste o sistema descrito pela equação (3.49). Foi utilizado como sinal de entrada para o sistema o sinal PRBS mostrado na figura 3.3 com tempo de amostragem igual a 0,02. As condições iniciais dos dados de saída são  $y(0) = 0,4$  e  $y'(0) = -1,6$ . Para avaliar o desempenho do método utilizou-se a norma do erro quadrático como descrito na equação (3.71). Foi testado apenas o método de Poisson normalizado não recursivo. O resultados para  $\lambda = 1$  e  $\lambda = 2$  podem ser vistos na figura 3.16 abaixo.



**Figura 3.16:** Comportamento da norma quadrática do erro em função do número de tempo máximos da função de Poisson de maior índice no caso a eliminação da influência das condições iniciais dos sinais de entrada e saída. Linha contínua  $\lambda = 1$  e linha pontilhada  $\lambda = 2$ .

Observando a figura acima podemos notar que o erro a partir de três tempos máximos é praticamente desprezível. Assim, para eliminar o efeito das condições iniciais pode-se utilizar os sinais de entrada e saída no intervalo  $[t_c, t_f]$ , onde  $t_f$  é o tempo final e  $t_c$  é dado por:

$$t_c = 3t_{\max} = \frac{3k}{\lambda} = \frac{3n_A}{\lambda} \quad (3.74)$$

onde  $n_A$  é a ordem do polinômio denominador da função de transferência. Observa-se também que quanto maior for o valor de  $\lambda$  mais rápida é a redução de erro.

É importante observar que a eliminação dos dados iniciais deve ser feita após os momentos de Poisson dos sinais de entrada e saída terem sido calculados. Os momentos de Poisson devem ser calculados com todos os dados.

Até agora sempre foram utilizados dados em formato de variação ou delta, isto é, os dados de entrada e saída do sistema devem ser expressos como variações em relação a um estado estacionário qualquer. Isso significa que os sinais de entrada e saída não contêm *off-set*. O método de Poisson como expresso até o momento não é capaz de lidar muito bem com a presença de *off-set* nos sinais. Para exemplificar isso vamos analisar o sistema dado pela equação (3.49). Aos dados gerados pela simulação desse sistema foram adicionados diferentes níveis de *off-set*. Para identificar esse sistema utilizou-se o método de Poisson e método ARX211. O modelo discreto foi convertido para contínuo utilizando a função `d2c` do MATLAB 5.3. Os resultados desse teste são apresentados na tabela 3.6 abaixo.

**Tabela 3.6:** Resultados da identificação com diferentes níveis de *off-set* nos sinais de entrada e saída.

	sem <i>off-set</i>		<i>off-set</i> = 0,01		<i>off-set</i> = 0,1	
	Poisson	ARX221	Poisson	ARX221	Poisson	ARX221
$a_1$	3,9962	4,0000	5,0539	5,0460	5,0731	4,3950
$a_0$	4,9999	5,0000	6,2279	6,2110	0,2284	1,1350
$b_1$	1,9983	2,0000	2,0114	2,0000	2,2180	2,0050
$b_0$	2,0001	2,0000	2,1289	2,1190	0,0020	0,3587

Como podemos ver pelos resultados apresentados na tabela acima pequenas quantidades de *off-set* presentes nos sinais de entrada prejudicam a identificação mesmo quando efetuada por métodos discretos e principalmente pelo método de Poisson. No caso dos métodos discretos é possível contornar essa dificuldade aumentando a ordem do modelo. Para o método de Poisson esse artifício não se mostrou eficaz.

Entretanto é possível com uma pequena modificação permitir ao método de Poisson levar em consideração a existência de *bias* nos sinais de entrada e saída do sistema. Para isso vamos considerar que os sinais de entrada e saída com *bias* podem ser descritos da seguinte forma.

$$\begin{aligned} y(t) &= \hat{y}(t) + y_b \\ u(t) &= \hat{u}(t) + u_b \end{aligned} \quad (3.75)$$

onde  $\hat{u}(t)$  e  $\hat{y}(t)$  representam os sinais de entrada e saída livres de *bias* respectivamente e  $y_b$  e  $u_b$  o *bias* constante presente nos sinais de entrada e saída respectivamente.

Para analisarmos como o método de Poisson pode levar em conta a presença de *bias* nos sinais de entrada e saída vamos considerar o seguinte sistema linear de segunda ordem com a presença de um zero.

$$\frac{d^2 y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_1 \frac{du(t)}{dt} + b_0 u(t) \quad (3.76)$$

Aplicando-se a equação (3.50) (considerando  $\lambda = \beta = 1$  e as condições iniciais dos sinais de entrada e saída nulas) ao lado esquerdo da equação (3.76) podemos calcular os momentos de Poisson do sinal de saída e de suas derivadas conforme:

$$\begin{aligned} M_2 \left\{ \frac{d^2 y}{dt^2} \right\} &= y_0^{t_f} - 2y_1^{t_f} + y_2^{t_f} \\ M_2 \left\{ \frac{dy}{dt} \right\} &= y_1^{t_f} - y_2^{t_f} \\ M_2 \{y\} &= y_2^{t_f} \end{aligned} \quad (3.77)$$

e os momentos de Poisson do sinal de entrada e sua derivada conforme:

$$\begin{aligned} M_2 \left\{ \frac{du}{dt} \right\} &= u_1^{t_f} - u_2^{t_f} \\ M_2 \{u\} &= u_2^{t_f} \end{aligned} \quad (3.78)$$

Como os dados de entrada e saída contêm *bias*, se aplicarmos a transformada de Poisson às equações definidas em (3.75) obteremos:

$$\begin{aligned} y_k^{t_f} &= \hat{y}_k^{t_f} + y_b M_k \{1\} \\ u_k^{t_f} &= \hat{u}_k^{t_f} + u_b M_k \{1\} \end{aligned} \quad (3.79)$$

Substituindo a equação (3.79) nas equações (3.77) e (3.78) obtemos:

$$\begin{aligned} M_2 \left\{ \frac{d^2 y}{dt^2} \right\} &= \hat{y}_0^{t_f} - 2\hat{y}_1^{t_f} + \hat{y}_2^{t_f} + y_b [M_0 \{1\} - 2M_1 \{1\} + M_2 \{1\}] \\ M_2 \left\{ \frac{dy}{dt} \right\} &= \hat{y}_1^{t_f} - \hat{y}_2^{t_f} + y_b [M_1 \{1\} - M_2 \{1\}] \\ M_2 \{y\} &= \hat{y}_2^{t_f} + y_b M_2 \{1\} \end{aligned} \quad (3.80)$$

e

$$M_2 \left\{ \frac{du}{dt} \right\} = \hat{u}_1^{t_f} - \hat{u}_2^{t_f} + u_b [M_1 \{1\} - M_2 \{1\}] \quad (3.81)$$

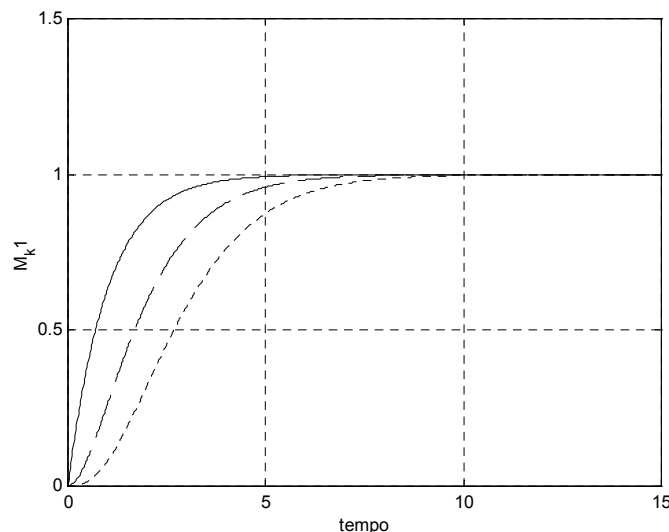
$$M_2 \{u\} = \hat{u}_2^{t_f} + u_b M_2 \{1\}$$

Analisando as equações acima vemos que os momentos de Poisson dos sinais de entrada e saída contendo *bias* são iguais aos momentos de Poisson do mesmo sinais livres de *bias* mais um termo que depende da magnitude do *bias* e dos momentos de Poisson da função constante igual a 1, que pode ser entendida como um degrau unitário no instante zero. Esse termo, da mesma forma que as funções pulso de Poisson, só são relevantes nos instantes iniciais, pois, como é de fácil compreensão, o sinal de saída da cadeia de filtros de Poisson quando o sinal de entrada é um degrau unitário no instante zero convergem para 1 quando o ganho de todos os elementos da cadeia de filtros é igual à unidade. Isso pode ser visto na figura 3.17 abaixo que mostra a saída dos três primeiros estágios da cadeia de filtros de Poisson com ganho unitário em todos os elementos quando o degrau unitário é aplicado na entrada da cadeia.

Conclui-se que os termos devido ao *bias* praticamente anulam-se a partir de um determinado momento de forma que os momentos de Poisson das derivadas do sinais de entrada e saída são iguais ao momentos de Poisson dos sinais correspondentes livres de *bias*. Isso é esperado uma vez que derivando as equações definidas em (3.75) obtemos.

$$\frac{d^{(n)}y}{dt^{(n)}} = \frac{d^{(n)}\hat{y}}{dt^{(n)}} \quad (3.82)$$

$$\frac{d^{(n)}u}{dt^{(n)}} = \frac{d^{(n)}\hat{u}}{dt^{(n)}}$$



**Figura 3.17:** Momentos de Poisson de um degrau unitário. Linha contínua –  $M_0\{1\}$ , linha tracejada  $M_1\{1\}$  e linha pontilhada  $M_2\{1\}$ .

Entretanto os momentos de Poisson do degrau unitário continuam a afetar os momentos de Poisson dos sinais de entrada e saída. Podemos então dizer que os momentos de Poisson dos sinais de entrada e saída livres de *bias* podem ser expressos da seguinte forma:

$$\begin{aligned} M_k \{ \hat{y}(t) \} &= M_k \{ y(t) \} - y_b M_k \{ 1 \} \\ M_k \{ \hat{u}(t) \} &= M_k \{ u(t) \} - u_b M_k \{ 1 \} \end{aligned} \quad (3.83)$$

Aplicando-se o método de Poisson normalizado à equação (3.76) (considerando  $\lambda = \beta = 1$  e as condições iniciais dos sinais de entrada e saída nulas) e considerando que o método de Poisson só funciona perfeitamente quando os sinais de entrada e saída são livres de *bias*, temos:

$$M_2 \left\{ \frac{d^2 \hat{y}(t)}{dt^2} \right\} + a_1 M_2 \left\{ \frac{d\hat{y}(t)}{dt} \right\} + a_0 M_2 \{ \hat{y}(t) \} = b_1 M_2 \left\{ \frac{d\hat{u}(t)}{dt} \right\} + b_0 M_2 \{ \hat{u}(t) \} \quad (3.84)$$

Como, desconsiderando os instantes iniciais, podemos dizer que os momentos de Poisson das derivada dos sinais livres de *bias* são iguais aos momentos de Poisson do sinais com *bias* e inserindo as equações definidas em (3.83) na equação (3.84), temos:

$$\begin{aligned} M_2 \left\{ \frac{d^2 y(t)}{dt^2} \right\} + a_1 M_2 \left\{ \frac{dy(t)}{dt} \right\} + a_0 M_2 \{ y(t) \} - a_0 y_b M_2 \{ 1 \} \\ = b_1 M_2 \left\{ \frac{du(t)}{dt} \right\} + b_0 M_2 \{ u(t) \} - b_0 u_b M_2 \{ 1 \} \end{aligned} \quad (3.85)$$

Essa equação pode se reescrita na forma de um regressor linear conforme:

$$a_1 M_2 \left\{ \frac{dy(t)}{dt} \right\} + a_2 M_2 \{ y(t) \} - b_1 M_2 \left\{ \frac{du(t)}{dt} \right\} - b_0 M_2 \{ u(t) \} + d M_2 \{ 1 \} = -M_2 \left\{ \frac{d^2 y(t)}{dt^2} \right\} \quad (3.86)$$

onde o parâmetro  $d$  é definido por:

$$d = b_0 u_b - a_0 y_b \quad (3.87)$$

A equação (3.86) pode ser reescrita na forma matricial da seguinte forma:

$$\mathbf{M}_2 \boldsymbol{\theta}^T = -M_2 \left\{ \frac{d^2 y(t)}{dt^2} \right\} \quad (3.88)$$

onde:

$$\mathbf{M}_2 = \left[ M_2 \left\{ \frac{dy(t)}{dt} \right\} \quad M_2 \{y(t)\} \quad -M_2 \left\{ \frac{du(t)}{dt} \right\} \quad -M_2 \{u(t)\} \quad M_2 \{1\} \right] \quad (3.89)$$

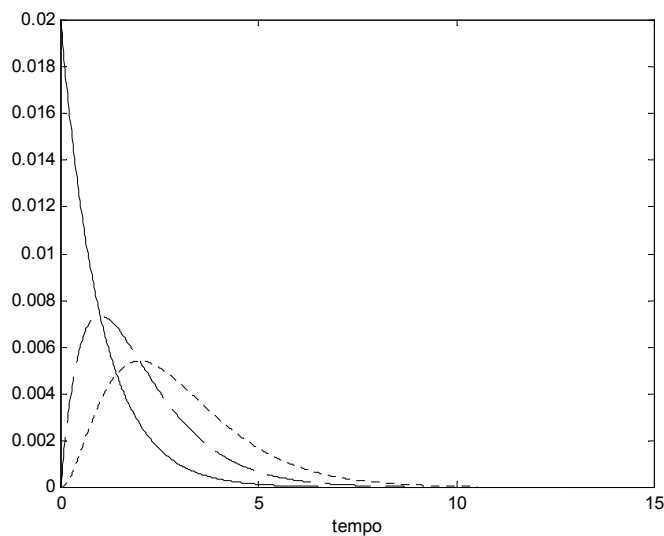
$$\theta = [a_1 \quad a_0 \quad b_1 \quad b_0 \quad d]$$

Assim podemos reescrever os elementos da equação (3.37) de maneira a levar em consideração a existência de *bias* nos sinais de entrada e saída, conforme:

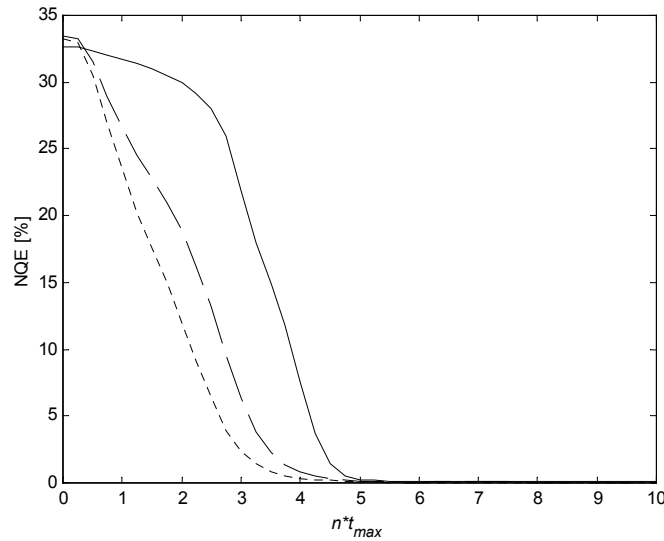
$$\begin{aligned} \mathbf{M}_k &= [\gamma_n(2:n,:)Y_n^{t_0} \mid -\gamma_m U_m^{t_0} \mid -\gamma_{(n-1)} P_n^{t_0} \mid M_n \{1\}] \\ \mathbf{y}_k^{t_0} &= -\gamma(1,:)Y_k^{t_0} \\ \theta &= [\underline{a} \mid \underline{b} \mid \underline{c} \mid d] \end{aligned} \quad (3.90)$$

Agora nos resta determinar qual a quantidade de dados que devemos desconsiderar para a construção da matriz de regressores. Observando que a derivada das repostas ao degrau da cadeia de filtro de Poisson são iguais às funções pulso de Poisson, conforme pode ser visto na figura 3.18, abaixo, podemos considerar uma parametrização em relação ao instante correspondente ao máximo da função pulso de maior índice conforme definido em (3.73).

Para determinar essa quantidade vamos tomar como exemplo novamente o sistema definido pela equação (3.49). Foi utilizado como sinal de entrada para o sistema o sinal PRBS mostrado na figura 3.3 com tempo de amostragem igual a 0,02. Ao sinal de entrada foi adicionado um *bias* igual a 2 e ao sinal de saída um *bias* igual a 3. A figura 3.19 mostra a evolução da norma quadrática do erro, definida em (3.71) em função de  $n$  vezes o tempo correspondente ao máximo da função pulso de Poisson de maior índice.



**Figura 3.18:** Derivadas das repostas da cadeia de filtros de Poisson ao degrau unitário.



**Figura 3.19:** Comportamento da norma quadrática do erro em função do número de tempo máximos da função de Poisson de maior índice no caso da eliminação de influência de *bias* nos sinais de entrada e saída. Linha contínua  $\lambda = 1$ , linha tracejada  $\lambda = 2$  e linha pontilhada  $\lambda = 3$ .

Observando a figura 3.18 podemos notar que a partir de 5 vezes o instante correspondente ao máximo a norma quadrática do erro é praticamente desprezível. Assim podemos dizer que é possível identificar um modelo linear invariante no tempo com sinais contendo *bias* usando o método de Poisson, desde que utilizemos os dados, após aplicação da cadeia de filtros de Poisson no intervalo  $[t_c, t_f]$ , onde  $t_f$  é o tempo final e  $t_c$  é dado por:

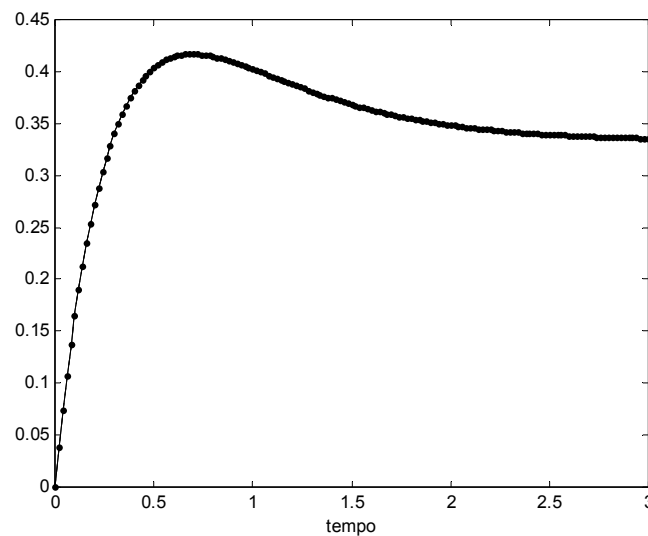
$$t_c = 5t_{\max} = \frac{5k}{\lambda} = \frac{5n_A}{\lambda} \quad (3.91)$$

A possibilidade de podermos levar em consideração a existência de *bias* nos sinais de entrada e saída é muito importante para aumentar a robustez do método de sintonia, pois, principalmente na presença de dados muito ruidosos é difícil determinar qual é o valor do *bias* que devemos descontar para expressar os sinais em forma de variações ou deltas.

### 3.9 Comparação entre a identificação contínua e discreta

É interessante comparar o desempenho do método de Poisson com os métodos de identificação discretos na identificação de modelos lineares quando há a presença de ruído branco nos sinais de entrada e saída do sistema. O método de Poisson como apresentado é equivalente à identificação de uma estrutura tipo ARX, como implementado no *system identification toolbox* do MATLAB [MATHWORKS, 1998], pois o ajuste dos parâmetros de modelos dessa estrutura é feito através de mínimos quadrados sem necessidade de otimização iterativa.

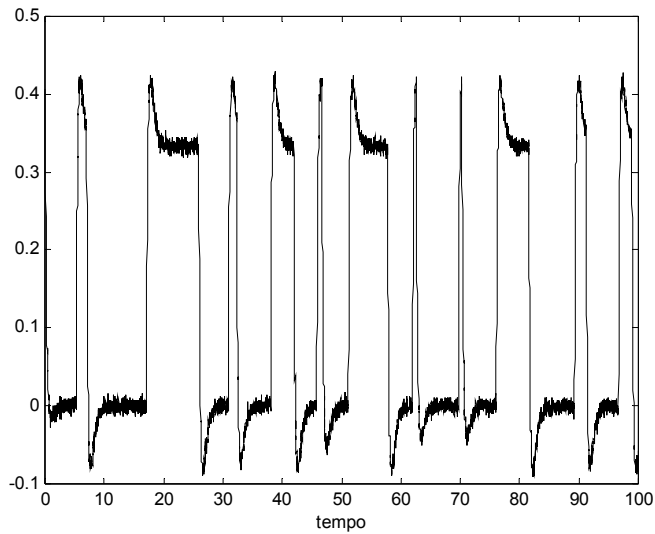
Para comparar o desempenho dos dois métodos foi utilizado o modelo descrito pela equação (3.49) perturbado por um sinal tipo PRBS conforme a figura 3.3. O tempo de amostragem utilizado foi de 0,02. Primeiramente foi feita a identificação utilizando dados sem distúrbios utilizando um modelo ARX com ordem 2 no denominador e numerador e tempo morto igual a 1. Essa é a ordem do modelo discreto gerado pela discretização do sistema em questão. Para o método de Poisson utilizou-se a mesma ordem do sistema, isto é, denominador de segunda ordem e numerador de primeira ordem com tempo morto igual a zero. O resultado dessa identificação pode ser visto na figura 3.20 onde é apresentado a resposta a um degrau unitário do sistema “real” e dos modelos identificados por ambos os métodos para o caso de um sinal sem ruído.



**Figura 3.20:** Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson ( $\lambda = 5$ ) - linha pontilhada e do modelo ARX221 - linha contínua com pontos.

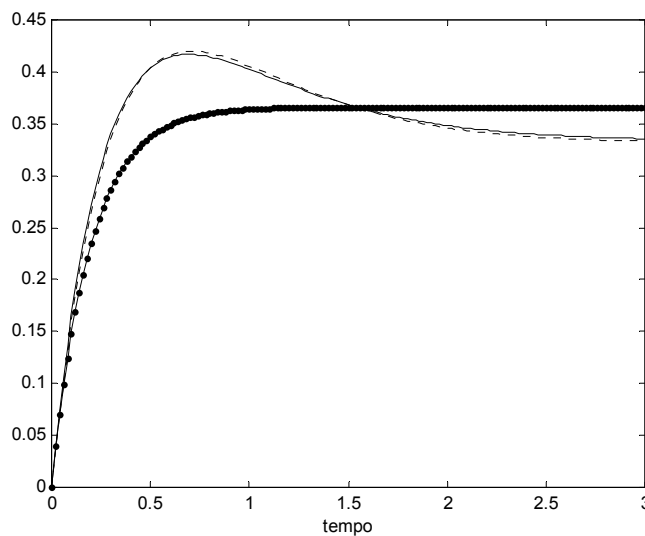
Pode-se observar que ambos os métodos tem desempenho idêntico e geram modelos praticamente perfeitos do sistema. Entretanto nos casos reais sempre existe a presença de ruído nos dados. Para simular a presença de ruído de medição, algo muito comum em casos reais, foi adicionado um ruído ao sinal de saída do sistema. A relação ruído sobre sinal, definida em (3.72), utilizada foi de aproximadamente 3%. O sinal de saída utilizado pode ser visto na figura 3.21, abaixo:





**Figura 3.21:** Sinal de saída com 3% de ruído branco utilizado na identificação do sistema definido na equação (3.49).

O resultado da identificação obtida mantendo a mesma ordem do modelo ARX utilizada na identificação com dados livres de ruído pode ser visto na figura 3.22, abaixo.

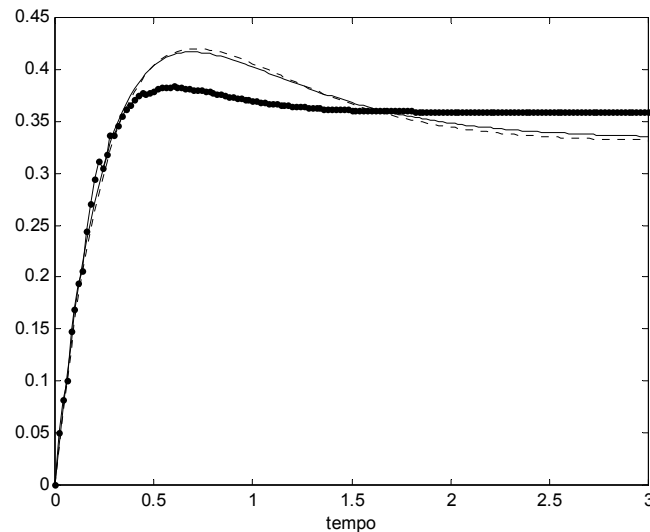


**Figura 3.22:** Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson ( $\lambda = 5$ ) - linha pontilhada e do modelo ARX221 - linha contínua com pontos. Modelos identificados com 3% de ruído no sinal de saída.

Observa-se que, nesse caso, a qualidade do modelo ARX degenerou ao ponto do modelo não conseguir mais reproduzir o *overshot*, ao passo que o modelo gerado pelo método de Poisson continua gerando um modelo muito bom, praticamente idêntico ao gerado com dados sem ruído. Para que a estrutura ARX consiga identificar esse sistema utilizando esse dados contaminados é necessário aumentar a ordem do denominador e do numerador desse. Para produzir um modelo semelhante ao gerado com dados sem ruído foi necessário utilizar

um modelo ARX com ordem 15 no numerador e no denominador, mantendo o tempo morto igual a um. A resposta ao degrau desse modelo é idêntica ao apresentado na figura 3.20.

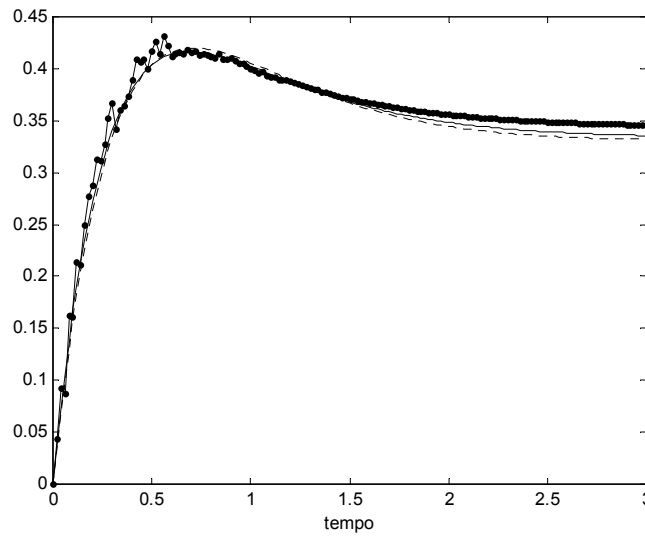
Entretanto, à medida que aumentamos a quantidade de ruído presente no sistema o modelo ARX 15 15 1 perde a capacidade de identificar corretamente a dinâmica do sistema mesmo aumentando a sua ordem. Por exemplo, se utilizarmos um sinal de saída com 15% de ruído e mantivermos a ordem do modelo ARX teremos o resultado apresentado na figura 3.23.



**Figura 3.23:** Resposta ao degrau unitário do sistema identificado - linha contínua, do modelo gerado pelo método de Poisson - linha pontilhada e do modelo ARX15 15 1 - linha contínua com pontos. Modelos identificados com 15% de ruído no sinal de saída.

Como podemos ver pela figura acima, o modelo ARX não consegue ter o mesmo desempenho que o método de Poisson. Se tentarmos aumentar ainda mais a ordem do modelo ARX para tentarmos compensar o ruído, por exemplo para 30 no numerador e no denominador obteremos o resultado apresentado na figura 3.24, a seguir.

Como podemos notar por essa figura, mesmo melhorando a qualidade do modelo esse apresenta um comportamento estranho na parte transiente. Ao mesmo tempo observamos que o método de Poisson continua apresentando um bom resultado. Com isso podemos concluir que modelos ARX não são apropriados para a identificação de sistemas com dados ruidosos. Para a identificação desse sistema utilizando modelos discretos seria necessário identificar um modelo de distúrbio, utilizando uma estrutura OE, ARMAX ou BJ.



**Figura 3.24:** Resposta ao degrau unitário do sistema identificado, linha contínua, do modelo gerado pelo método de Poisson, linha pontilhada e do modelo ARX30 30 1, linha contínua com pontos, identificados com 15% de ruído no sinal de saída.

[Zhu, 1998] apresenta um método de identificação de sistemas lineares baseado na utilização de modelos do tipo ARX de ordem elevada a fim de eliminar o efeito do ruído. Posteriormente, para evitar os comportamentos estranhos observados na fase transiente da resposta dos modelos (conforme visto nas figuras 3.23 e 3.24), esses têm a sua ordem reduzida através da utilização de um critério assintótico no domínio da frequência. A identificação de modelos contínuos gerados pelo método de Poisson gera modelos de melhor qualidade uma vez que não necessitamos utilizar modelos de ordem muito elevada e executar uma posterior redução de ordem como visto no exemplo acima.

### 3.10 Análise de uma forma alternativa de interpretar o método do momentos de Poisson.

Em 1994 foi apresentada uma visão alternativa do método de identificação de modelos lineares no espaço contínuo de tempo [Johansson, 1994]. A sua proposta é utilizar o seguinte operador linear.

$$\xi = f(p) = \frac{\lambda}{p + \lambda} = \frac{1}{\tau p + 1}, \quad \tau = \frac{1}{\lambda} \quad (3.92)$$

onde:

$$p = \frac{d}{dt} \quad (3.93)$$

é o operador de diferenciação. Resolvendo a equação (3.92) para  $p$  temos:

$$p = \frac{1-\xi}{\tau\xi} \quad (3.94)$$

Como a definição do operador  $p$  é possível reescrever a equação (3.26), que define um sistema de SISO linear e invariante no tempo da seguinte forma.

$$y^{(n)}(t) + \sum_{i=1}^{n-1} a_i p^i y(t) = \sum_{j=0}^m b_j p^j u(t) \quad (3.95)$$

Essa equação pode ser reescrita na forma de uma razão de polinômios de  $p$ , da seguinte forma:

$$\frac{y(t)}{u(t)} = G(p) = \frac{B(p)}{A(p)} = \frac{b_m p^m + b_{(m-1)} p^{m-1} + \dots + b_0}{p^n + a_{n-1} p^{n-1} + a_{n-2} p^{n-2} + \dots + a_0} \quad (3.96)$$

Aplicando-se a equação (3.94), que relaciona o operador  $p$  com o operador  $\xi$ , na equação acima obtemos a seguinte equação:

$$G^*(\xi) = \frac{A^*(\xi)}{B^*(\xi)} = \frac{\beta_m \xi^m + \beta_{(m-1)} \xi^{(m-1)} + \dots + \beta_1 \xi}{\alpha_n \xi^n + \alpha_{(n-1)} \xi^{(n-1)} + \alpha_{(n-2)} \xi^{(n-2)} + \dots + 1} \quad (3.97)$$

que por sua vez pode ser reescrita da seguinte forma:

$$y(t) = -\alpha_n \{\xi^n y(t)\} - \dots - \alpha_1 \{\xi y(t)\} + \beta_m \{\xi^m u(t)\} + \dots + \beta_1 \{\xi^1 u(t)\} \quad (3.98)$$

sendo escrito vetorialmente como

$$y(t) = \theta^* \Phi^*(t) \quad (3.99)$$

onde:

$$\begin{aligned} \theta^* &= [\alpha_n \quad \dots \quad \alpha_1 \quad \beta_m \quad \dots \quad \beta_1] \\ \Phi^*(t) &= [-\{\xi^n y(t)\} \quad \dots \quad -\{\xi y(t)\} \quad \{\xi^m u(t)\} \quad \dots \quad \{\xi u(t)\}]^T \end{aligned} \quad (3.100)$$

Analisando a equação (3.99) observamos que esta é um regressor linear da mesma forma que a equação (3.37). Assim podemos determinar a estimativa do vetor de parâmetros

do modelo,  $\hat{\theta}^*$ , através do método dos mínimos quadrados conforme descrito pela equação (3.39).

Entretanto o objetivo é identificar os parâmetros do vetor  $\theta$  e não os do vetor  $\hat{\theta}^*$ . Para isso devemos estabelecer uma relação entre esse dois vetores que pode ser descrita por:

$$\theta^* = F\theta + G \quad (3.101)$$

onde:

$$F = \begin{bmatrix} J_{n \times n}^* & 0_{n \times m} \\ 0_{m \times n} & J_{(m+1) \times (m+1)}^* \end{bmatrix} \quad (3.102)$$

onde a matriz  $J^*$  é definida por:

$$J^* = \begin{bmatrix} m_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix}, \quad m_{ij} = (-1)^{(i-j)} \binom{n-j}{i-j} \tau^j \quad (3.103)$$

O vetor  $G$  da equação (3.101) é definido por:

$$G = [g_1 \quad \cdots \quad g_n \quad 0_{1 \times n}]^T, \quad g_i = (-1)^i \binom{n}{i} \quad (3.104)$$

Observando as equações desse método notamos uma clara semelhança com o método de Poisson normalizado, a começar pela definição do operador  $\xi$  que nada mais é do que o filtro linear de primeira ordem que compõe a cadeia de filtros de Poisson, como mostrado na figura 3.1.

Dessa forma os elementos da matriz de regressores,  $\Phi^*(t)$ , são os momentos de Poisson dos sinais  $y(t)$  e  $u(t)$ . No método de Poisson utilizamos como regressores os momentos de Poisson das derivadas de  $y(t)$  e  $u(t)$  que são geradas através de uma transformação linear conforme definida pela equação (3.18). No método proposto por Johansson a transformação linear que gera o vetor de parâmetros do sistema original,  $\theta$ , é muito semelhante à transformação que gera os momentos de Poisson das derivadas de  $y(t)$  e  $u(t)$  quando consideramos que todas as condições iniciais de ambos os sinais de entrada e saída são nulas. Isso pode ser constatado comparando a definição da matriz  $J^*$  em (3.103) com a da matriz  $J_n$  definida em (3.14).

Para demonstrar que ambos os métodos são equivalentes iremos aplicar o método de Poisson normalizado ao seguinte sistema genérico e transformando-o na representação proposta por Johansson.

$$G(p) = \frac{B(p)}{A(p)} = \frac{b_2 p^2 + b_1 p + b_0}{p^3 + a_2 p^2 + a_1 p + a_0} \quad (3.105)$$

onde  $p$  é o operador definido em (3.93).

Esse modelo pode ser reescrito na forma de entrada-saída conforme:

$$A(p)y(t) = B(p)u(t) \quad (3.106)$$

Aplicando-se a transformação de Poisson ao sistema definido em (3.106) temos o seguinte:

$$\begin{aligned} M_k \{y^{(3)}\} + a_2 M_k \{y^{(2)}\} + a_1 M_k \{y^{(1)}\} + a_0 M_k \{y^{(0)}\} = \\ = b_2 M_k \{u^{(2)}\} + b_1 M_k \{u^{(1)}\} + b_0 M_k \{u^{(0)}\} \end{aligned} \quad (3.107)$$

Expandindo os momentos de Poisson das derivadas de  $y(t)$  e  $u(t)$  de acordo com a equação (3.50) para o método de Poisson normalizado ( $\lambda = \beta = \gamma$ ) e com o índice da cadeia de filtros de Poisson,  $k$ , igual a 3 temos:

$$\begin{aligned} \gamma^3 (y_0 - 3y_1 + 3y_2 - y_3) + a_2 \gamma^2 (y_1 - 2y_2 + y_3) + a_1 \gamma (y_2 - y_3) + a_0 y_3 = \\ = b_2 \gamma^2 (u_1 - 2u_2 + u_3) + b_1 \gamma (u_2 - u_3) + b_0 u_3 \end{aligned} \quad (3.108)$$

onde:

$f^{(k)}$  = derivada de ordem  $k$  de  $f(t)$  e  $f_k$  é o momento de ordem  $k$  de  $f(t)$  correspondente a saída do  $(k+1)$ -ésimo elemento da cadeia de filtros de Poisson.

Reagrupando os elementos da equação (3.108) pelas saídas da cadeia de filtros de Poisson temos:

$$\begin{aligned} \gamma^3 y_0 + (-3\gamma^3 + a_2 \gamma^2) y_1 + (3\gamma^3 - 2a_2 \gamma^2 + a_1 \gamma) y_2 + (-\gamma^3 + a_2 \gamma^2 - a_1 \gamma + a_0) y_3 = \\ = b_2 \gamma^2 u_1 + (-2b_2 \gamma^2 + b_1 \gamma) u_2 + (b_2 \gamma^2 - b_1 \gamma + b_0) u_3 \end{aligned} \quad (3.109)$$

Dividindo a equação (3.109) por  $\gamma^3$  e lembrando que  $\tau = 1/\gamma$ , temos:

$$\begin{aligned} y_0 + (-3 + a_2 \tau) y_1 + (3 - 2a_2 \tau + a_1 \tau^2) y_2 + (-1 + a_2 \tau - a_1 \tau^2 + a_0 \tau^3) y_3 = \\ = b_2 \tau u_1 + (-2b_2 \tau + b_1 \tau^2) u_2 + (b_2 \tau - b_1 \tau^2 + b_0 \tau^3) u_3 \end{aligned} \quad (3.110)$$

Agrupando os termos entre parênteses poderemos reescrever a equação (3.110) da seguinte forma:

$$y_0 + \alpha_1 y_1 + \alpha_2 y_2 + \alpha_3 y_3 = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 \quad (3.111)$$

onde:

$$\begin{aligned} \alpha_1 &= -3 + a_2 \tau \\ \alpha_2 &= 3 - 2a_2 \tau + a_1 \tau^2 \\ \alpha_3 &= -1 + a_2 \tau - a_1 \tau^2 + a_0 \tau^3 \end{aligned} \quad (3.112)$$

e

$$\begin{aligned} \beta_1 &= b_2 \tau \\ \beta_2 &= -2b_2 \tau + b_1 \tau^2 \\ \beta_3 &= b_2 \tau - b_1 \tau^2 + b_0 \tau^3 \end{aligned} \quad (3.113)$$

As equações definidas em (3.112) e (3.113) podem ser expressas de forma matricial da conforme:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \tau & 0 & 0 \\ -2\tau & \tau^2 & 0 \\ \tau & -\tau^2 & \tau^3 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} + \begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix} \quad (3.114)$$

e

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \tau & 0 & 0 \\ -2\tau & \tau^2 & 0 \\ \tau & -\tau^2 & \tau^3 \end{bmatrix} \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix} \quad (3.115)$$

Lembrando que no presente caso os vetores de parâmetros  $\theta$  e  $\theta^*$  são:

$$\begin{aligned} \theta &= [a_2 \quad a_1 \quad a_0 \quad b_2 \quad b_1 \quad b_0]^T \\ \theta^* &= [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \beta_1 \quad \beta_2 \quad \beta_3]^T \end{aligned} \quad (3.116)$$

podemos relacioná-los através de uma única expressão combinando as equações (3.114) e (3.115) da seguinte forma:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \tau & 0 & 0 & 0 & 0 & 0 \\ -2\tau & \tau^2 & 0 & 0 & 0 & 0 \\ \tau & -\tau^2 & \tau^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau & 0 & 0 \\ 0 & 0 & 0 & -2\tau & \tau^2 & 0 \\ 0 & 0 & 0 & \tau & -\tau^2 & \tau^3 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \\ b_2 \\ b_1 \\ b_0 \end{bmatrix} + \begin{bmatrix} -3 \\ 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3.117)$$

que é a exatamente a transformação linear definida em (3.101) proposta por Johansson.

Dessa forma demonstramos que o método proposto por Johansson é equivalente ao método do momentos de Poisson normalizado. A única diferença entre os dois métodos é que no método de Poisson aplicado ao exemplo acima o termo independente utilizado no ajuste por mínimos quadrados é  $y_0$  que é o momento de Poisson de ordem 0 do sinal de saída do sistema,  $y(t)$ , ao passo que no método proposto por Johansson é utilizado o próprio sinal de saída. Isso pode ser ajustado, isto é, fazer o método de Johansson utilizar o sinal de saída filtrado pelos primeiro estágio da cadeia de filtros de Poisson,  $y_0$ , reescrevendo a equação (3.97) a seguinte forma:

$$G^*(\xi) = \frac{A^*(\xi)}{B^*(\xi)} = \frac{\beta_m \xi^{(m+1)} + \beta_{(m-1)} \xi^m + \dots + \beta_1 \xi^2}{\alpha_n \xi^{(n+1)} + \alpha_n \xi^n + \alpha_{(n-1)} \xi^{(n-1)} + \dots + \xi} \quad (3.118)$$

Substituindo-se a definição de  $\xi$  em função de  $p$ , dado pela equação (3.92), podemos transformar a equação (3.118) para uma função do operador  $p$  e demonstrar que a relação entre os vetores  $\theta$  e  $\theta^*$  continua sendo a mesma que no caso original proposto nesse método.

No método de Poisson também foi visto que podemos utilizar o momento de Poisson da derivada de qualquer ordem de  $y(t)$  como termo independente no método dos mínimos quadrados. Isso significa que podemos utilizar, por exemplo, o erro quadrático da predição da derivada de  $y(t)$  em vez do erro quadrático da predição do próprio  $y(t)$  como função objetivo no processo de identificação. Os benefícios de tal escolha de função objetivo foi analisado em [Machado e Trierweiler, 2003]. O método proposto por Johansson não possui essa possibilidade uma vez que não utiliza o momento de Poisson das derivadas, apenas os de  $y(t)$  e  $u(t)$ . O máximo que podemos no método proposto por Johansson é escolher qual momento de Poisson de  $y(t)$  utilizar com termo independente no ajuste por mínimos quadrados. Para isso basta realizar uma transformação adicional nos parâmetros do sistema dada por:

$$\alpha_i^* = \frac{\alpha_i}{\alpha_j}, \quad \beta_i^* = \frac{\beta_i}{\alpha_j} \quad (3.119)$$

onde  $\alpha_j$  é o parâmetro correspondente ao momento de ordem  $j$  que desejamos utilizar como termo independente.



### 3.11 Referências Bibliográficas

- Green, H.S. e Messel, H., “On the expansion of functions in terms of their moments”, *Qart. Appli. Math.*, V 11, pp. 403-409, (1954)
- Piovosso, M.J. e Bolgiano Jr., L. P., “Digital simulation using Poisson transform sequences”, *Symposium on Computer Processing in Communications*, Polytechnic Institute of Brooklyn, pp. 505-513, (1969).
- Fairman F. W. e Shen, D. W. C., “Parameter identification for linear time-varying dynamic processes”, *PROC. IEE.* V. 17, n°10, pp. 2025- 2029, (1970).
- Saha, D. C. e Rao, G. P. “Identification of continuous dynamical systems – the poisson moment functional (PMF) approach”, *Lecture notes in control and information sciences*, Springer-Verlag, pp. 1-51, (1983).
- Unbehauen, H. e Rao, G. P., “Identification of Continuous Systems”, *North-Holland Systems and Control Series*, V.10, pp 167-199 (1987).
- Johansson, R., “Identification of continuous-time models”, *IEEE Transactions on signal processing*, V. 42, n° 4, pp-887-897, (1994).
- Garnier, H., “Identification de modeles parametriques continus par moments de Poisson”, *Tese de doutorado*, Universidade Henri Poincare, Nancy 1, França, (1995).
- Garnier, H. e Mensler, M., “CONTSID: a continuous-time system identification toolbox for MATLAB”, *5<sup>th</sup> European Control Conference*, Karlsruhe (Germany), (1999).
- Trierweiler, J. O e Machado, V. C.,”Which is the best criterion for identification of dynamic models”, *15<sup>th</sup> IFAC World Congress on Automatic Control* (Barcelona), (2004).
- MATHWORKS INC, THE, “System Identification Toolbox for use with MATLAB version 4.0.5”, *Natick/MA/EUA*, (1998).
- Zhu, Y., “Multivariable process identification for MPC: the asymptotic method and its applications”, *J. Proc. Control*, V. 8, N° 2, pp. 101-115, (1998).

## Capítulo 4

### Compensação de *bias* dos parâmetros

No capítulo anterior demonstramos os métodos dos mínimos quadrados combinados com o método dos momentos de Poisson para o ajuste de parâmetros de modelos lineares contínuos invariantes no tempo. O método dos mínimos quadrados por sua natureza quadrática possui solução analítica o que o torna um método atraente para ajuste de parâmetros. Entretanto o método dos mínimos quadrados possui a desvantagem de gerar parâmetros incorretos quando há a presença de distúrbios nos dados utilizados no ajuste. Uma das áreas de maior interesse da identificação de sistema é o desenvolvimento de métodos para eliminação do erro dos parâmetros ajustados.

Nesse capítulo iremos demonstrar o efeito da presença de distúrbios nos dados utilizados no ajuste sobre o valor dos parâmetros dos modelos. Iremos também descrever de forma analítica o efeito do tempo de amostragem dos dados e da constante  $\lambda$  da cadeia de filtros de Poisson sobre o erro dos parâmetros ajustados. Finalmente serão apresentados três métodos para a compensação do erro dos parâmetros. Esses métodos são:

- 1) Máxima verossimilhança;
- 2) Variável instrumental;
- 3) Compensação de bias através do uso de pré-filtros.

Como conclusão iremos aplicar esse três métodos na identificação de um sistema linear na presença de ruídos brancos e coloridos nos dados de saída e comparar o resultados.

No restante desse capítulo será usado o termo *bias*, amplamente empregado na literatura especializada, para se referir ao erro dos parâmetros estimados.

## 4.1 Análise do bias dos parâmetros ajustados pelo método dos mínimos quadrados.

Um sistema linear contínuo contaminado por distúrbios pode ser representado pelas seguintes equações:

$$\begin{aligned} z(t) &= \frac{B(p)}{A(p)} u(t) \\ y(t) &= z(t) + v(t) \end{aligned} \quad (4.1)$$

onde:

$u(t)$  é o sinal de entrada do sistema

$z(t)$  é o sinal de saída livre de distúrbios

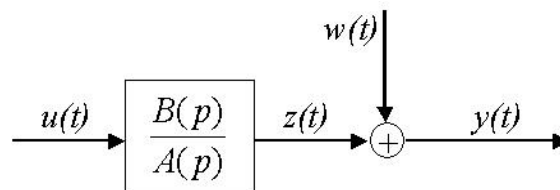
$y(t)$  é o sinal de saída contaminado por ruído

$v(t)$  é o ruído branco ou colorido que contamina os sinal de saída do sistema

$p$  é o operador de diferenciação  $d/dt$  e

$A$  e  $B$  são polinômios do operador de diferenciação  $p$ .

Essa equação pode ser representada pelo diagrama de blocos mostrado na figura 4.1, abaixo:



**Figura 4.1:** Sistema contínuo perturbado por um ruído genérico  $w(t)$ .

O sistema de equações descrito em (4.1) pode ser reescrito na forma de uma única equação conforme:

$$y(t)^T A(p) = u^T B(p) + v^T A(p) \quad (4.2)$$

Comparando a equação (4.2) com a figura 4.1 podemos definir o ruído genérico  $w(t)$  como sendo dado por:

$$w(t) = v(t)^T A(p) \quad (4.3)$$

Como visto no capítulo 3 se aplicarmos a transformação de Poisson à equação (4.2) podemos expressá-la na forma de um regressor linear da seguinte forma:

$$\mathbf{y}_k^{t_0} = \mathbf{M}_k \theta + \mathbf{w}_k^{t_0} \quad (4.4)$$

onde  $\mathbf{w}_k^{t_0}$  é a matriz composta pelos momentos de Poisson do sinal  $w(t)$ .

Se desconsideramos a existência do distúrbio  $w(t)$  na equação (4.4) já vimos no capítulo 3 que o vetor de parâmetros  $\theta$  pode ser calculado através do métodos dos mínimos quadrados dado por:

$$\hat{\theta} = (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \mathbf{M}_k^T \mathbf{y}_k^{t_0} = \left[ \sum_{l=1}^N \mathbf{M}_k^T(l) \mathbf{M}_k(l) \right]^{-1} \sum_{l=1}^N \mathbf{M}_k^T(l) \mathbf{y}_k^{t_0}(l) \quad (4.5)$$

Se substituirmos a equação (4.4) na equação (4.5) obtemos a seguinte expressão:

$$\hat{\theta} = \theta + (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \mathbf{M}_k^T \mathbf{w}_k^{t_0} \quad (4.6)$$

Assim podemos concluir que o método dos mínimos quadrados quando aplicado a dados contendo distúrbios gera uma estimativa do vetor de parâmetros diferente do vetor real. Essa diferença é usualmente conhecida como *bias* e é dado por:

$$\Delta_{\theta} = (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \mathbf{M}_k^T \mathbf{w}_k^{t_0} \quad (4.7)$$

A equação acima nos mostra que o *bias* é dado pelo produto de duas parcelas:

$$\Delta_{\theta} = \mathbf{R}_{\text{MM}}^{-1} \mathbf{R}_{\text{Mw}} \quad (4.8)$$

onde:

$$\mathbf{R}_{\text{MM}} = \mathbf{M}_k^T \mathbf{M}_k = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (\mathbf{M}_k^T(i) \mathbf{M}_k(i)) = E(\mathbf{M}_k^T \mathbf{M}_k) \quad (4.9)$$

é a matriz de covariância dos regressores e:

$$\mathbf{R}_{\text{Mw}} = \mathbf{M}_k^T \mathbf{w}_k^{t_0} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (\mathbf{M}_k^T(i) \mathbf{w}_k^{t_0}(i)) = E(\mathbf{M}_k^T \mathbf{w}_k^{t_0}) \quad (4.10)$$

é a matriz de variância cruzada entre a matriz de regressores e o ruído. Essa última parcela fornece uma medida da quantidade de informação sobre a dinâmica do processo contida no ruído  $w(t)$ .

Podemos concluir que a estimativa do vetor de parâmetros dado pelo método dos mínimos quadrados é convergente a medida que o número de observações tende ao infinito. De forma resumida essa estimativa é dada por:

$$\lim_{N \rightarrow \infty} \hat{\theta}(N) = \theta + \Delta_{\theta} \quad (4.11)$$

Demonstramos, portanto que o método dos mínimos quadrados fornece, sempre que algum distúrbio estiver presente nos dados, uma estimativa contendo um *bias*. De forma geral o *bias* provocado por ruído branco é menor do que o provocado por ruído colorido. No caso do método dos momentos de Poisson o ruído branco desvia para um ruído colorido devido a sua passagem pela cadeia de filtros de Poisson.

## 4.2 Análise do valor do *bias* em função do tempo de amostragem e da constante da cadeia de filtros para um sistema de 1ª ordem

No capítulo 2 vimos que o tempo de amostragem dos dados utilizados na identificação de um modelo linear afeta a estimativa dos parâmetros. No caso dos modelos no espaço de tempo discreto vimos que tanto um tempo de amostragem muito grande com um muito pequeno afetam a qualidade da estimação. No primeiro caso temos o efeito *aliasing* e no segundo temos problemas numéricos devido a dependência dos parâmetros do modelo discreto com o tempo de amostragem (a medida que o tempo de amostragem tende para zero todos os pólos e zeros do modelo tendem para o ponto (1,0) no plano complexo do espaço de tempo discreto).

No capítulo 3 vimos que a precisão da estimativa dos parâmetros de um modelo linear dado pelo método dos momentos de Poisson depende da constante  $\lambda$  da cadeia de filtros de Poisson. Entretanto não vimos uma correlação clara entre a precisão das estimativas com o valor dessa constante. Garnier em sua tese de doutorado [Garnier, 1995] demonstrou uma correlação analítica entre o erro da predição do método dos momentos de Poisson em função do tempo de amostragem e da constante  $\lambda$  para um sistema de primeira ordem quando contaminado tanto por ruído branco quanto por ruído colorido.

Considere o seguinte sistema de primeira ordem contaminado por ruído:

$$y(p) = \frac{b}{p+a} u(p) + v(p) \quad (4.12)$$

que também pode ser descrito por:

$$\begin{aligned} \frac{dz(t)}{dt} + az(t) &= bu(t) \\ y(t) &= z(t) + v(t) \end{aligned} \quad (4.13)$$

com  $v(t)$  sendo um ruído branco ou colorido.

O estudo das propriedades estatísticas do método dos momentos de Poisson em conjunto com o método dos mínimos quadrados no caso perfeitamente contínuo é muito complexo. Uma maneira de evitar essas complexidades e tratar o sistema como parcialmente contínuo, isto é, a parte determinística como contínua e a parte estocástica como discreta. Na realidade o fato de os dados de entrada e saída do sistema a ser identificado serem amostrados e um método numérico de integração a ser usado para calcular os momentos de Poisson desses, nos permite transpor o estudo de um sistema contínuo para uma perspectiva discreta.

Na prática não dispomos de  $u(t)$  e  $y(t)$  mas sim de amostragens desses,  $u(l.T_e)$  e  $y(l.T_e)$  com  $l \in \mathbf{I}$ , em um intervalo de tempo  $T_e$  constante. Normalmente trata-se os dados discretos utilizando um mantenedor (*hold*) de ordem zero, isto é, considera-se que os valores de  $u(t)$  e  $y(t)$  entre duas amostragens consecutivas possuem valores constante. Um modelo de primeira ordem discreto correspondente ao modelo dado pela equação (4.12) considerando um mantenedor de ordem zero é dado por:

$$y(z^{-1}) = \frac{b_d z^{-1}}{1 + a_d z^{-1}} u(z^{-1}) + v(z^{-1}) \quad (4.14)$$

com

$$\begin{aligned} a_d &= -e^{a.T_e} \\ b_d &= \frac{b}{a} (1 + e^{-a.T_e}) \end{aligned} \quad (4.15)$$

Foi analisado o caso particular quando ruído branco é adicionado ao sinal de saída. A representação do sistema anterior na forma de equações de diferenças é dada por:

$$y(l) = -a_d y(l-1) + b_d u(l-1) + e(l) + a_d e(l-1) \quad (4.16)$$

onde  $e(l)$  é um ruído branco.

No capítulo anterior não tratamos da questão da implementação numérica da cadeia de filtros de Poisson entretanto essa questão é importante uma vez que os momentos de Poisson dos sinais de entrada e saída são gerados a partir de medidas discretas destes mesmo. Iremos considerar nesse estudo que os momentos de Poisson dos dados de entrada e saída são gerados

a partir de uma discretização dos elementos da cadeia de filtros de Poisson considerando que os valores entre dois tempos de amostragem consecutivos são constantes. Dessa forma os momentos de Poisson de ordem 0 e 1 (os mínimos necessários para identificar um modelo de primeira ordem) dos sinais de entrada e saída nos diferentes instantes de amostragem são dados por:

$$u_0(l) = -a_1 u_0(l-1) + b_1 u(l-1) \quad (4.17)$$

$$u_1(l) = -a_1 u_1(l-1) + b_1 u_0(l-1) \quad (4.18)$$

$$y_0(l) = -a_1 y_0(l-1) + b_1 y(l-1) \quad (4.19)$$

$$y_1(l) = -a_1 y_1(l-1) + b_1 y_0(l-1) \quad (4.20)$$

com

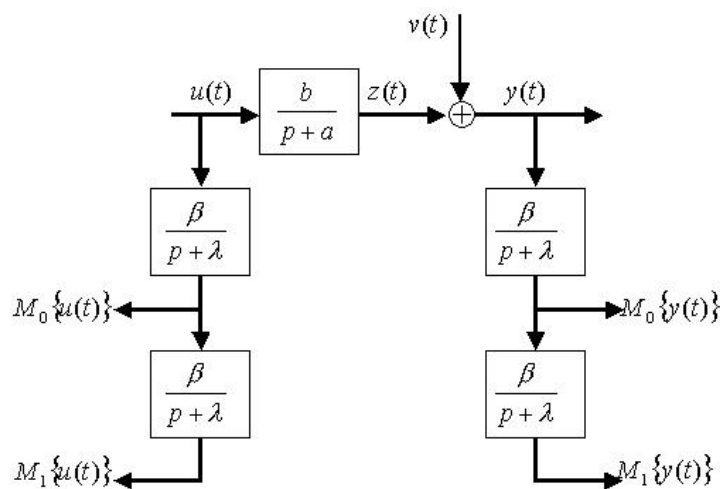
$$u_0(l) = M_0 \{u(l)\} \quad u_1(k) = M_1 \{u(l)\} \quad (4.21)$$

$$y_0(l) = M_0 \{y(l)\} \quad y_1(l) = M_1 \{y(l)\} \quad (4.22)$$

e

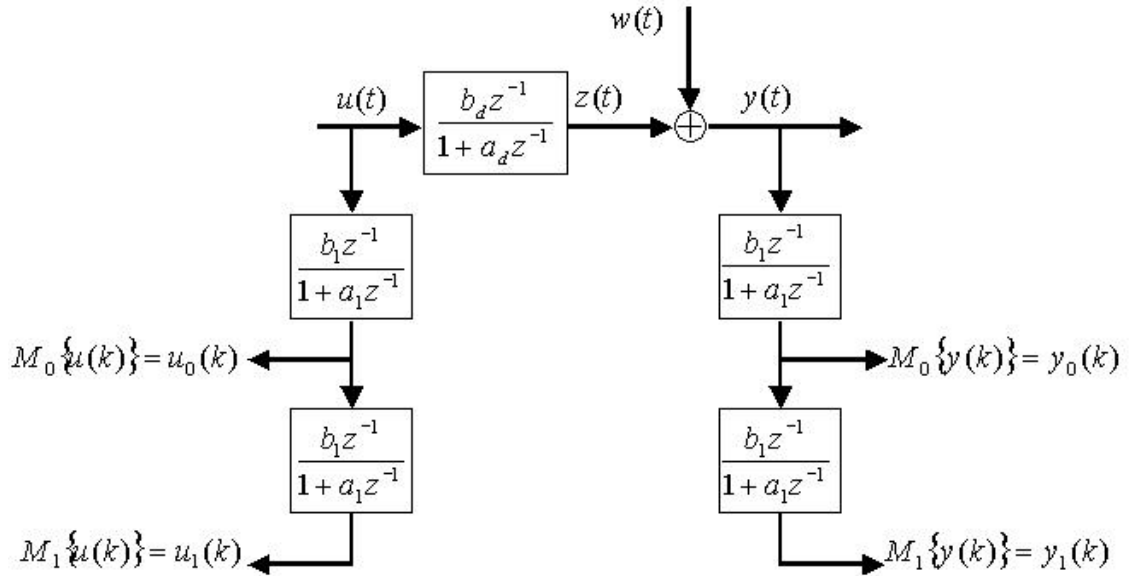
$$\begin{aligned} a_1 &= -e^{-\lambda T_e} \\ b_1 &= \frac{\beta}{\lambda} (1 - e^{-\lambda T_e}) \end{aligned} \quad (4.23)$$

O esquema de estimação de um modelo de primeira ordem contínuo é representado na figura 4.2, abaixo:



**Figura 4.2:** Esquema de identificação de um sistema contínuo de primeira ordem.

A versão discreta do mesmo esquema considerando invariância dos valores entre instantes de amostragem consecutivos é mostrado na figura 4.3, abaixo:



**Figura 4.3:** Esquema de identificação de um sistema contínuo na forma discreta equivalente.

Estabelecido essas definições consideremos o sistema de primeira ordem definido pela seguinte equação diferencial:

$$\frac{dy(t)}{dt} + ay(t) = bu(t) + w(t) \quad (4.24)$$

As condições iniciais são consideradas nulas e  $w(t)$  é um ruído contínuo que perturba a saída do sistema. Aplicando o método dos momentos de Poisson de ordem mínima à equação acima temos:

$$M_1 \left\{ \frac{dy(t)}{dT} \right\} + aM_1 \{y(t)\} = bM_1 \{u(t)\} + M_1 \{w(t)\} \quad (4.25)$$

com

$$M_1 \left\{ \frac{dy(t)}{dt} \right\} = \beta M_0 \{y(t)\} - \lambda M_1 \{y(t)\} \quad (4.26)$$

Os momentos de Poisson dos sinais de entrada e saída são calculados através das equações (4.17) à (4.20). A equação (4.4) para o momento  $k$ .  $T_e$  pode ser reescrita da seguinte forma:

$$\mathbf{y}_k^{t_0}(l) = \mathbf{M}_k(l)\theta + \mathbf{w}_k^{t_0}(l) \quad (4.27)$$



Para o exemplo do sistema de primeira ordem temos:

$$\mathbf{y}_k^{t_0}(l) = M_1 \left\{ \frac{dy(t)}{dt} \right\} (l) = \beta y_0(l) - \lambda y_1(l) \quad (4.28)$$

$$\mathbf{M}_k(l) = [-y_1(l) \quad u_1(l)] \quad (4.29)$$

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix} \quad (4.30)$$

Dispondo de  $N$  diferentes medidas dos sinais de entrada e saída podemos reescrever a expressão dos mínimos quadrados dado pela equação (4.4) da seguinte forma:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum_{l=1}^N (y_1^2(l)) & -\sum_{l=1}^N (y_1(l)u_1(l)) \\ -\sum_{l=1}^N (y_1(l)u_1(l)) & \sum_{l=1}^N (u_1^2(l)) \end{bmatrix}^{-1} \begin{bmatrix} -\beta \sum_{l=1}^N (y_1(l)y_0(l)) + \lambda \sum_{l=1}^N (y_1^2(l)) \\ \beta \sum_{l=1}^N (y_0(l)u_1(l)) - \lambda \sum_{l=1}^N (y_1(l)u_1(l)) \end{bmatrix} \quad (4.31)$$

O sistema estudado é suposto estável e o número de observações,  $N$ , infinito. Considerando isso os somatórios da equação (4.31) podem ser substituído pela expectativa matemática. Assim nesse caso, a equação anterior pode ser reescrita da seguinte forma:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \mathcal{E}(y_1^2(l)) & -\mathcal{E}(y_1(l)u_1(l)) \\ -\mathcal{E}(y_1(l)u_1(l)) & \mathcal{E}(u_1^2(l)) \end{bmatrix}^{-1} \begin{bmatrix} -\beta \mathcal{E}(y_1(l)y_0(l)) + \lambda \mathcal{E}(y_1^2(l)) \\ \beta \mathcal{E}(y_0(l)u_1(l)) - \lambda \mathcal{E}(y_1(l)u_1(l)) \end{bmatrix} \quad (4.32)$$

Supondo que o sinal de entrada e um ruído  $e$  são seqüências independentes com médias zeros e variância  $\sigma_u^2$  e  $\sigma_e^2$  temos:

$$\mathcal{E}(u(l)) = 0 \quad \mathcal{E}(u(i)u(j)) = \begin{cases} 0 & \text{se } i \neq j \\ \sigma_u^2 & \text{se } i = j \end{cases} \quad (4.33)$$

$$\mathcal{E}(e(l)) = 0 \quad \mathcal{E}(e(i)e(j)) = \begin{cases} 0 & \text{se } i \neq j \\ \sigma_e^2 & \text{se } i = j \end{cases} \quad (4.34)$$

Ainda, se consideramos que os sinais são estacionários podemos considerar a seguinte propriedades:

$$\mathcal{E}(y_1(l-1)y_1(l-1)) = \mathcal{E}(y_1^2(l)) \quad (4.35)$$

Para transformar a equação (4.32) em uma função analítica de  $\lambda$  e  $Te$  é necessário o conhecimento de uma série de expectâncias matemáticas que podem ser obtidas combinando as equações (4.17) à (4.20) com as equações (4.15) e (4.23) e considerando o sistema linear de primeira ordem discretizado como sendo dado pela equação (4.16) com uma pequena alteração, o parâmetro do filtro do ruído é independente e igual a  $c$ , conforme:

$$y(l) = -a_d y(l-1) + b_d u(l-1) + e(l) + ce(l-1) \quad (4.36)$$

Assim as expectâncias matemáticas necessárias são dadas pelas seguintes relações.

$$\mathcal{E}(y^2(l)) = \frac{b_d \sigma_u^2 + (1 - 2a_d c + c^2) \sigma_e^2}{1 - a_d^2} \quad (4.37)$$

$$\mathcal{E}(y_0(l)y(l)) = \frac{-a_d b_1 \mathcal{E}(y^2(l)) + b_1 c \sigma_e^2}{1 - a_1 a_d} \quad (4.38)$$

$$\mathcal{E}(y_0^2(l)) = \frac{-2a_1 b_1 \mathcal{E}(y_0(l)y(l)) + b_1^2 \mathcal{E}(y^2(l))}{1 - a_1^2} \quad (4.39)$$

$$\mathcal{E}(u_1^2(l)) = \frac{b_1^4 (1 + a_1^2)}{(1 - a_1^2)^3} \sigma_u^2 \quad (4.40)$$

$$\mathcal{E}(y_1(l)y(l)) = \frac{-a_d b_1}{1 - a_1 a_d} \mathcal{E}(y_0(l)y(l)) \quad (4.41)$$

$$\mathcal{E}(y_0(l)y_1(l)) = \frac{-a_1 b_1 \mathcal{E}(y_0^2(l)) - a_1 b_1 \mathcal{E}(y_1(l)y(l)) + b_1^2 \mathcal{E}(y_0(l)y(l))}{1 - a_1^2} \quad (4.42)$$

$$\mathcal{E}(y(l)u_0(l)) = \frac{b_1 b_d}{1 - a_1 a_d} \sigma_u^2 \quad (4.43)$$

$$\mathcal{E}(y_0(l)u_0(l)) = \frac{-a_1 b_1}{1 - a_1^2} \mathcal{E}(y(l)u_0(l)) \quad (4.44)$$

$$\mathcal{E}(y(l)u_1(l)) = \frac{-a_d b_1}{1 - a_1 a_d} \mathcal{E}(y(l)u_0(l)) \quad (4.45)$$

$$\mathcal{E}(y_0(l)u_1(l)) = \frac{-a_1 b_1 \mathcal{E}(y_0(l)u_0(l)) - a_1 b_1 \mathcal{E}(y(l)u_1(l)) + b_1^2 \mathcal{E}(y(l)u_0(l))}{1 - a_1^2} \quad (4.46)$$

$$\mathcal{E}(y_1(l)u_0(l)) = \frac{-a_1 b_1}{1 - a_1^2} \mathcal{E}(y_0(l)u_0(l)) \quad (4.47)$$

$$\mathcal{E}(y_1(l)u_1(l)) = \frac{-a_1 b_1 \mathcal{E}(y_1(l)u_0(l)) - a_1 b_1 \mathcal{E}(y_0(l)u_1(l)) + b_1^2 \mathcal{E}(y_0(l)u_0(l))}{1 - a_1^2} \quad (4.48)$$

$$\mathcal{E}(y_1^2(l)) = \frac{-2a_1 b_1 \mathcal{E}(y_0(l)y_1(l)) + b_1^2 \mathcal{E}(y_0^2(l))}{1 - a_1^2} \quad (4.49)$$

Considerando o parâmetro  $c$  como sendo igual a zero, temos o caso de um ruído branco adicionado ao sinal de saída. Com essa consideração em mente e substituindo as diferentes expectâncias matemáticas apresentadas acima é possível chegar as seguinte expressões para as estimativas dos parâmetros  $a$  e  $b$ .

$$\hat{a} = \frac{\lambda(-1 + e^{-aTe})}{-1 + e^{-\lambda Te}} \quad (4.50)$$

$$\hat{b} = \frac{b}{a} \frac{\lambda(-1 + e^{-aTe})}{-1 + e^{-\lambda Te}} = \frac{b}{a} \hat{a} \quad (4.51)$$

Se considerarmos que temos um tempo de amostragem suficientemente pequeno ( $Te \ll \min[1/a, 1/\lambda]$ ) podemos aproximar os termos exponenciais presentes nas equações acima por uma expansão em séries de Taylor truncadas nos termos de segunda ordem.

$$e^{-aTe} \approx 1 - aTe + \frac{(aTe)^2}{2} \quad (4.52)$$

$$e^{-\lambda Te} \approx 1 - \lambda Te + \frac{(\lambda Te)^2}{2} \quad (4.53)$$

Substituindo essas expansões nas equações (4.50) e (4.51) temos as seguintes expressões para as estimativas dos parâmetros  $a$  e  $b$ .

$$\hat{a} = a + \frac{Te(a - \lambda)a}{\lambda Te - 2} \quad (4.54)$$

$$\hat{b} = b + \frac{Te(a - \lambda)b}{\lambda Te - 2} \quad (4.55)$$

Observando as equações acima vemos que o *bias* dos parâmetros estimados pelo método de Poisson tende a zero a medida que o tempo de amostragem também tende a zero. É

possível observar também que escolhendo a constante da cadeia de filtros de Poisson,  $\lambda$ , igual ao parâmetro  $a$  o *bias* é nulo aumentando a medida que  $\lambda$  se afasta de  $a$ .

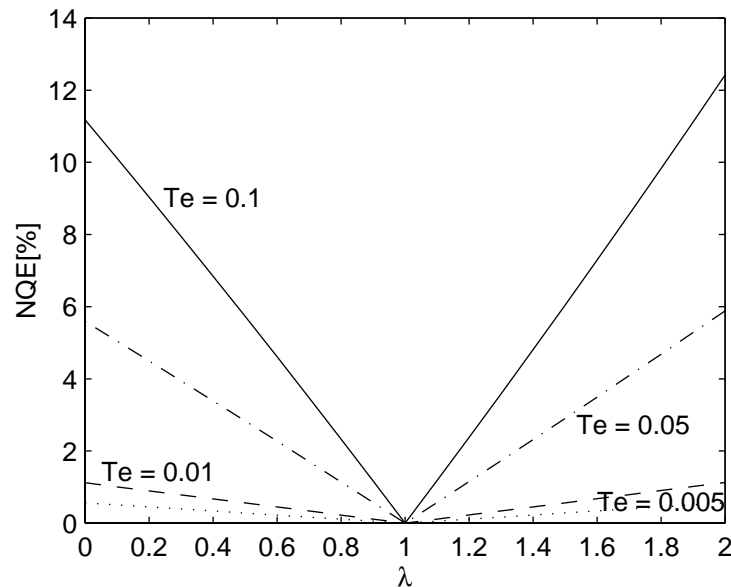
Com essas duas expressões é possível expressar a norma quadrática do erro como definido no capítulo 3 de forma analítica.

$$NQE[\%] = 100Te \left| \frac{a - \lambda}{\lambda Te - 2} \right| \sqrt{a^2 + b^2} \quad (4.56)$$

Para ilustrar o efeito do tempo de amostragem  $Te$  e da constante  $\lambda$  vamos tomar com exemplo o seguinte sistema:

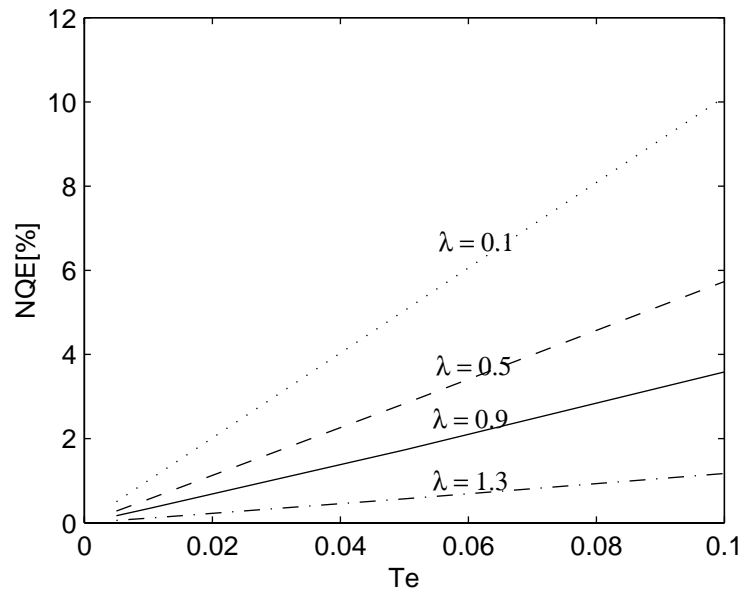
$$G(s) = \frac{b}{s + a} \quad (4.57)$$

com  $b = 2$  e  $a = 1$ . O gráfico da norma quadrática do erro em função da constante  $\lambda$  para vários tempos de amostragem,  $Te$ , é apresentado na figura 4.4, abaixo:



**Figura 4.4:** Norma quadrática do erro em função da constante  $\lambda$  para diferentes tempos e amostragem,  $Te$ .

Os mesmos dados podem ser vistos na figura 4.5 onde é apresentada a norma do erro quadrático em função do tempo de amostragem para diferentes valores de  $\lambda$ .



**Figura 4.5:** Norma quadrática do erro em função do tempo de amostragem,  $T_e$ , para diferentes valores da constante  $\lambda$ .

No caso de escolhermos o valor de  $c$  como sendo diferente de zero, por exemplo igual a  $a_d$ , temos um sistema perturbado por um sinal colorido. Nesse caso a dedução de uma equação analítica para a influência do tempo da amostragem,  $T_e$ , e da constante  $\lambda$  sobre o *bias* dos parâmetros estimados torna-se muito trabalhosa e a equação gerada muito complexa. Garnier em sua tese de doutorado apresenta a análise desse caso. Para sistema de ordem maior do que 1 a análise torna-se ainda mais complexa mas pode-se inferir que o valor ótimo de  $\lambda$  está de alguma forma relacionado a dinâmica dominante do sistema.

### 4.3 Métodos de compensação de *bias*

No capítulo 3 vimos que o método dos momentos de Poisson consegue lidar bem com a presença de ruído branco no sinal de saída até valores de razão ruído sobre sinal de mais ou menos 20%, o que já é uma grande vantagem sobre métodos discretos como ARX como também demonstrados no capítulo 3. Acima desse valor o método dos momentos de Poisson começa a gerar estimativas dos parâmetros com *bias* significativo.

Para lidar com esses casos e para aqueles casos onde a natureza do distúrbio do sistema não pode ser representada por um ruído branco devemos utilizar algum método de compensação de *bias*.

Existem diversos métodos para compensação ou redução de *bias*. Entre os mais conhecidos estão os métodos dos mínimos quadrados generalizados, o métodos da máxima verossimilhança, o método das variáveis instrumentais e o método de compensação de *bias*

por pré-filtro dos sinais de entrada e/ou saída. Nesse trabalho iremos analisar os três últimos métodos.

### 4.3.1 Método da máxima verossimilhança

O método da máxima verossimilhança foi amplamente estudado para métodos de identificação de modelos discretos principalmente por Ljung [Ljung, 1987] e aplicado no *Identification Toolbox* do MATLAB uma das ferramentas mais usadas na área de identificação de sistemas lineares. A sua aplicação para identificação de modelos contínuos foi primeiramente estudada por [Unbehauen e Rao, 1987].

Seja  $\psi$  uma variável randômica da qual  $N$  medidas estejam disponíveis. Essas medidas são estatisticamente independentes uma da outra. Considerando que um função de densidade de probabilidade ( $\mathcal{F}$ ) seja dependente de um conjunto de parâmetros  $p_1, p_2, \dots, p_n$  e descrita no estágio  $\psi = x_i$  por:

$$\mathcal{F}_\psi(x_i) = \mathcal{F}_\psi(x_i; p_1, p_2, \dots, p_n) \quad \text{para } i = 1, \dots, N \quad (4.58)$$

Visando estimar os valores dos parâmetros  $p_1, \dots, p_n$  utilizando as medidas disponíveis definimos uma função de verossimilhança,  $L$ , como:

$$\begin{aligned} L &= L(x_1, x_2, \dots, x_N; p_1, p_2, \dots, p_n) = \mathcal{F}_\psi(x_1, x_2, \dots, x_N; p_1, p_2, \dots, p_n) = \\ &= \prod_{i=1}^N \mathcal{F}_\psi(x_i; p_1, \dots, p_n) \end{aligned} \quad (4.59)$$

Se a natureza da função de função de densidade de probabilidade (por exemplo: Gaussiana, Poisson etc) for conhecida, então a função de verossimilhança pode ser considerada um função dos parâmetros desconhecidos. A idéia básica por trás do método da máxima verossimilhança é obter um estimativa,  $\hat{p}_i$ , dos parâmetros  $p_i$  de tal forma que a função de verossimilhança para as medidas  $x_1, \dots, x_N$  seja maximizada. De forma a maximizar a função de verossimilhança devemos resolver o seguinte conjunto de equações:

$$\frac{\partial L}{\partial p_i} = 0, \quad \text{para } i = 1, 2, \dots, n \quad (4.60)$$

A solução das equações definidas em (4.60) é  $p_i \equiv \hat{p}_i$ . Se a função de verossimilhança for monotônica também podemos maximizar o logaritmo de  $L$ . Assim:

$$\frac{\partial}{\partial p_i} \ln L = 0, \quad \text{para } i = 1, 2, \dots, n \quad (4.61)$$

Essas equações são conhecidas como equações de máxima verossimilhança. É possível demonstrar que para um número suficientemente grande de observações a estimativa dada pelo método da máxima verossimilhança é estatisticamente consistente, eficiente e exaustivo [Unbehauen e Rao, 1987].

Para demonstrar o uso do método da máxima verossimilhança vamos considerar o exemplo de um modelo contínuo equivalente a uma estrutura ARMAX para identificação de modelos discretos. Vamos referir esse modelo como ARMAX-contínuo. Esse modelo pode ser descrito por:

$$\mathbf{A.M}\{y(t)\}^T - \mathbf{B.M}\{u(t)\}^T = \mathbf{C.M}\{w(t)\}^T \quad (4.62)$$

Onde os componentes dos vetores  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  devem ser estimados e onde  $w(t)$  é ruído branco com uma distribuição gaussiana com média zero e desvio padrão  $\sigma_w$ . O vetor de parâmetros,  $\theta$ , é dado por:

$$\theta = [a_1 \quad \dots \quad a_n \quad | \quad b_0 \quad \dots \quad b_m \quad | \quad c_1 \quad \dots \quad c_l] \quad (4.63)$$

Nessa situação  $\sigma_w$  também deve ser estimada, isto é, é necessário determinar o sinal  $w(t)$ . Como estamos considerando que  $w(t)$  é um ruído branco com distribuição gaussiana podemos considerar que a função de densidade de probabilidade para a medida  $i$  seja dada por uma distribuição gaussiana de forma que:

$$F(M_k \{w(t)\}(i); \theta \sigma_w^2) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(-\frac{M_k^2 \{w(t)\}(i)}{2\sigma_w^2}\right) \quad (4.64)$$

Como o sinal  $w(t)$  é decorrelacionado, a função de verossimilhança para  $N$  observações pode ser expressa por:

$$L = L(M_k \{w(t)\}(i); \theta \sigma_w^2) = \frac{1}{(\sqrt{2\pi\sigma_w^2})^N} \exp\left(-\frac{1}{2\sigma_w^2} \sum_{i=1}^N M_k^2 \{w(t)\}(i)\right) \quad (4.65)$$

Além da função de verossimilhança devemos estabelecer uma função custo. A mais comum de todas as funções custo é o somatório dos erros quadráticos. No caso do método dos momentos de Poisson usamos o somatório do quadrado do momento de Poisson de ordem  $k$  do sinal  $w(t)$  dado por:

$$V(\theta) = \frac{1}{2} \sum_{i=1}^N M_k^2 \{w(t)\} \quad (4.66)$$

As equações de máxima verossimilhança são, então, dadas por:

$$\ln L = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma_w^2 - \frac{1}{\sigma_w^2} V(\theta) \quad (4.67)$$

e por:

$$\frac{\partial}{\partial a_i} \ln L = \frac{\partial V(\theta)}{\partial a_i} = 0 \quad (4.68)$$

$$\frac{\partial}{\partial b_i} \ln L = \frac{\partial V(\theta)}{\partial b_i} = 0 \quad (4.69)$$

$$\frac{\partial}{\partial c_i} \ln L = \frac{\partial V(\theta)}{\partial c_i} = 0 \quad (4.70)$$

$$\frac{\partial}{\partial \sigma_w^2} \ln L = -\frac{N}{2\sigma_w^2} + \frac{1}{\sigma_w^4} V(\theta) = 0 \quad (4.71)$$

A equação (4.71) dá a estimativa da variância do erro do modelo,  $w(t)$ .

$$\sigma_w^2 = \frac{2}{N} V(\hat{\theta}) \quad (4.72)$$

Para estimar os valores dos parâmetros do modelo é necessário a resolução das equações (4.68) a (4.70). Para determinar o valor das derivadas parciais de  $V(\theta)$  em função dos parâmetros do modelo devemos utilizar a equação (4.66). Para isso é necessário montar um expressão para o cálculo de  $M_k\{w(t)\}$ . Isso pode ser feito a partir da equação (4.62) que pode ser reescrita da seguinte forma:

$$M_k\{w(t)\} = M_k\{y(t)\} + \sum_{i=1}^n a_i M_k \left\{ \frac{d^i y(t)}{dt^i} \right\} - \sum_{i=0}^m b_i M_k \left\{ \frac{d^i u(t)}{dt^i} \right\} - \sum_{i=1}^l c_i M_k \left\{ \frac{d^i w(t)}{dt^i} \right\} \quad (4.73)$$

A equação (4.73) é linear em relação aos parâmetros  $a_i$  e  $b_i$  mas é não linear em relação a  $c_i$  pois o sinal  $w(t)$  é função dos parâmetros do modelos e é calculado por:

$$w(t) = \frac{A(p)}{C(p)} y(t) - \frac{B(p)}{C(p)} u(t) \quad (4.74)$$

A não linearidade da equação (4.73) em relação aos parâmetros  $c_i$  impede que a minimização de  $V(\theta)$  seja feita analiticamente. Dessa forma é necessário utilizar uma técnica de minimização numérica apropriada. Nas técnicas de minimização clássicas ambas as matrizes gradiente e hessiana de  $V(\theta)$  em relação a  $\theta$  devem ser calculadas. Por definição essas duas matrizes são dadas por:



$$V_{\theta} = \begin{bmatrix} \frac{\partial V}{\partial p_1} \\ \vdots \\ \frac{\partial V}{\partial p_{n+m+1}} \end{bmatrix} \quad (4.75)$$

$$V_{\theta\theta} = \begin{bmatrix} \frac{\partial^2 V}{\partial p_1 \partial p_1} & \dots & \frac{\partial^2 V}{\partial p_1 \partial p_{n+m+1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial p_{n+m+1} \partial p_1} & \dots & \frac{\partial^2 V}{\partial p_{n+m+1} \partial p_{n+m+1}} \end{bmatrix} \quad (4.76)$$

Para montar essas duas matrizes é necessário o cálculo das seguintes derivadas parciais:

$$\frac{\partial V}{\partial p_i} = \sum_{j=1}^N M_k \{w(t)\} (i) \frac{\partial M_k \{w(t)\}}{\partial p_i} \quad (4.77)$$

e

$$\frac{\partial^2 V}{\partial p_i \partial p_j} = \sum_{j=1}^N \frac{\partial M_k \{w(t)\}}{\partial p_i} \frac{\partial M_k \{w(t)\}}{\partial p_j} + \sum_{j=1}^N M_k \{w(t)\} \frac{\partial^2 M_k \{w(t)\}}{\partial p_i \partial p_j} \quad (4.78)$$

onde  $M_k \{w(t)\}$  é dado pela equação (4.73).

Se um método de minimização tipo gradiente for escolhido um algoritmo iterativo do tipo:

$$\hat{\theta}_{i+1} = \hat{\theta}_i + \delta\theta_i \quad (4.79)$$

pode ser utilizado, onde a parcela  $\delta\theta$  representa o incremento do vetor de parâmetros a cada iteração. Essa parcela tende a zero a medida que  $\hat{\theta}$  se aproxima de  $\theta$  e seu cálculo varia conforme o método de minimização utilizado. Por exemplo, para o método de Newton-Raphson a equação (4.79) fica:

$$\hat{\theta}_{i+1} = \hat{\theta}_i - \rho \left[ V_{\theta\theta}(\hat{\theta}_i) \right]^{-1} V_{\theta}(\hat{\theta}_i) \quad (4.80)$$

O processo iterativo de ajuste do modelo pode ser descrito pelos seguintes passos:

- 1) Selecionar a estimativa inicial mais apropriada para o vetor de parâmetros  $\hat{\theta}_0$ ;

- 2) Calcular  $I_\theta$  e  $I_{\theta\theta}$  para  $i = 0$  usando todas as  $N$  observações ;
- 3) Calcular a nova estimativa  $\hat{\theta}_{i+1}$  ;
- 4) Repetir os passos 2 e 3;
- 5) Para o processo iterativo se os critérios de otimização forem atingidos.

O grande esforço computacional associado a esse método é compensado pelo cálculo simultâneo da matriz  $V_{\theta\theta}$  que permite a estimação da variância dos parâmetros estimados que é dada por:

$$\sigma_{\theta_i}^2 = \sigma_w^2 \left[ V_{\theta\theta}(\hat{\theta}_i) \right]^{-1} \quad (4.81)$$

Apesar do método da máxima verossimilhança ser teoricamente consistente ele apresenta algumas dificuldades do ponto de vista prático. A primeira grande dificuldade é a necessidade da geração do ruído branco  $w(t)$  de acordo com a equação (4.74). Devido a essa equação é necessário garantir que o polinômio  $C(p)$  seja estável. Isso torna necessária a imposição de restrições no processo de otimização que nem sempre são simples de implementar uma vez que a restrição deve ser posta nas raízes do polinômio e não nos parâmetros propriamente dito.

Outra dificuldade de um método iterativo de busca como esse é a necessidade de um chute inicial para os parâmetros a serem estimados. Normalmente a escolha do chute inicial para o modelo do processo ( $A$  e  $B$ ) é estimada usando o método dos mínimos quadrados. A estimativa inicial dos parâmetros para o modelo do distúrbio é obtida também através do método dos mínimos quadrados utilizando como entrada do modelo a diferença entre a predição da saída da primeira estimativa do modelo do processo e como saída a própria saída do sistema.

O método da máxima verossimilhança possui poucas referências de aplicação na identificação de modelos contínuos. Referências são encontradas em [Unbehauen e Rao, 1987] e em [Johansson, 1994]. Foi implementada nesse trabalho a metodologia descrita por Unbehauen que é muito similar a implementada no *System Identification Toolbox* do MATLAB. Foi observado que ao contrário da versão discreta a versão contínua é muito mais sensível quanto a estrutura dos modelos, tanto do processo quanto do distúrbio, ou seja, para a identificação de modelos contínuos pelo método da máxima verossimilhança é necessário um bom conhecimento da estrutura do modelo o que é outra desvantagem desse método.

### 4.3.2 Método da variável instrumental

Outra classe de métodos de compensação de *bias* nos parâmetros devido a presença de ruído nos dados de entrada e saída é conhecida como métodos das variáveis instrumentais. Os métodos baseados nesse princípio foram desenvolvidos principalmente para compensar o *bias* nos parâmetros quando o sinal de saída é contaminado por ruído colorido. Essa classe de métodos teve um desenvolvimento considerável no início da década de 1970 porque eles são de fácil implementação e geram estimativa cujas propriedades assintóticas são bem conhecidas. Inicialmente esses métodos foram desenvolvidos para a identificação de sistemas discretos, mas a sua extensão para sistemas contínuos é direta.

Relembrando que no caso do método dos momentos de Poisson o modelo do sistema é descrito pela equação (4.27), reescrita abaixo:

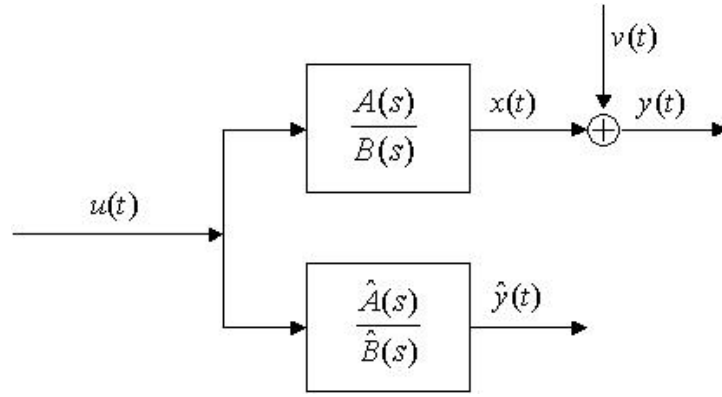
$$\mathbf{y}_k^{t_0}(l) = \mathbf{M}_k(l)\theta + \mathbf{w}_k^{t_0}(l) \quad (4.82)$$

Como o *bias* provocado pelo método dos mínimos quadrados é devido a correlação existente entre a matriz de regressores  $\mathbf{M}_k$  com o ruído, ou erro da equação,  $w(t)$ , o princípio do método da variável instrumental consiste em escolher ou gerar uma variável  $Z$ , chamada de variável instrumental, que atende as seguintes propriedades:

$$\mathcal{E}[Z(l)\mathbf{M}_k(l)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N Z(l)\mathbf{M}_k(l) \neq 0, \quad \text{é não singular} \quad (4.83)$$

$$\mathcal{E}[Z(l)\mathbf{w}_k^{t_0}(l)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N Z(l)\mathbf{w}_k^{t_0}(l) = 0 \quad (4.84)$$

Diversos tipos de variáveis instrumentais têm sido propostos na literatura de identificação de processos [Söderström e Stoica, 1983]. Nesse trabalho estudaremos o método da variável instrumental do modelo auxiliar (VIMA), método esse que foi primeiramente proposto em [Saha et al., 1991, a, b] e posteriormente estudado por [Garnier, 1995]. Esse princípio consiste em construir a variável instrumental,  $Z$ , a partir de um modelo auxiliar ou em paralelo conforme mostrado na figura 4.6, abaixo:



**Figura 4.6:** Método da variável instrumental baseado em modelo auxiliar (VIMA).

Os parâmetros do modelo auxiliar são, em um primeiro instante, estimados da mesma forma que o método da máxima verossimilhança, pelo método dos mínimos quadrados. O modelo gerado é então usado para gerar a predição  $\hat{y}(t)$ .

A variável instrumental,  $Z$ , é então construída da seguinte forma:

$$Z = [\mathbf{M}_{\hat{y}} \quad \mathbf{M}_u] \quad (4.85)$$

onde, conforme visto no capítulo 3:

$$\mathbf{M}_{\hat{y}} = -[\gamma(2:n, :) \hat{Y}_n^{t_0}] \quad (4.86)$$

$$\hat{Y}_n^{t_0} = [\hat{y}_0(t_0) \quad \hat{y}_1(t_0) \quad \cdots \quad \hat{y}_n(t_0)] \quad (4.87)$$

$$\mathbf{M}_u = [\gamma \mathcal{U}_n^{t_0}] \quad (4.88)$$

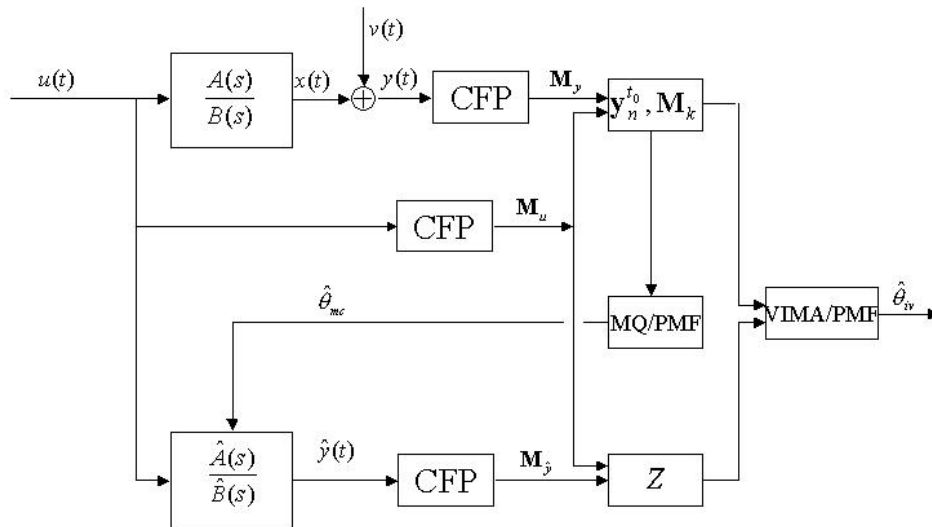
Dispondo de  $N$  observações dos sinais de entrada e saída do sistema podemos estimar os parâmetros do modelo de acordo com a seguinte expressão, que nada mais é do que a expressão do método dos mínimos quadrados onde uma das matrizes de regressores foi substituída por uma matriz com as variáveis instrumentais.

$$\hat{\theta} = \left[ \sum_{l=1}^N \mathbf{Z}(l) \mathbf{M}_k^T(l) \right]^{-1} \left[ \sum_{l=1}^N \mathbf{Z}(l) \mathbf{y}_k^{t_0}(l) \right] \quad (4.89)$$

O método da variável instrumental pode ser resumido pelas seguintes etapas:

- 1) Estimação do modelo auxiliar via o método dos mínimos quadrados;
- 2) Simulação do modelo auxiliar para gerar a variável instrumental  $Z$ ;

3) Estimação dos parâmetros via equação (4.89).



**Figura 4.7:** Diagrama de blocos representativo do método de variável instrumental combinado com o método de Poisson.

O método da variável instrumental é interessante para realizar a compensação de *bias* pois não é necessário o conhecimento de nenhuma informação sobre a natureza do ruído. Além disso esse é um método não recursivo ao contrário do método da máxima verossimilhança. O método da variável instrumental também pode ser expresso na forma recursiva conforme exposto em [Ljung, 1897]. Entretanto devemos observar que esse método, em comparação com o método dos mínimos quadrados, requer mais tempo de processamento uma vez que exige a utilização de duas seqüências de dados diferentes. O método da variável instrumental pode ser resumido pelo diagrama exposto na figura 4.7, acima:

### 4.3.3 Método de compensação de bias por pré-filtros

Outra técnica para a compensação do *bias* dos parâmetros estimados pelo método dos mínimos quadrados é através da utilização de pré-filtros apropriados nos sinais de entrada e/ou saída. Essa técnica é usualmente conhecida como compensação de *bias*. Ela consiste basicamente na determinação do valor do *bias* dos parâmetros e sua posterior eliminação do vetor de parâmetros [Sagara e Wada, 1977 e Stoica e Söderström, 1982]. Existem na literatura várias referências de diferentes variações desse método [Zheng e Feng, 1989/1990; Feng e Zeng 1991 e Nguyen et al., 1993a] para sistemas monovariáveis discretos e [Nguyen et al., 1993b] para sistemas multivariáveis.

A desvantagem desses métodos é que estes necessitam do conhecimento prévio de uma estimativa da variabilidade do ruído que contamina o sistema [Zhao et al., 1991 e Yang et al., 1993]. Existe, entretanto, outra abordagem que não apresenta essa restrição. A idéia central dessa abordagem é que um filtro aplicado no sinal de entrada do sistema pode fornecer informações a respeito do *bias* da estimação pelo método dos mínimos quadrados. Esse

método, inicialmente aplicado a sistemas discretos foi estendido aos sistemas contínuos por Zhao em [Zhao et al., 1992]. Garnier em sua tese [Garnier, 1995] analisa essa abordagem aplicada aos métodos dos momentos de Poisson.

Como visto anteriormente o *bias* dos parâmetros estimados é dados pela equação (4.7) reescrita abaixo:

$$\Delta_{\theta} = \mathbf{R}_{MM}^{-1} \mathbf{R}_{Mw} \quad (4.90)$$

Portanto para estimar o *bias* dos parâmetros estimados pelo método dos mínimos quadrados é necessário apenas o conhecimento de  $\mathbf{R}_{Mw}$ , a matriz de correlação entre o vetor de medidas, ou regressores, e o erro da equação, pois  $\mathbf{R}_{MM}$  é a matriz de covariância dos parâmetros do modelo. A matriz  $\mathbf{R}_{Mw}$  pode ser escrita da seguinte forma:

$$\mathbf{R}_{Mw} = \begin{bmatrix} \mathbf{R}_{yw}^T & \mathbf{R}_{uw}^T \end{bmatrix}^T \quad (4.91)$$

onde

$$\mathbf{R}_{yw} = \mathcal{E}[\mathbf{M}_y(l) \mathbf{w}_k^{t_0}(l)] \quad (4.92)$$

$$\mathbf{R}_{uw} = \mathcal{E}[\mathbf{M}_u(l) \mathbf{w}_k^{t_0}(l)] \quad (4.93)$$

Supondo que o ruído não apresente correlação com o sinal de entrada (sistema em malha aberta)  $\mathbf{R}_{Mw}$  pode ser reescrito da seguinte forma:

$$\mathbf{R}_{Mw} = \begin{bmatrix} \mathbf{R}_{yw}^T & | & 0 \end{bmatrix}^T \quad (4.94)$$

Um método para estima  $\mathbf{R}_{yw}$  é a aplicação de um filtro linear estável  $[F(p)]^{-1}$  contendo  $n$  raízes  $z_i$  ( $z_i < 0$ ,  $i=1,2, \dots, n$ ) no sinal de entrada do sistema. O filtro  $F(p)$  é definido por:

$$F(p) = \prod_{i=1}^n (p - z_i) = p^n + \sum_{i=1}^n f_i p^{n-i} \quad (4.95)$$

De agora em diante o símbolo “\*” será utilizado para representar as variáveis correspondentes ao modelo aumentado. O modelo aumentado pode ser então representado, desprezando as condições iniciais, por:

$$A(p)Y(p) = B^*(p)U^*(p) + A(p)V(p) \quad (4.96)$$

onde:

$$A(p) = \sum_{i=1}^n a_i p^{n-i}, \quad a_0 = 1 \quad (4.97)$$

$$B^*(p) = F(p)B(p) = \sum_{i=0}^{n+m} b_i^* p^{n+m-i} \quad (4.98)$$

$$U^*(p) = \frac{U(p)}{F(p)} \quad (4.99)$$

A equação diferencial correspondente ao modelo (4.96) é dada por:

$$\sum_{i=0}^n a_i \frac{d^{n-i} y(t)}{dt^{n-i}} = \sum_{i=0}^{n+m} b_i^* \frac{d^{n+m-i} u(t)}{dt^{n+m-i}} + \sum_{i=0}^n a_i \frac{d^{n-i} v(t)}{dt^{n-i}} \quad (4.100)$$

Aplicando a essa equação a transformação de Poisson de ordem mínima (que nesse caso é  $n+m$ ) com a introdução do filtro  $F(p)$  até o instante  $t_k$  ela pode ser representada de forma algébrica apropriada.

$$\mathbf{M}_{n+m} \{y(t)\}_{t_k}^T A = \mathbf{M}_{n+m} \{u^*(t)\}_{t_k}^T B^* + \mathbf{M}_{n+m} \{v(t)\}_{t_k}^T A \quad (4.101)$$

onde:

$$\mathbf{M}_{n+m} \{y(t)\}_{t_k}^T = \left[ M_{n+m} \left\{ \frac{d^n y(t)}{dt^n} \right\}_{t_k} \quad M_{n+m} \left\{ \frac{d^{n-1} y(t)}{dt^{n-1}} \right\}_{t_k} \quad \dots \quad M_{n+m} \{y(t)\}_{t_k} \right]^T \quad (4.102)$$

$$\mathbf{M}_{n+m} \{u^*(t)\}_{t_k}^T = \left[ M_{n+m} \left\{ \frac{d^{n+m} u^*(t)}{dt^{n+m}} \right\}_{t_k} \quad M_{n+m} \left\{ \frac{d^{n+m-1} u^*(t)}{dt^{n+m-1}} \right\}_{t_k} \quad \dots \quad M_{n+m} \{u^*(t)\}_{t_k} \right]^T \quad (4.103)$$

$$B^* = [b_0^* \quad b_1^* \quad \dots \quad b_{n+m}^*]^T \quad (4.104)$$

Como já anteriormente visto, a equação (4.101) pode ser escrita na forma de um regressor linear clássico:

$$\mathbf{y}_{n+m}^{t_k}(k) = \mathbf{M}_k(k) \theta^* + \mathbf{w}(k) \quad (4.105)$$

onde

$$\mathbf{y}_{n+m}^{t_k}(k) = M_{n+m} \left\{ \frac{d^n y(t)}{dt^n} \right\}_{t_k} = \gamma_n(1,:) Y_{n+m,n}^{t_k}(k) \quad (4.106)$$

$$\mathbf{M}_k(k) = \begin{bmatrix} [-\gamma_n(2:n+m,:)Y_{n+m,n}^{t_k}(k)]^T & [-\gamma_{n+m}U_{n+m,n+m}^{*t_k}(k)]^T \end{bmatrix} \quad (4.107)$$

$$\theta^* = \begin{bmatrix} A^T & B^{*T} \end{bmatrix}^T \quad (4.108)$$

$$Y_{n+m,n}^{t_k}(k) = \begin{bmatrix} M_m \{y(t)\}_{t_k} & M_{m+1} \{y(t)\}_{t_k} & \cdots & M_{m+n} \{y(t)\}_{t_k} \end{bmatrix}^T \quad (4.109)$$

$$U_{n+m,n}^{*t_k}(k) = \begin{bmatrix} M_0 \{u^*(t)\}_{t_k} & M_1 \{u^*(t)\}_{t_k} & \cdots & M_{m+n} \{u^*(t)\}_{t_k} \end{bmatrix}^T \quad (4.110)$$

$$\mathbf{w}(k) = M_{m+m} \{v(t)\}_{t_k}^T A(2:n+m) \quad (4.111)$$

Quando o número de observações é suficientemente grande, tendendo ao infinito, o vetor de parâmetros estimado pelo método dos mínimos quadrados,  $\hat{\theta}_{mq}^*$  do modelo aumentado tende assintoticamente para:

$$\lim_{N \rightarrow \infty} \hat{\theta}_{mq}^*(N) = \theta^* + \mathbf{R}_{M^*M^*}^{-1} \mathbf{R}_{M^*w} = \theta^* + \Delta_{\theta_{mq}^*} \quad (4.112)$$

Da mesma forma que a equação (4.94) podemos escrever  $\mathbf{R}_{M^*w}$  da seguinte forma:

$$\mathbf{R}_{M^*w} = G \mathbf{R}_{yw} \quad (4.113)$$

com

$$G = \begin{bmatrix} I_n & 0_{n+m+1,n} \end{bmatrix} \quad (4.114)$$

Agora, sabendo que  $z_i$  ( $i = 1, 2, \dots, n$ ) são as raízes do polinômio  $F(p)$  e que  $B^*(p)$  é dado pela equação (4.98) então podemos dizer que  $z_i$  ( $i = 1, 2, \dots, n$ ) também são raízes de  $B^*(p)$ , ou seja:

$$B^*(z_i) = F(z_i)B(z_i) = b_0^* z_i^{n+m} + b_1^* z_i^{n+m-1} + \cdots + b_{n+m-1}^* z_i + b_{n+m}^* = 0 \quad (4.115)$$

Essa equação pode ser reescrita na forma matricial de acordo com:

$$L^T \theta^* = 0 \quad (4.116)$$

onde  $L^T$  é dado por:



$$L^T = \begin{bmatrix} 0 & \cdots & 0 & z_1^{n+m} & \cdots & z_1 & 1 \\ 0 & \cdots & 0 & z_2^{n+m} & & z_2 & 1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & z_m^{n+m} & \cdots & z_m & 1 \end{bmatrix}_{2n+m+1, n} \quad (4.117)$$

Tendo essas considerações em mente, o termo de *bias*  $\Delta_{\theta^*}$  pode ser estimado através do seguinte raciocínio.

Se  $\hat{\theta}_{mq}^*(N)$  possui um *bias*, então a equação (4.116) não pode ser resolvida de forma que podemos escrever o seguinte:

$$L^T \hat{\theta}_{mq}^*(N) = \varepsilon(N) \quad (4.118)$$

Para um número finito de observações podemos calcular  $\varepsilon(N)$  multiplicando a equação (4.112) por  $L^T$  de forma que:

$$\varepsilon(N) = L^T \theta^* + L^T \hat{\mathbf{R}}_{M^*M^*}^{-1}(N) G \hat{\mathbf{R}}_{yw}(N) \quad (4.119)$$

Igualando as equações (4.118) e (4.119) podemos dizer que:

$$\hat{\mathbf{R}}_{yw}(N) = \left[ L^T \hat{\mathbf{R}}_{M^*M^*}^{-1}(N) G \right]^{-1} L^T \hat{\theta}_{mq}^*(N) \quad (4.120)$$

Assim, considerando as equações (4.90), (4.113) e (4.120) podemos concluir que o *bias* dos parâmetros pode ser calculado pela seguinte expressão:

$$\hat{\Delta}_{\theta_{mq}^*} = \hat{\mathbf{R}}_{M^*M^*}^{-1} G \left[ L^T \hat{\mathbf{R}}_{M^*M^*}^{-1}(N) G \right]^{-1} L^T \hat{\theta}_{mq}^*(N) \quad (4.121)$$

Os parâmetros do sistema sem *bias* são, portanto, dados por:

$$\hat{\theta}_{cb}^*(N) = \hat{\theta}_{mq}^*(N) - \hat{\Delta}_{\theta_{mq}^*}(N) \quad (4.122)$$

Os parâmetros do sistema original,  $\hat{\theta}_{cb}$ , podem ser obtidos através da divisão polinomial do vetor de parâmetros do sistema aumentado  $\hat{\theta}_{cb}^*$  pelo filtro  $F(p)$ .

A versão recursiva desse método é dada pelas seguintes equações [Garnier, 1995]:

$$\varepsilon^*(k) = y_k^{t_0} - \mathbf{M}_k^{*T}(k) \hat{\theta}_{mq}^*(k-1) \quad (4.123)$$

$$K^*(k) = P^*(k-1) \mathbf{M}_k^*(k) \left[ 1 + \mathbf{M}_k^{*T}(k) P^*(k-1) \mathbf{M}_k^*(k) \right]^{-1} \quad (4.124)$$

$$\hat{\theta}_{mq}^*(k) = \hat{\theta}_{mq}^*(k-1) + K^*(k)\varepsilon^*(k) \quad (4.125)$$

$$P^*(k) = P^*(k-1) - K^*(k)\mathbf{M}_k^*P^*(k-1) \quad (4.126)$$

$$\hat{\Delta}_{\theta_{cb}^*}^*(k) = P^*(k)G\left[L^T P^*(k)G\right]^{-1} L^T \theta_{mc}^*(k) \quad (4.127)$$

$$\alpha(k) = \alpha_0 \alpha(k-1) + (1 - \alpha_0) \quad (4.128)$$

$$\hat{\theta}_{cb}^*(k) = \hat{\theta}_{mc}^*(k) - \alpha(k)\hat{\Delta}_{\theta_{mq}^*}^*(k) \quad (4.129)$$

Um fator de ponderação  $\alpha$  é introduzido no algoritmo para evitar que a correção do *bias* seja totalmente levada em conta nas primeiras iterações para evitar oscilações dos valores dos parâmetros.

De forma geral podemos resumir o método de correção dos parâmetros estimados pelo método dos mínimos quadrados através das seguintes etapas:

- 1) Aplicação de um filtro linear estável no sinal de entrada do sistema;
- 2) Estimação dos parâmetros do sistema pelo método dos mínimos quadrados;
- 3) Estimação do *bias*,  $\hat{\Delta}_{\theta_{mq}^*}^*$  ;
- 4) Cálculo dos parâmetros livres de *bias* do modelo aumentado  $\hat{\theta}_{cb}^*$  ;
- 5) Cálculo dos parâmetros originais  $\hat{\theta}_{cb}$  ;

A condição necessária para que esse método funcione adequadamente é que a matriz  $\left[L^T \hat{\mathbf{R}}_{M^*M^*}^{-1}(N)G\right]$  seja não singular, isto é, que possa ser invertida. Para garantir isso é necessário satisfazer duas condições [Garnier, 1995]:

- 1) O filtro  $F(p)$  tenha  $n$  raízes estáveis (negativas) e distintas;
- 2) O sinal de entrada tenha movimentação suficiente para excitar o sistema na faixa de frequência apropriada.

A convergência desse método foi analisada em [Feng e Zeng, 1991] para o caso discreto podendo ser aplicado diretamente ao caso contínuo. Esse método possui como fator de síntese, além das constantes dos filtros da cadeia de filtros de Poisson, o pré-filtro aplicado no sinal de entrada do sistema. A escolha desse filtro representa um fator importante na eficiência alcançada por esse método. De forma geral deve-se escolher esse pré-filtro de forma que a sua largura de banda seja maior do que a do sistema a ser identificado [Garnier, 1995]. Isso implica na necessidade de um certo conhecimento do sistema a ser

identificado. No caso de não se possuir nenhuma informação uma variação desse método em os parâmetros do pré-filtro também são estimados foi desenvolvido em [Nguyen *et al.*, 1993] e em [Sibille *et al.*, 1994].

Esse método, como o método da variável instrumental apresenta a vantagem de não necessitar de iterações como no caso do método da máxima verossimilhança. Ele também é interessante pois é capaz de determinar diretamente o *bias* dos parâmetros, o que nos permite estimar a quantidade de distúrbio presente nos dados utilizados na identificação.

#### 4.4 Estudo comparativo dos métodos de compensação de *bias*

Para finalizar o estudo dos métodos de compensação de *bias* iremos aplicá-los a um sistema descrito pela seguinte função de transferência [Garnier, 1995]:

$$\frac{X(s)}{U(s)} = \frac{2}{s^2 + 3s + 2} \quad (4.130)$$

Dos métodos abordados analisaremos apenas o método da variável instrumental e da compensação de *bias* pelo uso de pré-filtros. O método da máxima verossimilhança não será abordado pois a implementação deste elaborada no curso desse trabalho apresentou os problemas já descritos na seção 4.3, como geração de modelos oscilatórios e forte sensibilidade quanto a ordem do modelo a ser estimado.

As condições experimentais empregadas nesse teste foram as seguintes:

- 1) O sinal de entrada é um PRBS variando entre  $-1$  e  $+1$ ;
- 2) As condições iniciais são todas nulas;
- 3) O período de amostragem usado é  $0,02$ ;
- 4) Os parâmetros da cadeia de filtros de Poisson utilizados foram  $\lambda = \beta = 2$ ;
- 5) O número de pontos amostrados foi de  $5000$ ;
- 6) Ao sinal determinístico  $x(t)$  foi adicionado um ruído. Foram testadas tanto a adição de ruído branco como de ruído colorido. O ruído colorido foi gerado através da filtragem do ruído branco com o seguinte filtro linear discreto.

$$v(k) = \frac{1 + 3,98q^{-1} - 0,8616q^{-2}}{1 - 0,6195q^{-1} + 0,09499q^{-2}} w(k) \quad (4.131)$$

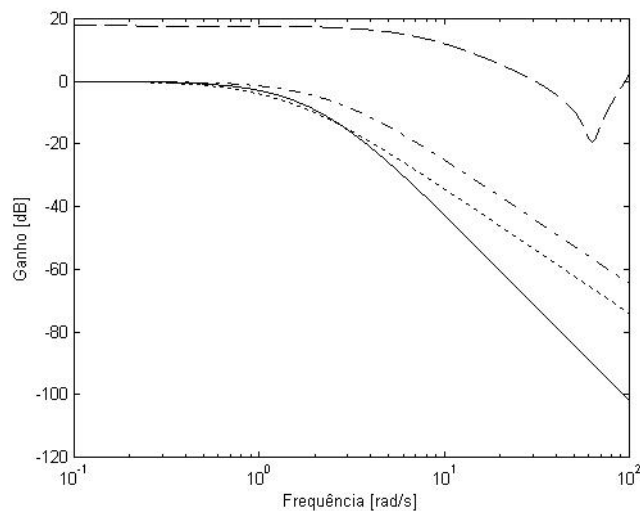
O modelo de geração de ruído foi escolhido de forma que sua largura de banda seja maior do que a do sistema a ser identificado, isto é, o ruído afeta o sistema também em frequências médias e altas. Ambos os tipos de ruídos foram adicionados de forma crescente aos dados determinísticos variando de a razão ruído sobre sinal (RRS) de 0% a 70%.

Os parâmetros do sistema foram estimados usando o método dos mínimos quadrado (MQ), Variável Instrumental (VI) e da compensação por pré-filtro (PF). No método de compensação por pré-filtro foi utilizado o seguinte pré-filtro:

$$F(p) = \frac{6}{(p+3)(p+2)} \quad (4.132)$$

As rotinas utilizadas nesse teste foram implementadas utilizando o MATLAB 5.3 e fazem parte do *Continuous Time System Identification Toolbox* (CONTSID), [Garnier e Mensler, 1999].

Os diagramas de Bode (apenas magnitude) do sistema a ser identificado, do modelo tipo ARMA gerador do ruído colorido, do filtro  $F(p)$  e da cadeia de Poisson de ordem 2 são apresentados juntos na figura 4.8, abaixo:



**Figura 4.8:** Magnitude do sistema - linha pontilhada, do modelo ARMA de geração de ruído - linha tracejada, do pré-filtro  $F(p)$  - linha ponto-tracejada e da cadeia de filtros de Poisson - linha contínua.

Os resultado das estimações para diferentes níveis de ruído são apresentados na tabela 4.1 para o caso de ruído branco e na tabela 4.2 para o caso de ruído colorido. Nessa tabelas são apresentados os valores dos parâmetros estimados em cada nível de ruído bem como os pólos correspondentes. Para facilitar a comparação dos resultados utilizou-se dois índices. O primeiro foi a norma quadrática do erro dos parâmetros estimados em relação aos parâmetros do sistema usado na simulação dos dados, chamado aqui de  $NQE\theta$ . Esse índice já foi

discutido no capítulo 3. O outro índice utilizado foi o erro percentual da área da resposta no domínio do tempo a um degrau na entrada que é dado por:

$$NQED[\%] = 100 \frac{\|A_S - A_M\|_2}{\|A_S\|_2} \quad (4.133)$$

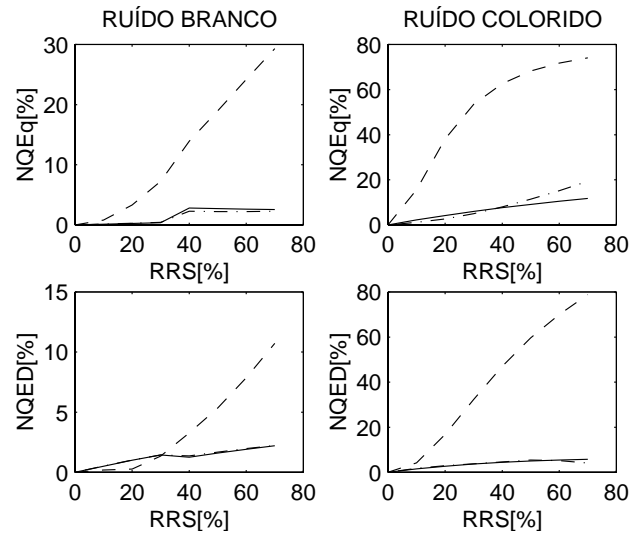
onde  $A_S$  é a área da resposta ao degrau do sistema real e  $A_M$  a área da resposta ao degrau do modelo identificado.

Para facilitar a interpretação dos dados estes são apresentados em forma gráfica na figura 4.9.

É interessante também analisarmos a variação desses índices em função da constante da cadeia e filtros de Poisson. Para isso mantivemos um nível de razão ruído sobre sinal de 30% e variamos a constante  $\lambda$  de 0,5 até 10. Os resultados dessa análise são apresentados na figura 4.10.

**Tabela 4.1:** Resultados da identificação com dados contaminados com ruído branco.

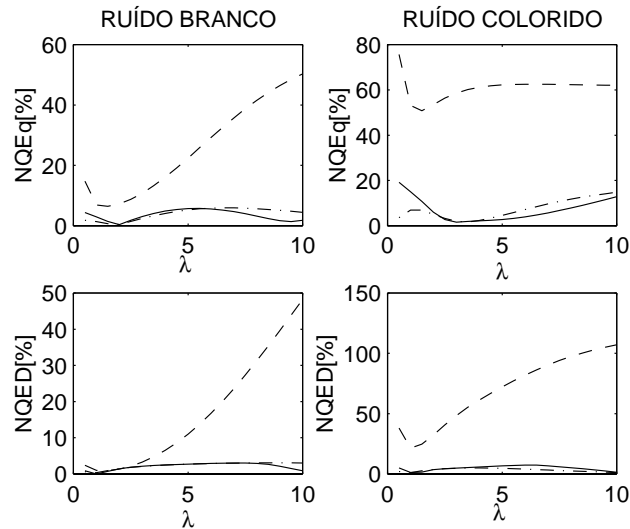
	RRS[%]	$a_1 (3)$	$a_2 (2)$	$b_0 (2)$	$p_1 (-2)$	$p_2 (-1)$	$NQE\theta(\%)$	$NQED(\%)$
MQ	0	3,0001	2,0000	2,0000	-2,0001	-0,9999	0,0015	0,0042
VI		3,0001	2,0000	2,0000	-2,0001	-0,9999	0,0015	0,0042
CB		3,0000	2,0000	2,0000	-2,0000	-1,0000	0,0000	0,0000
MQ	10	2,9728	1,9868	1,9865	-1,9583	-1,0146	0,8032	0,1860
VI		3,0022	2,0028	2,0052	-2,0016	-1,0006	0,1533	0,5157
CB		3,0011	2,0023	2,0047	-2,0000	-1,0011	0,1288	0,5076
MQ	20	2,8916	1,9441	1,9384	-1,8282	-1,0634	3,3134	0,2685
VI		3,0040	2,0054	2,0103	-2,0026	-1,0014	0,2971	1,0160
CB		3,0022	2,0045	2,0092	-1,9999	-1,0023	0,2554	1,0017
MQ	30	2,7655	1,8770	1,8615	-1,5696	-1,1959	7,2488	1,3489
VI		3,0052	2,0077	2,0150	-2,0026	-1,0025	0,4270	1,4459
CB		3,0033	2,0068	2,0138	-1,9997	-1,0035	0,3819	1,4345
MQ	40	2,5614	1,7521	1,7251	-1,28-0,33i	-1,28-0,33i	13,9201	3,2720
VI		2,9305	1,9508	1,9602	-1,9081	-1,0224	2,2793	1,3587
CB		2,9132	1,9424	1,9602	-1,8801	-1,0331	2,8039	1,2504
MQ	50	2,4002	1,6672	1,6257	-1,20+0,47i	-1,20-0,47i	18,9519	5,3441
VI		2,9302	1,9525	1,9640	-1,9056	-1,0246	2,2262	1,6791
CB		2,9142	1,9445	1,9542	-1,8797	-1,0345	2,7160	1,6028
MQ	60	2,2331	1,5793	1,5224	-1,12+0,58i	-1,12-0,58i	24,1706	7,8552
VI		2,9285	1,9534	1,9669	-1,9007	-1,0277	2,2210	1,9682
CB		2,9151	1,9466	1,9586	-1,8792	-1,0359	2,6324	1,9176
MQ	70	2,0688	1,4934	1,4208	-1,03+0,65i	-1,03-0,65i	29,2981	10,7260
VI		2,9249	1,9535	1,9687	-1,8929	-1,0320	2,2726	2,2299
CB		2,9159	1,9487	1,9629	-1,8786	-1,0373	2,5535	2,2013



**Figura 4.9:** Variação dos índices de performance em função do grau de ruído ( $\lambda=\beta=2$ ). Linha tracejada, mínimos quadrados, linha ponto-tracejada, variável instrumental, linha contínua, compensação por pré-filtro. Primeira linha – Norma quadrática do erro dos parâmetros, segunda linha – Norma do erro da resposta ao degrau.

**Tabela 4.2:** Resultados da identificação com dados contaminados com ruído colorido.

	RRS[%]	$a_1 (3)$	$a_2 (2)$	$b_0 (2)$	$p_1 (-2)$	$p_2 (-1)$	$NQE\theta(\%)$	$NQED(\%)$
MQ	0	3,0001	2,0000	2,0000	-2,0001	-0,9999	0,0015	0,0000
VI		3,0001	2,0000	2,0000	-2,0001	-0,9999	0,0015	0,0000
CB		3,0000	2,0000	2,0000	-2,0000	-1,0000	0,0000	0,0000
MQ	10	2,5014	1,7319	1,6957	-1.25-0.41i	-1.25+0.41i	15,5875	4,1572
VI		2,9570	1,9799	1,9881	-1,9325	-1,0246	1,1863	1,7590
CB		2,9229	1,9631	1,9673	-1,8771	-1,0458	2,2196	1,4768
MQ	20	1,7829	1,3499	1,2477	-0.89-0.74i	-0.89+0.74i	38,1160	16,8013
VI		2,9040	1,9561	1,9702	-1,8420	-1,0620	2,6614	2,8900
CB		2,8538	1,9303	1,9386	-1,7520	-1,1017	4,2024	2,6909
MQ	30	1,2896	1,0978	0,9437	-0.64-0.83i	-0.64+0.83i	53,4407	32,2805
VI		2,8215	1,9196	1,9345	-0,6448	-0,6448	5,0079	3,7668
CB		2,7909	1,9006	1,9345	-1,6764	-1,1450	5,9957	3,6523
MQ	40	0,9870	0,9552	0,7632	-0.493-0.844i	-0.493+0.844i	62,6536	47,0484
VI		2,7154	1,8744	1,8851	-1.36-0.176i	-1.36+0.176i	8,0430	4,6257
CB		2,7332	1,8737	1,8906	-1.36-0.078i	-1.36+0.078i	7,6360	4,4195
MQ	50	0,7985	0,8784	0,6573	-0.39- 0.84i	-0.39+ 0.84i	68,2010	59,6133
VI		2,5967	1,8262	1,8285	-1.29-0.37i	-1.29+0.37i	11,4340	5,4348
CB		2,6806	1,8493	1,8705	-1.34-0.23i	-1.34+0.23i	9,1237	5,0149
MQ	60	0,6735	0,8388	0,5934	-0.33-0.85i	-0.33+0.85i	71,7001	70,0025
VI		2,4673	1,7742	1,7647	-1.23-0.50i	-1.23+0.50i	15,1480	5,1432
CB		2,6324	1,8272	1,8527	-1.31-0.31i	-1.31+0.31i	10,4793	5,4627
MQ	70	0,5839	0,8208	0,5533	-0.29-0.85i	-0.29+0.85i	74,0475	78,8028
VI		2,3233	1,7151	1,6903	-1.16-0.60i	-1.16+0.60i	19,3260	4,1091
CB		2,5873	1,8066	1,8366	-1.29-0.36i	-1.29+0.36i	11,7434	5,7865



**Figura 4.10:** Variação dos índices de performance em função da constante da cadeia de filtros e Poisson  $\lambda$ , (RRS = 30%). Linha tracejada, mínimos quadrados, linha ponto-tracejada, variável instrumental, linha contínua, compensação por pré-filtro. Primeira linha – Norma quadrática do erro dos parâmetros, segunda linha – Norma do erro da resposta ao degrau.

Examinando os resultados apresentados na tabela 4.1 e na figura 4.9 podemos verificar que o desempenho dos métodos dos mínimos quadrados se deteriora rapidamente a medida que a quantidade de ruído presente nos dados aumenta. Esses resultados comprovam, como dito anteriormente que o método dos mínimos quadrados associado com os métodos dos momentos de Poisson fornece boas estimativas dos parâmetros quando a quantidade de ruído é baixa ( $< 20\%$ ) e esse ruído é do tipo branco. Isso se deve ao efeito de filtragem dos dados promovida pelo método de Poisson que atenua o ruído presente nos dados. Para níveis superiores de ruído, entretanto, os métodos de compensação de *bias* se tornam fundamentais. Também é possível verificar que o método de compensação de *bias* por pré-filtro das entradas apresenta um performance um pouco superior ao método da variável instrumental. Esses dois métodos permitem, como demonstrado, obter boas estimativas mesmo na presença de quantidades significativas de ruídos nos dados.

Quando o ruído é do tipo colorido a performance do método dos mínimos quadrados acoplado com o método dos momentos de Poisson se deteriora muito mais rapidamente do que no caso do ruído branco. Nessa situação esse método fornece estimativas equivalentes aos métodos de compensação de *bias* somente para níveis de ruído muito baixo (RRS $<10\%$ ). Isto considerando apenas o índice *NQED*. Considerando o índice *NQE $\theta$*  vemos que o método de compensação de *bias* por pré-filtro é superior ao método da variável instrumental para níveis de ruído superior a 40%. Isso se deve provavelmente ao fato que o modelo auxiliar do método da variável instrumental, estimado pelo método dos mínimos quadrados, possui um *bias* muito grande.

Analisando a figura 4.10 podemos analisar a influência da constante  $\lambda$  nos dois métodos de compensação de *bias*. Como no caso do método dos mínimos quadrados, o

---

desempenho dos dois métodos estudados para a compensação de *bias*, depende fortemente da constante  $\lambda$ . Considerando apenas o índice  $NQE\theta$  vemos que a sua influência, tanto no caso do método dos mínimos quadrados como no caso dos métodos de compensação de *bias* e para ambos tipos de ruído é similar. Agora, considerando o índice  $NQED$ , vemos que os métodos de compensação de *bias* apresentam menor sensibilidade do que o método dos mínimos quadrados tanto para o caso de ruído branco, como para o caso de ruído colorido. De uma forma geral o valor da constante  $\lambda$  deve ser escolhido de forma a ser levemente superior a frequência de corte do sistema,  $\omega_1 = 1$  rad/s e  $\omega_2 = 2$  rad/s para o presente sistema [Garnier, 1995].



## 4.5 Referências Bibliográficas

- Unbehauen, H. e Rao, G. P., “Identification of Continuous Systems”, North-Holland Systems and Control Series, V.10, pp 167-199 (1987).
- Ljung, L., “System Identification: Theory for the User”, Prentice Hall: Englewood Cliffs, NJ (1987).
- Garnier, H., “Identification de modeles parametriques continus par moments de Poisson”, Tese de doutorado, Universidade Henri Poincare, Nancy 1, França, (1995).
- Garnier, H., Sibille P., Nguyen H.L., Spott T., “A bias-compensating least-squares method for continuous-time system identification via Poisson moments functionals”, 10<sup>th</sup> IFAC Symposium on Identification and System Parameter Estimation, Copenhagen-Denmark, pp. 3675-3680, (1994).
- Sagara S., Wada K., “On line modified least-squares parameters estimation of linear discrete dynamic systems”, Int.J. Control, Vol. 25, n° 3, pp. 329-343, (1997).
- Stoica P., Soderstrom, T., “Bias correction in least-squares identification”, Int. J. Control, Vol. 25, n°3, pp. 449-457,(1892).
- Zheng W. X., Feng C. B., “Unbiased parameter estimation of linear systems in the presence of input and output noise”, Int. J. of Adaptive Control and Signal Processing, Vol. 3, pp. 231-251, (1989).
- Zheng W.X., Feng C.B., “Identification of stochastic time lag systems in the presence of colored noise”, Automatica, Vol. 26, n° 4, pp. 769-779, (1990).
- Feng C.B., Zheng, W.X., “Robust identification of stochastic linear systems with correlated noise”, Proc. of the IEE, Vol. 138, n° 5, pp. 484-492, (1991).
- Nguyen H. L., Sibille P., Garnier H., “A new bias-compensating least-squares method for identification of stochastic linear system in the presence of colored noise”, Conference on Decision and Control, San Antonio – USA, pp. 2038-2043, (1993).
- Nguyen H. L., Sibille P., Garnier H., “A modified on-line least-squares method for identification of linear multi-variable systems”. 2<sup>nd</sup> European Control Conference, Growingen – The Netherlands, Vol. 3, pp 1330-1335, (1993).
- Yang Z.J., Sagara S., Wada, K., “Identification of continuous-time systems from sampled input-output data using bias eliminating techniques”. Control Theory and Advanced Technology, Vol. 9, n° 1, pp. 53-75, (1993).
- Zhao, Z.Y. Sagara S., Wada K., “Bias-compensating least-squares method for identification of continuous-time systems from sampled data”. Int. J. Control, Vol. 53, n° 2, pp.445-461, (1991).

- Zhao, Z.Y. Sagara S., Tomizuka M., “A new bias-compensating LS method for continuous system identification in the presence of coloured noise”, *Int. J. Control*, Vol. 56, n° 6, pp. 1441-1452, (1992).
- Sibille P., Nguyen H.L., Garnier H., Richard A., “A new on-line bias compensating least-square method for identification of linear discrete-time systems”, 10<sup>th</sup> IFAC Symposium on Identification and System Parameter Estimation, Copenhagen-Denmark, pp. 3429-3434, (1994).
- Saha, D.C., Bapat, V.N., Roy, B.K., “The Poisson moment functional technique – some new results”, dans “Identification of continuous-time systems – Methodology and computer implementation”, N.K Sinha & G.P.Rao, Dordrecht, pp. 327-361, 1991a.
- Saha, D.C., Roy, B.K, Bapat, V.N., “Some generalizations and further uses of the Poisson Moment Functionals (PMF) approach to continuous-time model estimation, 9<sup>th</sup> IFAC Symposium on Identification and System Parameter Estimation, Budapest – Hungary, pp. 1324-1333, 1991b.
- Garnier, H. e Mensler, M., “CONTSID: a continuous-time system identification toolbox for MATLAB”, 5th European Control Conference, Karlsruhe (Germany), (1999).

## Capítulo 5

### Conclusões e sugestões de trabalhos futuros

Para finalizar essa dissertação, será resumido nesse capítulo as principais conclusões obtidas nos capítulos anteriores. Na seqüência serão apresentadas sugestões e temas para trabalhos futuros na área de identificação de modelos no espaço de tempo contínuo.

#### 5.1 Conclusões

O objetivo principal desse trabalho foi a análise detalhada do método dos momentos de Poisson para a identificação de modelos lineares no espaço contínuo. Para isso primeiramente introduzimos no capítulo 2 os conceitos básicos de identificação linear. Nesse capítulo analisamos inicialmente alguns métodos de identificação de modelos lineares discretos uma vez que estes são os mais amplamente empregados. O fato de os modelos de identificação discretos serem mais amplamente empregados se deve principalmente ao fato da reduzida carga computacional requerida para a sua utilização. Com o advento dos computadores digitais, e uma vez que os primeiros computadores comerciais possuíam uma limitada capacidade de processamento, os métodos discretos foram preferencialmente utilizados durante as décadas de 1970 e 1980.

Em seguida apresentamos o conceito de identificação de modelos contínuos e abordamos o problema da necessidade de se utilizar derivadas de sinais para identificação de modelos desse tipo. Foram apresentados de maneira sintética vários métodos para contornar esse problema como por exemplo métodos de funções de modulação, filtro lineares e análise espectral dos sinais de entrada e saída do sistema. Nesse capítulo, como conclusão, analisamos os principais problemas da identificação de modelos discretos que é a questão do tempo de amostragem dos dados. No Apêndice B é apresentado de forma clara o problema decorrente de tempos de amostragem muito grandes denominado efeito *aliasing*. Vimos

também que a utilização de tempos de amostragem muito pequenos leva a erros numéricos devido ao fato de que os parâmetros de modelos discretos são função do tempo de amostragem dos dados.

No capítulo 3 foi estudado especificamente o método dos momentos de Poisson para a identificação de modelos lineares contínuos. Primeiramente foi estudada a implementação do cálculo dos momentos de Poisson dos sinais de entrada e saída e o cálculo dos momentos de Poisson das derivadas dos sinais em função dos momentos de diferentes ordens do próprio sinal. Foi demonstrado que a aplicação da transformação de Poisson a uma equação diferencial ordinária a transforma em uma equação algébrica com os mesmos parâmetros da equação original.

Foi estudado também a influência das condições iniciais dos sinais de entrada e saída no processo de identificação. Vimos que é possível eliminá-la através do corte das primeiras amostragens dos sinais de entrada e saída. Similarmente demonstramos ser possível utilizar dados contendo *off-set* na identificação através do método de Poisson, bastando para isso adicionar um novo parâmetro ao vetor de parâmetros e uma coluna cujos valores são todos iguais a um na matriz de regressores.

Foi analisado, de forma empírica, a influência do valor da constante da cadeia de filtros de Poisson no resultado da identificação, análise essa que no capítulo 4 foi feita de forma analítica para sistemas de primeira ordem com ruído branco. Essa análise mostrou que para sistemas desse tipo o melhor resultado é obtido quando se utiliza o valor da constante  $\lambda$  igual ao valor do pólo do sistema. Para sistema mais complexos vimos que, apesar de o valor de  $\lambda$  ter uma influência importante sobre o desempenho do método de Poisson, a correlação com a dinâmica do sistema não é tão simples.

Ainda no capítulo 3 estudamos também uma versão iterativa do método de Poisson que permite a atualização da estimação dos parâmetros do modelo a medida que um novo conjunto de medidas dos sinais de entrada e saída é obtido. Nesse capítulo também realizamos uma comparação entre a identificação contínua e a identificação discreta comparando o método de Poisson e o método ARX. Nesse exemplo o método de Poisson se mostrou superior ao método de identificação ARX quando há a presença de ruído nos sinais do sistema.

Finalmente comparamos o método dos momentos de Poisson com um outro método de identificação de modelos contínuos proposto por Johansson demonstrando serem ambos os métodos idênticos.

Até o capítulo 3 lidamos apenas com a identificação de modelos lineares sem levar em consideração a presença de distúrbios nos dados de entrada e saída. Para esse tipo de identificação utilizamos apenas o método dos mínimos quadrados. No capítulo 4 demonstramos que esse método produz estimativas errôneas quando há a presença de ruído nos dados de entrada e saída.

Para compensar esse desvios apresentamos três técnicas encontradas na literatura. A primeira e mais comum, a da máxima verossimilhança, é a mais amplamente empregada. Exemplos do seu emprego podem ser encontrados no *Identification Toolbox* do MATLAB onde os modelos tipo ARMAX e BJ a utilizam variando apenas a estrutura do modelo identificado. Essa técnica apresenta um custo computacional elevado visto que é iterativa e necessita do cálculo do gradiente e da hessiana da função objetivo em cada paço.

A segunda técnica apresentada foi o método da variável instrumental. Essa técnica também está implementada no *Identification Toolbox*. Esta técnica necessita de menos capacidade computacional uma vez que usa o método dos mínimos quadrados para obter uma modelo auxiliar, simula esse modelo auxiliar para obter uma variável, chamada na literatura de instrumental, e por fim, mais um último passo algébrico similar ao método dos mínimos quadrados para a identificação do modelo final.

A terceira técnica apresentada é a técnica de compensação de *bias* pela pré-filtragem do sinal de entrada por um filtro linear. Essa é a técnica que apresenta quase a mesma necessidade computacional que o método da variável instrumental. Entretanto o aspecto mais interessante desta é que ela permite o cálculo explícito do erro (*bias*) de cada parâmetro.

Finalmente comparamos as últimas duas técnicas de compensação de *bias* através de um exemplo simulado. Demonstramos com esse exemplo que a performance da técnica de compensação de *bias* por pré-filtro é levemente superior a da variável instrumental.

A identificação de modelos lineares no espaço de tempo contínuo tem recebido um grande interesse nos últimos dez anos e o método dos momentos de Poisson é um dos mais elegantes e eficientes e tem sido amplamente estudado em diferentes aplicações. A identificação de modelos contínuos faz muito mais sentido para o engenheiro químico pois a maiorias dos sistemas com os quais lidamos são contínuos. Além disso, como já comentado, o aumento alucinante que temos presenciado na capacidade dos computadores disponíveis no mercado elimina a necessidade de grande quantidade de cálculos como uma restrição para qualquer método numérico a ser aplicado na prática.

## 5.2 Sugestões

A seguir listaremos algumas sugestões de assuntos que por falta de tempo não foram abordados nesse trabalho.

- 1) **Identificação com restrições:** Uma característica interessante dos modelos no domínio de tempo contínuo é que os seu parâmetros geralmente estão relacionados com características conhecidas. Dessa forma seria interessante se pudéssemos realizar um identificação utilizando informações já disponíveis sobre o sistema, como por exemplo, limites dentro dos quais certos parâmetros devem ser mantidos. Uma característica particular dos modelos lineares é que características dinâmicas como resposta inversa e *overshoot* estão relacionadas com a posição relativa de pólos e zeros do sistema e estes por sua vez estão diretamente relacionados com os parâmetros do modelo. Assim também seria possível forçar, na identificação, um determinado comportamento conhecido sobre o sistema fixando restrições quanto a posição relativa dos pólos e zeros do sistema.
- 2) **Identificação de modelos variantes no tempo:** Uma característica sobre a transformação de Poisson que não foi abordada nesse trabalho é que ela também é aplicável a sistemas variantes no tempo em que os parâmetros de um equação diferencial ordinária são funções polinomiais do tempo. As primeiras referências sobre o método de Poisson abordam essa propriedade, entretanto pouco tem se visto nos último anos nesse sentido. Não se tem nenhuma referência de funções desenvolvidas para a identificação de modelos variantes no tempo utilizando o método de Poisson.
- 3) **Identificação de sistema distribuídos:** Outra característica interessante do método de Poisson é que ele também é aplicável a equações diferenciais parciais lineares. É possível portanto aplicá-lo na identificação de sistema variantes não só no tempo mas também no espaço.
- 4) **Identificação de sistema em malha fechada:** Esse assunto é com certeza um dos grandes objetivos da identificação de sistemas. Aqueles que trabalham na industria sabem que nem sempre é possível abrir uma malha de controle para a realização de testes a fim de obter dados para a identificação de um determinado sistema. Algumas vezes não podemos abrir as malhas porque o sistema é instável em malha aberta, outras vezes porque não podemos tirar determinada variável controlada de posição a fim de evitar tirar o produto de especificação. Portanto um método que permita identificar os sistema sem a necessidade de abrir as malhas e utilizando dados naturais de processo seria de grande valor prático. Muito se tem feito nessa área, mas pouco coisa combinada com identificação no espaço de tempo contínuo e, apesar do muito já feito, ainda não se obteve um método confiável e prático com os já disponíveis para identificação em malha aberta.

- 5) **Identificação de tempo morto:** Um outro assunto interessante na identificação de sistemas é a determinação de tempo morto. A maioria das técnicas convencionais de identificação necessitam que o tempo morto do sistema seja informado previamente a identificação. O desenvolvimento de um método eficiente de identificação do tempo morto seria de grande utilidade prática uma vez que nem sempre é possível inferi-lo manualmente.
- 6) **Análise teórica do valor ótimo de  $\lambda$ :** Como vimos nesse trabalho, o valor da constante  $\lambda$  tem grande influência no desempenho do método dos momentos de Poisson. Entretanto, com exceção de sistemas de primeira ordem com ruído branco, não existe uma correlação entre as características do sistema e o valor ótimo dessa constante. A pesquisa desse assunto se mostra como um excelente trabalho de pesquisa teórica.
- 7) **Identificação multivariável com perturbações simultâneas:** Esse é um assunto de grande interesse prático, uma vez que a maioria dos sistemas reais são multivariáveis e as perturbações normalmente utilizadas para a sua identificação são simultâneas, isto é, mais de uma variável de entrada tem o seu valor manipulado ao mesmo tempo.

# Apêndice A

## Cálculo do gradiente de $y$ predito para modelo ARMAX

Considere o modelo ARMAX dado pela seguinte equação:

$$C(q)\hat{y}(t/\theta) = B(q)u(t) + [C(q) - A(q)]y(t) \quad (\text{A.1})$$

onde o vetor de parâmetros  $\theta$  é dado por:

$$\theta = [a_k \quad b_k \quad c_k]$$

Diferenciando a equação (A.1) em relação a  $a_k$  temos:

$$C(q)\frac{\partial}{\partial a_k}\hat{y}(t/\theta) = -q^{-k}y(t) \quad (\text{A.2})$$

da mesma forma, diferenciando a equação (A.1) em relação a  $b_k$  e  $c_k$  temos, respectivamente:

$$C(q)\frac{\partial}{\partial b_k}\hat{y}(t/\theta) = -q^{-k}u(t) \quad (\text{A.3})$$

e

$$q^{-k}\hat{y}(t/\theta) + C(q)\frac{\partial}{\partial c_k}\hat{y}(t/\theta) = q^{-k}y(t) \quad (\text{A.4})$$

Com o vetor  $\varphi(t/\theta)$  dado pela equação (2.57) reescrita abaixo:

$$\varphi(t) = [-y(t-1) \quad \dots \quad -y(t-n) \quad u(t-1) \quad \dots \quad u(t-m)] \quad (\text{A.5})$$

podemos mostrar que:

$$C(q)\psi(t, \theta) = \varphi(t, \theta) \quad (\text{A.6})$$

que pode ser reescrita, caso  $C(q)$  seja inversível, da seguinte forma:

$$\psi(t, \theta) = \frac{1}{C(q)}\varphi(t, \theta) \quad (\text{A.7})$$



o que demonstra a equação (2.64), onde  $\psi(t/\theta)$  é dado por:

$$\psi(t/\theta) = \left[ \frac{\partial}{\partial a_k} \hat{y}(t/\theta) \quad \frac{\partial}{\partial b_k} \hat{y}(t/\theta) \quad \frac{\partial}{\partial c_k} \hat{y}(t/\theta) \right] \quad (\text{A.8})$$

# Apêndice B

## Efeito *Aliasing*

O teorema de Nyquist nos diz que um sinal periódico só pode ser reconstruído a partir de amostragens somente quando o tempo de amostragem é menor do que a metade do seu período, ou seja:

$$T_s < \frac{P}{2}$$

O efeito *aliasing* é o efeito de distorção provocado pela amostragem inadequada do sinal. Um exemplo clássico desse efeito é a ilusão que temos quando vemos a roda de um carro em movimento em um filme. Temos a impressão que a roda se move no sentido contrário. Isso é causado pelo efeito *aliasing* pois a velocidade circular da roda é muito maior do que a frequência com que as cenas de um filme normal são gravadas. Para evitar essa ilusão seria necessária a utilização de filmes de alta velocidade. Quanto maior a velocidade do veículo maior a velocidade do filme.

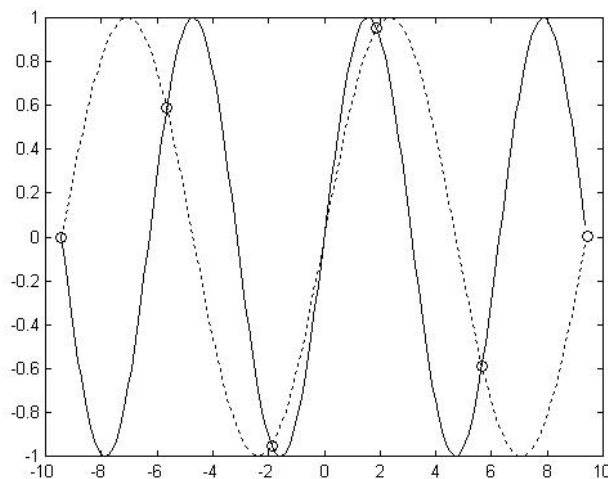
Por exemplo, tomemos a função

$$y = \sin(x)$$

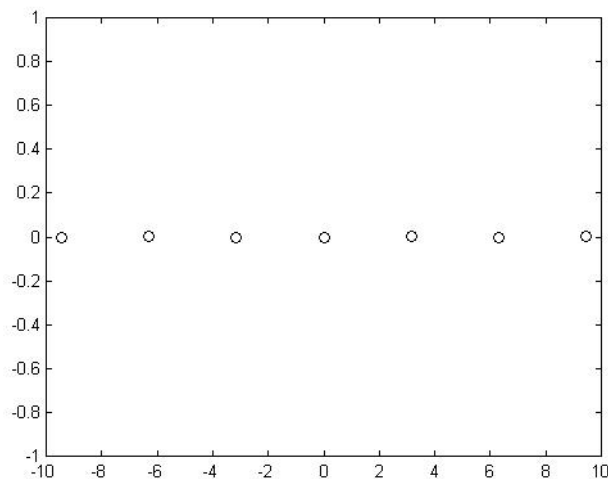
cujo período é  $2\pi$ . Portanto, segundo o teorema de Nyquist, para que possamos reconstruir essa função de forma única devemos ter um tempo de amostragem menor do que  $\pi$ .

Se tomarmos um tempo de amostragem igual a  $1,2\pi$  teremos uma amostragem inferior ao mínimo necessário. Se tomarmos essas amostras e tentarmos ajustar uma curva do tipo  $y = \sin(ax)$  obteremos infinitas soluções para  $a$ . Duas soluções possíveis são  $a = 1$  e  $a = 2/3$ , como podemos ver na figura B.1 , abaixo.

Caso tomemos o tempo de amostragem como sendo exatamente igual  $\pi$  teremos o caso em que os ponto amostrados possuem a mesma magnitude, como mostrado na figura 2.2, abaixo. Vemos que podemos dizer que a função que esses pontos representam é  $y = 0$ .

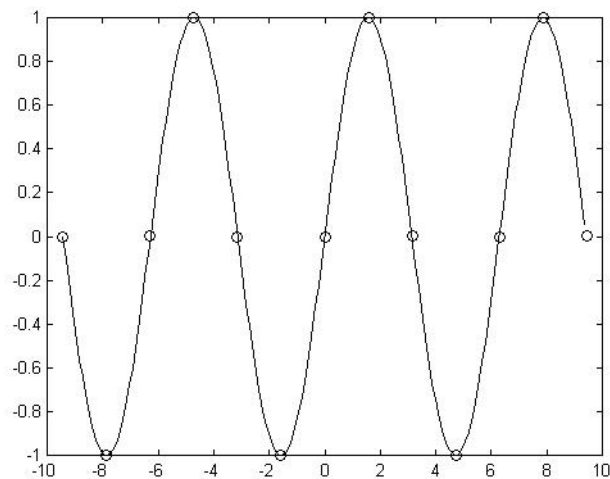


**Figura B.1:** Diferentes funções do tipo  $y = \sin(ax)$  ajustadas a amostragem da função  $y = \sin(x)$  com tempo de amostragem igual a  $T_S = 1,2\pi$ . Linha pontilhada  $a = 2/3$ , linha contínua  $a = 1$ .



**Figura B.2:** Amostragem da função  $y = \sin(x)$  quando o tempo de amostragem é exatamente igual a metade do seu período, ou seja,  $T_S = \pi$ .

Agora, se tomarmos um tempo de amostragem adequado, no presente caso isto significa  $T_S < \pi$ , a única função capaz de se ajustar a esse dados é a própria  $y = \sin(x)$ , com o pode ser visto na figura 2.3, abaixo, na qual utilizamos um tempo e amostragem igual a  $\pi/2$ .



**Figura B.3:** Amostragem da função  $y = \sin(x)$  quando o tempo de amostragem é igual a ,  $T_S = \pi/2$ .

No caso da identificação de sistemas o efeito *aliasing* provoca a perda de informações das componentes de frequência mais altas. Para exemplificar esse exemplo tomemos com exemplo o sinal dado pela seguinte composição de funções senoidais.

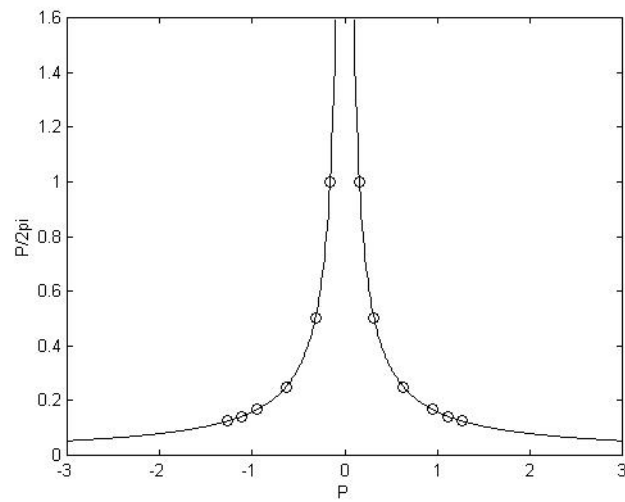
$$y(\omega) = \sin(\omega) + \frac{1}{2} \sin(2\omega) + \frac{1}{4} \sin(4\omega) + \frac{1}{6} \sin(6\omega) + \frac{1}{7} \sin(7\omega) + \frac{1}{8} \sin(8\omega)$$

Podemos ver claramente que a regra de formação de cada parcela desse sinal é dada por:

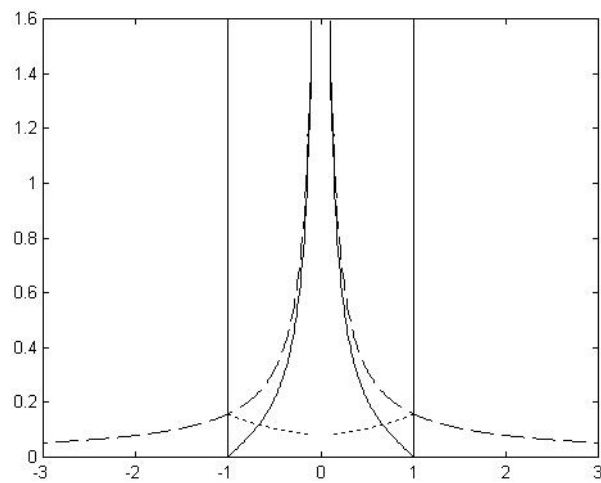
$$y_i = \left( \frac{P}{2\pi} \right) \sin \left( \frac{2\pi}{P} \omega \right)$$

A figura 2.4 abaixo, mostra o espectro do sinal em função de  $\omega$  para  $P$  variando no intervalo  $[-3,3] - \{0\}$  com as componentes utilizadas em destaque. Se amostrarmos esse sinal em uma determinada frequência as informações das componentes superiores a essa frequência distorcerão as componentes inferiores conforme mostrado na figura B.5.

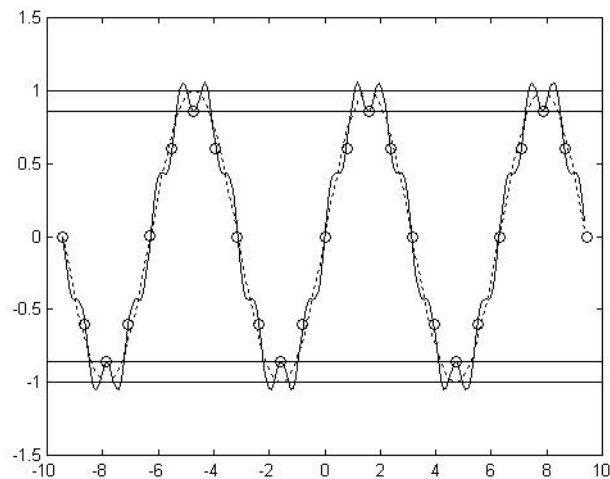
Se amostrarmos esse sinal com um tempo de amostragem igual metade do período da terceira componente,  $T_S = \pi/4$  perderemos completamente as informações dessa componente e as informações das três últimas componentes distorcerão as duas primeiras componente. Esse efeito pode ser observados através do gráfico temporal da primeira e da quarta componentes combinadas e da segunda e da quinta componentes combinadas, como pode ser visto nas figuras 2.6 e 2.7, respectivamente.



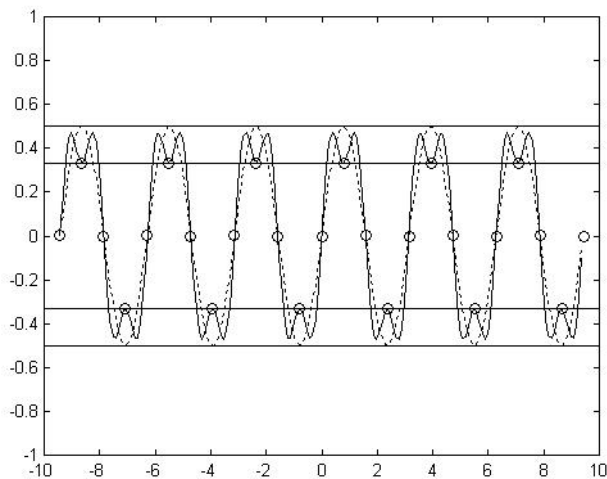
**Figura B.4:** Espectro de magnitude de um sinal periódico.



**Figura B.5:** Espectro de magnitudes de um sinal periódico distorcido.



**Figura B.6:** Distorção da primeira componente pelo rebatimento da quinta componente em torno da frequência  $2/\pi$ . Linha pontilhada – componente original, Linha contínua, componente distorcida.



**Figura B.7:** Distorção da segunda componente pelo rebatimento da quarta componente em torno da frequência  $2/\pi$ . Linha pontilhada – componente original, Linha contínua, componente distorcida.

Como podemos ver a partir das figuras acima a amostragem faz com que as componentes com frequência menor que a frequência de amostragem tenha a informação sobre a potência de seu sinal avaliada de forma errônea apesar da componente distorcida em si ter uma magnitude similar a da componente original.

## Símbolos e Nomenclatura

$t$	Tempo.
$k$	Variável amostragem discreta.
$N$	Numero de amostragens utilizadas na identificação.
$T_e$	Tempo de amostragem.
$u$	Sinal de entrada de um sistema contínuo.
$x$	Estado contínuo do sistema
$y$	Sinal de saída de um sistema contínuo.
$e$	Diferença entre saída do sistema real e predição do modelo.
$w$	Sinal tipo ruído branco.
$v$	Sinal tipo ruído colorido.
$\dot{f}$	Deriva do sinal $f$ .
$\hat{y}$	Predição do modelo
$\delta(t)$	Resposta de um sistema linear a um degrau unitário no sinal de entrada
$\xi$	Operador linear genérico
$p$	Operador linear de diferenciação.
$q$	Operador linear discreto de deslocamento unitário.
$s$	Variável complexa da transformada de Laplace
$z$	Variável complexa discreta.
$\theta_f$	Vetor de parâmetros do polinômio $F$ .
$\theta$	Vetor de parâmetros do sistema real.
$a, b, c, d$	Parâmetros dos modelos lineares
$\hat{\theta}$	Vetor de parâmetros estimados.
$G$	Modelo do sistema.
$H$	Modelo do distúrbio que afeta o sistema $G$ .
$\underline{a}, \underline{b}, \underline{c}, \underline{d}$	Polinômios dos modelos lineares.
<b>A, B, C, D</b>	Matrizes do modelos de espaço de estados linear
$I_n$	Matriz identidade de ordem $n$
$\varphi(t)$	Vetor de regressores.
$\psi(t)$	Gradiente da predição de $y$ em função dos parâmetros do modelo.
$\psi'(t)$	Hessiana da predição de $y$ em função dos parâmetros do modelo.
$V$	Função objetivo de minimização do erro quadrático.
$V'$	Gradiente da função objetivo.
$V''$	Hessiana da função objetivo.
$L$	Função de verossimilhança
$A^T$	Matriz transposta de $A$ .
$\mathcal{L}$	Transformação linear genérica.
$\mathcal{L}$	Transformada de Laplace
$\omega(t)$	Função de modulação.
$\Psi$	Função ortogonal.
$M_k\{f(t)\}$	Momento de Poisson de ordem $k$ do sinal $f(t)$ .
$f_k^{t_0}$	Momento de Poisson de ordem $k$ do sinal $f(t)$ calculado no instante $t_0$ .
$p_k(t)$	Função pulso de Poisson de ordem $k$ .
$\lambda, \beta$	Constantes da cadeia de filtros de Poisson.

$\gamma_n$	Matriz de coeficientes binomiais para cálculo dos momentos de Poisson das derivadas.
$F_n^{t_0}$	Vetor dos momentos de Poisson de ordem 0 até $n$ do sinal $f(t)$ calculado no instante $t_0$ .
$F(0)$	Vetor de condições iniciais do sinal $f(t)$ .
$P_n^{t_0}$	Vetor das funções pulso de Poisson de ordem 0 até $n$ do sinal calculadas no instante $t_0$ .
$\mathbf{M}_f$	Vetor dos momentos de Poisson das derivada do sinal $f(t)$ .
$\mathbf{M}_k$	Matriz de momentos de Poisson de ordem $k$ dos regressores do sistema.
$\mathbf{y}_k^{t_0}$	Vetor do momento de Poisson da derivada de mais alta ordem do sistema.
$\Delta_\theta$	Erro ( <i>bias</i> ) da estimação dos parâmetros.
$\mathbf{R}_{MM}$	Matriz de covariância dos regressores.
$\mathbf{R}_{Mw}$	Matriz de variância cruzada entre a matriz de regressores e o ruído
$\mathbf{R}_{yM}$	Matriz de variância cruzada entre a matriz de regressores e o sinal de saída.
$\mathbf{R}_{uM}$	Matriz de variância cruzada entre a matriz de regressores e o sinal de entrada.
$E(w)$	Esperança matemática do sinal $w(t)$ .
$\sigma_w^2$	Desvio padrão do sinal $w(t)$ .
$Z(t)$	Variável auxiliar do método da variável instrumental.
AR	<i>Autoregressive model</i>
ARX	<i>Autoregressive with eXogeneous input model</i>
ARMAX	<i>Autoregressive Moving Average with eXogeneous input model</i>
OE	<i>Output Error model</i>
BJ	<i>Box-Jenkins model</i>
PMF	<i>Poisson Moment Function</i>
MDPMF	<i>Multi Dimensional Poisson Moment Function</i>
PCBF	<i>Piecewise Constant Basis Function</i>
WF	<i>Walsh Function</i>
BPF	<i>Block-Pulse Function</i>
HF	<i>Haar Functions</i>
$NQE\theta$	Norma quadrática do erro dos parâmetros estimados.
$NQED$	Norma quadrática do erro da resposta ao degrau unitário.