

# CAPÍTULO 9

## ESTRUTURAS DE MODELOS DE SISTEMAS LINEARES INVARIANTES NO TEMPO

Um modelo de um sistema é uma descrição de (algumas de) suas propriedades, adequado a um certo propósito. O modelo não precisa ser uma descrição precisa do sistema para servir a seus propósitos (LJUNG, 1999).

A identificação de sistemas trata de construir e selecionar modelos de sistemas dinâmicos para atender a certos propósitos. Um primeiro passo para isso é determinar uma classe de modelos dentro da qual a busca do modelo mais adequado possa ser efetuada. Neste capítulo se discutem tais classes de modelos para sistemas lineares invariantes no tempo. Há diversas classes de estrutura de modelos (SÖDERSTRÖM; STOICA, 1989):

- a. Uma estrutura geral de modelo linear

$$A(q) \cdot y(t) = \frac{B(q)}{F(q)} u(t) + \frac{C(q)}{D(q)} e(t) \quad E[e^2(t)] = \sigma^2$$

Neste caso, escolhida a estrutura do modelo, deve-se selecionar os graus dos polinômios  $A(q)$ ,  $B(q)$ ,  $C(q)$ ,  $D(q)$  e  $F(q)$ , isto é,  $n_a$ ,  $n_b$ ,  $n_c$ ,  $n_d$  e  $n_f$ .

- b. Modelos em espaço de estados

$$\begin{cases} x(t+1) = A(\theta) \cdot x(t) + B(\theta) \cdot u(t) + w(t) \\ y(t) = C(\theta) \cdot x(t) + v(t) \end{cases}$$

onde  $w(t)$  corresponde a “ruído no processo” e  $v(t)$  a “ruído na medição”.

Neste caso, a seleção da estrutura do modelo é feita escolhendo-se os parâmetros da mesma, isto é, a dimensão do vetor de estado e a parametrização propriamente dita, a saber, a forma como  $\theta$  entra nas matrizes  $A(\theta)$ ,  $B(\theta)$  etc.

### 9.1 MODELOS DE SISTEMAS LINEARES INVARIANTES NO TEMPO

Um modelo linear invariante no tempo é especificado através da resposta ao impulso  $\{g(k)\}_1^\infty$ , do espectro  $\Phi_v(\omega) = \sigma \cdot |H(e^{j\omega})|^2$  da perturbação adicionada e, possivelmente, da

função densidade de probabilidade (FDP) da perturbação  $e(t)$ . Um modelo completo é então dado por (LJUNG, 1999):

$$y(t) = G(q)u(t) + H(q)e(t) \quad f_e(x) = \text{FDP de } e \quad (9.1)$$

$$\text{com: } G(q) = \sum_{k=1}^{\infty} g(k) q^{-k} \quad H(q) = 1 + \sum_{k=1}^{\infty} h(k) q^{-k}$$

Assim, um determinado modelo corresponde à especificação de três funções:  $G$ ,  $H$  e  $f_e$ . É, na maioria dos casos, inviável efetuar esta especificação enumerando as sequências infinitas  $\{g(k)\}$ ,  $\{h(k)\}$  junto com a função  $f_e(x)$ . Ao invés disso, prefere-se trabalhar com estruturas que permitam a especificação de  $G$  e  $H$  em termos de um número finito de valores numéricos. Funções de transferência racionais e descrições de espaço de estados de dimensão finita são exemplos típicos disso. Além disso, na maioria das vezes, a FDP  $f_e$  não é especificada como uma função, mas descrita em termos de umas poucas características numéricas, tipicamente o 1º e o 2º momentos (LJUNG, 1999):

$$\begin{cases} E[e(t)] = \int x f_e(x) dx = 0 \\ E[e^2(t)] = \int x^2 f_e(x) dx = \sigma^2 \end{cases} \quad (9.2)$$

É comum também se assumir que  $e(t)$  seja gaussiana, caso em que a FDP é completamente especificada por (9.2).

A especificação de (9.1) em termos de um número finito de coeficientes tem uma importante consequência para a identificação de sistemas. Normalmente, não é possível determinar esses coeficientes *a priori*, a partir do conhecimento dos mecanismos físicos que regem o comportamento do sistema. Ao invés disso, a determinação de todos ou de alguns deles deve ser feita por procedimentos de estimação. Isso significa que esses coeficientes entram no modelo (9.1) como parâmetros a serem estimados, sendo denotados pelo vetor  $\theta$ , de forma que a descrição geral da estrutura do modelo se torna (LJUNG, 1999):

$$y(t) = G(q, \theta) \cdot u(t) + H(q, \theta) \cdot e(t) \quad (9.3a)$$

$$f_e(x, \theta), \quad \text{a FDP de } e(t), \quad \{e(t)\} = \text{ruído branco} \quad (9.3b)$$

A Equação (9.3) descreve um modelo linear geral, sendo  $y(t)$  a saída com dimensão  $n_y$  no instante  $t$  e  $u(t)$  a entrada  $n_u$ -dimensional.  $e(t)$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero. Tal sequência é intitulada ruído branco. Tem-se então que  $E[e(t)e^T(s)] = \Sigma(\theta) \cdot \delta(t-s)$ .  $G(q, \theta)$  é um filtro com dimensão  $n_y \times n_u$  e  $H(q, \theta)$  é um filtro com dimensão  $n_y \times n_v$ . Na maioria dos casos, os filtros  $G(q, \theta)$  e  $H(q, \theta)$  são de ordem finita. Eles são então funções racionais de  $q$ .

Os filtros  $G(q, \theta)$  e  $H(q, \theta)$  bem como a matriz de covariância do ruído  $\Sigma(\theta)$ , são funções do vetor de parâmetros  $\theta$ . Note que (9.3) não é mais um modelo, mas sim um conjunto de modelos e cabe ao procedimento de estimação escolher o membro do conjunto que seja o mais adequado para o propósito em questão. Um conjunto parametrizado de modelos como (9.3) é chamado uma **estrutura de modelo** e é denotado por  $M$ . O modelo específico associado ao valor do vetor de parâmetros  $\theta$  é denotado por  $M(\theta)$  (LJUNG, 1999).

## 9.2 ERRO DA EQUAÇÃO, RESÍDUO E ERRO DE PREDIÇÃO

Seja um processo descrito pela seguinte equação:

$$y(t) + a_1 \cdot y(t-1) + \dots + a_{n_a} \cdot y(t-n_a) = b_1 \cdot u(t-1) + \dots + b_{n_b} \cdot u(t-n_b) + e(t) \quad (9.4)$$

Suponha o modelo para o processo da Equação (9.4) dado pela seguinte equação de diferenças:

$$y(t) + \hat{a}_1 \cdot y(t-1) + \dots + \hat{a}_{n_a} \cdot y(t-n_a) = \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) + \varepsilon(t) \quad (9.5)$$

Há uma diferença importante entre a perturbação  $e(t)$  em (9.4) e o termo  $\varepsilon(t)$  em (9.5), embora ambos surjam na mesma posição nessas equações. O termo  $e(t)$  responde pelo efeito de todas as perturbações não-medidas e ruídos, agindo na saída do processo, ao passo que o termo  $\varepsilon(t)$  é usado para considerar os erros do modelo, que são função de seus parâmetros, isto é,  $\varepsilon(t) = \varepsilon(t, \theta)$ . Note que em (9.4) os parâmetros  $\theta$  são valores reais, fixos e conhecidos, enquanto que na Equação (9.5) eles são valores a serem estimados.

A razão para se incluir  $\varepsilon(t)$  no modelo (9.5) é que seria improvável que com  $\varepsilon(t) \equiv 0$  (para todo  $t$ ) resultasse uma descrição perfeita de um processo real. Portanto,  $\varepsilon(t)$  descreve o desvio nos dados entre o processo real e o sistema simulado, sendo que este último se trata de um sistema linear perfeito (determinístico) (SÖDERSTRÖM; STOICA, 1989). Como  $\varepsilon(t)$  aparece como um erro direto na equação de diferenças, indicando que um modelo linear não consegue descrever com perfeição um processo real, é comum chamá-lo de **erro da equação** ou **resíduo**.

Introduzindo-se:

$$\begin{cases} \hat{A}(q) = 1 + \hat{a}_1 \cdot q^{-1} + \dots + \hat{a}_{n_a} \cdot q^{-n_a} \\ \hat{B}(q) = \hat{b}_1 \cdot q^{-1} + \dots + \hat{b}_{n_b} \cdot q^{-n_b} \end{cases}$$

a Equação (9.5) pode ser reescrita, adotando-se assim a mesma notação da Expressão (2.17):

$$\hat{A}(q) \cdot y(t) = \hat{B}(q) \cdot u(t) + \varepsilon(t) \quad (9.6)$$

Portanto, o erro da equação ou resíduo pode ser calculado por:

$$\varepsilon(t) = \hat{A}(q) \cdot y(t) - \hat{B}(q) \cdot u(t) \quad (9.7)$$

A Figura 9.1 mostra o diagrama de blocos de geração do resíduo ou erro da equação, conforme a Equação (9.7).

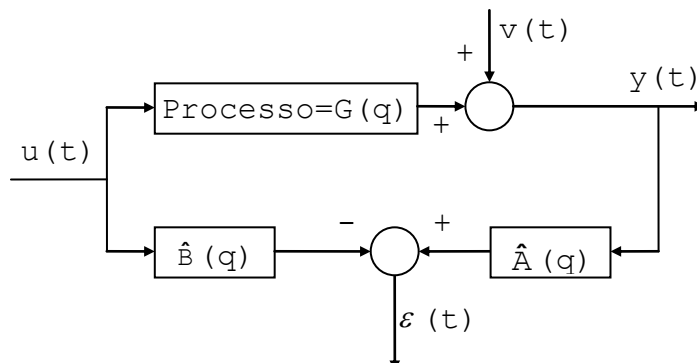


Fig. 9.1 Diagrama de blocos para cálculo do resíduo ou erro da equação.

O modelo (9.5) pode também ser expresso na forma de uma regressão linear:

$$y(t) = \varphi^T(t) \hat{\theta} + \varepsilon(t) \quad (9.8)$$

onde:  $\varphi^T(t) = [-y(t-1) \dots -y(t-n_a) \ u(t-1) \dots u(t-n_b)]$

$$\hat{\theta} = [\hat{a}_1 \dots \hat{a}_{n_a} \ \hat{b}_1 \dots \hat{b}_{n_b}]^T$$

Pode-se reescrever a Expressão (9.6) na seguinte forma:

$$y(t) = y(t) - \hat{A}(q) \cdot y(t) + \hat{B}(q) \cdot u(t) + \varepsilon(t) = [1 - \hat{A}(q)] \cdot y(t) + \hat{B}(q) \cdot u(t) + \varepsilon(t) \quad (9.9)$$

ou ainda como:

$$y(t) = \frac{\hat{B}(q)}{\hat{A}(q)} u(t) + \frac{1}{\hat{A}(q)} \varepsilon(t) = \hat{G}(q) \cdot u(t) + \hat{H}(q) \cdot \varepsilon(t) \quad (9.10)$$

$$\text{onde: } \hat{G}(q) = \frac{\hat{B}(q)}{\hat{A}(q)} \quad \text{e} \quad \hat{H}(q) = \frac{1}{\hat{A}(q)} \quad (9.11)$$

Vê-se que (9.10) equivale a (9.3a). Substituindo-se  $\hat{G}(q)$  e  $\hat{H}(q)$  em (9.9), resulta:

$$y(t) = [1 - \hat{H}^{-1}(q)] \cdot y(t) + \hat{H}^{-1}(q) \cdot \hat{G}(q) \cdot u(t) + \varepsilon(t) \quad (9.12)$$

Caso se despreze o erro da equação  $\varepsilon(t)$  na Equação (9.12), resulta:

$$\hat{y}(t/t-1) = \hat{H}^{-1}(q) \cdot \hat{G}(q) \cdot u(t) + [1 - \hat{H}^{-1}(q)] \cdot y(t) \quad (9.13)$$

A Equação (9.13) corresponde ao preditor um passo à frente para este modelo, conforme visto na Expressão (5.9). Esta forma do preditor não depende de  $f_e(x, \theta)$ . Poder-se-ia chegar em (9.6) através de considerações que não sejam probabilísticas. Então a especificação (9.3b) não se aplica. É usado o termo **modelos preditores** para modelos que somente especifiquem  $G$  e  $H$  como em (9.3a) ou na forma (9.6). Analogamente, **modelos probabilísticos** significam descrições (9.3) que forneçam uma caracterização completa das propriedades probabilísticas do sistema.

A Equação (9.13) representa o preditor 1 passo à frente desse modelo. Assim, caso se despreze o erro da equação  $\varepsilon(t)$  nas Equações (9.5), (9.8) e (9.9), pode-se prever a saída no instante  $t$ , usando dados de entrada e saída do processo em  $t-1$ ,  $t-2$ , ... como:

$$\hat{y}(t/t-1) = -\hat{a}_1 \cdot y(t-1) - \dots - \hat{a}_{n_a} \cdot y(t-n_a) + \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) = \varphi^T(t) \hat{\theta} \quad (9.14)$$

Pode-se ainda substituir (9.11) em (9.13) para computar o preditor um passo à frente:

$$\hat{y}(t/t-1) = [1 - \hat{A}(q)] \cdot y(t) + \hat{B}(q) \cdot u(t) \quad (9.15)$$

Assim, a equação de erro:

$$\varepsilon(t) = y(t) - \hat{y}(t/t-1) \quad (9.16)$$

pode ser interpretada também como um **erro de predição** um passo à frente. Veja o diagrama de blocos da Figura 9.2 para essa interpretação (ZHU; BACKX, 1993).

Deve-se ter em mente, ao analisar a Figura 9.2, que o sistema que representa o processo real pode ser dado por:

$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t) = G(q) \cdot u(t) + H(q) \cdot e(t) = G(q) \cdot u(t) + v(t) \quad (9.17)$$

Pode-se também escrever o erro de predição como:

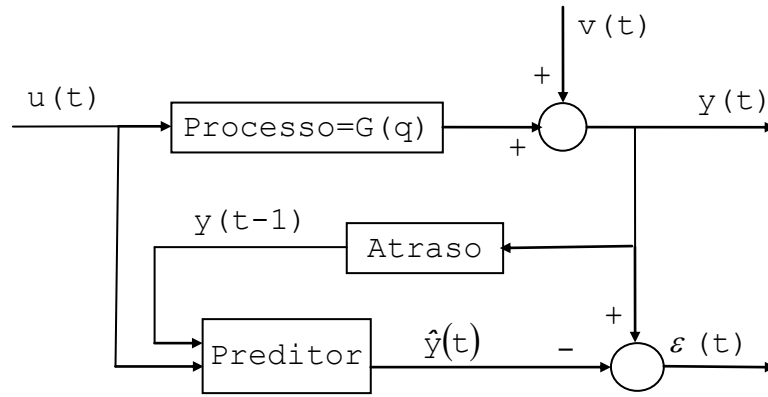


Fig. 9.2 Diagrama de blocos para cálculo do erro de predição.

$$\varepsilon(t, \hat{\theta}) = y(t) - \hat{y}(t/t-1) = y(t) - \hat{y}(t/t-1, \hat{\theta}) = y(t) - \varphi^T(t) \hat{\theta} \quad (9.18)$$

Em (9.18),  $\hat{y}(t/t-1, \hat{\theta})$  denota a predição de  $y(t)$ , supondo fornecidos os dados de entrada e saída até o instante  $t-1$ , isto é,  $y(t-1)$ ,  $u(t-1)$ ,  $y(t-2)$ ,  $u(t-2)$ , ... e baseado no vetor de parâmetros do modelo  $\hat{\theta}$ .

Portanto, os termos **resíduo**, **erro de predição** e **erro da equação** são equivalentes, sendo calculados pela Equação (9.18).

### 9.3 UMA FAMÍLIA DE MODELOS DE FUNÇÕES DE TRANSFERÊNCIA

Nesta seção, diferentes modos de descrever (9.3a) em termos de  $\theta$  são analisados, especificando-se como  $G(q, \theta)$ ,  $H(q, \theta)$  e  $\Sigma(\theta)$  dependem do vetor de parâmetros  $\theta$ , isto é, diferentes formas de parametrizar o conjunto de modelos são discutidas. O modo mais imediato de parametrizar  $G$  e  $H$  é representá-los como funções racionais e deixar os parâmetros serem os coeficientes do numerador e do denominador. Nesta seção são descritas várias maneiras de efetuar tais parametrizações. Tais estruturas de modelo são também conhecidas como modelos **caixa-preta** (LJUNG, 1999).

#### 9.3.1 Estrutura de modelo de erro da equação - estrutura ARX

Provavelmente, a relação mais simples de entrada/saída é obtida escrevendo-se o modelo como uma equação de diferenças linear, conforme visto na Equação (9.5). Assim:

$$\varepsilon(t) = y(t) + \hat{a}_1 \cdot y(t-1) + \dots + \hat{a}_{n_a} \cdot y(t-n_a) - \hat{b}_1 \cdot u(t-1) - \dots - \hat{b}_{n_b} \cdot u(t-n_b) \quad (9.19)$$

Como visto na Seção 9.2,  $\varepsilon(t)$  é normalmente chamado de erro da equação. Assim, o modelo (9.19) é intitulado **modelo de erro da equação** (LJUNG, 1999). Os parâmetros ajustáveis são:

$$\hat{\theta} = [\hat{a}_1 \ \dots \ \hat{a}_{n_a} \ \hat{b}_1 \ \dots \ \hat{b}_{n_b}]^T$$

A seguinte estrutura, mostrada em (9.10) e repetida a seguir, é do tipo **ARX** (*AutoRegressive with eXogenous input*):

$$y(t) = \frac{\hat{B}(q)}{\hat{A}(q)} u(t) + \frac{1}{\hat{A}(q)} \varepsilon(t) \quad (9.10)$$

onde AR se refere à estrutura do modelo de perturbação  $1/\hat{A}(q)$  e  $X$  à entrada extra  $u(t)$  (chamada de variável exógena em econometria). O fluxo de sinais da estrutura ARX pode ser visto na Figura 9.3.

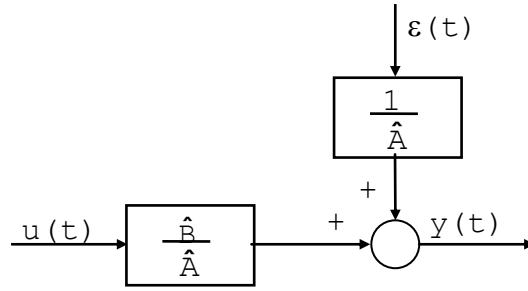


Fig. 9.3 A estrutura de modelo ARX.

Analisando-se a Figura 9.3, nota-se que o modelo (9.10) não é a representação mais natural do ponto de vista físico, pois se assume que o ruído branco passe através da dinâmica do denominador do sistema antes de ser somado à saída. No entanto, este modelo tem uma propriedade importante, que o torna uma escolha preferencial em muitas aplicações: o preditor um passo à frente define uma regressão linear, conforme a Equação (9.14).

O preditor um passo à frente para a estrutura ARX dada em (9.10) é mostrado na Equação (9.15). Sem uma estrutura estocástica, o preditor (9.15) é uma escolha natural, se o termo  $\varepsilon(t)$  em (9.10) for considerado “insignificante” ou “difícil de avaliar”. É assim perfeitamente natural trabalhar com a Expressão (9.14) também para modelos “determinísticos”.

O vetor  $\varphi(t)$ , extraído de (9.14), também conhecido como vetor de observações, é dado por:

$$\varphi(t) = [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n_a) \ u(t-1) \ u(t-2) \ \dots \ u(t-n_b)]^T$$

Percebe-se que  $\varphi(t)$  contém valores de dados coletados da entrada e da saída do processo até o instante  $t-1$ . O modelo ARX descreve a variável observada  $y(t)$  como uma combinação linear desconhecida dos componentes do vetor de observações  $\varphi(t)$  mais a perturbação. Esta é uma propriedade importante de (9.14). O preditor um passo à frente é o produto escalar entre um vetor de dados conhecido  $\varphi(t)$  e o vetor de parâmetros  $\theta$ . Tal modelo é chamado uma regressão linear em estatística.

A saída do processo real é dada por:

$$y(t) = \theta^T \varphi(t) + v(t) = \varphi^T(t) \theta + v(t)$$

Se a característica do termo de perturbação  $v(t)$  não for especificada, pode-se considerar  $\hat{y}(t/t-1, \theta) = \hat{\theta}^T \varphi(t) = \varphi^T(t) \hat{\theta}$  como uma “estimativa” do que  $y(t)$  vai ser no próximo instante de amostragem, tendo os valores anteriores de  $y(t)$  e  $u(t)$ ,  $t = t-1, t-2, \dots$ . Esta estimativa, é claro, depende também dos parâmetros do modelo  $\theta$ . A estimativa pode ter um senso estatístico exato, se  $\{v(t)\}$  for uma sequência de variáveis aleatórias (randômicas) independentes com média zero (isto é, ruído branco).

Se não há nenhuma entrada ( $n_b = 0$ ) e  $v(t)$  é ruído branco, pode-se escrever:

$$y(t) + \hat{a}_1 \cdot y(t-1) + \dots + \hat{a}_{n_a} \cdot y(t-n_a) = v(t)$$

### 9.3.2 Estrutura de modelo de erro da equação - estrutura FIR

No caso especial em que  $n_a = 0$ ,  $y(t)$  é modelada como uma resposta finita ao impulso (*finite impulse response* - **FIR**). Esta estrutura de modelo é particularmente comum em aplicações de processamento de sinal. O modelo FIR pode também ser chamado de modelo de função-peso truncada, sendo descrito pela seguinte equação:

$$y(t) = \hat{B}(q) \cdot u(t) + \varepsilon(t) \quad (9.20)$$

Expandindo-se (9.20), tem-se que:

$$y(t) = [\hat{b}_1 \cdot q^{-1} + \hat{b}_2 \cdot q^{-2} + \dots + \hat{b}_{n_b} \cdot q^{-n_b}] \cdot u(t) + \varepsilon(t)$$

$$y(t) = \hat{b}_1 \cdot u(t-1) + \hat{b}_2 \cdot u(t-2) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) + \varepsilon(t)$$

Pode-se escrever:

$$y(t) = \sum_{j=1}^{n_b} \hat{b}_j \cdot u(t-j) + \varepsilon(t) \quad (9.21)$$

Desprezando-se o termo  $\varepsilon(t)$  e aplicando-se um pulso unitário em  $t=0$ , resulta:

$$y(0) = 0$$

$$y(1) = \hat{b}_1 \cdot u(0) = \hat{b}_1$$

$$y(2) = \hat{b}_1 \cdot u(1) + \hat{b}_2 \cdot u(0) = \hat{b}_2$$

$$y(3) = \hat{b}_1 \cdot u(2) + \hat{b}_2 \cdot u(1) + \hat{b}_3 \cdot u(0) = \hat{b}_3$$

$\vdots$

Uma resposta finita ao impulso pode ser obtida convoluindo-se  $g(t)$  com  $u(t) = \delta(t)$ :

$$y(t) = \sum_{j=0}^t g(t-j) \cdot u(j) \cdot T = \sum_{j=0}^t g(t-j) \cdot \delta(j) \cdot T$$

Considerando-se o atraso causado pelo segurador de ordem zero, reescreve-se a equação anterior como:

$$y(t) = \sum_{j=0}^t g(t-j) \cdot u(j-1) \cdot T = \sum_{j=0}^t g(t-j) \cdot \delta(j-1) \cdot T$$

Portanto:

$$y(0) = g(0) \cdot \delta(-1) \cdot T = g(0) \cdot 0 \cdot T = 0$$

$$y(1) = [g(1) \cdot \delta(-1) + g(0) \cdot \delta(0)] \cdot T = g(0) \cdot T$$

$$y(2) = [g(2) \cdot \delta(-1) + g(1) \cdot \delta(0) + g(0) \cdot \delta(1)] \cdot T = g(1) \cdot T$$

$$y(3) = [g(3) \cdot \delta(-1) + g(2) \cdot \delta(0) + g(1) \cdot \delta(1) + g(0) \cdot \delta(2)] \cdot T = g(2) \cdot T$$

$\vdots$

Esta resposta é conhecida como FIR e, por analogia com a Equação (9.21), intitulou-se esta estrutura de FIR. Nota-se que  $\hat{b}_i$  é igual a  $g(i-1) \cdot T$ ,  $i = 1$  a  $n_b$ . Na estrutura FIR, o ruído branco é somado diretamente na saída, definindo um modelo do tipo erro na saída. Por outro lado, tanto a função de transferência do processo  $\hat{B}(q)$  como a da perturbação têm o polinômio  $\hat{A}(q) = 1$  como fator comum. Neste sentido, a estrutura FIR também é do tipo erro na equação.

### 9.3.3 Estrutura de modelo de erro da equação - estrutura ARMAX

A desvantagem básica do modelo (9.10) é a falta de liberdade para descrever as propriedades do termo de perturbação. Pode-se acrescentar flexibilidade a isso, descrevendo-se o erro da equação como uma média móvel. Para tal, pode-se estender a estrutura de modelo (9.5) para:

$$y(t) + \hat{a}_1 \cdot y(t-1) + \dots + \hat{a}_{n_a} \cdot y(t-n_a) = \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) + \varepsilon(t) + \hat{c}_1 \cdot \varepsilon(t-1) + \dots + \hat{c}_{n_c} \cdot \varepsilon(t-n_c) \quad (9.22)$$

com  $\hat{C}(q) = 1 + \hat{c}_1 \cdot q^{-1} + \dots + \hat{c}_{n_c} \cdot q^{-n_c}$

A Expressão (9.22) pode ser reescrita como:

$$\hat{A}(q) \cdot y(t) = \hat{B}(q) \cdot u(t) + \hat{C}(q) \cdot \varepsilon(t) \quad (9.23)$$

ou

$$y(t) = \frac{\hat{B}(q)}{\hat{A}(q)} u(t) + \frac{\hat{C}(q)}{\hat{A}(q)} \varepsilon(t) \quad (9.23a)$$

Definindo-se a correspondência com (9.3), resulta:

$$\hat{G}(q, \theta) = \frac{\hat{B}(q)}{\hat{A}(q)} \quad \hat{H}(q, \theta) = \frac{\hat{C}(q)}{\hat{A}(q)} \quad (9.24)$$

O modelo de perturbação é um processo estocástico do tipo ARMA, dado por  $\hat{C}(q)/\hat{A}(q)$ . Em vista disso, a Estrutura (9.23) é chamada **ARMAX**, a qual é uma ferramenta padrão em controle e econometria. Esta estrutura foi proposta por Åström e Bohlin (1965).

Qualquer sistema linear de ordem finita com perturbações estacionárias possuindo uma densidade espectral de potência racional, pode ser descrito por um modelo ARMAX (SÖDERSTRÖM; STOICA, 1989).

O preditor um passo à frente para a estrutura ARMAX em (9.23) é obtido inserindo-se (9.24) em (5.9). Resulta:

$$\hat{y}(t/t-1, \hat{\theta}) = \frac{\hat{B}(q)}{\hat{C}(q)} u(t) + \left[ 1 - \frac{\hat{A}(q)}{\hat{C}(q)} \right] \cdot y(t) \quad (9.25)$$

$$\text{ou} \quad \hat{C}(q) \cdot \hat{y}(t/t-1, \hat{\theta}) = \hat{B}(q) \cdot u(t) + [\hat{C}(q) - \hat{A}(q)] \cdot y(t) \quad (9.25a)$$

Isto significa que a predição é obtida filtrando-se  $u$  e  $y$  através de um filtro com dinâmica do denominador determinada por  $\hat{C}(q)$ .

O preditor (9.25) pode ser reescrito em analogia com (9.15), adicionando-se  $[1 - \hat{C}(q)] \cdot \hat{y}(t/t-1, \hat{\theta})$  a ambos os lados de (9.25a):

$$\hat{y}(t/t-1, \hat{\theta}) = \hat{B}(q) \cdot u(t) + \hat{C}(q) \cdot [y(t) - \hat{y}(t/t-1, \hat{\theta})] + \hat{y}(t/t-1, \hat{\theta}) - \hat{A}(q) \cdot y(t)$$

Adicionando-se a parcela nula  $y(t) - y(t)$  ao segundo membro da equação anterior, resulta:

$$\hat{y}(t/t-1, \hat{\theta}) = \hat{B}(q) \cdot u(t) + [1 - \hat{A}(q)] \cdot y(t) + [\hat{C}(q) - 1] \cdot [y(t) - \hat{y}(t/t-1, \hat{\theta})] \quad (9.26)$$

O erro de predição é dado por:

$$\varepsilon(t, \hat{\theta}) = y(t) - \hat{y}(t/t-1, \hat{\theta})$$



Substituindo-se este valor em (9.26):

$$\hat{y}(t/t-1, \hat{\theta}) = \hat{B}(q) \cdot u(t) + [1 - \hat{A}(q)] \cdot y(t) + [\hat{C}(q) - 1] \cdot \varepsilon(t, \hat{\theta})$$

Substituindo-se  $\hat{A}(q)$ ,  $\hat{B}(q)$  e  $\hat{C}(q)$  por seus respectivos polinômios:

$$\begin{aligned} \hat{y}(t/t-1, \hat{\theta}) = & \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) - \hat{a}_1 \cdot y(t-1) + \dots - \hat{a}_{n_a} \cdot y(t-n_a) + \\ & + \hat{c}_1 \cdot \varepsilon(t-1, \hat{\theta}) + \dots + \hat{c}_{n_c} \cdot \varepsilon(t-n_c, \hat{\theta}) \end{aligned} \quad (9.27)$$

Portanto, (9.27) pode ser reescrita como:

$$\hat{y}(t/t-1, \hat{\theta}) = \varphi^T(t, \hat{\theta}) \hat{\theta} \quad (9.28)$$

onde:

$$\begin{aligned} \varphi^T(t, \hat{\theta}) = & [-y(t-1) \dots -y(t-n_a) \quad u(t-1) \dots u(t-n_b) \quad \varepsilon(t-1, \hat{\theta}) \dots \varepsilon(t-n_c, \hat{\theta})] \\ \hat{\theta} = & (\hat{a}_1 \dots \hat{a}_{n_a} \quad \hat{b}_1 \dots \hat{b}_{n_b} \quad \hat{c}_1 \dots \hat{c}_{n_c})^T \end{aligned}$$

Note a semelhança com a regressão linear (9.14). A Equação (9.28) não corresponde, no entanto, a uma regressão linear, devido ao efeito não-linear de  $\hat{\theta}$  no vetor  $\varphi(t, \hat{\theta})$ . Para enfatizar a semelhança com (9.14), ela é chamada uma **regressão pseudolinear**.

Caso se deseje, o vetor de parâmetros pode ser complementado com a variância do ruído:

$$\Sigma(\hat{\theta}) = \sigma^2 = E[\varepsilon^2(t)]$$

de forma que:

$$\hat{\theta}' = (\hat{\theta}^T \sigma^2)^T$$

pode ser usado como um novo vetor de parâmetros de dimensão  $n_a + n_b + n_c + 1$ .

### 9.3.4 Estrutura de modelo de erro da equação - estrutura ARIMAX

Uma versão com uma integração imposta na descrição do sistema é o modelo **ARIMA(X)** (I para integração, com ou sem a variável externa  $u$ ), o qual é útil para descrever processos com perturbações não-estacionárias. Pode-se obtê-lo substituindo-se  $y(t)$  em (9.23) por  $\Delta y(t) = y(t) - y(t-1)$  ou, equivalentemente,  $\hat{A}(q)$  por  $(1 - q^{-1}) \cdot \hat{A}(q)$ .

Seja um modelo ARMA dado por:

$$\hat{A}(q) \cdot y(t) = \hat{C}(q) \cdot \varepsilon(t)$$

Portanto:

$$y(t) = \frac{\hat{C}(q)}{\hat{A}(q)} \varepsilon(t)$$

Trata-se de um caso particular do modelo ARMAX, quando não há sinais exógenos, isto é, quando  $u(t) = 0$ . O polinômio  $\hat{A}(q)$  pode ser fatorado em termos de suas raízes  $\lambda_i$ ,  $i = 1, 2, \dots, n_a$ :

$$\hat{A}(q) = (1 - \lambda_1 \cdot q^{-1}) \cdot (1 - \lambda_2 \cdot q^{-1}) \cdot \dots \cdot (1 - \lambda_{n_a} \cdot q^{-1})$$

Assuma que uma das raízes seja unitária, p. ex.  $\lambda_1 = 1$ . As demais podem ser reais ou surgem em pares conjugados e satisfazem  $|\lambda_i| < 1$ ,  $i = 2, 3, \dots, n_a$ . Pode-se então escrever:

$$y(t) = \frac{\hat{C}(q)}{\hat{A}(q)} \varepsilon(t) = \frac{\hat{C}(q)}{(1 - \lambda_1 \cdot q^{-1}) \cdot (1 - \lambda_2 \cdot q^{-1}) \cdot \dots \cdot (1 - \lambda_{n_a} \cdot q^{-1})} \varepsilon(t)$$

Como se assumiu que  $\lambda_1 = 1$ , resulta:

$$y(t) = \frac{\hat{C}(q)}{\hat{A}(q)} \varepsilon(t) = \frac{\hat{C}(q)}{(1 - q^{-1}) \cdot (1 - \lambda_2 \cdot q^{-1}) \cdot \dots \cdot (1 - \lambda_{n_a} \cdot q^{-1})} \varepsilon(t) \quad (9.29)$$

Nesta equação se percebe a presença de um polo em  $q = 1$ , que equivale a um integrador. Portanto, a entrada  $\varepsilon(t)$  afeta a saída  $y(t)$  por meio de um integrador. Por essa razão, este modelo é chamado de ARIMA (*AutoRegressive Integrated Moving Average*). Multiplicando-se  $y(t)$  por  $(1 - q^{-1})$ , resulta:

$$y(t) \cdot (1 - q^{-1}) = \Delta y(t) = \frac{\hat{C}(q)}{(1 - \lambda_2 \cdot q^{-1}) \cdot \dots \cdot (1 - \lambda_{n_a} \cdot q^{-1})} \varepsilon(t)$$

Redefinindo-se  $\hat{A}'(q)$ , não mais com ordem  $n_a$  mas sim com ordem  $(n_a - 1)$ , tem-se:

$$\hat{A}'(q) = (1 - \lambda_2 \cdot q^{-1}) \cdot (1 - \lambda_3 \cdot q^{-1}) \cdot \dots \cdot (1 - \lambda_{n_a} \cdot q^{-1})$$

Reescrevendo-se a equação original para  $y(t)$ :

$$y(t) = \frac{\hat{C}(q)}{(1 - q^{-1}) \cdot \hat{A}'(q)} \varepsilon(t)$$

Portanto:

$$(1 - q^{-1}) \cdot y(t) = \Delta y(t) = \frac{\hat{C}(q)}{\hat{A}'(q)} \varepsilon(t) \quad (9.30)$$

Sabe-se que um modelo ARMA com parâmetros constantes, com todas as raízes dentro do círculo de raio unitário, isto é, satisfazendo  $|\lambda_i| < 1$ ,  $i = 2, 3, \dots, n_a$  como em (9.29) representa um processo estocástico estacionário. Entretanto, a saída do modelo (9.30) é  $\Delta y(t)$ , de modo que para se recuperar  $y(t)$  é necessário realizar uma “integração discreta”, ou seja, calcular o seguinte somatório (AGUIRRE, 2007):

$$y(t) = \sum_{i=1}^t \Delta y(i) \quad \text{com } y(0) = 0$$

A estrutura ARIMA serve para modelar séries temporais não estacionárias, geradas por processos com uma raiz unitária, ou seja, processos com integração. Séries temporais desse tipo podem ser tornadas estacionárias, tomando-se a primeira diferença da saída, sendo que o resultado pode então ser modelado como um processo ARMA. Para voltar ao sinal de saída original, deve-se integrar a saída do modelo ARMA. Este procedimento é denominado modelagem ARIMA, em que o “I” indica o processo de integração discreta usado.

### 9.3.5 Outras estruturas de modelo do tipo erro da equação

Ao invés de modelar o erro da Equação (9.7) como uma média móvel, como foi feito em (9.23a), ele pode ser descrito como uma auto-regressão. Isso resulta no seguinte conjunto de estruturas de modelos (LJUNG, 1999):

$$\hat{A}(q) \cdot y(t) = \hat{B}(q) \cdot u(t) + \frac{1}{\hat{D}(q)} \varepsilon(t) \quad (9.31)$$

com  $\hat{D}(q) = 1 + \hat{d}_1 \cdot q^{-1} + \dots + \hat{d}_{n_d} \cdot q^{-n_d}$

Este modelo, seguindo a terminologia prévia, pode ser chamado **ARARX**. Esta estrutura de modelo foi publicada pela primeira vez na literatura de controle por Clarke (1967). Mais genericamente, poder-se-ia usar uma descrição ARMA do erro da equação, gerando uma estrutura **ARARMAX**:

$$\hat{A}(q) \cdot y(t) = \hat{B}(q) \cdot u(t) + \frac{\hat{C}(q)}{\hat{D}(q)} \varepsilon(t) \quad (9.32)$$

A Estrutura (9.32) contém (9.10) (ARX), (9.23) (ARMAX) e (9.31) (ARARX) como casos especiais. Ela forma a família de modelos do tipo erro da equação. Estas estruturas de modelo equivalem a descrições onde as funções de transferência  $G$  e  $H$  possuem o polinômio  $A$  como um fator comum nos denominadores. Seu esquema é visto na Figura 9.4.

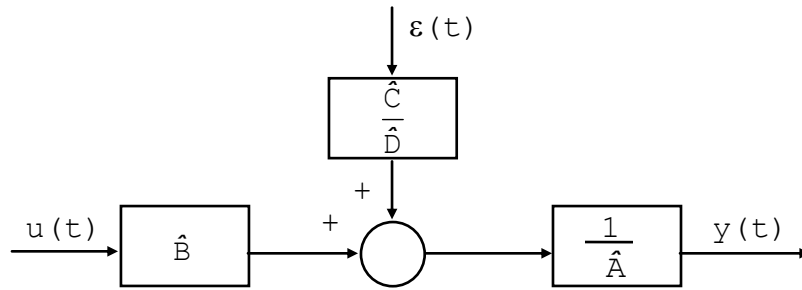


Fig. 9.4 A família de modelos do tipo erro da equação.

### 9.3.6 Estrutura de modelo de erro na saída

Do ponto de vista físico, pode parecer mais natural parametrizar as funções de transferência  $G$  e  $H$  independentemente, sem um fator comum no denominador. Caso se suponha que a relação entre entrada e saída não perturbada  $x$  possa ser escrita como uma equação linear de diferenças e que as perturbações consistam em ruído branco na medição, então se obtém a seguinte descrição (LJUNG, 1999):

$$x(t) + \hat{f}_1 \cdot x(t-1) + \dots + \hat{f}_{n_f} \cdot x(t-n_f) = \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) \quad (9.33a)$$

$$y(t) = x(t) + \varepsilon_{OE}(t) \quad (9.33b)$$

Pode-se escrever o modelo como:

$$y(t) = \frac{\hat{B}(q)}{\hat{F}(q)} u(t) + \varepsilon_{OE}(t) \quad (9.34)$$

com  $\hat{F}(q) = 1 + \hat{f}_1 \cdot q^{-1} + \dots + \hat{f}_{n_f} \cdot q^{-n_f}$

O fluxo de sinal desse modelo é mostrado na Figura 9.5.

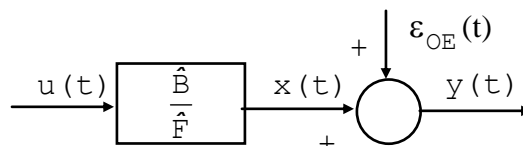


Fig. 9.5 A estrutura do modelo de erro na saída.

A Equação (9.34) é chamada uma estrutura de modelo de **erro na saída** (OE - *output error*). O nome provém do fato de se poder escrever (9.34) na seguinte forma:

$$\varepsilon_{OE}(t) = y(t) - \frac{\hat{B}(q)}{\hat{F}(q)} u(t) = y(t) - \hat{G}(q) \cdot u(t) = y(t) - \hat{y}(t) \quad (9.35)$$

sendo  $\varepsilon_{OE}(t)$  o erro na saída, isto é, a diferença entre a saída medida  $y(t)$  e a saída do modelo livre de ruído  $[\hat{B}(q)/\hat{F}(q)] \cdot u(t)$ . O diagrama de blocos para gerar a estrutura de modelo de erro na saída é mostrado na Figura 9.6.

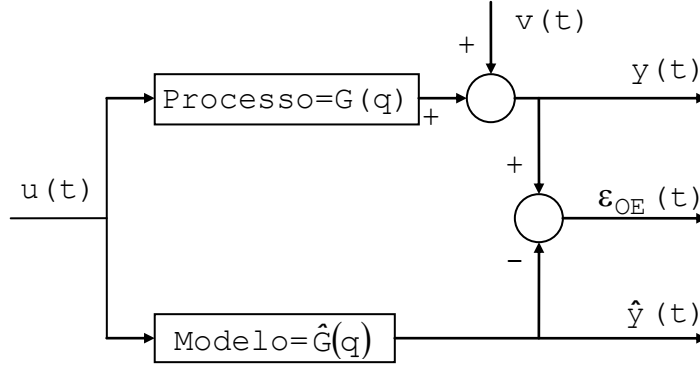


Fig. 9.6 Geração do erro na saída  $\varepsilon_{OE}$ .

Deve-se enfatizar a diferença entre as Figuras 9.1 e 9.6. Na Figura 9.1 calcula-se o erro da equação ou erro de predição  $\varepsilon(t)$ , ao passo que na Figura 9.6 calcula-se o erro de simulação ou erro de saída (*output error*  $\varepsilon_{OE}(t)$ ). A relação entre os dois erros pode ser calculada como indicado a seguir. O erro da equação ou erro de predição  $\varepsilon(t)$  é calculado a partir de (9.13), conforme mostrado a seguir:

$$\begin{aligned} \varepsilon(t) &= y(t) - \hat{y}(t) = y(t) - \hat{H}^{-1}(q) \cdot \hat{G}(q) \cdot u(t) - [1 - \hat{H}^{-1}(q)] \cdot y(t) = y(t) - \hat{B}(q) \cdot u(t) - [1 - \hat{A}(q)] \cdot y(t) \\ \therefore \varepsilon(t) &= \hat{A}(q) \cdot y(t) - \hat{B}(q) \cdot u(t) \end{aligned} \quad (9.36)$$

Já o erro de saída  $\varepsilon_{OE}(t)$  é dado por:

$$\varepsilon_{OE}(t) = y(t) - \hat{y}(t) = y(t) - \hat{G}(q) \cdot u(t) = y(t) - \frac{\hat{B}(q)}{\hat{A}(q)} u(t) \quad (9.37)$$

Caso se multiplique ambos os lados de (9.37) por  $\hat{A}(q)$ :

$$\hat{A}(q) \cdot \varepsilon_{OE}(t) = \hat{A}(q) \cdot y(t) - \hat{B}(q) \cdot u(t) \quad (9.38)$$

Portanto, de (9.36) e (9.38) se conclui que:

$$\varepsilon(t) = \hat{A}(q) \cdot \varepsilon_{OE}(t) \quad (9.39)$$

Quando se emprega um modelo com estrutura FIR, como  $\hat{A}(q) = 1$ , resulta que os erros da equação e de saída são idênticos. O vetor de parâmetros a ser determinado é:

$$\hat{\theta} = (\hat{b}_1 \ \hat{b}_2 \ \dots \ \hat{b}_{n_b} \ \hat{f}_1 \ \dots \ \hat{f}_{n_f})^T$$

Como  $x(t)$  em (9.33a) nunca é observado, ele deveria levar um índice  $\hat{\theta}$ , visto ser construído a partir de  $u$  usando (9.33a) (LJUNG, 1999):

$$x(t, \theta) + \hat{f}_1 \cdot x(t-1, \hat{\theta}) + \dots + \hat{f}_{n_f} \cdot x(t-n_f, \hat{\theta}) = \hat{b}_1 \cdot u(t-1) + \dots + \hat{b}_{n_b} \cdot u(t-n_b) \quad (9.40)$$

Comparando-se (9.34) com (9.3) verifica-se que  $\hat{H}(q, \hat{\theta}) = 1$ , o que gera o seguinte preditor um passo à frente:

$$\hat{y}(t/t-1, \hat{\theta}) = \frac{\hat{B}(q)}{\hat{F}(q)} u(t) = x(t, \hat{\theta}) \quad (9.41)$$

Com auxílio do vetor:

$$\varphi^T(t, \hat{\theta}) = [u(t-1) \cdots u(t-n_b) \quad -x(t-1, \hat{\theta}) \cdots -x(t-n_f, \hat{\theta})] \quad (9.42)$$

a expressão (9.41) pode ser reescrita como:

$$\hat{y}(t/t-1, \hat{\theta}) = \varphi^T(t, \hat{\theta}) \hat{\theta} \quad (9.43)$$

A expressão (9.43) se assemelha ao preditor ARMAX (9.28), pois o preditor também é uma regressão pseudo-linear, uma vez que o vetor de regressores  $\varphi(t, \theta)$  da equação (9.28) também é influenciado, de forma não-linear, por  $\theta$ . Perceba que em (9.40) os  $x(t-k, \theta)$  não são observados, mas usando (9.41) eles podem ser computados:

$$x(t-k, \hat{\theta}) = \hat{y}(t-k/t-k-1, \hat{\theta}) \quad k=1, 2, \dots, n_f$$

### 9.3.7 Estrutura de modelo Box-Jenkins

Uma evolução natural da estrutura de erro na saída (9.34) é modelar também as propriedades do erro na saída. Descrevendo-se isto como um modelo ARMA, resulta:

$$y(t) = \frac{\hat{B}(q)}{\hat{F}(q)} u(t) + \frac{\hat{C}(q)}{\hat{D}(q)} \varepsilon(t) \quad (9.44)$$

Esta é a mais natural parametrização de dimensão finita a partir da descrição (9.3): as funções de transferência  $G$  e  $H$  são parametrizadas independentemente como funções racionais, de modo que esta estrutura provê tanto um modelo do processo como da perturbação. A Estrutura (9.44) foi tratada por Box e Jenkins (1970) e seu diagrama de blocos é mostrado na Figura 9.7.

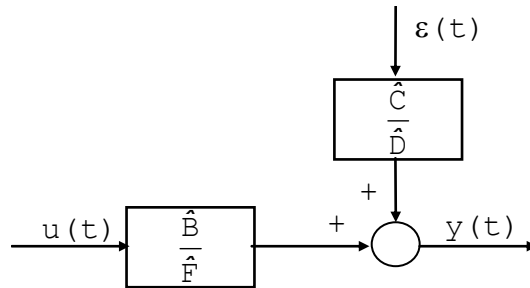


Fig. 9.7 A estrutura de modelo Box-Jenkins (BJ).

De acordo com (5.9), o preditor um passo à frente para (9.44) é:

$$\hat{y}(t/t-1, \hat{\theta}) = \frac{\hat{D}(q) \hat{B}(q)}{\hat{C}(q) \hat{F}(q)} u(t) + \frac{\hat{C}(q) - \hat{D}(q)}{\hat{C}(q)} y(t) \quad (9.45)$$

### 9.3.8 Uma família genérica de estruturas de modelos lineares

As estruturas discutidas nesta seção podem gerar 32 diferentes modelos, dependendo de qual dos cinco polinômios  $A$ ,  $B$ ,  $C$ ,  $D$  e  $F$  é usado. Até aqui foram mostradas algumas das possibilidades. Diversas dessas estruturas de modelo pertencem às mais usadas na prática. Apresenta-se, a seguir, uma estrutura generalizada de modelos (LJUNG, 1999):

$$\hat{A}(q) \cdot y(t) = \frac{\hat{B}(q)}{\hat{F}(q)} u(t) + \frac{\hat{C}(q)}{\hat{D}(q)} \varepsilon(t) \quad E[\varepsilon^2(t)] = \sigma^2 \quad (9.46)$$

Esta estrutura generalizada foi publicada pela primeira vez em (LJUNG, 1987).

Algumas vezes, a dinâmica de  $u$  para  $y$  contém um atraso de  $n_k$  amostras, assim alguns coeficientes de  $\hat{B}(q)$  são zero:

$$\hat{B}(q) = \hat{b}_{n_k} \cdot q^{-n_k} + \hat{b}_{n_k+1} \cdot q^{-n_k-1} + \dots + \hat{b}_{n_k+n_b-1} \cdot q^{-n_k-n_b+1} = q^{-n_k} \cdot \bar{B}(q) \quad b_{n_k} \neq 0$$

$$\text{onde } \bar{B}(q) = \hat{b}_{n_k} + \hat{b}_{n_k+1} \cdot q^{-1} + \dots + \hat{b}_{n_k+n_b-1} \cdot q^{-n_b+1}$$

Expressando-se explicitamente esse atraso:

$$\hat{A}(q) \cdot y(t) = q^{-n_k} \frac{\bar{B}(q)}{\hat{F}(q)} u(t) + \frac{\hat{C}(q)}{\hat{D}(q)} \varepsilon(t) \quad (9.47)$$

Para facilidade de notação, usa-se aqui  $n_k = 1$  em (9.46). A partir dessa expressão pode-se sempre derivar as correspondentes a (9.47), substituindo-se  $u(t)$  por  $u(t - n_k + 1)$ .

A estrutura (9.46) é geral demais para a maioria das aplicações práticas sendo que, normalmente, um ou mais dos cinco polinômios deve ser igualado à unidade para um determinado processo. A partir de (5.9) sabe-se que o preditor um passo à frente para (9.46) é (LJUNG, 1999):

$$\hat{y}(t/t-1, \hat{\theta}) = \frac{\hat{D}(q) \hat{B}(q)}{\hat{C}(q) \hat{F}(q)} u(t) + \left[ 1 - \frac{\hat{D}(q) \hat{A}(q)}{\hat{C}(q)} \right] y(t) \quad (9.48)$$

Os casos comumente usados de (9.46) são descritos na Tabela 9.1 (LJUNG, 1999).

Tab. 9.1 Alguns modelos comuns SISO como casos especiais de (9.46).

Polinômios usados em (9.46)	Nome da estrutura do modelo
$B$	FIR ( <i>finite impulse response</i> )
$A, B$	ARX
$A, B, C$	ARMAX
$A, C$	ARMA
$A, B, D$	ARARX
$A, B, C, D$	ARARMAX
$B, F$	OE (output error)
$B, C, D, F$	BJ (Box-Jenkins)

## 9.4 CASO MULTIVARIÁVEL

Seja agora o caso onde a entrada  $u(t)$  seja um vetor de dimensão  $m$  e a saída  $y(t)$  seja um vetor  $p$ -dimensional. O caso mais simples é a generalização do conjunto de erros da equação (9.5) (LJUNG, 1999):

$$y(t) + A_1 y(t-1) + \dots + A_{n_a} y(t-n_a) = B_1 u(t-1) + \dots + B_{n_b} u(t-n_b) + \varepsilon(t) \quad (9.49)$$

onde  $A_i$  são matrizes  $p \times p$  e  $B_i$  são matrizes  $p \times m$ .

Pode-se introduzir os seguintes polinômios:

$$\begin{cases} \mathbf{A}(q) = \mathbf{I} + \mathbf{A}_1 q^{-1} + \dots + \mathbf{A}_{n_a} q^{-n_a} \\ \mathbf{B}(q) = \mathbf{B}_1 q^{-1} + \dots + \mathbf{B}_{n_b} q^{-n_b} \end{cases}$$

Tem-se então polinômios matriciais em  $q^{-1}$ , significando que  $\mathbf{A}(q)$  seja uma matriz cujos elementos são polinômios em  $q^{-1}$ . O sistema é dado por (LJUNG, 1999):

$$\mathbf{y}(t) = \mathbf{G}(q, \theta) \mathbf{u}(t) + \mathbf{H}(q, \theta) \varepsilon(t)$$

$$\text{com } \mathbf{G}(q, \theta) = \mathbf{A}^{-1}(q) \mathbf{B}(q) \quad \mathbf{H}(q, \theta) = \mathbf{A}^{-1}(q)$$

$\mathbf{G}(q, \theta)$  é uma matriz  $p \times m$ , cujos elementos são funções racionais de  $q^{-1}$  (ou  $q$ ). A fatorização em termos de dois polinômios matriciais é também chamada de descrição fracionária da matriz (*matriz fraction description* - MFD).

A parametrização de (9.49) (isto é, quais elementos das matrizes deveriam ser incluídos no vetor de parâmetros  $\theta$ ) poderia ser feita em analogia direta com (9.8). Suponha todos os elementos a serem incluídos em  $\theta$ . Pode-se então definir a matriz  $[n_a \cdot p + n_b \cdot m] \times p$  (LJUNG, 1999):

$$\theta = (\mathbf{A}_1 \ \mathbf{A}_2 \ \mathbf{A}_3 \ \dots \ \mathbf{A}_{n_a} \ \mathbf{B}_1 \ \dots \ \mathbf{B}_{n_b})^T$$

e o vetor coluna de regressores com dimensão  $[n_a \cdot p + n_b \cdot m]$ :

$$\varphi(t) = \begin{bmatrix} -\mathbf{y}(t-1) \\ \vdots \\ -\mathbf{y}(t-n_a) \\ \mathbf{u}(t-1) \\ \vdots \\ \mathbf{u}(t-n_b) \end{bmatrix}$$

Reescreve-se (9.49) como:

$$\mathbf{y}(t) = \theta^T \varphi(t) + \varepsilon(t)$$

Trata-se de uma analogia com a regressão linear (9.8). Isto pode ser encarado como  $p$  diferentes regressões lineares, escritas umas sobre as outras, todas com o mesmo vetor de regressão.

É fácil visualizar uma série de variantes para o caso MIMO, como a equação de diferenças vetorial (LJUNG, 1999):

$$\begin{aligned} \mathbf{y}(t) + \mathbf{A}_1 \mathbf{y}(t-1) + \dots + \mathbf{A}_{n_a} \mathbf{y}(t-n_a) &= \mathbf{B}_1 \mathbf{u}(t-1) + \dots + \mathbf{B}_{n_b} \mathbf{u}(t-n_b) + \\ &+ \varepsilon(t) + \mathbf{C}_1 \varepsilon(t-1) + \dots + \mathbf{C}_{n_c} \varepsilon(t-n_c) \end{aligned}$$

$$\text{ou } \mathbf{G}(q, \theta) = \mathbf{A}^{-1}(q) \mathbf{B}(q) \quad \mathbf{H}(q, \theta) = \mathbf{A}^{-1}(q) \mathbf{C}(q)$$

que é a extensão natural do modelo ARMAX.

Um modelo Box-Jenkins multivariável assume a forma:

$$\mathbf{G}(q, \theta) = \mathbf{F}^{-1}(q) \mathbf{B}(q) \quad \mathbf{H}(q, \theta) = \mathbf{D}^{-1}(q) \mathbf{C}(q)$$

## 9.5 EXEMPLOS DE DIFERENTES ESTRUTURAS DE MODELOS

### 9.5.1 Exemplo 1

Seja o processo de 1ª ordem  $S$  dado pela seguinte equação:

$$y(t) + a_0 \cdot y(t-1) = b_0 \cdot u(t-1)$$

Assuma que esse processo seja afetado por perturbações, de forma que o sistema que melhor o descreva seja dado pela seguinte equação:

$$y(t) + a_0 \cdot y(t-1) = b_0 \cdot u(t-1) + e(t) + c_0 \cdot e(t-1)$$

onde  $e(t)$  é ruído branco. Dois diferentes conjuntos de valores de parâmetros são usados:

$$S_1 : a_0 = -0,8 \quad b_0 = 1,0 \quad c_0 = 0,0$$

$$S_2 : a_0 = -0,8 \quad b_0 = 1,0 \quad c_0 = -0,8$$

Os sistemas  $S_1$  e  $S_2$  podem ser escritos como:

$$S_1 : y(t) - 0,8 \cdot y(t-1) = 1,0 \cdot u(t-1) + e(t)$$

$$S_2 : \begin{cases} x(t) - 0,8 \cdot x(t-1) = 1,0 \cdot u(t-1) \\ y(t) = x(t) + e(t) \end{cases}$$

O ruído branco  $e(t)$  entra no sistema de diferentes modos. Para o sistema  $S_1$  ele surge como uma perturbação no sistema enquanto que para  $S_2$  ele é adicionado à saída, como visto na Figura 9.8. Note que para o sistema  $S_2$ , o sinal  $x(t)$  pode ser interpretado como a saída determinística ou livre de ruído.

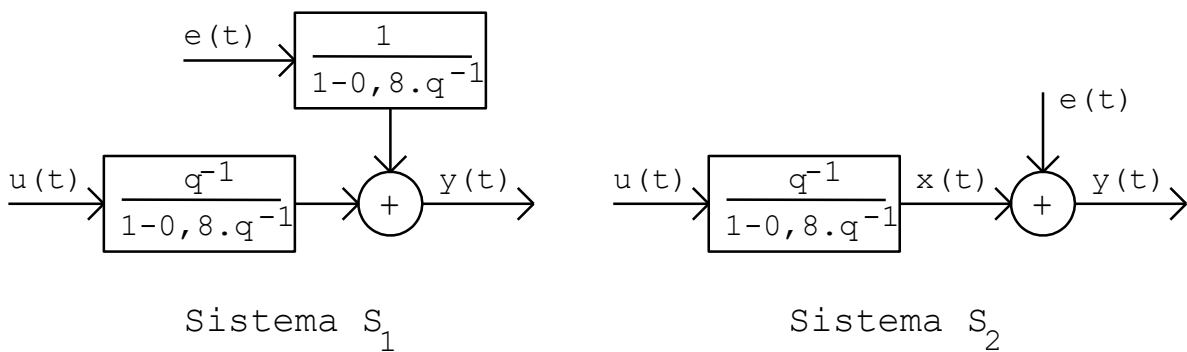


Fig. 9.8 Representação esquemática dos sistemas  $S_1$  e  $S_2$ .

O sistema  $S_1$  corresponde a uma estrutura de modelo do tipo ARX, onde:

$$A(q) = 1 + a_0 \cdot q^{-1} = 1 - 0,8 \cdot q^{-1} \quad B(q) = b_0 \cdot q^{-1} = q^{-1}$$

O sistema  $S_2$  equivale a uma estrutura de modelo do tipo erro na saída (OE) com:

$$B(q) = b_0 \cdot q^{-1} = q^{-1} \quad F(q) = 1 + f_0 \cdot q^{-1} = 1 - 0,8 \cdot q^{-1}$$

**OBSERVAÇÃO:**

Ao se analisar o sistema  $S_2$ , pode-se pensar em classificá-lo como uma estrutura de modelo do tipo ARMAX, onde:

$$A(q) = 1 + a_0 \cdot q^{-1} = 1 - 0,8 \cdot q^{-1} \quad B(q) = b_0 \cdot q^{-1} = q^{-1} \quad C(q) = 1 + c_0 \cdot q^{-1} = 1 - 0,8 \cdot q^{-1}$$

Percebe-se então que, neste caso,  $A(q) = C(q)$ , o que torna:



$$H(q, \theta) = \frac{C(q)}{A(q)} = 1$$

Resulta:

$$y(t) = G(q, \theta) \cdot u(t) + H(q, \theta) \cdot e(t) = \frac{B(q)}{A(q)} u(t) + \frac{C(q)}{A(q)} e(t) = \frac{q^{-1}}{1 - 0,8 \cdot q^{-1}} u(t) + e(t)$$

Assim, neste caso, não se trata de uma estrutura ARMAX, mas sim de erro na saída.

Para realizar a identificação paramétrica, considere a estrutura de modelo  $M$  dada pela equação de diferenças:

$$y(t) + a \cdot y(t-1) = b \cdot u(t-1) + \varepsilon(t) \quad (9.50)$$

Esta estrutura de modelo é um conjunto de modelos temporais discretos lineares de 1ª ordem. O vetor de parâmetros é definido como:

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}$$

Na Equação (9.47)  $\varepsilon(t)$  é chamado de erro da equação ou resíduo e descreve o desvio nos dados de um sistema linear de 1ª ordem perfeito (determinístico).  $\varepsilon(t)$  pode ser encarado como função do vetor de parâmetros  $\theta$ :

$$\varepsilon(t) = y(t) + a \cdot y(t-1) - b \cdot u(t-1) = y(t) - [-y(t-1) \quad u(t-1)]^T \theta = y(t) - \varphi(t)^T \theta$$

Percebe-se que estender o modelo (9.47) para um sistema linear de ordem  $n$  é simples, bastando adicionar termos  $a_i \cdot y(t-i)$  e  $b_i \cdot u(t-i)$  para  $i = 2, \dots, n$ .

Ao efetuar a identificação, a saída resultante do modelo  $x_m(t)$  é dada por:

$$x_m(t) + a \cdot x_m(t-1) = b \cdot u(t-1)$$

Seria de se esperar que  $x_m(t)$  fosse próximo à saída  $x(t)$  sem ruído. No entanto, este sinal não é conhecido ao usuário na maioria das vezes. Ao invés disso pode-se comparar  $x_m(t)$  ao sinal medido de saída  $y(t)$ . Para um bom modelo, o sinal  $x_m(t)$  deveria explicar todos os padrões em  $y(t)$ , que são devidos à entrada. No entanto,  $x_m(t)$  não deve ser igual a  $y(t)$ , visto que a saída também tem um componente que é gerado pelas perturbações. Esse componente estocástico deveria aparecer na saída do modelo  $x_m(t)$ .

Processos estocásticos são aqueles afetados por ruído aleatório. Todos os processos práticos são estocásticos, mas quando o ruído é desprezível, os processos são normalmente considerados como determinísticos (SÖDERSTRÖM; STOICA, 1989).

### 9.5.2 Exemplo 2

Classifique as estruturas de cada um dos três sistemas apresentados na Seção 4.2. O sistema  $S_1$  é descrito pelo seguinte modelo:

$$y(t) = \frac{0,1463 \cdot q^{-5}}{1 - 0,9512 \cdot q^{-1}} u(t) + \frac{1}{1 - 0,9512 \cdot q^{-1}} e(t) = G(q) \cdot u(t) + H_1(q) \cdot e(t)$$

Tem-se que:

$$G(q) = \frac{0,1463 \cdot q^{-5}}{1 - 0,9512 \cdot q^{-1}} \quad \text{e} \quad H_1(q) = \frac{1}{1 - 0,9512 \cdot q^{-1}}$$

Percebe-se que se trata de uma estrutura do tipo ARX. O sistema  $S_2$  é dado por:

$$y(t) = \frac{0,1463 \cdot q^{-5}}{1 - 0,9512 \cdot q^{-1}} u(t) + \frac{1 + 0,5 \cdot q^{-1}}{1 - 0,9512 \cdot q^{-1}} e(t) = G(q) \cdot u(t) + H_2(q) \cdot e(t)$$

Neste caso, reconhece-se uma estrutura ARMAX. O sistema  $S_3$  é descrito por:

$$y(t) = \frac{0,1463 \cdot q^{-5}}{1 - 0,9512 \cdot q^{-1}} u(t) + \frac{1 - 0,9512 \cdot q^{-1}}{1 - 0,9512 \cdot q^{-1}} e(t) = \frac{0,1463 \cdot q^{-5}}{1 - 0,9512 \cdot q^{-1}} u(t) + e(t) = G(q) \cdot u(t) + e(t)$$

Verifica-se, neste caso, que a estrutura resultante é do tipo OE.

## 9.6 MODELOS NA FORMA DE ESPAÇO DE ESTADOS

Um sistema dinâmico pode ser descrito usando modelos internos ou externos. Modelos internos, como por exemplo, os modelos de espaço de estados, descrevem todos os acoplamentos internos entre as variáveis do sistema. Modelos externos, como por exemplo, funções de transferência, fornecem apenas a relação entre a entrada e a saída do sistema (ÅSTRÖM; WITTENMARK, 1997).

Como já se viu, a relação entrada/saída para um sistema linear genérico pode ser expressa por uma função de resposta ao impulso ou função-peso. Foi visto que cálculos usando operadores de deslocamento ( $q$ ) podem ser usados para gerar relações de entrada/saída, que levam à caracterização do comportamento de entrada/saída em termos de operadores de transferência de pulso (ÅSTRÖM; WITTENMARK, 1997). Na forma de espaço de estados, a relação entre entrada, ruído e saída é escrita como um sistema de 1ª ordem de equações diferenciais ou de diferenças, usando um vetor auxiliar de estados  $x(t)$ . Esta forma é especialmente útil para melhor visualizar o que ocorre em sistemas físicos, sendo, nesse aspecto, mais apropriada que os modelos descritos na seção anterior.

### 9.6.1 Modelos de tempo contínuo baseados no enfoque físico

Para a maioria dos sistemas físicos é mais fácil construir modelos com enfoque físico em tempo contínuo que em tempo discreto, isto porque a maioria das leis da Física (leis de Newton, relações em circuitos elétricos etc) são expressas em tempo contínuo. Isso implica que a modelagem normalmente conduz a uma representação na forma (LJUNG, 1999):

$$\dot{x}(t) = F(\theta) x(t) + G(\theta) u(t) \quad (9.51)$$

$F$  e  $G$  são matrizes com dimensão  $n \times n$  e  $n \times m$ , respectivamente, para um estado  $n$ -dimensional e uma entrada  $m$ -dimensional.  $\theta$  é um vetor de parâmetros que tipicamente corresponde a valores desconhecidos dos coeficientes físicos, constantes de material etc. O modelamento é normalmente efetuado em termos de variáveis de estado  $x$  que possuem um significado físico (posição, velocidade etc) e então as saídas medidas são combinações conhecidas dos estados. Seja  $\eta(t)$  as medições que seriam obtidas com sensores ideais, livres de ruído (LJUNG, 1999):

$$\eta(t) = H x(t) \quad (9.52)$$

Usando  $p$  como operador de diferenciação, (9.51) pode ser escrita como:

$$[pI - F(\theta)] x(t) = G(\theta) u(t)$$

o que significa que o operador de transferência de  $u$  para  $\eta$  em (9.52) é:

$$\eta(t) = G_C(p, \theta) u(t)$$

$$G_C(p, \theta) = H[pI - F(\theta)]^{-1} G(\theta)$$

Obteve-se assim um modelo do sistema através de uma função de transferência em tempo contínuo, que é parametrizada em termos de coeficientes físicos.

Em verdade, uma versão com alguma corrupção por ruído de  $\eta(t)$  é obtida, resultante de imperfeições na medição e perturbações em (9.51). Há diversas opções diferentes para descrever esses efeitos do ruído e da perturbação. Aqui é mostrada a abordagem mais simples. Assuma que as medições sejam amostradas nos instantes  $t = k \cdot T$ ,  $k = 1, 2, \dots$  e os efeitos das perturbações nesses instantes sejam  $v_T(k \cdot T)$ . Então a saída medida é:

$$y(k \cdot T) = Hx(k \cdot T) + v_T(k \cdot T) = G_C(p, \theta) u(t) + v_T(k \cdot T) \quad (9.53)$$

### 9.6.2 Amostrando a função de transferência

Conforme visto na Seção 2.1, há diversos modos de transformar  $G_C(p, \theta)$  em uma representação que seja explicitamente de tempo discreto. Suponha que a entrada seja constante ao longo do intervalo de amostragem  $T$  como em (2.12):

$$u(t) = u_k = u(k \cdot T) \quad k \cdot T \leq t < (k+1) \cdot T \quad (9.54)$$

Então, a equação diferencial (9.51) pode ser facilmente resolvida de  $t = k \cdot T$  a  $t = k \cdot T + T$ , produzindo (LJUNG, 1999):

$$x(k \cdot T + T) = A_T(\theta) x(k \cdot T) + B_T(\theta) u(k \cdot T) \quad (9.55)$$

onde:

$$\begin{cases} A_T(\theta) = e^{F(\theta)T} \\ B_T(\theta) = \int_{\tau=0}^T e^{F(\theta)\tau} G(\theta) d\tau \end{cases} \quad (9.56)$$

Introduzindo-se  $q$  para o deslocamento à frente de  $T$  unidades de tempo, pode-se reescrever (9.55) como:

$$[qI - A_T(\theta)] x(kT) = B_T(\theta) u(kT)$$

ou

$$\eta(kT) = G_T(q, \theta) u(kT)$$

$$G_T(q, \theta) = H[qI - A_T(\theta)]^{-1} B_T(\theta)$$

Assim, (9.53) pode equivalentemente ser fornecida na forma de dados amostrados:

$$y(t) = G_T(q, \theta) u(t) + v_T(t) \quad t = T, 2T, 3T \dots$$

Quando (9.54) é válida, nenhuma aproximação é envolvida nesta representação. No entanto, em vista de (9.56),  $G_T(q, \theta)$  pode ser uma função bastante complicada de  $\theta$ .

As Equações (9.55) e (9.53) constituem um modelo no espaço de estados em tempo discreto. Por simplicidade, faz-se  $T=1$ . Introduce-se ainda uma parametrização arbitrária da matriz que relaciona  $x$  a  $\eta$ :  $H = C(\theta)$ . Obtém-se (LJUNG, 1999):

$$x(t+1) = A(\theta) x(t) + B(\theta) u(t) \quad (9.57a)$$

$$y(t) = C(\theta) x(t) + v(t) \quad (9.57b)$$

correspondente a:

$$y(t) = G(q, \theta) u(t) + v(t) \quad (9.58)$$

$$G(q, \theta) = C(\theta) [qI - A(\theta)]^{-1} B(\theta)$$

Embora amostrar uma descrição em tempo contínuo seja uma forma natural de obter o modelo (9.57), poder-se-ia também, para certas aplicações, gerá-la diretamente em tempo discreto, com as matrizes  $A$ ,  $B$  e  $C$  diretamente parametrizadas em termos de  $\theta$ , ao invés de indiretamente através de (9.56).

### 9.6.3 Representação da perturbação em modelos de espaço de estados

Na representação (9.57) e (9.58) poder-se-ia também modelar as propriedades do ruído  $\{v(t)\}$ . Uma opção seria postular um modelo para o ruído na forma (LJUNG, 1999):

$$v(t) = H(q, \theta) e(t)$$

sendo  $\{e(t)\}$  ruído branco com variância  $\lambda^2$ . Os parâmetros  $\theta$  em  $H(q, \theta)$  poderiam ser parcialmente em comum com aqueles de  $G(q, \theta)$  ou serem parâmetros adicionais do modelo de perturbação.

Para descrições de espaço de estados é, contudo, mais comum dividir o termo de ruído concentrado em contribuições do **ruído de medição**  $v(t)$  e do **ruído de processo**  $w(t)$  agindo nos estados, de modo que (9.57) é escrito (LJUNG, 1999):

$$\begin{cases} x(t+1) = A(\theta) x(t) + B(\theta) u(t) + w(t) \\ y(t) = C(\theta) x(t) + v(t) \end{cases} \quad (9.59)$$

Aqui,  $\{w(t)\}$  e  $\{v(t)\}$  são assumidos como sendo sequências de variáveis aleatórias independentes com valores de média nulos e covariâncias:

$$\begin{aligned} E[w(t) w^T(t)] &= R_1(\theta) \\ E[v(t) v^T(t)] &= R_2(\theta) \\ E[w(t) v^T(t)] &= R_{12}(\theta) \end{aligned} \quad (9.60)$$

As perturbações  $w(t)$  e  $v(t)$  podem frequentemente ser sinais cujas origens físicas sejam conhecidas. Neste caso, pode não ser realista assumi-las como ruído branco. Para se chegar a (9.59) e (9.60) requer uma modelagem extra e uma extensão do vetor de estado.

Verificou-se que os modelos de espaço de estados provêm uma gama de possibilidades de modelamento: pode-se usar o modelamento físico em tempo contínuo com ou sem uma correspondente descrição de ruído em tempo contínuo para obter estruturas com parâmetros físicos  $\theta$ . Pode-se usar parametrização física da parte dinâmica combinada com uma parametrização de caixa-preta das propriedades de ruído. Finalmente, pode-se usar estruturas de espaço de estados de caixa-preta. Estas têm a vantagem sobre

a caixa-preta de entrada/saída que a flexibilidade na escolha da representação pode assegurar propriedades numéricas melhores da parametrização (LJUNG, 1999).

### REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas** – técnicas lineares e não-lineares aplicadas a sistemas reais. 3ed. Belo Horizonte, Editora UFMG, 2007.
- ÅSTRÖM, K. J. **Introduction to stochastic control theory**. New York, Academic Press, 1970.
- ÅSTRÖM, K. J.; BOHLIN, T. Numerical identification of linear dynamic systems from normal operating records. In: **Proceedings** of the IFAC Symposium on Self-Adaptive Systems, Teddington, England, 1965.
- ÅSTRÖM, K. J.; WITTENMARK, B. **Computer-controlled systems** - theory and design. 3.ed. Englewood Cliffs, Prentice Hall, 1997.
- BOX, G. E. P.; JENKINS, C. W. **Time series analysis: forecasting and control**. São Francisco, Holden-Day, 1970.
- CLARKE, D. W. Generalized least squares estimation of parameters of a dynamic model. In: **Proceedings** of the First IFAC Symposium on Identification in Automatic Control Systems, Prague, paper 3.17, 1967.
- LJUNG, L. **System identification: theory for the user**. Englewood Cliffs, Prentice Hall, 1987.
- LJUNG, L. **System identification: theory for the user**. 2.ed., Englewood Cliffs, Prentice Hall, 1999.
- SÖDERSTRÖM, T.; STOICA, P. **System identification**. Hemel Hempstead, U.K., Prentice Hall International, 1989.