CAPÍTULO 11

PROPRIEDADES ESTATÍSTICAS DO ESTIMADOR DOS MÍNIMOS QUADRADOS

O método dos mínimos quadrados é um caso especial do método de identificação do erro de predição (PEM), descrito na Subseção 10.2.2. Neste capítulo, analisam-se suas propriedades estatísticas, quando aplicado para estimar os parâmetros de diversas estruturas de modelos lineares.

11.1 PROPRIEDADES ESTATÍSTICAS DE ESTIMADORES

Como a estimativa $\hat{\boldsymbol{\theta}}$ pode ser considerada uma variável aleatória, sua precisão pode ser medida por diversas propriedades estatísticas, tais como consistência, polarização (*bias*) e eficiência.

Diz-se que uma estimativa $\hat{\theta}_N$ é **consistente** (*consistent*) se (SÖDERSTRÖM; STOICA, 1989):

$$\hat{\theta}_N \to \theta \qquad \text{conforme } N \to \infty$$

onde θ corresponde ao conjunto de parâmetros verdadeiros do sistema.

Pode-se entender que a consistência implica que o efeito das perturbações na estimação dos parâmetros pode ser "eliminado", tomando-se mais pontos de medição da entrada e saída.

Uma estimativa $\hat{\theta}_N$ é **polarizada** (*biased*) se seu valor esperado desvia do valor verdadeiro θ, isto é (SÖDERSTRÖM; STOICA, 1989):

$$\mathrm{E}\left(\hat{\theta}_N\right) \neq \theta$$

A diferença $\mathrm{E}\left(\hat{\theta}_N\right) - \theta$ é chamada de *bias*. Caso se tenha:

$$\mathrm{E}\left(\hat{\theta}_N\right) = \theta$$

diz-se que $\hat{\theta}_N$ é não polarizado (*unbiased*).

Diz-se que um estimador é **eficiente** ou de **mínima variância** se para todos os estimadores não polarizados $\hat{\theta}^*$, tem-se que:

$$\mathrm{cov}\left(\hat{\theta}\right) \leq \mathrm{cov}\left(\hat{\theta}^*\right) \qquad \text{ou}$$

$$\det\left[\operatorname{cov}(\hat{\theta})\right] - \det\left[\operatorname{cov}(\hat{\theta}^*)\right] \leq 0$$

onde a matriz de covariância é definida por:

$$\operatorname{cov}(\hat{\theta}) = \mathrm{E}\left\{\left[\hat{\theta} - \mathrm{E}(\hat{\theta})\right]^{\mathrm{T}}\left[\hat{\theta} - \mathrm{E}(\hat{\theta})\right]\right\}$$

Se esta propriedade se aplica apenas se $N \to \infty$, então ela é chamada eficiência assintótica.

O desejável é que a estimativa $\hat{\theta}_N$ seja consistente, não polarizada e eficiente.

11.2 SISTEMA IDENTIFICÁVEL

Uma questão fundamental na identificação de sistemas se refere ao conceito de **identificabilidade**, que analisa se as características em malha aberta do sistema podem ou não ser obtidas conforme o número de dados coletados tenda a infinito. De um modo informal, a identificabilidade se relaciona com o problema de saber se o procedimento de identificação irá produzir um valor único do vetor de parâmetros $\hat{\boldsymbol{\theta}}$ e/ou se o modelo resultante é igual ao sistema real. A identificabilidade está relacionada com o fato do conjunto de dados (as condições experimentais) ser suficientemente informativo para distinguir entre diferentes modelos, bem como com propriedades da própria estrutura do modelo: se os dados são suficientemente informativos para distinguir entre modelos diferentes, então a questão é se diferentes valores de $\hat{\boldsymbol{\theta}}$ podem gerar modelos com comportamento igual (LJUNG, 1999).

Há diferentes definições na literatura sobre o conceito de identificabilidade. A mais comum é relacionar a propriedade de identificabilidade à consistência da estimativa do parâmetro $\hat{\theta}_N$. De um modo informal, um sistema é dito **identificável** (em um dado conjunto de modelos) se a estimativa dos parâmetros do modelo for consistente. As propriedades de identificabilidade de um dado sistema dependem da estrutura de modelo escolhida *M*, do método de identificação *I* e das condições experimentais *H*.

11.3 ANÁLISE DAS PROPRIEDADES DO ESTIMADOR DOS MÍNIMOS QUADRADOS APLICADO À ESTRUTURA DE MODELO DA FAMÍLIA ERRO NA EQUAÇÃO

Analisam-se, a seguir, as propriedades do estimador dos mínimos quadrados quando aplicado à estrutura de modelo da família erro na equação.

11.3.1 Análise da consistência da estimativa dos mínimos quadrados quando aplicada à estrutura de modelo da família erro na equação

Assuma que os dados coletados de um processo SISO com estrutura do tipo erro na equação tenham sido gerados por um sistema descrito pela seguinte equação:

$$y(t) = \varphi^{\mathrm{T}}(t)\,\theta + v(t) \qquad \text{(forma de regressão linear)} \tag{11.1}$$

onde $\varphi^{\mathrm{T}}(t) = \left[-y(t-1) \ \cdots \ -y(t-n_a) \quad u(t-1) \ \cdots \ u(t-n_b)\right]$

$$\theta^{\mathrm{T}} = \left[a_1 \ \cdots \ a_{n_a} \ \ b_1 \ \cdots \ b_{n_b}\right]$$

sendo que $v(t)$ pode assumir diferentes formas, dependendo da estrutura escolhida, como por exemplo:

$$v(t) = e(t) \quad \text{(estrutura ARX)} \qquad\qquad v(t) = C(q)\cdot e(t) \quad \text{(estrutura ARMAX)}$$

ou, equivalentemente:

$$A(q)\cdot y(t) = B(q)\cdot u(t) + v(t) \qquad \text{(forma de função de transferência)} \qquad (11.1a)$$

onde $A(q) = 1 + a_1 \cdot q^{-1} + \ldots + a_{n_a} \cdot q^{-n_a}$ e $B(q) = b_1 \cdot q^{-1} + \ldots + b_{n_b} \cdot q^{-n_b}$

sendo que $A(q)$ e $B(q)$ são os polinômios reais do denominador e numerador, respectivamente e $v(t)$ é a perturbação de saída, que se supõe ser um processo estocástico estacionário de média nula. Pode-se pensar em θ como um "valor real" do vetor de parâmetros.

Suponha que os dados sejam coletados de um experimento em malha aberta, de modo que a perturbação $v(t)$ e a entrada $u(t)$ sejam independentes. Inserindo-se (11.1) e (10.13) em (10.12), resulta (LJUNG, 1999):

$$\hat{\theta}_N = [\boldsymbol{R}(N)]^{-1}\left\{\sum_{t=1}^{N}\varphi(t)\left[\varphi^{\mathrm{T}}(t)\,\theta + v(t)\right]\right\} = \theta + [\boldsymbol{R}(N)]^{-1}\left[\sum_{t=1}^{N}\varphi(t)\cdot v(t)\right] \qquad (11.2)$$

Nota-se primeiro que se $v(t)$ em (11.2) for pequeno quando comparado a $\varphi(t)$, então o termo de erro:

$$[\boldsymbol{R}(N)]^{-1}\left[\sum_{t=1}^{N}\varphi(t)\cdot v(t)\right]$$

será pequeno e assim $\hat{\theta}_N$ estará próximo de θ.

Para analisar o que ocorre quando N tende a infinito, assume-se que $\{v(t)\}$ seja uma realização de um processo estocástico estacionário e se atenha à situação da Equação (10.9). Assumindo que a entrada u seja um processo estocástico estacionário, somatórios do tipo:

$$\hat{R}_u(\tau) = \sum_{t=1}^{N} u(t)\cdot u(t-\tau) \quad \to \quad R_u(\tau) = \mathrm{E}[u(t)\cdot u(t-\tau)]$$

convergem quando $N \to \infty$. Então a matriz $\boldsymbol{R}(N)$ (que é constituída por tais somatórios e outros relacionados) irá convergir:

$$\boldsymbol{R}(N) \to \boldsymbol{R}^* \quad \left(= \mathrm{E}\left[\varphi(t)\cdot\varphi^{\mathrm{T}}(t)\right]\right) \qquad \text{conforme } N \to \infty$$

Também:

$$\sum_{t=1}^{N}\varphi(t)\cdot v(t) \ \to \ \boldsymbol{h}^* \qquad (= \mathrm{E}[\varphi(t)\cdot v(t)]) \qquad \text{conforme } N \to \infty$$

Assim:

$$\hat{\boldsymbol{\theta}}_N \ \to \ \boldsymbol{\theta} + \left(\boldsymbol{R}^*\right)^{-1}\boldsymbol{h}^* \qquad \text{conforme } N \to \infty$$

desde que $\boldsymbol{R}^*$ seja não singular.

Para que a estimativa dos mínimos quadrados seja consistente, deve-se exigir que (LJUNG, 1999; SÖDERSTRÖM; STOICA, 1989):

a. $\boldsymbol{R}^*$ seja não singular

Isto ocorre se $\{u(t)\}$ e $\{v(t)\}$ forem independentes e se a entrada $u(t)$ for persistentemente excitante de ordem $n_a + n_b$. $\boldsymbol{R}^*$ é não singular na maioria dos casos. As exceções para que $\boldsymbol{R}^*$ seja não singular são:

- a entrada não seja persistentemente excitante de ordem $n_a + n_b$;
- os dados sejam completamente livres de ruído ($v(t) \equiv 0$) e a ordem do modelo seja escolhida muito alta (o que implica que $A(q)$ e $B(q)$ tenham fatores comuns); e
- a entrada $u(t)$ seja gerada por uma realimentação linear de baixa ordem da saída.

b. $\boldsymbol{h}^* = 0$

Diferente da condição “a”, esta condição não é satisfeita na maioria dos casos. Ela é satisfeita apenas se:

b1. $\{v(t)\}$ for uma sequência de variáveis aleatórias independentes com média nula (ruído branco). Então $v(t)$ não depende do que tenha ocorrido até o instante $t-1$ e, sendo não correlacionada com todos os dados do passado e, em particular, com $\varphi(t)$, resulta $\mathrm{E}[\varphi(t)\cdot v(t)] = 0$; ou

b2. a sequência de entrada $\{u(t)\}$ for independente da sequência $\{v(t)\}$ com média zero e $n_a = 0$ em (10.9) (estrutura FIR). Então $\varphi(t)$ contém apenas termos u e daí $\mathrm{E}[\varphi(t)\cdot v(t)] = 0$.

Quando $n_a > 0$, de modo que $\varphi(t)$ contenha $y(k)$, $t-1 \le k \le t-n_a$ e $\{v(t)\}$ não seja ruído branco, então normalmente $\mathrm{E}[\varphi(t)\cdot v(t)] \ne 0$. Isso ocorre porque $\varphi(t)$ contém $y(t-1)$, que por sua vez contém o termo $v(t-1)$, que é correlacionado com $v(t)$. Portanto, pode-se esperar consistência apenas nos casos “b1” e “b2”. Assim, a estimação consistente de θ requer que $v(t)$ seja ruído branco (LJUNG, 1999; SÖDERSTRÖM; STOICA, 1989). No entanto, assumir que a perturbação $v(t)$ seja ruído branco é muito restritivo e não realista (ZHU; BACKX, 1993).

11.3.2 Análise da polarização da estimativa dos mínimos quadrados quando aplicada à estrutura de modelo da família erro na equação

Analisa-se, a seguir, como a estimativa dos mínimos quadrados pode ser não polarizada, quando aplicada à estrutura de modelo da família erro na equação. Optou-se por lidar com notação matricial nesta subseção. Neste caso, substituindo-se (11.1) e (10.18) em (10.17), resulta:

$$\hat{\boldsymbol{\theta}}_N = \boldsymbol{\theta} + \left(\boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{\Phi}_N\right)^{-1} \boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{V}$$

Tomando-se a expectância de ambos os lados desta equação:

$$\mathrm{E}\left(\hat{\boldsymbol{\theta}}_N\right) = \mathrm{E}(\boldsymbol{\theta}) + \mathrm{E}\left[\left(\boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{\Phi}_N\right)^{-1}\boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{V}\right] \tag{11.3}$$

Considere as seguintes hipóteses:

a. o termo $v(t)$ é um processo estocástico estacionário com média nula, isto é,

$\mathrm{E}[v(t)] = 0$; e

b. o termo $v(t)$ é não correlacionado com os sinais do vetor de dados, isto é, $\mathrm{E}\left[\boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{V}\right] = 0$.

As hipóteses "a" e "b" indicam que $v(t)$ seja ruído branco. Aplicando-se a hipótese "b" a (11.3), tem-se que:

$$\mathrm{E}\left(\hat{\boldsymbol{\theta}}_N\right) = \mathrm{E}(\boldsymbol{\theta}) + \mathrm{E}\left[\left(\boldsymbol{\Phi}_N{}^{\mathrm{T}}\boldsymbol{\Phi}_N\right)^{-1}\boldsymbol{\Phi}_N{}^{\mathrm{T}}\right] \cdot \mathrm{E}(\boldsymbol{V})$$

Aplicando-se a hipótese "a" à equação acima, resulta:

$$\mathrm{E}\left(\hat{\boldsymbol{\theta}}_N\right) = \boldsymbol{\theta}$$

Portanto, sob as hipóteses "a" e "b", isto é, que a perturbação $v(t)$ seja ruído branco, o estimador dos mínimos quadrados é não polarizado.

Pergunta-se então sob que condições o estimador dos mínimos quadrados de uma estrutura de modelo da família de erro na equação é não polarizado e consistente? Para responder esta pergunta, considere a expressão (11.2), repetida a seguir, onde se considera que a ordem dos polinômios $A(q)$ e $B(q)$ seja n:

$$\hat{\theta}_N - \theta = [\boldsymbol{R}(N)]^{-1}\left[\sum_{t=1}^{N}\varphi(t)\cdot v(t)\right] \tag{11.2}$$

Assuma que: $y(t) = y^{\circ}(t) + v(t)$ onde $y^{\circ}(t)$ é a saída livre de ruído.

De (11.2) resulta:

$$\hat{\theta}_N - \theta = [\mathbf{R}(N)]^{-1}\left\{\sum_{t=1}^{N}\begin{bmatrix}-y^{\circ}(t-1)\\ \vdots \\ -y^{\circ}(t-n)\\ u(t-1)\\ \vdots \\ u(t-n)\end{bmatrix}\cdot v(t) + \sum_{t=1}^{N}\begin{bmatrix}-v(t-1)\\ \vdots \\ -v(t-n)\\ 0\\ \vdots \\ 0\end{bmatrix}\cdot v(t)\right\}$$

Quando $N \to \infty$, o primeiro termo do lado direito da igualdade tende a zero, quando se trata de um experimento em malha aberta e o segundo termo tende a:

$$[\boldsymbol{R}(N)]^{-1}\begin{bmatrix}\mathrm{E}[v(t-1)\cdot v(t)]\\ \vdots \\ \mathrm{E}[v(t-n)\cdot v(t)]\\ 0\\ \vdots \\ 0\end{bmatrix} \tag{11.4}$$

Esta é a polarização assintótica do estimador dos mínimos quadrados, a qual é nula apenas se a perturbação $v(t)$ for ruído branco. Esta é uma condição fortemente restritiva, que não é satisfeita em situações práticas, sendo portanto não realista, pois é extremamente improvável encontrar tal condição. Nota-se que o vetor de dados $\varphi(t)$ está correlacionado com a entrada $v(t)$, pois a saída $y(t)$ contém a perturbação $v(t)$. Isto implica que a condição "b" após a Equação (11.3) não é válida. Portanto, em geral, o estimador dos mínimos quadrados de modelos com estrutura da família de erro na equação é polarizado e não consistente. No entanto, tentar obter resíduos do tipo ruído branco é um caminho para um estimador não polarizado e consistente (ZHU; BACKX, 1993).

11.3.3 Distribuição do erro no domínio da frequência do estimador dos mínimos quadrados aplicado à estrutura de modelo da família erro na equação

Até aqui, foram estudadas as propriedades do método dos mínimos quadrados no que tange aos parâmetros de modelos com estrutura de erro na equação. No entanto, em aplicações práticas, se está mais preocupado com o comportamento do modelo no domínio da frequência, pois ao se lidar com controle e simulação, o vetor de parâmetros é um meio para se obter um bom modelo. Se a ordem do modelo diferir da ordem do processo real, o que é muito frequente na prática, não faz sentido falar sobre erros em parâmetros. O que realmente importa é o erro gerado pelo modelo com relação ao sistema real.

Assuma que um experimento em malha aberta tenha sido realizado. A função-custo da estimativa dos mínimos quadrados é dada por:

$$V_N = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t)^2 = \frac{1}{N}\sum_{t=1}^{N} \left[\hat{A}(q)\cdot y(t) - \hat{B}(q)\cdot u(t)\right]^2$$

Quando $N \to \infty$:

$$V_N \to \mathrm{E}\left[\hat{A}(q)\cdot y(t) - \hat{B}(q)\cdot u(t)\right]^2$$

Suponha que o sistema real seja descrito por:

$$y(t) = G(q)\cdot u(t) + v(t)$$

Substituindo-se esta expressão na equação de V_N, resulta:

$$V_N \to \mathrm{E}\left\{\hat{A}(q)\cdot\left[G(q)\cdot u(t)+v(t)\right] - \hat{B}(q)\cdot u(t)\right\}^2 = \mathrm{E}\left\{\hat{A}^2(q)\cdot\left[\left(G(q)-\frac{\hat{B}(q)}{\hat{A}(q)}\right)\cdot u(t)+v(t)\right]^2\right\}$$

Como se assume independência entre $u(t)$ e $v(t)$, isto é, $\mathrm{E}\left[u(t)\cdot v(t)\right]=0$, resulta:

$$V_N \to \mathrm{E}\left[\hat{A}^2(q)\cdot\left(G(q)-\frac{\hat{B}(q)}{\hat{A}(q)}\right)^2\cdot u^2(t)\right] + \mathrm{E}\left[\hat{A}^2(q)\cdot v^2(t)\right]$$

Aplicando-se a fórmula de Parseval, obtém-se a função-custo assintótica da estimativa dos mínimos quadrados no domínio da frequência:

$$V_N \to \frac{1}{2\cdot\pi}\int_{-\pi}^{\pi}\left[G\left(\mathrm{e}^{\mathrm{j}\omega T}\right)-\frac{\hat{B}\left(\mathrm{e}^{\mathrm{j}\omega T}\right)}{\hat{A}\left(\mathrm{e}^{\mathrm{j}\omega T}\right)}\right]^2 \hat{A}^2\left(\mathrm{e}^{\mathrm{j}\omega T}\right)\Phi_u(\omega)\,d\omega + \frac{1}{2\cdot\pi}\int_{-\pi}^{\pi}\hat{A}^2\left(\mathrm{e}^{\mathrm{j}\omega T}\right)\Phi_v(\omega)\,d\omega \quad (11.5)$$

onde $G\left(\mathrm{e}^{\mathrm{j}\omega T}\right)$ é a função de transferência do sistema real e $\Phi_u(\omega)$ e $\Phi_v(\omega)$ são os espectros de potência da entrada e da perturbação, respectivamente.

O primeiro termo de (11.5) mostra que na função-custo V_N, o filtro $\hat{A}^2(q)$ é parte da ponderação na aproximação do modelo. Para a maioria dos processos industriais, $\hat{A}(q)$ é um filtro passa-altas, visto que $1/\hat{A}(q)$ é um filtro passa-baixas. Portanto, o efeito de $\hat{A}^2(q)$ é impor uma forte ponderação em altas frequências. Isto certamente não é de-sejável em aplicações de modelos, tais como controle e simulação (ZHU; BACKX, 1993).

11.4 ANÁLISE DA CONSISTÊNCIA E DA POLARIZAÇÃO DO ESTIMADOR DOS MÍNIMOS QUADRADOS APLICADO À ESTRUTURA DE MODELO FIR

Na estimação do modelo FIR, assume-se que os dados sejam gerados pela Expressão (11.1), onde:

$$\varphi^{\mathrm{T}}(t) = [u(t-1) \;\; u(t-2) \;\; \cdots \;\; u(t-n_b)] \qquad \theta^{\mathrm{T}} = [b_1 \;\; b_2 \;\; \cdots \;\; b_{n_b}]$$

ou, equivalentemente:

$$y(t) = B(q) \cdot u(t) + e(t)$$

Assumindo-se que o sinal de entrada seja persistentemente excitante de ordem n_b, pode-se afirmar que a estimativa dos mínimos quadrados do modelo FIR seja não polarizada e consistente se a perturbação $e(t)$ for indepedente da entrada $u(t)$. Para isto ocorrer, não é necessário que a perturbação $e(t)$ seja ruído branco. Na prática, a independência entre a entrada $u(t)$ e a perturbação $e(t)$ é consequência de experimentos de identificação em malha aberta, isto é, em que não haja realimentação de $y(t)$ para $u(t)$. Isto está de acordo com a condição “b2” da Subseção 11.3.1. Assim, se os dados são coletados em um ensaio em malha fechada, a estimativa FIR é polarizada (ZHU, BACKX, 1993).

11.5 ANÁLISE DA CONSISTÊNCIA E DA POLARIZAÇÃO DO ESTIMADOR DOS MÍNIMOS QUADRADOS APLICADO À ESTRUTURA DE MODELO DE ERRO NA SAÍDA (OE)

Na estrutura de modelo de erro na saída, a estimação dos parâmetros pelo método dos mínimos quadrados é realizada minimizando-se a seguinte função-custo:

$$V_{OE,N} = \frac{1}{N} \sum_{t=1}^{N} [\varepsilon_{OE}(t)]^2 \qquad (11.6)$$

onde $\varepsilon_{OE}(t)$ é dado pela Expressão (9.38), repetida a seguir:

$$\varepsilon_{OE}(t) = y(t) - \frac{\hat{B}(q)}{\hat{F}(q)} u(t) \qquad (9.38)$$

Percebe-se que o erro na saída $\varepsilon_{OE}(t)$ é não linear nos parâmetros do polinômio $\hat{F}(q)$, pois conforme se nota em (9.38), $F(q)$ se encontra no denominador. A consequência desta não linearidade é que não existe uma solução analítica para este problema de minimização. Assim, um algoritmo de busca numérica é necessário para encontrar um mínimo. Problemas como mínimo local e não convergência podem ocorrer.

Estudam-se, a seguir, as propriedades estatísticas do método dos mínimos quadrados aplicado à estrutura de erro na saída, para dados obtidos de experimentos em malha aberta. Suponha que o processo real seja dado por:

$$y(t) = \frac{B(q)}{F(q)} u(t) + e(t)$$

onde $e(t)$ é um processo estocástico estacionário com média nula, a ordem do processo é n, a entrada é persistentemente excitante de ordem $2 \cdot n$ e a minimização de (11.6) converge para um mínimo global para qualquer N. Denote o erro do modelo por:

$$\Delta G(q) = G(q) - \hat{G}(q)$$

Então, se $N \to \infty$:

$$V_{OE,N} \to \mathrm{E}(V_{OE,N}) = \mathrm{E}[\Delta G(q) \cdot u(t) + e(t)]^2$$

$$\mathrm{E}(V_{OE,N}) = \Delta G(q)^2 \cdot \mathrm{E}\left[u(t)^2\right] + 2 \cdot \Delta G(q) \cdot \mathrm{E}[u(t) \cdot e(t)] + \mathrm{E}\left[e(t)^2\right]$$

Como os dados são oriundos de um experimento em malha aberta, em que a entrada não é correlacionada com a perturbação na saída, resulta:

$$V_{OE,N} \to \Delta G(q)^2 \cdot \mathrm{E}\left[u(t)^2\right] + \mathrm{E}\left[e(t)^2\right] \tag{11.7}$$

Se a ordem do modelo for correta e a minimização encontrar o mínimo global, então:

$$\Delta G(q)^2 \to 0 \qquad \text{conforme } N \to \infty$$

o que implica que:

$$\hat{G}(q) \to G(q) \qquad \text{conforme } N \to \infty$$

Como a ordem de $\hat{G}(q)$ e $G(q)$ são assumidas iguais, a expressão anterior é equivalente a dizer que:

$$\hat{\theta} \to \theta \qquad \text{conforme } N \to \infty$$

Portanto, o método do erro na saída é consistente sob condições relativamente fracas. Em outras palavras, ao se usar a estrutura OE, o modelo é não polarizado e o efeito da perturbação na saída $e(t)$ pode ser "eliminado", desde que a entrada não seja correlacionada com a perturbação (o que ocorre em experimentos em malha aberta). Se a ordem do modelo for inferior à real, o que pode frequentemente ocorrer em aplicações práticas, o modelo será polarizado.

11.5.1 Análise do erro em função da frequência do estimador dos mínimos quadrados aplicado à estrutura de modelo de erro na saída (OE)

A seguir, verifica-se como o erro do modelo obtido pela estrutura de erro na saída varia com a frequência. Aplicando-se a fórmula de Parseval em (11.7) para $N \to \infty$, resulta:

$$V_{OE,N} \to \frac{1}{2 \cdot \pi} \int_{-\pi}^{\pi} \left[G\left(e^{j\omega T}\right) - \frac{\hat{B}\left(e^{j\omega T}\right)}{\hat{A}\left(e^{j\omega T}\right)} \right]^2 \Phi_u(\omega)\, d\omega + \frac{1}{2 \cdot \pi} \int_{-\pi}^{\pi} \Phi_e(\omega)\, d\omega$$

Percebe-se que o erro do modelo é ponderado pelo espectro de potência da entrada, o qual pode ser manipulado pelo projeto do sinal de entrada.

Em resumo, em operação em malha aberta, a estrutura de erro na saída é consistente, se a ordem do modelo for correta. Se a ordem do modelo for inferior à real, o erro da estimativa do modelo é ponderado pelo espectro de potência do sinal de entrada. Em outras palavras, se a ordem do modelo for correta, o modelo identificado será preciso e se a ordem do modelo for baixa, o erro do modelo poderá ser afetado pelo projeto do sinal de entrada. No entanto, uma pré-condição fundamental para estas propriedades favoráveis, é que ocorra a convergência para o mínimo global do algoritmo numérico de busca utilizado. Em situações práticas, isto nem sempre é garantido (ZHU; BACKX, 1993).

11.6 ANÁLISE DA CONSISTÊNCIA DO ESTIMADOR DOS MÍNIMOS QUADRADOS APLICADO À ESTRUTURA DE MODELO BOX-JENKINS (BJ)

Os parâmetros do modelo da estrutura BJ, estimados pelo método dos mínimos quadrados, são determinados minimizando-se a seguinte função-custo:

$$V_N = \frac{1}{N}\sum_{t=1}^{N}\left\{\frac{\hat{D}(q)}{\hat{C}(q)}\left[y(t)-\frac{\hat{B}(q)}{\hat{A}(q)}u(t)\right]\right\}^2$$

As estimativas obtidas pela estrutura BJ são consistentes, tanto para identificação em malha aberta como em malha fechada.

11.7 ANÁLISE DA CONSISTÊNCIA E DA POLARIZAÇÃO DO ESTIMADOR DOS MÍNIMOS QUADRADOS APLICADO À ESTRUTURA GERAL DE MODELOS LINEARES

Suponha que o processo real seja descrito por:

$$y(t) = G(q)\cdot u(t) + H(q)\cdot e(t) \tag{11.8}$$

onde $G(q)$ é a função de transferência estável do processo, $H(q)$ é o filtro de perturbação, que é estável e de fase mínima (isto é, seu inverso $[H(q)]^{-1}$ também é estável e $e(t)$ é ruído branco de média nula e variância σ^2. Assume-se ainda que:

$$\begin{aligned} G(0) &= 0 \qquad (G(q)\ \text{tem pelo menos um atraso}) \\ H(0) &= 1 \qquad (H(0)\ \text{é mônico}) \end{aligned} \tag{11.9}$$

Assuma que a estrutura de modelo correta seja usada. O preditor para esse modelo é:

$$\hat{y}(t/t-1) = \frac{\hat{G}(q,\theta)}{\hat{H}(q,\theta)}u(t) + \left[1-\frac{1}{\hat{H}(q,\theta)}\right]\cdot y(t)$$

onde θ é o vetor que contém os parâmetros dos modelos do processo e da perturbação.

O erro de predição é dado por:

$$\varepsilon(t,\theta) = \frac{1}{\hat{H}(q,\theta)}\left[y(t)-\hat{G}(q,\theta)\cdot u(t)\right] \tag{11.10}$$

O vetor de parâmetros θ é determinado minimizando-se:

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N}\varepsilon(t,\theta)^2 = \frac{1}{N}\sum_{t=1}^{N}\left\{\frac{1}{\hat{H}(q,\theta)}\left[y(t)-\hat{G}(q,\theta)\cdot u(t)\right]\right\}^2$$

Quando $N\to\infty$, tem-se que:

$$V_N(\theta) \to \mathrm{E}\left[\varepsilon(t,\theta)^2\right] = \mathrm{E}\left\{\frac{1}{\hat{H}(q,\theta)}\left[y(t) - \hat{G}(q,\theta)\cdot u(t)\right]\right\}^2 \tag{11.11}$$

Substituindo-se (11.8) em (11.10), resulta:

$$\varepsilon(t,\theta) = \frac{1}{\hat{H}(q,\theta)}\left[G(q)\cdot u(t) + H(q)\cdot e(t) - \hat{G}(q,\theta)\cdot u(t)\right] = \frac{1}{\hat{H}(q,\theta)}\left[G(q) - \hat{G}(q,\theta)\right]\cdot u(t) + \frac{H(q)}{\hat{H}(q,\theta)}\,e(t) \tag{11.12}$$

Portanto:

$$\varepsilon(t,\theta) = e(t) + \text{um termo independente de } e(t) \tag{11.13}$$

Esta igualdade é devida às condições mostradas na Expressão (11.9).

Portanto, substituindo-se (11.13) em (11.11), considerando $N \to \infty$:

$$V_{N\to\infty}(\theta) \to \mathrm{E}\left[\varepsilon(t,\theta)^2\right] \geq \mathrm{E}\left[e(t)^2\right] = \sigma^2 \tag{11.14}$$

De (11.12):

$$\mathrm{E}[\varepsilon(t,\theta)]^2 = \frac{1}{\hat{H}(q,\theta)^2}\,\Delta G(q,\theta)^2 \cdot \mathrm{E}[u(t)]^2 + \frac{H(q)^2}{\hat{H}(q,\theta)^2}\,\sigma^2 \tag{11.15}$$

onde $\Delta G(q,\theta) = G(q) - \hat{G}(q,\theta)$.

Se o mínimo global for atingido para todo N e se a entrada for persistentemente excitante de ordem suficientemente alta, então analisando-se (11.15), nota-se que a restrição inferior da função-custo assintoticamente dada em (11.14) é alcançada quando:

$$\Delta G(q,\theta)^2 = 0 \qquad \text{e} \qquad \frac{H(q)^2}{\hat{H}(q,\theta)^2} = 1$$

Isto implica que quando $N \to \infty$:

$$\hat{G}(q,\theta) = G(q) \qquad \text{e} \qquad \hat{H}(q,\theta) = H(q)$$

Portanto, a estimativa dos mínimos quadrados é consistente. Ressalta-se que não se assumiu em nenhum ponto que os dados fossem coletados de experimentos em malha aberta. Deve-se enfatizar que, quando o processo estiver operando em malha fechada, a ordem dos modelos do processo e da perturbação devem ser corretas, para se chegar a uma estimativa consistente. Para dados coletados em malha aberta, esta condição pode ser um pouco relaxada. Se os modelos do processo e da perturbação forem parametrizados independentemente, isto é, $\hat{A}(q) = 1$ na expressão a seguir:

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

então pode-se obter uma estimativa consistente do modelo do processo para dados em malha aberta, mesmo quando a ordem do modelo de perturbação for muito baixa. Esta propriedade é desejável na prática, pois a precisão do modelo do processo é mais importante que a do modelo de perturbação. Quando a estrutura do modelo estiver incorreta, então o erro de predição do modelo é polarizado.

Pode-se gerar uma expressão assintótica no domínio da frequência para a função-custo (11.11) para um experimento em malha aberta, aplicando-se a fórmula de Parseval:

$$V_{N\to\infty} = \frac{1}{2\cdot\pi}\int_{-\pi}^{\pi}\left[G\left(e^{j\omega T}\right)-\frac{\hat{B}\left(e^{j\omega T}\right)}{\hat{F}\left(e^{j\omega T}\right)}\right]^2 \frac{\Phi_u(\omega)}{\hat{H}\left(e^{j\omega T}\right)^2}\,d\omega + \frac{1}{2\cdot\pi}\int_{-\pi}^{\pi}\frac{\Phi_v(\omega)}{\hat{H}\left(e^{j\omega T}\right)^2}\,d\omega$$

Percebe-se que o modelo de perturbação é parte da ponderação de frequência para a aproximação da função de transferência do processo.

11.8 AVALIAÇÃO GERAL DO ESTIMADOR DOS MÍNIMOS QUADRADOS

Delineiam-se, a seguir, as principais características do estimador dos mínimos quadrados (ZHU; BACKX, 1993):

a. o método dos mínimos quadrados é simples de entender e fácil de usar, devido à existência de uma solução analítica;

b. a estimativa de modelos FIR com dados coletados de um experimento em malha aberta é não polarizada e consistente, o mesmo não ocorrendo quando os dados são coletados em malha fechada;

c. a estimativa dos mínimos quadrados é não polarizada e eficiente (preditor de mínima variância) para uma função de transferência com estrutura ARX se, além de operar em malha aberta, a perturbação $v(t)$ seja ruído branco de média nula com variância constante σ^2. No entanto, na prática, a perturbação $v(t)$ normalmente não é ruído branco, o que causa a polarização da estimativa dos mínimos quadradaos. Deve-se enfatizar que, em alguns casos, a não consistência pode ser tolerável, por exemplo, se a relação sinal/ruído for grande, pois, neste caso, a polarizaçaõ é pequena.

d. o método dos mínimos quadrados pode gerar um bom modelo na forma de uma função de transferência se o nível de perturbações é baixo e a ordem do modelo coincide com a ordem do processo real. Caso contrário, a estimativa de modelos na forma de funções de transferência é polarizada e não consistente e o critério enfatiza um melhor ajuste da função de transferência em altas frequências, em detrimento de um bom ajuste nas médias e baixas frequências, quando o nível de perturbação cresce ou se a ordem do modelo é inferior à ordem verdadeira do processo ou ainda quando o experimento é realizado em malha fechada, o desempenho do método dos mínimos quadrados pode ser insatisfatório. Para melhorar estas estimativas, seria necessário eliminar a polarização e alterar a ponderação de frequência. Para tal, diferentes métodos foram propostos, tais como método generalizado dos mínimos quadrados, método das variáveis instrumentais, método do erro na saída e método da predição do erro.

11.9 EXEMPLOS DE ESTIMATIVAS CONSISTENTES E POLARIZADAS

Neste exemplo se usam os sistemas S_1, S_2 e S_3 da Seção 4.2, dados por:

$$y(t) = a\cdot y(t-1) + b\cdot u(t-5) + e(t) + c\cdot e(t-1)$$

Suponha que se deseje identificar estes processos usando o método dos mínimos quadrados. Propõe-se então empregar a seguinte estrutura ARX de modelo *M*:

$$y(t) = \hat{a} \cdot y(t-1) + \hat{b} \cdot u(t-5) + \varepsilon(t)$$

a. Obtenha a expressão $V_N(\hat{\theta})$ para o modelo dado, minimize-a e encontre o sistema de equações nas incógnitas $\hat{a}$ e $\hat{b}$.

b. Assuma que se aproxime a expectância E pelo seguinte somatório:

$$\mathrm{E}(\bullet) = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} (\bullet)$$

A seguinte aproximação pode ser feita:

$$\mathrm{E}\left[y^2(t-1)\right] \approx \frac{1}{N} \sum_{t=1}^{N} y^2(t-1)$$

A vantagem de se usar expectâncias ao invés de somatórios é que a análise pode lidar com um problema determinístico, ou mais precisamente, com um problema que não dependa de uma realização específica dos dados. Além disso, como se supõe estar trabalhando com processos estocásticos estacionários, resulta:

$$\mathrm{E}\left[y^2(t-1)\right] = \mathrm{E}\left[y^2(t)\right]$$

Aproxime o somatório pela média para estimar $\hat{a}$ e $\hat{b}$ para os 3 sistemas, com o processo excitado por um sinal PRBS de amplitude A=1 e afetado por ruído branco com variâncias σ^2=0; 0,001 e 1. Como os momentos de 1ª e 2ª ordens do PRBS são similares aos do ruído branco, assuma o sinal de entrada com média nula e variância λ^2.

c. Apresente as respostas ao degrau unitário dos modelos obtidos para os sistemas S_1, S_2 e S_3 e compare-as com a resposta do sistema sem perturbação.

d. Colete 1000 e 1.000.000 de pontos dos três sistemas excitados pelo sinal PRBS da alínea "b" e afetados pelo ruído dessa alínea. Aplicando o método dos mínimos quadrados da alínea "b", estime os coeficientes $\hat{a}$ e $\hat{b}$, com 1000 e 1.000.000 de pontos. Estime-os ainda via *toolbox* de Identificação do Matlab. Compare os valores obtidos com os valores da alínea "b" quando $N \to \infty$ e com os valores reais de a e b.

e. Repita a alínea "d", mas com uma excitação em degrau unitário, aplicado em t=0.

f. Repita a alínea "e" para o sistema S_1 com uma entrada pulso unitário aplicado em t=0.

g. Com os dados coletados na alínea "d", aplique o *toolbox* de Identificação do Matlab para gerar modelos com estruturas ARX, ARMAX e OE similares à dos próprios sistemas.

a. Cálculo da expressão $V_N(\theta)$ para o modelo dado e minimização da mesma para encontrar o sistema de equações nas incógnitas $\hat{a}$ e $\hat{b}$

Pode-se escrever:

$$\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \qquad \varphi(t) = \begin{pmatrix} y(t-1) \\ u(t-5) \end{pmatrix}$$

Então:

$$\hat{y}(t) = \varphi^{\mathrm{T}}(t)\,\hat{\theta}$$

$$\varepsilon(t) = y(t) - \hat{y}(t) = y(t) - \varphi^{\mathsf{T}}(t)\,\hat{\theta}$$

$$V_N(\hat{\theta}) = \frac{1}{2N}\sum_{t=1}^{N} \varepsilon^2(t) = \frac{1}{2N}\sum_{t=1}^{N} [y(t) - \hat{y}(t)]^2 = \frac{1}{2N}\sum_{t=1}^{N} \left[y(t) - \varphi^{\mathsf{T}}(t)\,\hat{\theta}\right]^2$$

Derivando-se $V_N(\hat{\theta})$ na expressão anterior, com relação aos parâmetros $\hat{a}$ e $\hat{b}$ e igualando-a a zero para determinar o mínimo, tem-se:

$$\frac{\partial V_N(\hat{\theta})}{\partial \hat{a}} = \frac{1}{N}\sum_{t=1}^{N} \left[y(t) - \varphi^{\mathsf{T}}(t)\cdot\hat{\theta}\right]\cdot\varphi^{\mathsf{T}}(t)\begin{bmatrix}1\\0\end{bmatrix} = 0$$

$$\frac{\partial V_N(\hat{\theta})}{\partial \hat{b}} = \frac{1}{N}\sum_{t=1}^{N} \left[y(t) - \varphi^{\mathsf{T}}(t)\cdot\hat{\theta}\right]\cdot\varphi^{\mathsf{T}}(t)\begin{bmatrix}0\\1\end{bmatrix} = 0$$

Das expressões anteriores resulta:

$$\sum_{t=1}^{N} \left[y(t) - \varphi^{\mathsf{T}}(t)\cdot\hat{\theta}\right]\cdot y(t-1) = 0 \qquad \sum_{t=1}^{N} \left[y(t) - \varphi^{\mathsf{T}}(t)\cdot\hat{\theta}\right]\cdot u(t-5) = 0$$

Isolando-se os somatórios com $\hat{\theta}$, obtém-se:

$$\left[\sum_{t=1}^{N} \varphi^{\mathsf{T}}(t)\cdot y(t-1)\right]\cdot\hat{\theta} = \sum_{t=1}^{N} y(t)\cdot y(t-1) \qquad \left[\sum_{t=1}^{N} \varphi^{\mathsf{T}}(t)\cdot u(t-5)\right]\cdot\hat{\theta} = \sum_{t=1}^{N} y(t)\cdot u(t-5)$$

Escrevendo-se as expressões acima como uma equação matricial:

$$\begin{bmatrix} \sum_{t=1}^{N} \varphi^{\mathsf{T}}(t)\cdot y(t-1) \\ \sum_{t=1}^{N} \varphi^{\mathsf{T}}(t)\cdot u(t-5) \end{bmatrix}\cdot\hat{\theta} = \begin{bmatrix} \sum_{t=1}^{N} y(t)\cdot y(t-1) \\ \sum_{t=1}^{N} y(t)\cdot u(t-5) \end{bmatrix}$$

Escrevendo-se os componentes do vetor $\varphi^{\mathsf{T}}(t)$:

$$\begin{bmatrix} \sum_{t=1}^{N} y^2(t-1) & \sum_{t=1}^{N} u(t-5)\cdot y(t-1) \\ \sum_{t=1}^{N} y(t-1)\cdot u(t-5) & \sum_{t=1}^{N} u^2(t-5) \end{bmatrix}\cdot\hat{\theta} = \begin{bmatrix} \sum_{t=1}^{N} y(t)\cdot y(t-1) \\ \sum_{t=1}^{N} y(t)\cdot u(t-5) \end{bmatrix}$$

Isolando-se o vetor $\hat{\theta}$:

$$\hat{\theta} = \begin{bmatrix} \sum_{t=1}^{N} y^2(t-1) & \sum_{t=1}^{N} u(t-5)\cdot y(t-1) \\ \sum_{t=1}^{N} y(t-1)\cdot u(t-5) & \sum_{t=1}^{N} u^2(t-5) \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{N} y(t)\cdot y(t-1) \\ \sum_{t=1}^{N} y(t)\cdot u(t-5) \end{bmatrix} \qquad (11.16)$$

b. Emprego das aproximações do somatório pela média para avaliar $\hat{a}$ e $\hat{b}$, supondo que a entrada $u(t)$ seja um sinal do tipo PRBS

Substituindo-se em (11.16) os somatórios pelas expectâncias, resulta:

$$\hat{\theta} = \begin{bmatrix} \mathrm{E}\left[y^2(t-1)\right] & \mathrm{E}\left[y(t-1)\cdot u(t-5)\right] \\ \mathrm{E}\left[y(t-1)\cdot u(t-5)\right] & \mathrm{E}\left[u^2(t-5)\right] \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{E}\left[y(t)\cdot y(t-1)\right] \\ \mathrm{E}\left[y(t)\cdot u(t-5)\right] \end{bmatrix}$$

Como se supõe estar trabalhando com processos estocásticos estacionários, resulta:

$\mathrm{E}\left[y^2(t-1)\right] = \mathrm{E}\left[y^2(t)\right]$ e $\mathrm{E}\left[u^2(t-5)\right] = \mathrm{E}\left[u^2(t)\right]$

Portanto:

$$\hat{\theta} = \begin{bmatrix} \mathrm{E}\left[y^2(t)\right] & \mathrm{E}\left[y(t-1)\cdot u(t-5)\right] \\ \mathrm{E}\left[y(t-1)\cdot u(t-5)\right] & \mathrm{E}\left[u^2(t)\right] \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{E}\left[y(t)\cdot y(t-1)\right] \\ \mathrm{E}\left[y(t)\cdot u(t-5)\right] \end{bmatrix} \tag{11.17}$$

Cada um dos elementos da equação matricial (11.17) é calculado a seguir. Neste caso, supõe-se conhecer o modelo exato que descreve o processo:

$$y(t) = a\cdot y(t-1) + b\cdot u(t-5) + e(t) + c\cdot e(t-1)$$

Portanto:

$$y^2(t) = a^2\cdot y^2(t-1) + b^2\cdot u^2(t-5) + e^2(t) + c^2\cdot e^2(t-1) + 2\cdot a\cdot b\cdot y(t-1)\cdot u(t-5) +$$
$$+ 2\cdot a\cdot y(t-1)\cdot e(t) + 2\cdot a\cdot c\cdot y(t-1)\cdot e(t-1) + 2\cdot b\cdot u(t-5)\cdot e(t) + 2\cdot b\cdot c\cdot u(t-5)\cdot e(t-1) +$$
$$+ c\cdot e(t)\cdot e(t-1)$$

O valor esperado de $y^2(t)$ é dado por:

$$\mathrm{E}\left[y^2(t)\right] = a^2\cdot \mathrm{E}\left[y^2(t)\right] + b^2\cdot \mathrm{E}\left[u^2(t)\right] + \mathrm{E}\left[e^2(t)\right] + c^2\cdot \mathrm{E}\left[e^2(t)\right] + 2\cdot a\cdot b\cdot \mathrm{E}\left[y(t-1)\cdot u(t-5)\right] +$$
$$+ 2\cdot a\cdot \mathrm{E}\left[y(t-1)\cdot e(t)\right] + 2\cdot a\cdot c\cdot \mathrm{E}\left[y(t-1)\cdot e(t-1)\right] + 2\cdot b\cdot \mathrm{E}\left[u(t-5)\cdot e(t)\right] +$$
$$+ 2\cdot b\cdot c\cdot \mathrm{E}\left[u(t-5)\cdot e(t-1)\right] + c\cdot \mathrm{E}\left[e(t)\cdot e(t-1)\right] \tag{11.18}$$

Tem-se que:

$$\mathrm{E}\left[u^2(t)\right] = \lambda^2 \tag{11.19}$$

$$\mathrm{E}\left[e^2(t)\right] = \sigma^2$$

$$\mathrm{E}\left[y(t-1)\cdot u(t-5)\right] = \mathrm{E}\left[y(t)\cdot u(t-4)\right] = 0 \tag{11.20}$$

O resultado da Equação (11.20) ocorre porque o sinal de entrada $u(t)$ só afeta a saída após 5 intervalos de tempo.

$\mathrm{E}\left[y(t-1)\cdot e(t)\right] = 0$ (assume-se que o ruído branco $e(t)$ não esteja correlacionado com a saída deslocada no tempo $y(t-1)$)

$$\mathrm{E}\left[y(t-1)\cdot e(t-1)\right] = \mathrm{E}\left[y(t)\cdot e(t)\right] = \mathrm{E}\left\{\left[a\cdot y(t-1) + b\cdot u(t-5) + e(t) + c\cdot e(t-1)\right]\cdot e(t)\right\} =$$
$$= \mathrm{a}\cdot \mathrm{E}\left[y(t-1)\cdot e(t)\right] + b\cdot \mathrm{E}\left[u(t-5)\cdot e(t)\right] + \mathrm{E}\left[e^2(t)\right] + c\cdot \mathrm{E}\left[e(t-1)\cdot e(t)\right] = \sigma^2$$

$\mathrm{E}\left[u(t-5)\cdot e(t)\right] = 0$ (assume-se que o ruído branco $e(t)$ não esteja correlacionado com a entrada $u(t)$ em nenhum instante de tempo)

$$\mathrm{E}\left[u(t-5)\cdot e(t-1)\right] = 0$$

$$\mathrm{E}\left[e(t)\cdot e(t-1)\right] = 0$$

Retornando-se à Equação (11.18), resulta:

$$\left(1-a^2\right)\cdot \mathrm{E}\left[y^2(t)\right]=b^2\cdot\lambda^2+\sigma^2+c^2\cdot\sigma^2+2\cdot a\cdot c\cdot\sigma^2$$

$$\mathrm{E}\left[y^2(t)\right]=\frac{b^2\cdot\lambda^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2} \qquad (11.21)$$

O valor esperado de $y(t)\cdot y(t-1)$ é dado por:

$$\mathrm{E}\left[y(t)\cdot y(t-1)\right]=\mathrm{E}\{[a\cdot y(t-1)+b\cdot u(t-5)+e(t)+c\cdot e(t-1)]\cdot y(t-1)\}=a\cdot \mathrm{E}\left[y^2(t)\right]+$$

$$+b\cdot \mathrm{E}[u(t-5)\cdot y(t-1)]+\mathrm{E}[e(t)\cdot y(t-1)]+c\cdot \mathrm{E}[e(t)\cdot y(t)]=$$

$$=a\,\frac{b^2\cdot\lambda^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2}+c\cdot\sigma^2=\frac{a\cdot b^2\cdot\lambda^2+[(a+c)\cdot(1+a\cdot c)]\cdot\sigma^2}{1-a^2} \qquad (11.22)$$

Por fim, tem-se que:

$$\mathrm{E}[y(t)\cdot u(t-5)]=\mathrm{E}\{[a\cdot y(t-1)+b\cdot u(t-5)+e(t)+c\cdot e(t-1)]\cdot u(t-5)\}=b\cdot \mathrm{E}\left[u^2(t)\right]=b\cdot\lambda^2 \qquad (11.23)$$

Substituindo-se de (11.19) a (11.23) na equação (11.17), tem-se que:

$$\hat{\theta}=\begin{bmatrix}\dfrac{b^2\cdot\lambda^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2} & 0\\ 0 & \lambda^2\end{bmatrix}^{-1}\begin{bmatrix}\dfrac{a\cdot b^2\cdot\lambda^2+[(a+c)\cdot(1+a\cdot c)]\cdot\sigma^2}{1-a^2}\\ b\cdot\lambda^2\end{bmatrix} \qquad (11.17a)$$

De (11.17a), resulta que os parâmetros estimados são:

$$\hat{a}=a+\frac{c\cdot\left(1-a^2\right)\cdot\sigma^2}{b^2\cdot\lambda^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2} \qquad \hat{b}=b$$

Tem-se então que:

- Sistema S_1:

  Como c=0, resulta $\hat{a}=a$ e $\hat{b}=b$, para σ^2=0; 0,001 e 1, significando que a identificação é consistente. Como o modelo proposto é ARX, da família erro da equação, isto só ocorreu pois se assumiu que a perturbação $e(t)$ é ruído branco, como preconizado na condição "b1" da Subseção 11.3.1.

- Sistema S_2:

  Neste caso, c=0,5. Para σ^2=0 resulta $\hat{a}=a$ e $\hat{b}=b$. Para σ^2=0,001 resulta:

  $\hat{a}=0{,}9532\neq a$ e $\hat{b}=b$

  Para $\sigma^2=1$ resulta:

  $\hat{a}=0{,}9726\neq a$ e $\hat{b}=b$

  A identificação não é consistente e há um *bias* no cálculo do parâmetro $\hat{a}$. O motivo disto é que a estrutura sugerida do modelo (ARX) não é a mesma estrutura do sistema (ARMAX). Nota-se ainda que o *bias* é tanto maior quanto maior for a variância do ruído branco $e(t)$.

- Sistema S_3:

  Para $c = -0{,}9512$ e $\sigma^2 = 0$ resulta $\hat{a} = a$ e $\hat{b} = b$. Para $\sigma^2 = 0{,}001$, resulta:

  $\hat{a} = 0{,}9470 \neq a$ e $\hat{b} = b$

  Para $\sigma^2 = 1$ resulta:

  $\hat{a} = 0{,}1747 \neq a$ e $\hat{b} = b$

  As conclusões para este caso são as mesmas que para o sistema S_2, exceto que a estrutura do sistema é OE e que o valor de $\hat{a}$ ficou com um *bias* bem maior.

c. Resposta ao degrau unitário dos modelos obtidos para os sistemas S_1, S_2 e S_3

A Figura 11.1 mostra o resultado das simulações realizadas. O sistema S_1 gera resultados iguais aos do sistema original, por isso não é mostrado nessa figura.

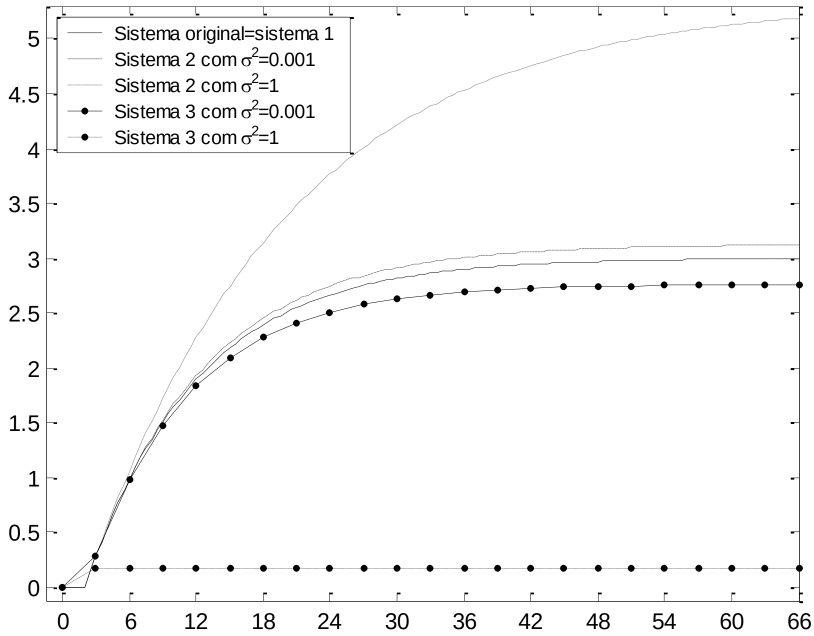

Fig. 11.1 Resposta dos modelos dos sistemas S_1, S_2 e S_3 com estrutura ARX com variância do ruído $\sigma^2 = 0$, $\sigma^2 = 0{,}001$ e $\sigma^2 = 1$.

Os modelos obtidos com ruído de baixa variância geraram respostas razoáveis, próximas do sistema sem ruído, ao passo que o resultado dos modelos criados com ruído de alta variância foi ruim, com respostas diferentes da esperada.

d. Cálculo dos parâmetros $\hat{a}$ e $\hat{b}$ para os modelos ARX dos sistemas S_1, S_2 e S_3, excitados por sinal PRBS, usando o método dos mínimos quadrados da alínea "b" e o comando ARX do Matlab, para *N*=1000 e *N*=1.000.000 bem como para σ^2=0; 0,001 e 1

Aplica-se a Expressão (11.16) a *N*=1000 e *N*=1.000.000 de pontos. Emprega-se o seguinte procedimento em Matlab:

```
% Cálculo com comando ARX do MATLAB
z = [y u];
theta = arx(z, [1 1 5])
```

Resultam os parâmetros apresentados na tabela a seguir.

- Para S_1 $(c = 0)$:

Como a estrutura do sistema e do modelo é a mesma e como a perturbação é ruído branco, quando $N\to\infty$, com qualquer variância do ruído, os parâmetros estimados são iguais aos reais. Com N=1.000.000, como a estimativa é consistente, os parâmetros se aproximam dos reais. Com N=1000 os parâmetros se afastam um pouco dos reais. Para N=1.000.000 e N=1000, quanto maior a variância da perturbação, mais distantes ficam os parâmetros dos reais. O método dos mínimos quadrados da alínea “b” produz resultados iguais aos do comando ARX do Matlab, que usa esse mesmo método.

	$\hat{a}$	$\hat{b}$
Parâmetros reais	a=0,9512	b=0,1463
$N \to \infty$ e $\sigma^2 = 0$	0,9512	0,1463
$N \to \infty$ e $\sigma^2 = 0{,}001$	0,9512	0,1463
$N \to \infty$ e $\sigma^2 = 1$	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9513	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9513	0,1463
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9510	0,1459
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9510	0,1459
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9540	0,1468
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9540	0,1468
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9523	0,1631
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9523	0,1631

- Para S_2 $(c = 0{,}5)$:

	$\hat{a}$	$\hat{b}$
Parâmetros reais	a=0,9512	b=0,1463
$N \to \infty$ e $\sigma^2 = 0$	0,9512	0,1463
$N \to \infty$ e $\sigma^2 = 0{,}001$	0,9532	0,1463
$N \to \infty$ e $\sigma^2 = 1$	0,9726	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9533	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9533	0,1463
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9725	0,1463
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9725	0,1463
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463

$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9574	0,1457
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9574	0,1457
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,9733	0,1334
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9734	0,1334

O parâmetro $\hat{a}$ tem *bias*, mesmo com perturbação do tipo ruído branco e com $N\to\infty$, pois a estrutura do modelo não é a mesma do sistema real. Com N=1000 e 1.000.000, ao aumentar a variância da perturbação, o *bias* dos parâmetros cresce.

- Para S_3 ($c = -0{,}9512$):

	$\hat{a}$	$\hat{b}$
Parâmetros reais	a=0,9512	b=0,1463
$N \to \infty$ e $\sigma^2 = 0$	0,9512	0,1463
$N \to \infty$ e $\sigma^2 = 0{,}001$	0,9470	0,1463
$N \to \infty$ e $\sigma^2 = 1$	0,1747	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9470	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9470	0,1463
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,1733	0,1450
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,1733	0,1450
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9466	0,1490
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9466	0,1490
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,1975	0,1830
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,1974	0,1830

Este foi o caso com piores resultados, pois aqui também a estrutura do sistema (OE) não coincide com a estrutura proposta do modelo (ARX).

e. Repetir a alínea "d", mas com uma excitação em degrau unitário aplicada em t=0

A Equação (11.18) apresenta o valor esperado de $y^2(t)$. Como o sinal de entrada $u(t)$ é um degrau de amplitude A, resulta que:

$$\mathrm{E}\left[u^2(t)\right] = A^2$$

O ganho K do sistema em malha aberta em regime estacionário é dado por:

$$y(t) - a\cdot q^{-1}\cdot y(t) = b\cdot q^{-5}\cdot u(t) \quad \to \quad \left(1 - a\cdot q^{-1}\right)\cdot y(t) = b\cdot q^{-5}\cdot u(t)$$

$$G(q) = \frac{y(t)}{u(t)} = \frac{b\cdot q^{-5}}{1 - a\cdot q^{-1}}$$

Portanto:

$$K = \lim_{q\to 1} G(q) = \frac{b}{1-a} = \frac{0{,}1463}{1-0{,}9512} = 3$$

Este valor, em verdade, já havia sido fornecido na Seção 4.2. Tem-se então que:

$$\mathrm{E}[y(t-1)\cdot u(t-5)] = K\cdot A^2$$

Retornando-se à Equação (11.18):

$$\left(1-\mathrm{a}^2\right)\cdot \mathrm{E}\left[y^2(t)\right] = b^2\cdot A^2 + \sigma^2 + c^2\cdot\sigma^2 + 2\cdot a\cdot b\cdot K\cdot A^2 + 2\cdot a\cdot c\cdot\sigma^2$$

Portanto:

$$\mathrm{E}\left[y^2(t)\right] = \frac{\left(b^2+2\cdot a\cdot b\cdot K\right)\cdot A^2 + \left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2}$$

O valor esperado de $y(t)\cdot y(t-1)$ é dado por:

$$\mathrm{E}[y(t)\cdot y(t-1)] = \mathrm{E}\{[a\cdot y(t-1)+b\cdot u(t-5)+e(t)+c\cdot e(t-1)]\cdot y(t-1)\} = a\cdot\mathrm{E}\left[y^2(t)\right]+$$

$$+b\cdot\mathrm{E}[u(t-5)\cdot y(t-1)]+\mathrm{E}[e(t)\cdot y(t-1)]+c\cdot\mathrm{E}[e(t)\cdot y(t)]=$$

$$= a\,\frac{\left(b^2+2\cdot a\cdot b\cdot K\right)\cdot A^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2}+b\cdot K\cdot A^2+c\cdot\sigma^2 =$$

$$= \frac{\left(a\cdot b^2+2\cdot a^2\cdot b\cdot K+b\cdot K-a^2\cdot b\cdot K\right)\cdot A^2+\left(a+a\cdot c^2+a^2\cdot c+c\right)\cdot\sigma^2}{1-a^2}$$

Por fim, tem-se que:

$$\mathrm{E}[y(t)\cdot u(t-5)] = \mathrm{E}[y(t-1)\cdot u(t-5)] = K\cdot A^2$$

Substituindo-se os valores obtidos anteriormente na Equação (11.17), resulta:

$$\hat{\theta} = \begin{bmatrix} \dfrac{\left(b^2+2\cdot a\cdot b\cdot K\right)\cdot A^2+\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2} & K\cdot A^2 \\ K\cdot A^2 & A^2 \end{bmatrix}^{-1}\cdot$$

$$\cdot\begin{bmatrix} \dfrac{\left(a\cdot b^2+2\cdot a^2\cdot b\cdot K+b\cdot K-a^2\cdot b\cdot K\right)\cdot A^2+\left(a+a\cdot c^2+a^2\cdot c+c\right)\cdot\sigma^2}{1-a^2} \\ K\cdot A^2 \end{bmatrix} \qquad (11.17b)$$

Para os valores dados dos parâmetros, resulta:

- Sistema S_1 ($c=0$):

	$\hat{a}$	$\hat{b}$
Parâmetros reais	$a=0{,}9512$	$b=0{,}1463$
$N\to\infty$ e $\sigma^2=0$	1,250	0,0
$N\to\infty$ e $\sigma^2=0{,}001$	0,9512	0,1463
$N\to\infty$ e $\sigma^2=1$	0,9512	0,1463
$N=1.000.000$ e $\sigma^2=0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N=1.000.000$ e $\sigma^2=0$ (comando ARX do Matlab)	0,9512	0,1463
$N=1.000.000$ e $\sigma^2=0{,}001$ (mínimos quadrados alínea "b")	0,9510	0,1469
$N=1.000.000$ e $\sigma^2=0{,}001$ (comando ARX do Matlab)	0,9510	0,1469

$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,9510	0,1464
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9510	0,1464
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9545	0,1380
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9545	0,1379
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,9467	0,2111
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9468	0,2110

A estimativa é consistente, pois a perturbação é ruído branco. Ocorre algo inesperado com $N \to \infty$ e $\sigma^2 = 0$: a matriz sendo invertida em (11.17b) é singular quando não há ruído. Isto ocorre pois o sinal de excitação (degrau) não é persistentemente excitante de ordem adequada. Substituindo-se em (11.17b) o valor dos parâmetros resulta:

$$\hat{\theta} = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

Portanto, a solução deste sistema não tem uma única solução.

- Sistema S_2 $(c = 0{,}5)$:

	$\hat{a}$	$\hat{b}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$
$N \to \infty$ e $\sigma^2 = 0$	1,250	0,0
$N \to \infty$ e $\sigma^2 = 0{,}001$	0,9728	0,08146
$N \to \infty$ e $\sigma^2 = 1$	0,9728	0,08146
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9726	0,08206
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9726	0,08206
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,9727	0,08132
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9727	0,08132
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea "b")	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea "b")	0,9592	0,1247
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9592	0,1247
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea "b")	0,9700	0,1333
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9700	0,1332

Neste caso, a identificação não é consistente e existe um *bias* no cálculo dos parâmetros $\hat{a}$ e $\hat{b}$. O motivo disto é que a estrutura sugerida do modelo (ARX) não é a mesma do sistema (ARMAX). Percebe-se ainda que o *bias* é tanto maior quanto maior for a variância do ruído branco $e(t)$.

- Sistema S_3 $(c = -0{,}9512)$:

	$\hat{a}$	$\hat{b}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$
$N \to \infty$ e $\sigma^2 = 0$	1,250	0,0
$N \to \infty$ e $\sigma^2 = 0{,}001$	0,0	3,0
$N \to \infty$ e $\sigma^2 = 1$	0,0	3,0
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9510	0,1469
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9510	0,1469
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9510	0,1464
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9510	0,1464
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9545	0,1380
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9545	0,1379
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9467	0,2111
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9468	0,2110

As conclusões para este caso são as mesmas que para o sistema S_2, exceto que a estrutura do sistema é OE. Nota-se ainda que, em todos os casos vistos, sempre que não há ruído ($\sigma^2 = 0$), exceto para $N \to \infty$, a estimação dos parâmetros é sem erro, pois se elimina a parte estocástica do sistema e se mantém apenas a parte determinística. Como se assumiu que a estrutura da parte determinística do modelo seja idêntica à dos sistemas, resulta a ausência de erro no cálculo dos parâmetros. A consistência é obtida, pois se está lidando com perturbação do tipo ruído branco.

f. Repita a alínea “e” apenas para o sistema S_1, considerando uma entrada na forma de um pulso unitário aplicado em $t = 0$

O termo $\mathrm{E}\left[y^2(t)\right]$ mostrado na Equação (11.18) é reapresentado a seguir:

$$\mathrm{E}\left[y^2(t)\right] = a^2 \cdot \mathrm{E}\left[y^2(t)\right] + b^2 \cdot \mathrm{E}\left[u^2(t)\right] + \mathrm{E}\left[e^2(t)\right] + c^2 \cdot \mathrm{E}\left[e^2(t)\right] + 2 \cdot a \cdot b \cdot \mathrm{E}\left[y(t-1) \cdot u(t-5)\right] +$$
$$+ 2 \cdot a \cdot \mathrm{E}\left[y(t-1) \cdot e(t)\right] + 2 \cdot a \cdot c \cdot \mathrm{E}\left[y(t-1) \cdot e(t-1)\right] + 2 \cdot b \cdot \mathrm{E}\left[u(t-5) \cdot e(t)\right] +$$
$$+ 2 \cdot b \cdot c \cdot \mathrm{E}\left[u(t-5) \cdot e(t-1)\right] + c \cdot \mathrm{E}\left[e(t) \cdot e(t-1)\right] \tag{11.18}$$

O cálculo de seus termos é feito conforme se segue:

$$\mathrm{E}\left[u^2(t)\right] = 0$$

$$\mathrm{E}\left[e^2(t)\right] = \sigma^2$$

$$\mathrm{E}\left[y(t-1) \cdot u(t-5)\right] = \mathrm{E}\left[y(t) \cdot u(t-4)\right] = 0$$

$$\mathrm{E}\left[y(t-1) \cdot e(t)\right] = 0$$

$$\mathrm{E}\left[y(t-1) \cdot e(t-1)\right] = \mathrm{E}\left[y(t) \cdot e(t)\right] = \sigma^2$$

$$\mathrm{E}\left[u(t-5) \cdot e(t)\right] = 0$$

$$\mathrm{E}[u(t-5)\cdot e(t-1)]=0$$

$$\mathrm{E}[e(t)\cdot e(t-1)]=0$$

Substituindo-se esses termos na Equação (11.18), resulta:

$$\mathrm{E}\left[y^2(t)\right]=\frac{\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2}$$

O valor esperado de $y(t)\cdot y(t-1)$ é dado por:

$$\mathrm{E}[y(t)\cdot y(t-1)]=\mathrm{E}\{[a\cdot y(t-1)+b\cdot u(t-5)+e(t)+c\cdot e(t-1)]\cdot y(t-1)\}=a\cdot\mathrm{E}\left[y^2(t)\right]+$$

$$+b\cdot\mathrm{E}[u(t-5)\cdot y(t-1)]+\mathrm{E}[e(t)\cdot y(t-1)]+c\cdot\mathrm{E}[e(t)\cdot y(t)]=$$

$$=a\,\frac{\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2}+c\cdot\sigma^2=\frac{[(a+c)\cdot(1+a\cdot c)]\cdot\sigma^2}{1-a^2}$$

Por fim, tem-se que:

$$\mathrm{E}[y(t)\cdot u(t-5)]=0$$

Substituindo-se os valores obtidos acima na equação (11.17), resulta:

$$\hat{\theta}=\begin{bmatrix}\dfrac{\left(1+c^2+2\cdot a\cdot c\right)\cdot\sigma^2}{1-a^2} & 0\\ 0 & 0\end{bmatrix}^{-1}\begin{bmatrix}\dfrac{[(a+c)\cdot(1+a\cdot c)]\cdot\sigma^2}{1-a^2}\\ 0\end{bmatrix}\qquad(11.17c)$$

A solução desta equação não é única, pois a matriz sendo invertida em (11.17c) é singular. Aplica-se então a Expressão (11.16) para N=1000 e $N=1.000.000$ de pontos e o comando ARX do Matlab. Resultam os parâmetros apresentados na tabela a seguir.

- Para S$_1$ $(c=0)$:

	$\hat{a}$	$\hat{b}$
Parâmetros reais	$a=0{,}9512$	$b=0{,}1463$
$N=1.000.000$ e $\sigma^2=0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N=1.000.000$ e $\sigma^2=0$ (comando ARX do Matlab)	0,9512	0,1463
$N=1.000.000$ e $\sigma^2=0{,}001$ (mínimos quadrados alínea “b”)	0,9510	0,1392
$N=1.000.000$ e $\sigma^2=0{,}001$ (comando ARX do Matlab)	0,9510	0,1392
$N=1.000.000$ e $\sigma^2=1$ (mínimos quadrados alínea “b”)	0,9510	-0,07925
$N=1.000.000$ e $\sigma^2=1$ (comando ARX do Matlab)	0,9510	-0,07925
$N=1000$ e $\sigma^2=0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N=1000$ e $\sigma^2=0$ (comando ARX do Matlab)	0,9512	0,1463
$N=1000$ e $\sigma^2=0{,}001$ (mínimos quadrados alínea “b”)	0,9475	0,1392
$N=1000$ e $\sigma^2=0{,}001$ (comando ARX do Matlab)	0,9475	0,1392
$N=1000$ e $\sigma^2=1$ (mínimos quadrados alínea “b”)	0,9500	-0,07931
$N=1000$ e $\sigma^2=1$ (comando ARX do Matlab)	0,9500	-0,07931

O parâmetro $\hat{a}$ é estimado com bastante precisão, ao passo que a estimação de $\hat{b}$ não é boa, exceto na ausência de perturbação $\left(\sigma^2=0\right)$, em que a estimativa é perfeita, pois se está identificando um sistema com estrutura e ordem corretas. Esta estimação imprecisa de $\hat{b}$ era esperada, pois a entrada fornece uma pequena contribuição à saída e é apenas através da influência de u em y que as informações sobre $\hat{b}$ são obtidas (SÖDERSTRÖM; STOICA, 1989). Por outro lado, o parâmetro $\hat{a}$ também des-creve o efeito do ruído na saída. Como o ruído está presente nos dados, exceto com $\sigma^2 = 0$, é natural que $\hat{a}$ seja estimado mais precisamente que $\hat{b}$. Esta falha para esti-mar $\hat{b}$ ocorre, mesmo usando um modelo com estrutura ARX igual ao do sistema S_1.

O desvio de $\hat{b}$ depende da realização do ruído, de modo que ele terá valores bem diferentes dependendo dos dados empregados, conforme se verifica na tabela a seguir, onde a semente usada para o ruído nas simulações do Matlab foi alterada.

	$\hat{a}$	$\hat{b}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$
$N = 1.000.000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9511	0,1747
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9511	0,1747
$N = 1.000.000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9511	1,045
$N = 1.000.000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9511	1,045
$N = 1000$ e $\sigma^2 = 0$ (mínimos quadrados alínea “b”)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0$ (comando ARX do Matlab)	0,9512	0,1463
$N = 1000$ e $\sigma^2 = 0{,}001$ (mínimos quadrados alínea “b”)	0,9181	0,1762
$N = 1000$ e $\sigma^2 = 0{,}001$ (comando ARX do Matlab)	0,9181	0,1762
$N = 1000$ e $\sigma^2 = 1$ (mínimos quadrados alínea “b”)	0,9118	1,099
$N = 1000$ e $\sigma^2 = 1$ (comando ARX do Matlab)	0,9117	1,099

O desempenho do método dos mínimos quadrados pode ser explicado, pois a matriz em (11.17c) tende a ser singular conforme $N \to \infty$. A despeito disso, a estimativa pôde ser computada, embora o valor do parâmetro $\hat{b}$ seja dependente da forma da perturbação. No entanto, a estimativa converge para uma variável aleatória e não para um valor constante. Assim, o método dos mínimos quadrados falhou em gerar estimativas consistentes. Isto ocorreu pois o sinal de entrada não é persistentemente excitante de ordem n_a+n_b=2, como citado na condição “a” da Subseção 11.3.1.

Como conclusão para as alíneas “d”, “e” e “f”, viu-se que para se obter estimativas consistentes dos parâmetros, a entrada deve ser do tipo PRBS. Se ela for degrau ou impulsiva, o método dos mínimos quadrados falha em gerar estimativas consistentes. Isto porque essas entradas não são persistentemente excitantes de ordem 2. Para garantir consistência na estimação, a entrada deve excitar suficientemente o processo, devendo ser, neste caso, pelo menos persistentemente excitante de ordem 2.

g. Use o *toolbox* de Identificação do Matlab e gere modelos com estruturas ARX, ARMAX e OE, iguais à dos sistemas S_1, S_2 e S_3, com sinal de excitação PRBS com A=1

- Sistema S_1 $(c = 0)$:

	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{f}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$	$c = 0$	$f = 0{,}9512$
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,3298	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9513	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9513	0,1463	$-5{,}787 \cdot 10^{-5}$	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1462	----	0,9514
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARX)	0,9510	0,1459	----	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9512	0,1463	$-2{,}524 \cdot 10^{-5}$	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,1439	----	0,9561
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,7523	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9540	0,1468	----	----
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9540	0,1470	-0,00114	----
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1445	----	0,9540
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARX)	0,9523	0,1631	----	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9513	0,1691	0,00303	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,1470	----	1,002

Sem perturbação, as três estruturas estimam exatamente os parâmetros $\hat{a}$ e $\hat{b}$. No entanto, com perturbação, como o sistema tem estrutura ARX, resulta que o modelo com essa estrutura gera os melhores resultados e com estrutura OE os piores.

- Sistema S_2 $(c = 0{,}5)$:

	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{f}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$	$c = 0{,}5$	$f = 0{,}9512$
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,3298	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9533	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9513	0,1463	0,5007	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1462	----	0,9515
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARX)	0,9725	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9510	0,1469	0,5009	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,1429	----	0,9583
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,7523	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9574	0,1457	----	----

$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9556	0,1463	0,5124	----
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1445	----	0,9556
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARX)	0,9734	0,1334	----	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9484	0,1401	0,5203	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,2071	----	1,002

O sistema S_2 tem estrutura ARMAX. Assim, as melhores estimativas são as geradas com esta estrutura. Observa-se ainda que, sem perturbação, a própria estrutura ARMAX estima um valor incorreto para $\hat{c}$, pois faltam informações para estimá-lo.

- Sistema S_3 $(c = -0{,}9512)$:

	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{f}$
Parâmetros reais	$a = 0{,}9512$	$b = 0{,}1463$	$c = -0{,}9512$	$f = 0{,}9512$
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,3298	----
$N = 1.000.000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9470	0,1463	----	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9512	0,1463	-0,9515	----
$N = 1.000.000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARX)	0,1733	0,1450	----	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9515	0,1461	-0,9517	----
$N = 1.000.000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,1460	----	0,9516
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARX)	0,9512	0,1463	----	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura ARMAX)	0,9512	0,1463	0,7523	----
$N = 1000$ e $\sigma^2 = 0$ (estrutura OE)	----	0,1463	----	0,9512
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARX)	0,9466	0,1490	----	----
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura ARMAX)	0,9511	0,1469	-0,9526	----
$N = 1000$ e $\sigma^2 = 0{,}001$ (estrutura OE)	----	0,1469	----	0,9511
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARX)	0,1974	0,1830	----	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura ARMAX)	0,9487	0,1646	-0,9506	----
$N = 1000$ e $\sigma^2 = 1$ (estrutura OE)	----	0,1668	----	0,9476

Quando o número disponível de pontos para a identificação é menor e há ruído, existe uma tendência do modelo com estrutura ARMAX ter seus parâmetros mais próximos dos parâmetros estimados pela estrutura OE, que é a estrutura do sistema S_3.

REFERÊNCIAS BIBLIOGRÁFICAS

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.