

CAPÍTULO 12

VALIDAÇÃO DE MODELOS

Uma vez que se tenha um modelo em mãos, a questão crucial é saber se ele é bom o bastante para sua aplicação pretendida. Testar essa adequação é conhecido como **validação** do modelo. Este capítulo visa descrever métodos que possam ser aplicados na verificação da qualidade de um modelo, a qual está associada à sua aplicação. Na prática, é recomendável usar diversos métodos e tentar validar um modelo, considerando um conjunto de indicadores, ao invés de se usar apenas um indicador (AGUIRRE, 2007).

Nos exemplos de identificação vistos até aqui neste livro, assume-se que a ordem do modelo seja conhecida e que somente os parâmetros devam ser estimados. Isto raramente ocorre na prática, onde a ordem do processo dificilmente é conhecida, pelo menos de forma exata. A seleção da ordem do modelo está intimamente relacionada com a sua validação. Na prática, pode-se realizar a identificação para ordens crescentes do modelo e parar quando um modelo seja satisfatório, sob algum critério de validação, como é mostrado a seguir neste capítulo (ZHU; BACKX, 1993).

Na identificação de sistemas, a seleção da ordem do modelo é relevante. Um modelo sobre-ajustado (ou sobre-parametrizado) pode levar a computações desnecessariamente complicadas para estimar os parâmetros e para aplicar o modelo estimado. Por outro lado, uma estrutura sub-ajustada (ou sub-parametrizada) pode ser muito imprecisa. A escolha da ordem do modelo é fortemente influenciada pelo uso pretendido do mesmo. É claro que não se consegue modelar de modo exato qualquer conjunto de dados real por um modelo linear de ordem finita. Mesmo assim, tais modelos frequentemente fornecem boas aproximações da dinâmica verdadeira. No entanto, os métodos para encontrar a ordem “correta” do modelo são baseados na hipótese que os dados provenham de um sistema real contido na classe de modelos considerada (SÖDERSTRÖM; STOICA, 1989). Ao se buscar a ordem “correta” do modelo, uma pergunta cabível é: qual estrutura de modelo dentre duas ou mais candidatas deve ser escolhida?

Este capítulo descreve métodos que podem ser usados para comparar duas ou mais estruturas de modelo. Para efetuar tais comparações um critério é necessário. Abranger-se-ão apenas os casos em que um método de erro de predição (PEM) seja usado para estimar os parâmetros. Lembre-se que para um método PEM, a estimativa dos parâmetros $\hat{\theta}_N$ é obtida como o elemento minimizante de uma função custo $V_N(\theta)$.

Ao expandir a estrutura de modelo, de modo a incluir mais termos no vetor de parâmetros θ , o valor de $V_N(\theta)$ diminui, pois novos graus de liberdade são adicionados ao problema da otimização. A comparação de estruturas de modelo pode ser interpretada como um teste para um decréscimo significativo nos valores mínimos da função custo, associados com as estruturas de modelo em questão (SÖDERSTRÖM; STOICA, 1989).

12.1 VALIDAÇÃO DE MODELOS ATRAVÉS DE SIMULAÇÃO LIVRE

O modo mais natural de se validar um modelo é comparar a saída da simulação livre do modelo com dados medidos do processo real. Neste caso, deseja-se saber quão bem o modelo reproduz, ao longo do tempo, os dados observados. Assim, uma forma simples mas eficiente de validar modelos é simulá-los com algum sinal de entrada e comparar sua saída com a saída medida do processo, como indicado na Figura 12.1 (ZHU; BACKX, 1993).

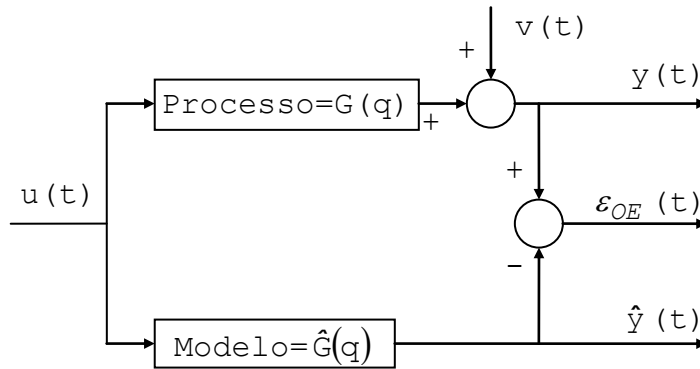


Fig. 12.1 Validação de modelo por meio de simulação livre.

Para a validação de modelos usa-se o erro de saída ao invés do erro da equação, porque $\varepsilon_{OE}(t)$ é o erro de simulação mais as perturbações, isto é, ele corresponde ao mau ajuste do modelo mais as perturbações de saída. Da figura 12.1 tem-se que:

$$\varepsilon_{OE}(t) = y(t) - \hat{y}(t) = G(q) \cdot u(t) + v(t) - \hat{G}(q) \cdot u(t) = [G(q) - \hat{G}(q)] \cdot u(t) + v(t)$$

onde $G(q)$ corresponde à função de transferência do sistema e $\hat{G}(q)$ equivale à função de transferência do modelo que representa o sistema.

A qualidade do ajuste é medida através do somatório dos quadrados do erro de saída:

$$V_N = \frac{1}{2N} \sum_{t=1}^N [\varepsilon_{OE}(t)]^2 \quad \text{onde } \varepsilon_{OE}(t) = y(t) - \hat{y}(t) = y(t) - \frac{\hat{B}(q)}{\hat{A}(q)} u(t).$$

Em geral, a função custo V_N decresce conforme a ordem do modelo n cresça. A redução de V_N deixa de ser significativa quando a ordem do modelo for grande o suficiente para o propósito da simulação. Baseado nisso, um procedimento para a seleção da ordem é simplesmente computar as estimativas dos mínimos quadrados e as correspondentes funções custo baseadas no erro de saída, para uma sequência de modelos de ordens $n=1, 2, 3, \dots$. A ordem apropriada pode ser escolhida como aquela em que V_N pare de decrescer significativamente. Quando o nível de perturbação é baixo, a ordem verdadeira

do processo pode ser encontrada usando este procedimento. No entanto, conforme o nível de perturbação cresça, a ordem encontrada por este procedimento pode ser superior à ordem real. A ordem mais alta do modelo busca modelar as perturbações, visando superar a polarização (*bias*) do estimador dos mínimos quadrados. Nas seções que se seguem, os critérios para parar o crescimento da ordem do modelo são apresentados.

Uma observação importante é que a validação do modelo e a seleção de sua ordem deveriam ser realizadas em um conjunto de dados que não tenha sido usado para a estimação do modelo, o que é normalmente intitulado validação cruzada. A razão é que um bom ajuste da saída do modelo aos dados que foram usados para estimar o modelo não é suficiente para provar a qualidade do modelo. No caso extremo, quando a ordem do modelo se iguala ao número de pontos coletados N , então o erro de ajuste é nulo, no entanto, a qualidade do modelo pode ser bastante pobre, devido à variância do ruído. Este assunto é discutido mais adiante neste capítulo.

Dados de estimação equivalem ao conjunto de dados que foi utilizado para estimar os parâmetros do modelo e **dados de validação** ao conjunto de dados usado para comparar com as respostas das simulações. Ambos os conjuntos de dados devem ser coletados sob condições que sejam as mais próximas possível das condições operacionais do processo. Se dois testes distintos para coletar esses dois conjuntos de dados não forem possíveis, deve-se então dividir o único conjunto disponível de dados em duas partes, não necessariamente de mesmo tamanho.

12.2 VALIDAÇÃO CRUZADA E AUTOVALIDAÇÃO

Os testes do modelo determinam quão bem o modelo é capaz de reproduzir os dados de validação. Não é surpreendente que um modelo possa ter um bom desempenho quando avaliado com o próprio conjunto de dados com que foi gerado (dados de estimação). O teste real é verificar se ele é capaz de descrever o comportamento do processo com base em um conjunto novo de dados (dados de validação). Uma forma sugerida de comparar modelos diferentes é avaliar seu desempenho quando aplicados a um conjunto de dados para o qual nenhum deles foi ajustado. Seleciona-se então o modelo que mostre o melhor desempenho. Tal procedimento é conhecido como **validação cruzada**.

A medida do desempenho para efeitos comparativos poderia ser, por exemplo, o somatório dos erros de predição ao quadrado ou a diferença entre saídas reais e saídas simuladas. Uma característica importante da validação cruzada é seu caráter pragmático. A comparação faz sentido sem quaisquer argumentos probabilísticos e sob quaisquer hipóteses sobre o processo real. Sua única desvantagem é a necessidade de se manter um conjunto de dados coletados para as comparações, de forma que não se pode usar todas as informações disponíveis para gerar os modelos (LJUNG, 1999).

Quando não se pode reservar dados novos para efetuar a validação cruzada, deve-

se comparar os modelos usando o conjunto de dados para o qual os modelos foram ajustados, realizando a chamada **autovalidação**. Este fato complica consideravelmente a comparação. O modelo obtido na estrutura de modelos de grau mais alto produzirá automaticamente um menor valor do critério de ajuste, pois o valor da função custo V_N diminui, visto que o modelo capta mais das características relevantes dos dados. Mas mesmo após uma estrutura de modelo que gere uma descrição adequada do sistema ter sido alcançada, o valor de V_N continua a diminuir, porque os parâmetros adicionais (desnecessários) se ajustam a características particulares da perturbação. Este efeito é conhecido como sobreajuste (ou sobreparametrização) e esse ajuste mais aprimorado é sem valor, pois irá se aplicar o modelo a dados com diferentes realizações de ruído (LJUNG, 1999).

12.3 DETERMINAÇÃO DA DIMENSÃO DO MODELO

O problema aqui abordado consiste na determinação de uma dimensão apropriada para o modelo. A dimensão de um modelo é definida como o número de parâmetros a ser estimado no mesmo e é designada por p .

Considere a situação onde haja uma sequência de estruturas de modelo de dimensão crescente. Com mais parâmetros livres em uma estrutura de modelo, um melhor ajuste é obtido para os dados observados. O fato importante a ser analisado é se a melhora no ajuste é significativa ou não. Considere inicialmente um caso ideal. Assuma que os dados sejam livres de perturbações e ruído ou que o número N de pontos coletados tenda a infinito. Suponha que haja uma estrutura de modelo M^* , de forma que, com valores adequados dos parâmetros, ela possa descrever o sistema de forma exata. Então a relação entre função custo e estrutura de modelo é mostrada na Figura 12.2, onde p^* denota a dimensão do modelo M^* (SÖDERSTRÖM; STOICA, 1989).

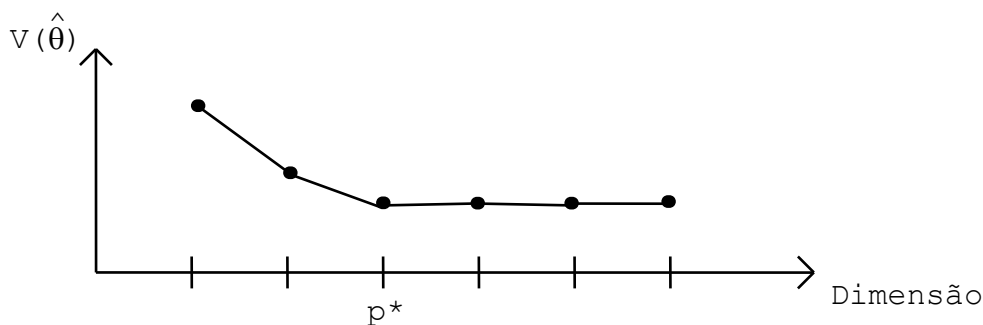


Fig. 12.2 Valor da função custo versus estrutura do modelo para caso ideal (sem perturbações ou $N \rightarrow \infty$).

Neste caso ideal, a estrutura de modelo M^* com dimensão p^* corresponde ao processo (sistema verdadeiro). A função custo $V(\hat{\theta})$ fica constante enquanto M tem pelo menos a mesma dimensão de M^* . O caso real equivale à situação quando os dados são afetados por perturbações e $N \ll \infty$. São analisadas duas condições: autovalidação e validação cruzada. Na autovalidação a função custo $V(\hat{\theta})$ decresce lentamente com o aumento da dimensão de M , como visto na Figura 12.3 (SÖDERSTRÖM; STOICA, 1989).

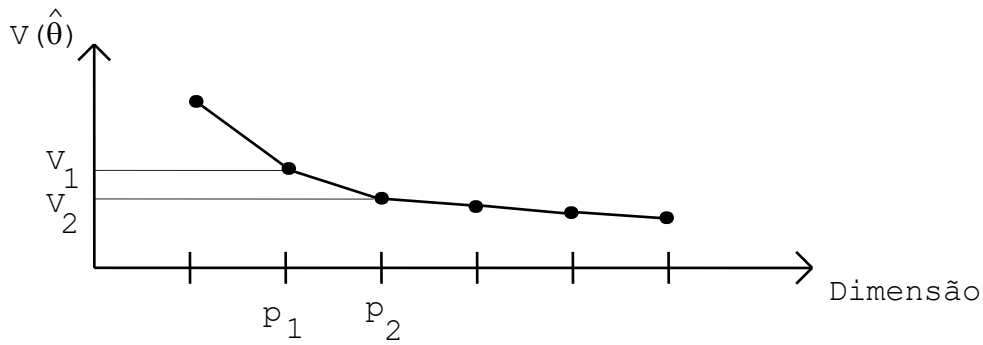


Fig. 12.3 Valor da função custo versus estrutura do modelo para caso real de autovalidação (dados com perturbação e $N \ll \infty$).

O problema é decidir se o decréscimo $\Delta V = V_1 - V_2$ é “pequeno” ou não. Se ele for “pequeno”, então a estrutura de modelo M_1 deve ser escolhida, caso contrário deve-se continuar a busca. O que se busca é o “joelho” da curva da Figura 12.3. Vários procedimentos foram criados para formalizar e atuar como diretriz na busca do “joelho”. Para um teste quantitativo normalizado parece razoável se considerar $\Delta V = (V_1 - V_2)/V_2$. Se o sistema real pode ser descrito pela estrutura de modelo M_1 , então o decréscimo ΔV deveria tender a zero conforme o número de pontos de dados N tenda a infinito. É de se esperar então que $N \cdot (V_1 - V_2)/V_2$ seja um fator apropriado para teste. Para valores “pequenos” desse fator, a estrutura M_1 deve ser escolhida, de outro modo a estrutura M_2 deve ser preferida.

O caso real com validação cruzada é mostrado na Figura 12.4.

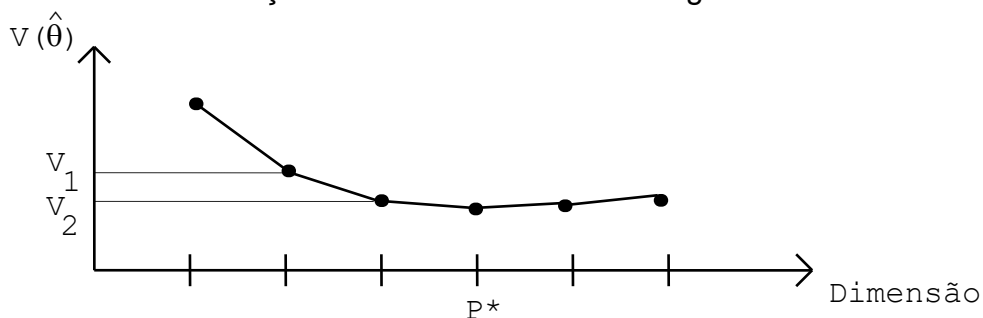


Fig. 12.4 Valor da função custo versus estrutura do modelo para caso real de validação cruzada (dados com perturbação e $N \ll \infty$).

Neste caso, a tendência é que a função custo atinja um mínimo e comece lentamente a crescer novamente.

12.4 PRINCÍPIO DA PARCIMÔNIA

O princípio da parcimônia é uma regra útil para determinar uma dimensão apropriada para o modelo. Este princípio diz que “dentre dois ou mais modelos que expliquem bem os dados coletados, o modelo com o menor número de parâmetros deve ser escolhido”. Essa frase segue a seguinte filosofia: não use parâmetros extras para descrever um sistema dinâmico se eles não forem necessários (BOX; JENKINS, 1976; SÖDERSTRÖM; STOICA, 1989). O uso do princípio da parcimônia é aqui ilustrado em um caso simples. Considere um sistema com saída única. Assuma que $\hat{\theta}_N$ denote um vetor de parâmetros, estimado a

partir dos dados disponíveis, usando o método do erro de predição (PEM). Assuma também que o sistema pertença à estrutura de modelo considerada. Isto significa que exista um vetor de parâmetros reais θ_0 (não necessariamente único).

Uma medida escalar é usada para avaliar a qualidade do modelo associado com $\hat{\theta}_N$. Tal medida deve ser uma função de $\hat{\theta}_N$ e é denotada por $W(\hat{\theta}_N)$ ou W_N , onde a dependência com o número de dados é enfatizada. A medida da avaliação $W(\theta)$ deve ser minimizada pelo vetor de parâmetros reais θ_0 (vide Figura 12.5) (SÖDERSTRÖM; STOICA, 1989).

$$W(\theta) \geq W(\theta_0) \quad \forall \theta$$

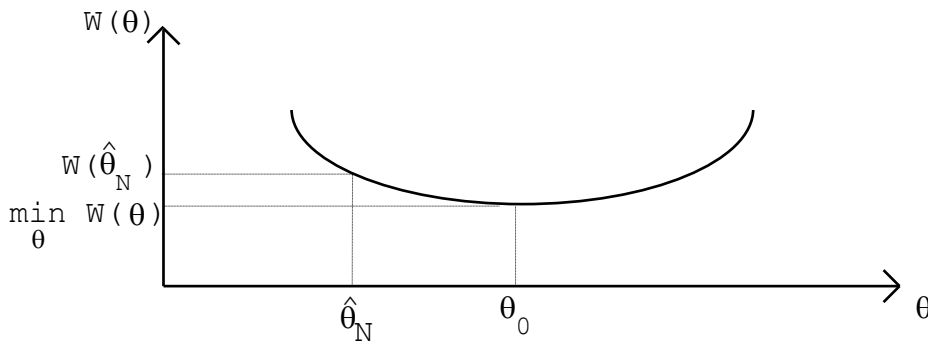


Fig. 12.5 Um critério de avaliação.

Quando a estimativa $\hat{\theta}_N$ se desviar um pouco de θ_0 , o critério W cresce um pouco acima de seu valor mínimo $W(\theta_0)$. Esse crescimento $[W(\hat{\theta}_N) - W(\theta_0)]$ é tomado como uma medida escalar do desempenho do modelo.

O critério de avaliação pode ser escolhido de diferentes maneiras, incluindo as seguintes (SÖDERSTRÖM; STOICA, 1989):

- a variância dos erros de predição um passo à frente, quando o modelo é aplicado para prever saídas futuras;
- a variância dos erros de predição múltiplos-passos-à-frente;
- o desvio entre a função de transferência estimada e o processo real. Tal desvio pode ser expresso no domínio da frequência; etc.

Neste caso, seja W_N a variância do erro de predição um passo à frente quando o modelo correspondente a $\hat{\theta}_N$ é usado para prever saídas futuras. Isto significa que (SÖDERSTRÖM; STOICA, 1989):

$$W_N = E[\varepsilon^2(t, \hat{\theta}_N)] = V_N \quad (12.1)$$

Se as estimativas $\hat{\theta}_N$ fossem exatas, isto é, $\hat{\theta}_N = \theta_0$, então os erros de predição $\{\varepsilon(t, \hat{\theta}_N)\}$ seriam ruído branco $\{e(t)\}$ e teriam variância σ^2 .

Considere agora quanto a variância do erro de predição W_N cresce devido ao desvio de $\hat{\theta}_N$ com relação ao valor real θ_0 , incluindo-se uma penalização para a dimensão p do modelo. Suponha que as mesmas condições experimentais sejam válidas para os dados passados e futuros. Resulta o seguinte valor para a expectância ou esperança matemática de W_N (SÖDERSTRÖM; STOICA, 1989):

$$E(W_N) \cong \sigma^2 \cdot \left(1 + \frac{p}{N}\right) \quad \text{onde } p = \dim(\theta) \quad (12.2)$$

Para gerar a Expressão (12.2) não se assumiu nenhuma estrutura específica de modelo. Ela informa que a variância esperada dos erros de predição cresce de p/N . Assim há uma penalidade em usar modelos com um número desnecessariamente alto de parâmetros. Isto pode ser considerado como uma expressão formal do princípio da parcimônia (SÖDERSTRÖM; STOICA, 1989).

12.5 CRITÉRIOS PARA SELEÇÃO DA DIMENSÃO DE MODELOS

O critério aqui proposto propõe imitar a ideia da validação cruzada, isto é, usar o valor de V_N como base de uma estimativa do que se teria, caso se tivesse aplicado a avaliação a um conjunto novo de dados, precavendo-se assim dos efeitos do sobreajuste. O critério proposto penaliza a redução da função custo $V_N(\hat{\theta}_N)$ com o aumento da dimensão do modelo. A estrutura de modelo que provê o menor valor deste critério é escolhida. Uma forma geral desse critério é a seguinte (SÖDERSTRÖM; STOICA, 1989):

$$W_N = V_N(\hat{\theta}_N) [1 + \beta(N, p)] \quad (12.3)$$

onde $\beta(N, p)$ é uma função de N e do número de parâmetros p no modelo, que deve crescer com p , de modo a penalizar estruturas de modelo complexas demais (sobrepametrizadas) de acordo com o princípio da parcimônia, mas deve tender a zero conforme $N \rightarrow \infty$, de modo a garantir que o termo de penalização em (12.3) não ofusque o decréscimo da função custo $V_N(\hat{\theta}_N)$ com estruturas de modelo subparametrizadas. Uma escolha típica é (SÖDERSTRÖM; STOICA, 1989):

$$\beta(N, p) = \frac{2p}{N}$$

Uma alternativa é usar o critério (SÖDERSTRÖM; STOICA, 1989):

$$W_N = \ln[V_N(\hat{\theta}_N)] + \beta(N, p) \quad (12.4)$$

onde o termo $\beta(N, p)$ deve penalizar modelos de dimensão elevada. A escolha $\beta(N, p) = 2p/N$ gera o **critério de informação de Akaike** (AIC - Akaike's information criterion) (SÖDERSTRÖM; STOICA, 1989):

$$W_N = \ln[V_N(\hat{\theta}_N)] + \frac{2p}{N} = AIC \quad (12.5)$$

A Expressão (12.5) pode ser aproximada por:

$$AIC \approx \ln\left[\left(1 + \frac{2p}{N}\right) \cdot V_N(\hat{\theta}_N)\right] \quad (12.6)$$

Não é difícil verificar que os Critérios (12.3) e (12.4) são assintoticamente equivalentes. Assim, para N grande, tem-se a partir de (12.3) (SÖDERSTRÖM; STOICA, 1989):

$$\ln\{V_N(\hat{\theta}_N) [1 + \beta(N, p)]\} = \ln[V_N(\hat{\theta}_N)] + \ln[1 + \beta(N, p)] \approx \ln[V_N(\hat{\theta}_N)] + \beta(N, p)$$

Como $\beta(N, p) = 2p/N$ é normalmente um valor pequeno, não ultrapassando 0,02, pode-se assumir que $\ln[1 + \beta(N, p)] \cong \beta(N, p)$. Por exemplo, para $\beta(N, p) = 0,02$ resulta que $\ln(1 + 0,02) = 0,0198$.

Resulta que os critérios das Equações (12.3) e (12.4) são minimizados pelas mesmas estruturas de modelo. Há outras propostas na literatura para o termo $\beta(N, p)$ em (12.3), além da sugestão de Akaike $[\beta(N, p) = 2p/N]$. Outro critério bastante usado é visto a seguir (SÖDERSTRÖM; STOICA, 1989):

$$W_N = V_N(\hat{\theta}_N) \frac{1 + \frac{p}{N}}{1 - \frac{p}{N}} = FPE \quad (12.7)$$

Este critério é conhecido como erro final de predição (FPE - *final prediction error*) e também foi proposto por Akaike. Ele reflete a variância do erro de predição que se obtém, na média, quando o modelo $M(\hat{\theta}_N)$ é aplicado como um preditor a conjuntos de dados diferentes daquele que foi usado para a estimação dos parâmetros (LJUNG, 1999). Para N grande, a Expressão (12.7) pode ser aproximada por (SÖDERSTRÖM; STOICA, 1989):

$$W_N = V_N(\hat{\theta}_N) \left[1 + \frac{2p/N}{1 - p/N} \right] \approx V_N(\hat{\theta}_N) \left[1 + \frac{2p}{N} \right]$$

que é da forma (12.3) com $\beta(N, p) = 2p/N$. Assim, para $p/N = 0,01$, resulta que:

$$\left[1 + \frac{2p/N}{1 - p/N} \right] = 1,0202 \quad \text{e} \quad \left[1 + \frac{2p}{N} \right] = 1,02$$

Portanto, segundo a teoria de Akaike, em um conjunto de diferentes modelos, escolha aquele com menor valor de *FPE* ou *AIC*.

12.5.1 Análise de consistência dos critérios FPE e AIC

Os critérios FPE e AIC não provêm estimativas consistentes da dimensão do modelo, isto é, existe um risco diferente de zero, mesmo para um grande número N de dados coletados, de se escolher uma dimensão muito alta para o modelo.

OBS.: uma regra de seleção da dimensão de um modelo é dita consistente se a probabilidade de escolher uma dimensão errada tenda a zero conforme o número de pontos de dados coletados tenda a infinito.

Os Critérios (12.3) e (12.4) podem ser interpretados como um teste χ^2 (vide apêndice E), com um nível pré-especificado de significância, como mostrado em (12.8) e (12.9) (SÖDERSTRÖM; STOICA, 1989):

$$\chi^2_\alpha(p_2 - p_1) = N \frac{\beta(N, p_2) - \beta(N, p_1)}{1 + \beta(N, p_1)} \quad (12.8)$$

$$\chi^2_\alpha(p_2 - p_1) = N \left\{ \exp \left[\frac{\beta(N, p_2) - \beta(N, p_1)}{N} \right] - 1 \right\} \quad (12.9)$$

Considere agora que se deseje especificar a Equação (12.8) para o critério FPE. Então, a partir de (12.3) e (12.7), resulta:

$$\beta(N, p) = \frac{2p}{N - p} \quad (12.10)$$

As Equações (12.8) e (12.9) mostram que certos níveis de significância podem ser associados com os critérios FPE e AIC quando eles selecionam duas estruturas de modelo. Substituindo-se (12.10) em (12.8), tem-se (SÖDERSTRÖM; STOICA, 1989):

$$\chi_{\alpha}^2(p_2 - p_1) = N \frac{\frac{2p_2}{N-p_2} - \frac{2p_1}{N-p_1}}{1 + \frac{2p_1}{N-p_1}} = N \frac{2N(p_2 - p_1)}{(N+p_1)(N+p_2)} \approx 2(p_2 - p_1) \quad (12.11)$$

onde a aproximação \approx vale para valores grandes de N .

Se, em particular, $p_2 - p_1 = 1$, então para valores grandes de N , $\chi_{\alpha}^2(1) = 2$, que resulta em $\alpha = 0,177$ (interpole linearmente na tabela D.1). Isto significa que o risco de selecionar a estrutura maior M_2 quando M_1 (com $p_2 - p_1 = 1$) é mais apropriada é assintoticamente igual a 17,7% (SÖDERSTRÖM; STOICA, 1989).

No caso do critério AIC, tem-se:

$$\beta(N, p) = 2p \quad (12.12)$$

Substituindo-se (12.12) em (12.9), resulta:

$$\chi_{\alpha}^2(p_2 - p_1) = N \left\{ \exp \left[\frac{2(p_2 - p_1)}{N} \right] - 1 \right\} \approx 2(p_2 - p_1) \quad (12.13)$$

onde a aproximação \approx vale para valores grandes de N . Quando $p_2 = p_1 + 1$, o risco de sobreparametrizar é de 17,7%, exatamente como no caso do FPE, conforme já era esperado, pois os critérios AIC e FPE são assintoticamente equivalentes. Repare que o risco de sobreparametrização não desaparece quando $N \rightarrow \infty$, o que significa que ambas as estimativas da dimensão (AIC e FPE) não são consistentes. Este fato não é um problema muito sério. Ambos os critérios gozam de certas propriedades no importante caso prático em que o processo não pertence à classe de estruturas de modelo considerada. Mais precisamente, tanto o critério AIC quanto o FPE definem bons modelos de predição, não importando se o sistema real pertence ou não ao conjunto de modelos.

12.6 PENALIZAÇÃO PELA COMPLEXIDADE DO MODELO

O critério AIC de Akaike pode ser encarado da seguinte forma (LJUNG, 1999):

$$\hat{\theta}_{N,AIC} = \arg \min_{\theta} \left[V_N(\hat{\theta}_N) + \frac{p}{N} \right] \quad (12.14)$$

Os métodos AIC e FPE podem ser considerados, de um ponto de vista pragmático, como sendo critérios para determinar a estrutura do modelo e do valor dos parâmetros dentro da estrutura. Conceitualmente, eles poderiam ser escritos como (LJUNG, 1999):

$$W_N(\theta_N, M) = V_N(\theta_N) + U_N(M) \quad (12.15)$$

onde M corresponde à estrutura do modelo, V_N é o critério de predição de erro dentro de certas estruturas de modelo M e $U_N(M)$ é uma função que mede alguma “complexidade” da estrutura do modelo.

Nos casos tratados até aqui [vide (12.14)], essa medida é relacionada com a dimensionalidade de θ (LJUNG, 1999):

$$U_N(M) = \frac{p}{N} \quad (12.16)$$

Os métodos do tipo PEM foram direcionados para prover descrições do sistema que gerem o menor erro médio quadrático. Um modelo que aparentemente forneça um menor erro médio quadrático de predição será escolhido, mesmo se for muito complexo. Na prática, pode-se desejar adicionar uma penalidade extra a (12.16) para a complexidade do modelo, refletindo o custo de usá-lo: “caso se aceite um modelo mais complexo, ele deve provar ser significativamente melhor” (LJUNG, 1999).

O que um modelo complexo significa e qual penalidade associar com ele são usualmente assuntos subjetivos. Um enfoque para esse problema adotado por Rissanen (1978) é que o principal objetivo da identificação é encontrar a menor descrição possível dos dados observados. Isso leva a um critério do tipo (12.15) com (LJUNG, 1999):

$$U_N(M) = p \frac{\log(N)}{N} \quad (12.17)$$

Este princípio é intitulado **descrição de mínimo comprimento** (MDL - *minimum description length*). Este princípio foi denominado BIC - **critério de informação de Bayes** (*Bayes' information criterion*) por Akaike.

12.7 ÍNDICES EMPREGADOS NA AVALIAÇÃO DE MODELOS

Em aplicações industriais, é frequentemente desejável ter-se um modelo que seja tão simples quanto possível, embora preciso o suficiente para propósitos de controle ou simulação. Com esse objetivo, pode-se determinar a ordem do modelo usando a função custo do erro de saída relativo, definida por (ZHU; BACKX, 1993):

$$RV = \frac{\sum_{t=1}^N [\varepsilon_{OE}(t)]^2}{\sum_{t=1}^N [y(t)]^2} \quad (12.18)$$

A regra é buscar a menor ordem, tal que RV seja menor que algum valor predeterminado. Em (ZHU, BACKX, 1993) sugere-se que um valor adequado esteja entre 10% a 20%, quando o objetivo do modelo é a simulação ou o projeto de controladores por realimentação. Dois outros índices usados são definidos a seguir (AGUIRRE, 2007):

$$RMSE = \frac{\sqrt{\sum_{t=1}^N [y(t) - \hat{y}(t)]^2}}{\sqrt{\sum_{t=1}^N [y(t) - \bar{y}]^2}} = \frac{\|y(t) - \hat{y}(t)\|}{\|y(t) - \bar{y}\|} \quad (12.19)$$

onde RMSE significa *root mean squared error*, sendo que $\hat{y}(t)$ é a simulação livre do sinal e \bar{y} é o valor médio do sinal medido $y(t)$, com a média calculada pelos dados medidos.

$$RMSE(i) = \frac{\sqrt{\sum_{t=1}^N [y(t) - \hat{y}(t)]^2}}{\sqrt{\sum_{t=1}^N [y(t) - y(t-i)]^2}} = \frac{\|y(t) - \hat{y}(t)\|}{\|y(t) - y(t-i)\|} \quad (12.20)$$

sendo que i indica o horizonte de predição, medido em intervalos de amostragem.

No caso do índice da Equação (12.19), comparam-se as predições do modelo com a média temporal do sinal, ou seja, a média usada como um preditor trivial. Analogamente, $RMSE(i)$ compara o modelo com o sinal original deslocado de i períodos de amostragem à frente. Em ambos os casos, valores inferiores a 1 indicam um melhor desempenho em relação ao preditor padrão considerado (média ou i passos à frente). Evidentemente, há outros índices, como o somatório dos valores absolutos dos erros de predição ou o somatório dos erros de predição ao quadrado (função custo) (AGUIRRE, 2007).

O MATLAB calcula o índice de ajuste (fit), dado por:

$$fit(\%) = 100 \cdot \left(1 - \frac{\|\hat{y}(t) - y(t)\|}{\|y(t) - \bar{y}\|} \right) \quad (12.21)$$

Em termos de comandos disponíveis no Matlab, a Expressão (12.21) se torna:

$$fit(\%) = 100 \cdot \left\{ 1 - \frac{\text{norm}[\hat{y}(t) - y(t)]}{\text{norm}[y(t) - \text{mean}(y(t))]} \right\} \quad (12.21a)$$

12.8 EXEMPLO DO EFEITO DA DIMENSÃO DO MODELO NA AUTOVALIDAÇÃO E NA VALIDAÇÃO CRUZADA

Para este exemplo emprega-se o seguinte processo:

$G(s) = \frac{3}{8 \cdot s + 1} \frac{2}{10 \cdot s + 1} \frac{1}{12 \cdot s + 1} = \frac{6}{960 \cdot s^3 + 296 \cdot s^2 + 30 \cdot s + 1}$ Discretizando-o com um segurador de ordem zero e período de amostragem $T = 2$ s, chega-se ao seguinte processo de terceira ordem:

$$G(q) = \frac{0,007149 \cdot q^{-1} + 0,02453 \cdot q^{-2} + 0,005252 \cdot q^{-3}}{1 - 2,444 \cdot q^{-1} + 1,99 \cdot q^{-2} - 0,5397 \cdot q^{-3}}$$

Dessa forma, tem-se que:

$$A(q) = 1 - 2,444 \cdot q^{-1} + 1,99 \cdot q^{-2} - 0,5397 \cdot q^{-3} \quad (12.22)$$

$$B(q) = 0,007149 \cdot q^{-1} + 0,02453 \cdot q^{-2} + 0,005252 \cdot q^{-3}$$

Montam-se então dois sistemas: um com estrutura ARX e outro com estrutura ARMAX, conforme indicado a seguir.

$$H_{ARX}(q) = \frac{1}{1 - 2,444 \cdot q^{-1} + 1,99 \cdot q^{-2} - 0,5397 \cdot q^{-3}}$$

$$H_{ARMAX}(q) = \frac{1 + 0,75 \cdot q^{-1} + 0,5 \cdot q^{-2} + 0,25 \cdot q^{-3}}{1 - 2,444 \cdot q^{-1} + 1,99 \cdot q^{-2} - 0,5397 \cdot q^{-3}}$$

Propõe-se simular estes sistemas em Simulink por 4000 s, coletando 2000 pontos de entrada e saída, dos quais 1000 são usados como dados de estimação e 1000 como dados de validação. São realizados ensaios com diferentes níveis de perturbação, com as seguintes variâncias σ^2 do ruído branco $e(t)$: 0 (sem perturbação), 10^{-8} (apenas ruído de medição), 10^{-6} (pouca perturbação), 10^{-5} (perturbação média) e 10^{-4} (muita perturbação). A frequência do ruído branco é 0,5 Hz. Apenas para se ter uma ideia do que representa na saída do sistema estes diferentes níveis de perturbação, exibe-se na Figura 12.6 a resposta do sistema apenas às perturbações e ao ruído de medição.

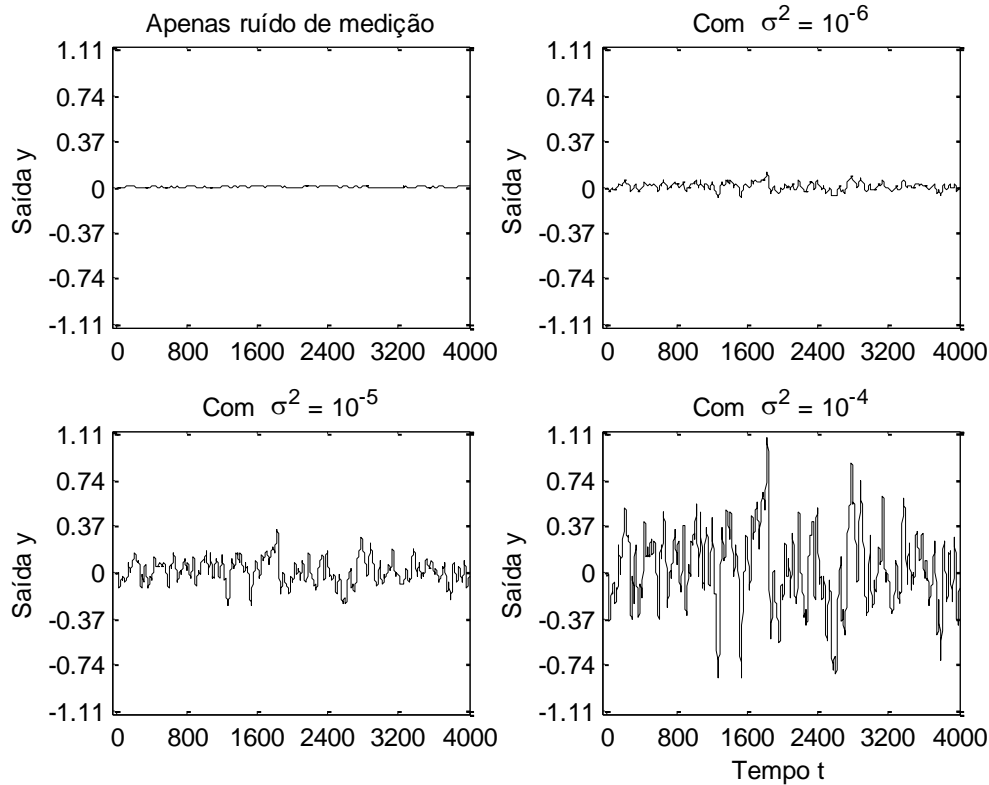


Fig. 12.6 Resposta do sistema de 3ª ordem aos diferentes níveis de perturbação.

12.8.1 Variações da função custo em termos da dimensão do modelo

O sinal de entrada para identificar este sistema de 3ª ordem, em que é preciso estimar 6 parâmetros, é um sinal PRBS de amplitude 1 e período $4 \cdot T=8$ s. Aplica-se um modelo com estrutura ARX. Supõe-se que se vá aumentando a ordem do modelo e que se faça $n_a=n_b=n$ e $n_k=1$ e que a dimensão do modelo seja $p=n_a+n_b=2 \cdot n$.

Calcula-se o valor da função custo V_N , para cada um dos quatro níveis de perturbação da Seção 12.8, considerando-se tanto o caso de autovalidação como de validação cruzada. O procedimento em Matlab para esse cálculo é mostrado a seguir.

```
% Cálculo da função custo
ordem = 6;
for i=1:ordem,
    NN(i,1) = i;
    NN(i,2) = i;
    NN(i,3) = 1;
end
```

```

ide = 1000;
val = 2000;
N = ide;
eixo_x = [1:ordem];
ze = [y(1:ide) u(1:ide)];
zv = [y(ide+1:val) u(ide+1:val)];
Va = arxstruc(ze,ze,NN); % Autovalidação
Vc = arxstruc(ze,zv,NN); % Validação cruzada
nn = selstruc(Vc,0);
nn_AIC = selstruc(Vc,'AIC');
nn_MDL = selstruc(Vc,'MDL');
nn_5 = selstruc(Vc,5);
plot(eixo_x,Va(1,1:dim),'k', eixo_x,Vc(1,1:dim),'k:')
ll = legend('Autovalidação','Validação cruzada');
set(ll,'FontSize',9)
xlabel('Ordem n=n_a=n_b do modelo')
ylabel('Função custo V')
hold

```

A Figura 12.7 mostra os gráficos da função custo com ordem n de 1 a 6. O comportamento dos quatro gráficos da Figura 12.7 é muito parecido, com a resposta da auto validação e da validação cruzada sendo também muito similares. Com ordem $n_a = n_b = n = 1$ a função custo é elevada, ao passar para ordem $n=2$ ela cai bastante e de $n=3$ em diante, a tendência da função custo é de uma queda quase imperceptível.

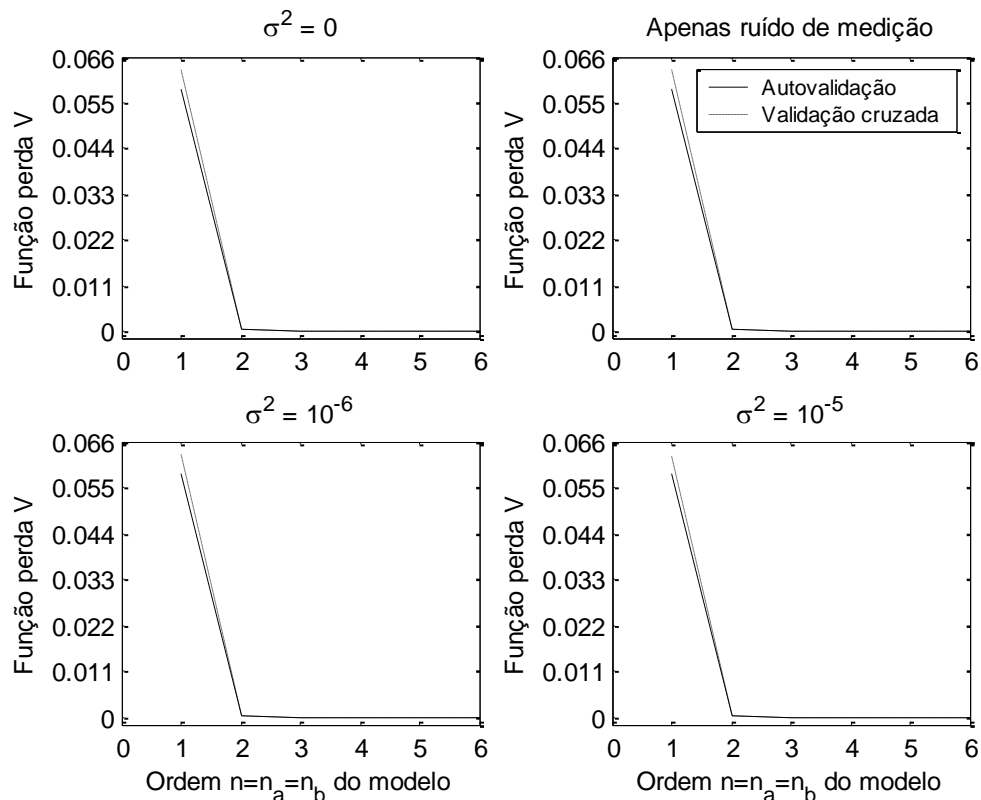


Fig. 12.7 Teste da ordem do modelo ARX para diferentes níveis de perturbação em sistema com estrutura ARX.

Mostra-se na Figura 12.8 uma ampliação dos gráficos da Figura 12.7.

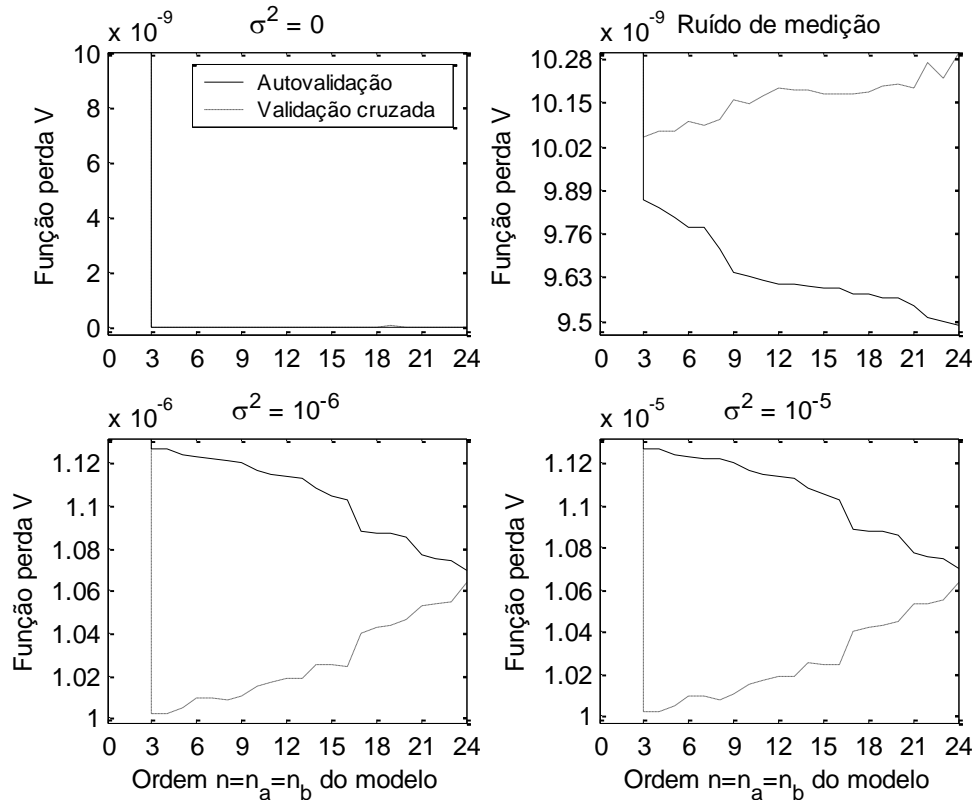


Fig. 12.8 Ampliação dos gráficos da Figura 12.7.

No caso da autovalidação na Figura 12.8, a função custo tende a reduzir lentamente, conforme a dimensão do modelo cresce, significando que quanto mais complexo é o modelo, melhor ele descreve os dados de saída que o gerou. Na validação cruzada ocorre um efeito contrário, com uma lenta tendência de crescimento da função custo. No entanto, em ambos os casos, a taxa de redução ou aumento é bastante reduzida.

A validação está relacionada com a escolha da ordem do modelo. No Matlab pode-se buscar o melhor modelo por critérios distintos, através do comando `nn=selstruc(Vc,c)`, onde V_c é a função custo da validação cruzada e c pode assumir diferentes valores:

- se $c = 0$:
`nn = selstruc(Vc,0)` retorna a estrutura de modelo com o melhor ajuste (maior valor de *fit*, conforme a Equação (12.21).
- se $c = \text{'AIC'}$:
`nn=selstruc(Vc,'AIC')` escolhe a estrutura com o menor valor do critério AIC de Akaike.
- se $c = \text{'MDL'}$:
`nn=selstruc(Vc,'MDL')` escolhe a estrutura com o menor valor do critério MDL de Rissanen.
- se c é um número inteiro, a estrutura é escolhida minimizando-se $(1 + c \cdot p/N) \cdot V_c$, onde p é a dimensão do modelo (número de parâmetros estimados), V_c é a função custo dos dados de validação e N é o número de pontos usado na validação.

Aplicando-se este comando aos modelos desta subseção, resulta a Tabela 12.1.

Tab. 12.1 Seleção da melhor ordem do modelo.

| c | $\sigma^2 = 0$ | $\sigma^2 = 10^{-8}$ | $\sigma^2 = 10^{-6}$ | $\sigma^2 = 10^{-5}$ | $\sigma^2 = 10^{-4}$ |
|--------------------|---------------------|----------------------|----------------------|----------------------|----------------------|
| $c = 0$ | $n_a = 4 \ n_b = 4$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ |
| $c = \text{'AIC'}$ | $n_a = 4 \ n_b = 4$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ |
| $c = \text{'MDL'}$ | $n_a = 4 \ n_b = 4$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ |
| $c = 5$ | $n_a = 4 \ n_b = 4$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ | $n_a = 3 \ n_b = 3$ |

O único resultado não esperado na Tabela 12.1 se refere a $\sigma^2 = 0$. Neste caso, o valor calculado pelo Matlab para a função custo para as várias ordens do modelo é:

$$V_c(n=1) = 0,06287$$

$$V_c(n=2) = 6,9464 \cdot 10^{-4}$$

$$V_c(n=3) = 1,8637 \cdot 10^{-15}$$

$$V_c(n=4) = 2,2204 \cdot 10^{-16}$$

$$V_c(n=5) = 2,2204 \cdot 10^{-16}$$

$$V_c(n=6) = 1,2114 \cdot 10^{-14}$$

Baseado nestes valores, nota-se que com base em números muito pequenos, o Matlab escolhe a ordem $n=4$, pois com ordem $n=3$ a função custo ainda não havia atingido o seu valor mínimo. Na Figura 12.9 mostram-se os gráficos para o sistema com estrutura ARMAX sendo identificado por um modelo com estrutura ARX. Os gráficos são similares aos da Figura 12.7, embora a amplitude da função custo tenha crescido um pouco, pois agora o processo e o modelo têm estruturas diferentes.

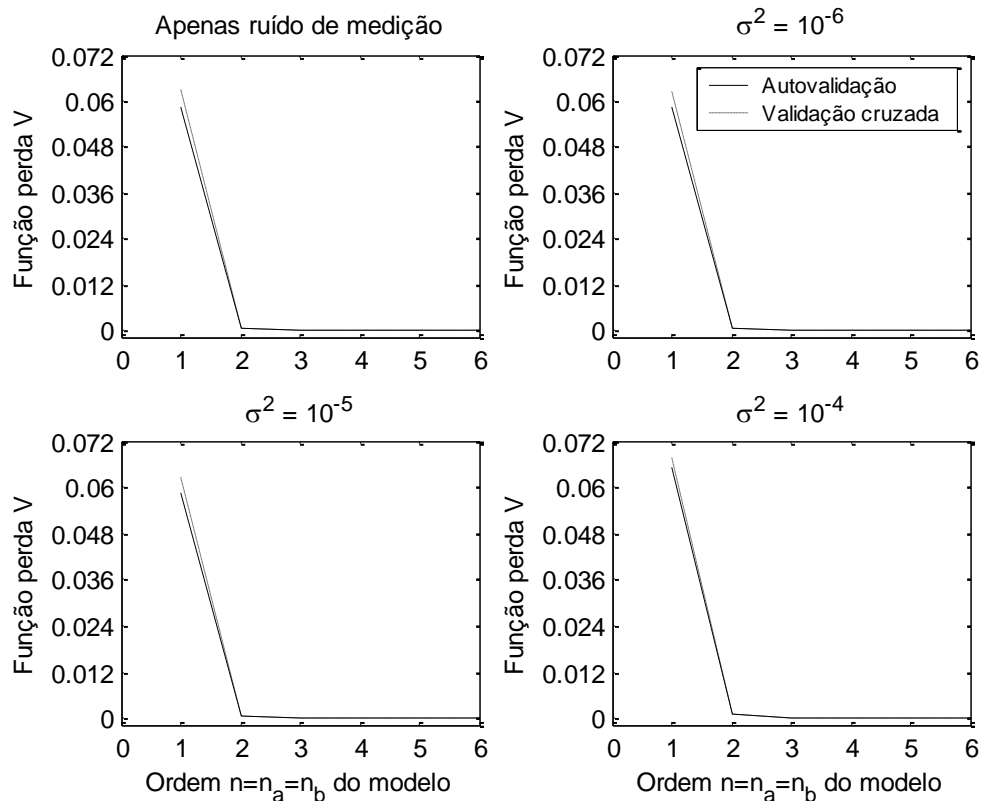


Fig. 12.9 Teste da ordem do modelo ARX para diferentes níveis de perturbação quando se emprega sistema com estrutura ARMAX.

A Figura 12.10 apresenta uma ampliação dos gráficos da Figura 12.9.

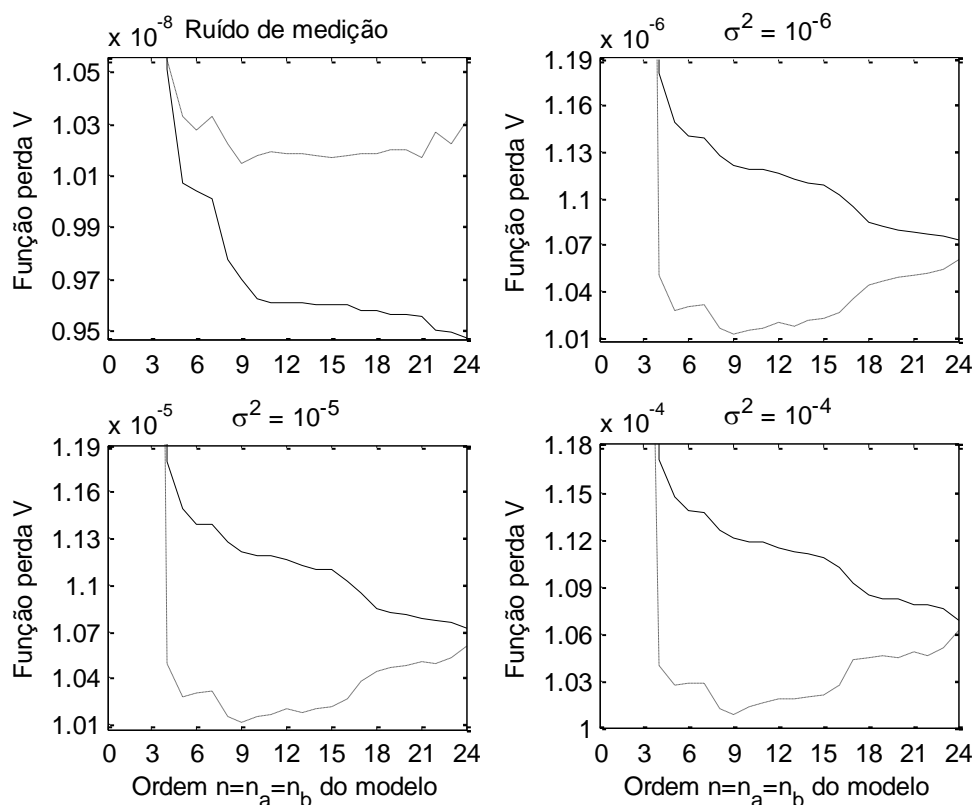


Fig. 12.10 Ampliação dos gráficos da Figura 12.9.

Como a estrutura do sistema não é a mesma do modelo identificado, o valor da função custo na Figura 12.10 cresce com relação ao da Figura 12.8. Percebe-se que o valor mínimo da função perda na validação cruzada ocorre para $n=9$.

Apresenta-se na Tabela 12.2 a seleção da melhor ordem do modelo para o modelo ARX identificando uma estrutura ARMAX.

Tab. 12.2 Seleção da melhor ordem do modelo.

| c | $\sigma^2 = 0$ | $\sigma^2 = 10^{-8}$ | $\sigma^2 = 10^{-6}$ | $\sigma^2 = 10^{-5}$ | $\sigma^2 = 10^{-4}$ |
|--------------------|---------------------|----------------------|----------------------|----------------------|----------------------|
| $c = 0$ | $n_a = 4 \ n_b = 4$ | $n_a = 9 \ n_b = 9$ | $n_a = 9 \ n_b = 9$ | $n_a = 9 \ n_b = 9$ | $n_a = 9 \ n_b = 9$ |
| $c = \text{'AIC'}$ | $n_a = 4 \ n_b = 4$ | $n_a = 9 \ n_b = 9$ | $n_a = 5 \ n_b = 5$ | $n_a = 9 \ n_b = 9$ | $n_a = 8 \ n_b = 8$ |
| $c = \text{'MDL'}$ | $n_a = 4 \ n_b = 4$ | $n_a = 5 \ n_b = 5$ | $n_a = 5 \ n_b = 5$ | $n_a = 5 \ n_b = 5$ | $n_a = 4 \ n_b = 4$ |
| $c = 5$ | $n_a = 4 \ n_b = 4$ | $n_a = 5 \ n_b = 5$ | $n_a = 5 \ n_b = 5$ | $n_a = 5 \ n_b = 5$ | $n_a = 5 \ n_b = 5$ |

Comparando-se os valores da Tabela 12.2 com os da Tabela 12.1, percebe-se que na Tabela 12.2 a ordem dos modelos tende a ser maior, por se estar identificando um sistema com uma determinada estrutura com um modelo com outra estrutura.

Exibem-se na Figura 12.11 os gráficos da função perda do sistema com estrutura ARX, caso se force um erro no tempo morto do modelo, fazendo-se $n_k=2$ em vez de $n_k=1$.

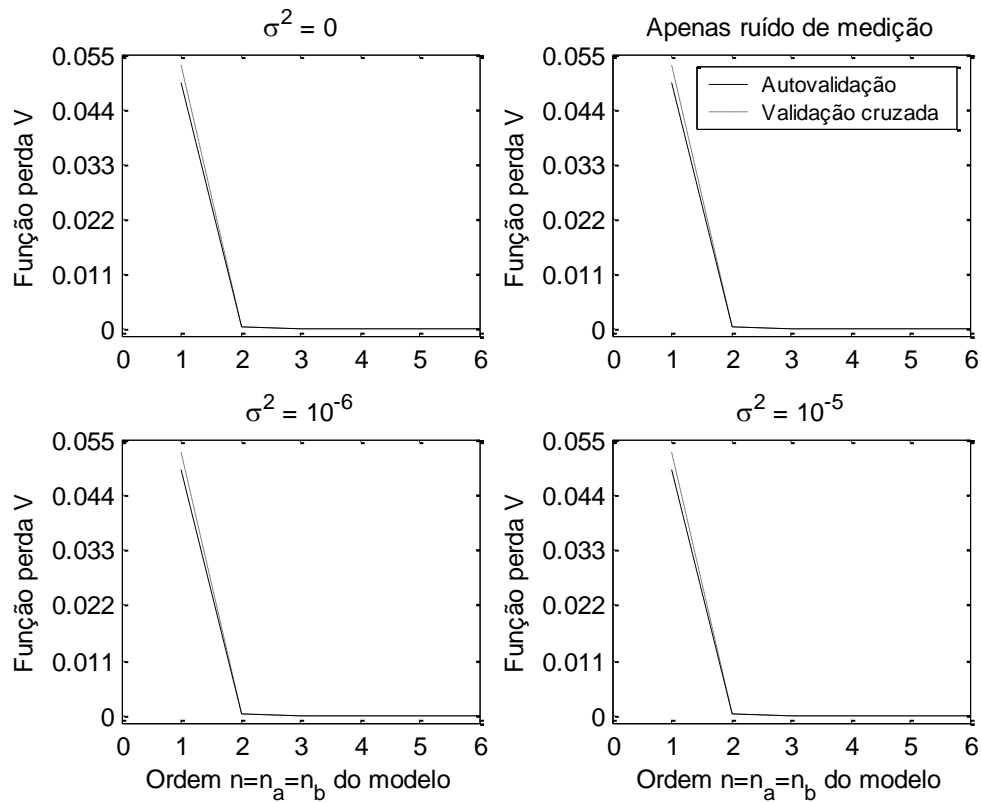


Fig. 12.11 Teste da ordem do modelo ARX com $n_k = 2$ para diferentes níveis de perturbação.

As Figuras 12.7 e 12.11 aparentemente são similares. Ampliando-se os gráficos da Figura 12.11, resulta a Figura 12.12.

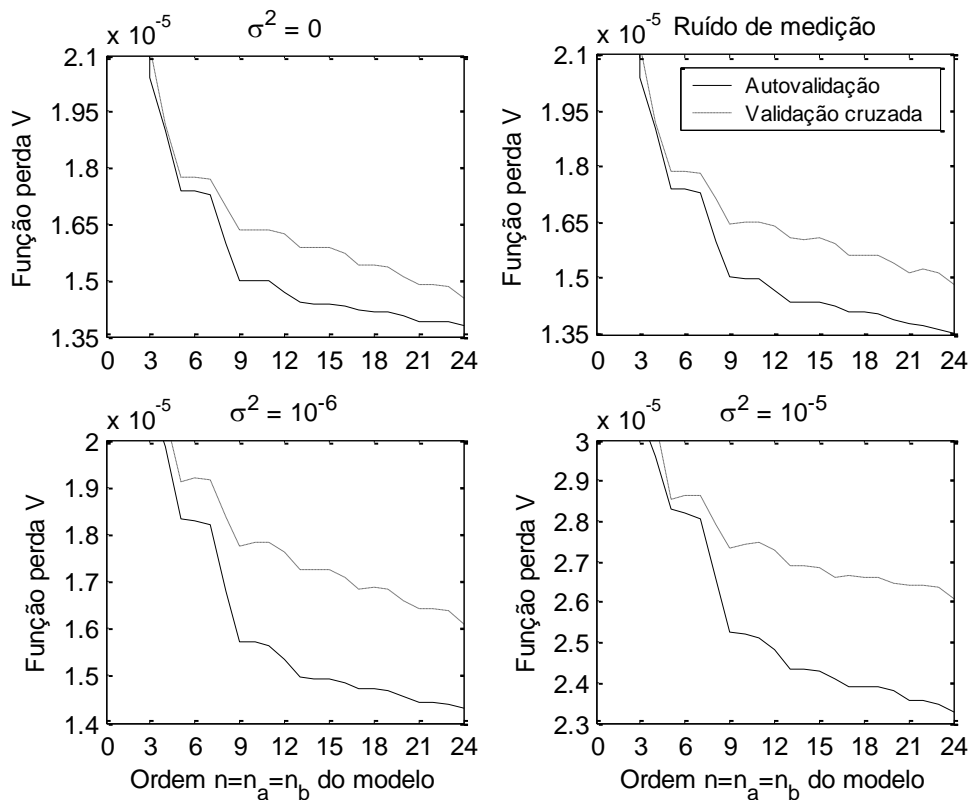


Fig. 12.12 Ampliação dos gráficos da Figura 12.11.

Comparando-se as Figuras 12.8 e 12.12, nota-se que a função custo tende a se estabilizar com o tempo morto incorreto em valores mais altos, que quando usando seu valor correto n_k . Além disso, a tendência de crescimento da função custo em função da ordem do modelo para a validação cruzada não aparece nos gráficos da Figura 12.12.

Conclui-se que a função custo varia com a ordem do modelo e é afetada por diferentes relações entre sinal e perturbação. A tendência é a função custo diminuir conforme a ordem do modelo cresça na autovalidação e aumentar conforme a ordem do modelo cresça na validação cruzada. Além disso, se o nível de ruído for muito baixo, a verdadeira ordem pode ser encontrada. Conforme o nível de perturbações aumenta, há uma tendência de se escolher um modelo de ordem mais elevada, quando se usa o método dos mínimos quadrados, como se vê na Tabela 12.2, onde a estrutura do modelo não é idêntica à do sistema, o que normalmente ocorre na prática. Isto é devido à polarização (*bias*) da estimativa dos mínimos quadrados e não devido à regra para selecionar a ordem do modelo. Assim, se uma perturbação elevada estiver presente, os modelos ARX serão de ordem elevada para descrever simultaneamente as características da perturbação e a dinâmica do sistema. Normalmente, não é desejável usar um modelo com ordem superior à real, pois um modelo de ordem mais alta demandará mais poder computacional, o que pode causar problemas numéricos no projeto de controladores (ZHU, BACKX, 1993).

12.8.2 Variação dos índices AIC e FPE em função da dimensão do modelo

Mostra-se aqui a variação dos índices de análise de desempenho AIC e FPE, para cada um dos quatro níveis de perturbação da Seção 12.8, considerando-se apenas a autovalidação. O procedimento em Matlab para cálculo desses índices é visto a seguir.

```
% Cálculo dos índices AIC e FPE
ordem = 6;
for i=1:ordem,
    NN(i,1) = i;
    NN(i,2) = i;
    NN(i,3) = 1; % 2
end
ide = 1000;
eixo_x = [1:ordem];
peso_FPE = 10;
peso_AIC = 10;
ze = [y(1:ide) u(1:ide)];
Va = arxstruc(ze,ze,NN);
for i=1:ordem,
    theta = arx(ze,[i i 1]);
    FPE_Matlab(i) = fpe(theta);
    AIC_Matlab(i) = aic(theta);
```

```

p = 2*i;
FPE1(i) = (1+peso_FPE*p/N)*Va(1,i)/(1-peso_FPE*p/N);
AIC1(i) = log(Va(1,i))+2*peso_AIC*p/N;
end
figure(1)
plot(eixo_x,FPE_Matlab,'k',eixo_x,FPE1,'k:')
ll = legend('FPE original', 'FPE com 10*p');
set(ll,'FontSize',8)
xlabel('Ordem n=n_a=n_b do modelo')
ylabel('FPE')
figure(2)
plot(eixo_x,AIC_Matlab,'k',eixo_x,AIC1,'k:')
ll = legend('AIC original','AIC com 10*2*p');
set(ll,'FontSize',8)
xlabel('Ordem n=n_a=n_b do modelo')
ylabel('AIC')

```

A penalização pelo aumento da dimensão do modelo é muito pequena nos índices AIC e FPE, como se vê nas Figuras 12.13 e 12.15. Para aumentá-la, usa-se um peso de 10 no cálculo do AIC e do FPE. O efeito desses pesos é visto nas Figuras 12.13 e 12.15.

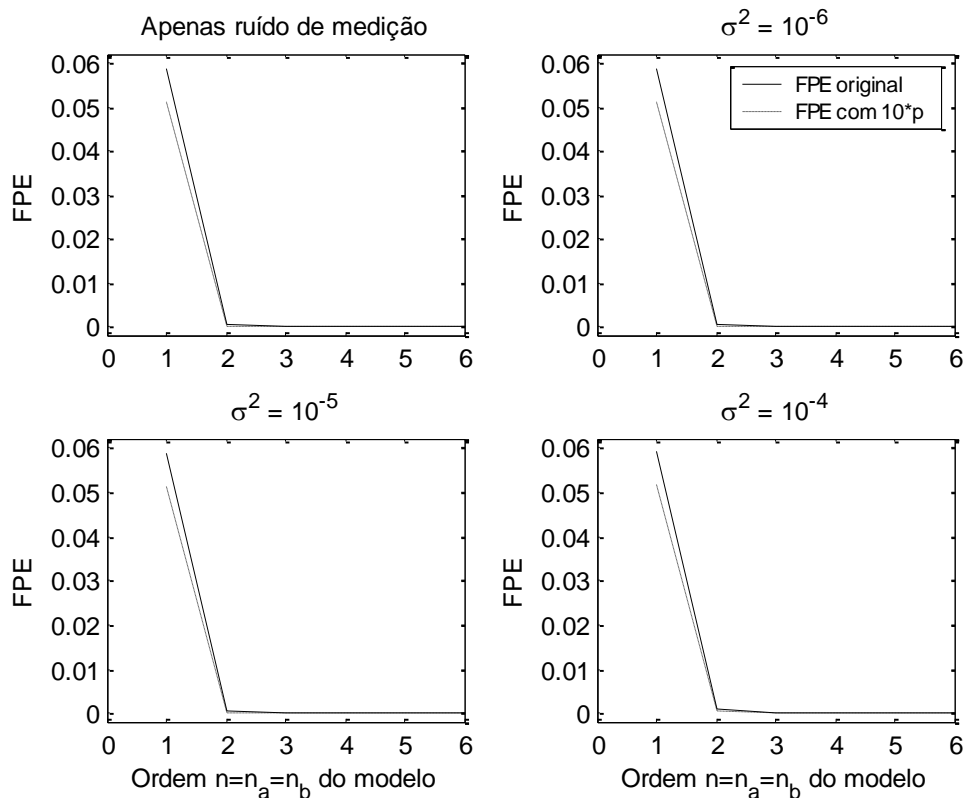


Fig. 12.13 Variação de FPE conforme se varia a ordem do modelo e o nível de perturbação.

Aplicando-se uma ampliação nos gráficos da Figura 12.13 resulta a Figura 12.14.

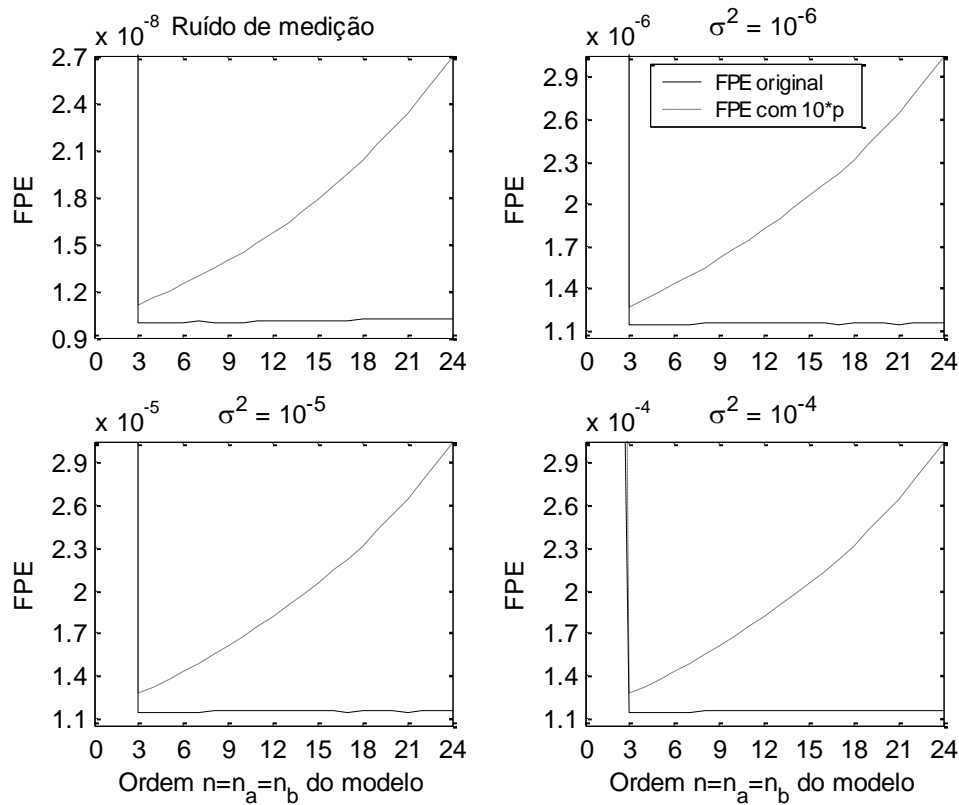


Fig. 12.14 Ampliação nos gráficos da Figura 12.13.

Nota-se nos gráficos da Figura 12.14, que com auxílio do FPE e, principalmente do FPE com ganho de 10, há um crescimento da função custo a partir de um valor mínimo, o qual se situa próximo à ordem “correta” do modelo.

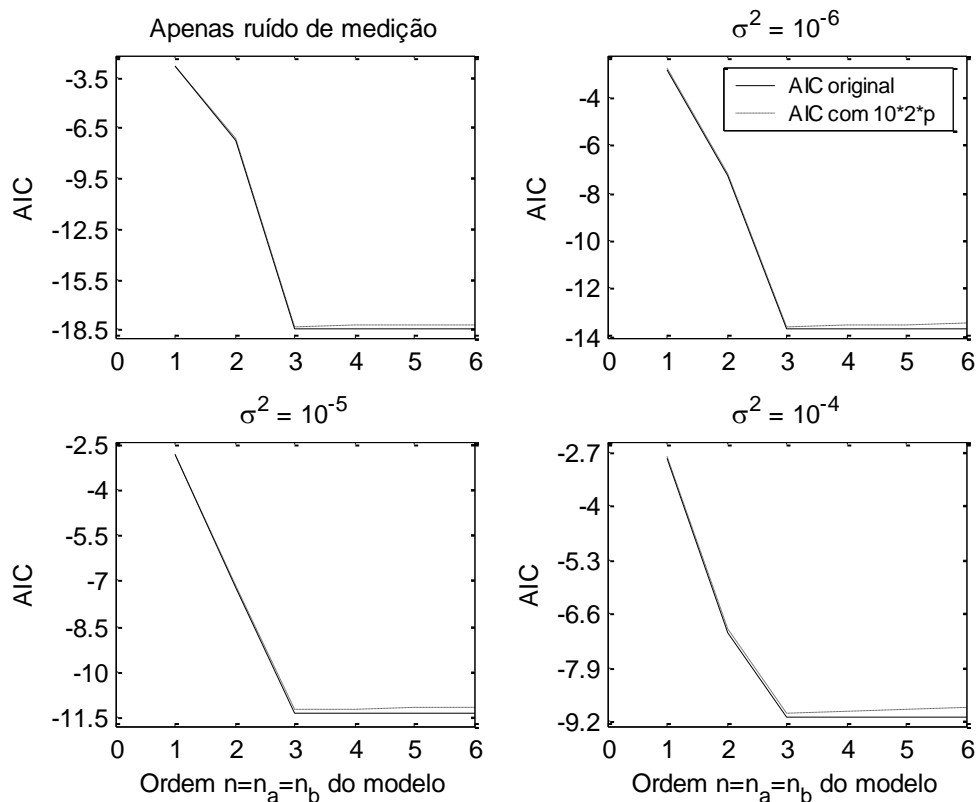


Fig. 12.15 Variação de AIC segundo a ordem do modelo e o nível de perturbação.

Aplicando-se uma ampliação nos gráficos da Figura 12.15 resulta a Figura 12.16.

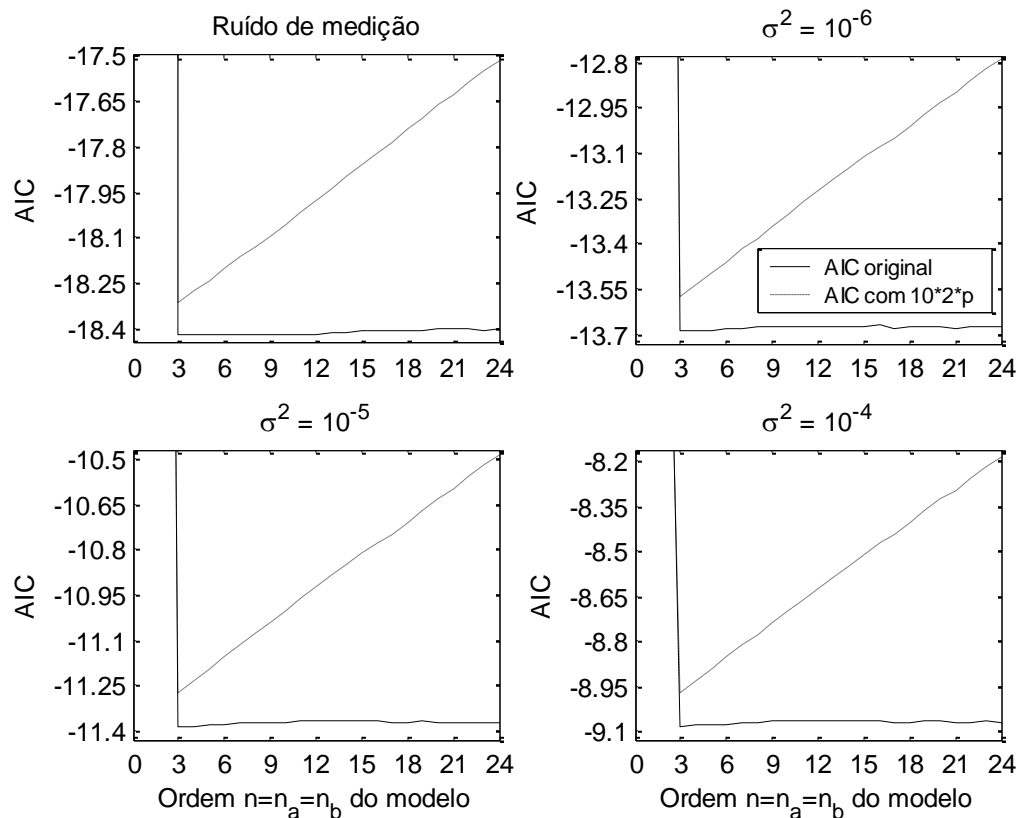


Fig. 12.16 Ampliação nos gráficos da Figura 12.15.

Como já visto na Figura 12.14, nota-se que com o uso do AIC e, principalmente dele com ganho, há um valor mínimo da função custo próximo à ordem “correta” do modelo.

12.9 COMPARAÇÃO DA RESPOSTA DO MODELO E DO SISTEMA E EMPREGO DO ÍNDICE DE AJUSTE “FIT”

A visualização da resposta do modelo e da resposta do sistema com os dados de validação é normalmente importante para se verificar o grau de coincidência entre ambas. Para tal, apresenta-se, a seguir, o procedimento em Matlab que permite realizar essa comparação, incluindo-se o valor do índice *fit*, conforme a Equação (12.21).

```
% Comparação da resposta real e simulada
T = 2;
ide = 1000;
val = 2000;
ze = iddata(y(1:ide),u(1:ide),T);
zv = [y(ide+1:val) u(ide+1:val)];
theta = arx(ze,[3 3 1]);
figure(1)
compare(zv,theta) % Simulação livre
figure(2)
compare(zv,theta,1) % Predição 1 passo à frente
figure(3)
compare(zv,theta,10) % Predição 10 passos à frente
```

A Figura 12.17 exibe o resultado do Matlab ao se usar o modelo ARX com $n_a = 3$, $n_b = 3$ e $n_k = 1$ para representar o sistema ARX com esta mesma ordem, com variância do ruído branco igual a $\sigma^2 = 10^{-5}$. Usa-se inicialmente a simulação livre com validação cruzada, com o modelo e o sistema excitados por um sinal PRBS de amplitude unitária e período 8 s. Foram usados 1000 pontos para gerar o modelo e 1000 pontos para validá-lo.

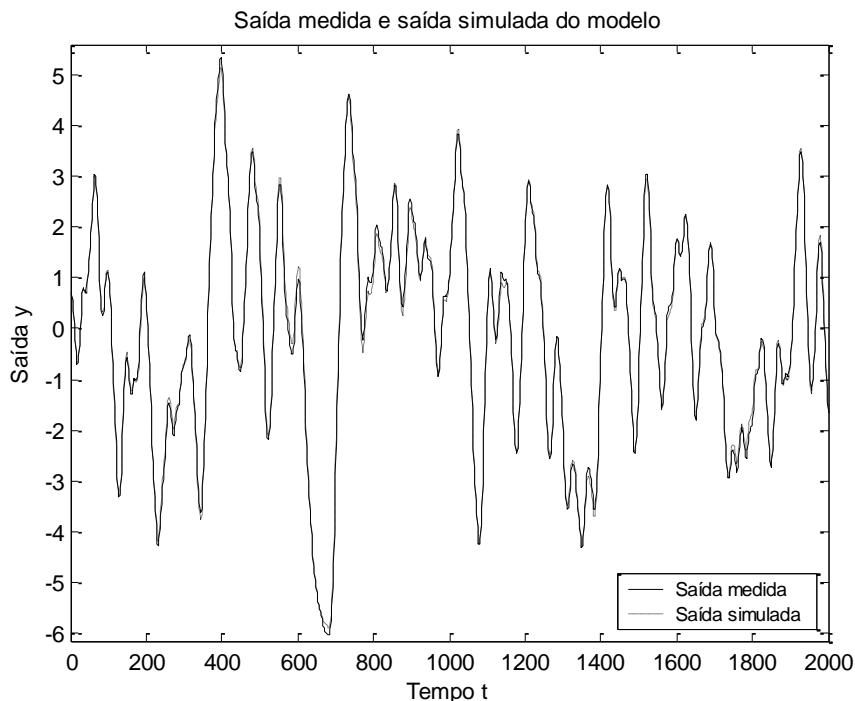


Fig. 12.17 Simulação livre com dados de validação para modelo e sistema ARX com $n_a = 3$, $n_b = 3$ e $n_k = 1$ e variância do ruído branco $\sigma^2 = 10^{-5}$.

O índice de ajuste é de 95,45%. Os parâmetros estimados pelo modelo são:

$$\begin{aligned} A(q) &= 1 - 2,444 \times q^{-1} + 1,991 \times q^{-2} - 0,5406 \times q^{-3} \\ B(q) &= 0,007 \times q^{-1} + 0,02476 \times q^{-2} + 0,005105 \times q^{-3} \end{aligned} \quad (12.23)$$

Os valores reais desses parâmetros são apresentados em (12.22), aqui repetidos para comodidade do leitor.

$$\begin{aligned} A(q) &= 1 - 2,444 \cdot q^{-1} + 1,99 \cdot q^{-2} - 0,5397 \cdot q^{-3} \\ B(q) &= 0,007149 \cdot q^{-1} + 0,02453 \cdot q^{-2} + 0,005252 \cdot q^{-3} \end{aligned} \quad (12.22)$$

Percebe-se que os valores estimados e reais estão muito próximos.

A Tabela 12.3 mostra o valor do coeficiente de ajuste *fit* ao se usar um modelo ARX para representar o sistema ARX, com variância do ruído branco de $\sigma^2 = 10^{-5}$, quando se fazem previsões de 1, 10, 100, 1000 e infinitos passos à frente, com validação cruzada.

Tab. 12.3 Valor do índice *fit* para previsões de 1, 10, 100, 1000 e infinitos passos à frente.

| | 1 passo | 10 passos | 100 passos | 1000 passos | Infinitos passos |
|-------------------|---------|-----------|------------|-------------|------------------|
| Índice <i>fit</i> | 99,85% | 96,89% | 95,42% | 95,42% | 95,45% |

Da Tabela 12.3 se deduz que:

- a simulação livre e as previsões com 100 e 1000 passos à frente têm um comportamento

bastante similar, pois como o número de passos à frente é elevado, essas predições passam a ter um comportamento similar ao da simulação livre; e

- a simulação com 1 passo à frente apresentou um ajuste (*fit*) de 99,85%, indicando que esta não é uma boa forma de validar um modelo, pois o resultado é normalmente muito bom. Conforme se passa a ter predições à frente com mais passos, aí sim a validação cruzada passa a ter validade e o ajuste tende a piorar.

12.10 INFLUÊNCIA DE DIVERSOS FATORES NA QUALIDADE DO MODELO OBTIDO

Analisa-se, a seguir, a influência de alguns fatores na qualidade do modelo obtido.

12.10.1 Influência da variância da perturbação na qualidade do modelo obtido

Apresenta-se na Tabela 12.4 o índice *fit* calculado com simulação livre para um modelo ARX com $n_a = 3$, $n_b = 3$ e $n_k = 1$, com perturbação com variâncias $\sigma^2 = 10^{-8}$ (só ruído de medição), $\sigma^2 = 10^{-6}$, $\sigma^2 = 10^{-5}$ e $\sigma^2 = 10^{-4}$, usando validação cruzada.

Tab. 12.4 Índice de ajuste para diferentes níveis de perturbação do sistema.

| σ^2 | <i>fit</i> (%) |
|------------|----------------|
| 10^{-8} | 99,85 |
| 10^{-6} | 98,56 |
| 10^{-5} | 95,45 |
| 10^{-4} | 85,88 |

A Tabela 12.4 indica que quanto menor for a variância da perturbação, tanto melhor será o ajuste do modelo gerado ao sistema real.

12.10.2 Influência da ordem na qualidade do modelo obtido

Caso se optasse por realizar a simulação livre do modelo com validação cruzada, mas empregando-se ordens maiores, com $\sigma^2 = 10^{-5}$, resulta a Tabela 12.5.

Tab. 12.5 Índice de ajuste para diferentes ordens do modelo.

| Ordem do modelo | <i>fit</i> (%) |
|-------------------------|----------------|
| $n_a = 3$, $n_b = 3$ | 95,45 |
| $n_a = 4$, $n_b = 4$ | 95,44 |
| $n_a = 5$, $n_b = 5$ | 95,48 |
| $n_a = 6$, $n_b = 6$ | 95,45 |
| $n_a = 10$, $n_b = 10$ | 95,45 |

O aumento da ordem do modelo teve um efeito praticamente nulo no coeficiente *fit*, pois se faz validação cruzada, com o sistema e o modelo com estruturas idênticas.

12.10.3 Influência da quantidade de pontos coletados na qualidade do modelo obtido

Este teste é realizado com variância do ruído branco $\sigma^2 = 10^{-5}$ e $\sigma^2 = 10^{-6}$, mas empregando-se agora não mais 1000 pontos de estimação e 1000 de validação, mas sim 10.000 e 100.000 pontos para a estimação e 1000 para a validação. A ideia é verificar o efeito do acréscimo do número de pontos no resultado da identificação. A ordem empregada é $n_a = 3$, $n_b = 3$ e $n_k = 1$. Os resultados são mostrados na Tabela 12.6.

Tab. 12.6 Índice de ajuste para diferentes números de pontos para estimação do modelo e para diferentes variâncias do ruído branco.

| σ^2 | $N_e = 1000$ pontos | $N_e = 10000$ pontos | $N_e = 100000$ pontos |
|------------|---------------------|----------------------|-----------------------|
| 10^{-5} | 95,45% | 95,58% | 96,42% |
| 10^{-6} | 98,56% | 98,61% | 98,87% |

A Tabela 12.6 indica que um aumento substancial no número de pontos gerou uma pequena melhora no fator de ajuste *fit* para ambas as variâncias do ruído branco, pois com um maior número de pontos para estimação, tende-se a gerar um modelo melhor.

12.10.4 Influência do período do PRBS na qualidade do modelo obtido

Neste teste, realizado com $\sigma^2 = 10^{-5}$ e $\sigma^2 = 10^{-6}$, altera-se o período do PRBS, que aumenta de $T_b = 4 \cdot T = 8$ s para $T_b = 16 \cdot T = 32$ s e diminui para $T_b = T = 2$ s, usando simulação livre com validação cruzada, como mostrado na Tabela 12.7.

Tab. 12.7 Índice de ajuste para três valores do período do PRBS e para duas variâncias do ruído branco.

| σ^2 | $T_b = T = 2$ s | $T_b = 4 \cdot T = 8$ s | $T_b = 16 \cdot T = 32$ s |
|------------|-----------------|-------------------------|---------------------------|
| 10^{-5} | 90,64% | 95,45% | 97,81% |
| 10^{-6} | 97,04% | 98,56% | 99,31% |

Verifica-se que, para este caso, o crescimento do período do PRBS foi benéfico, ao passo que sua diminuição foi prejudicial.

12.11 ANÁLISE DE RESÍDUOS

Os testes de validação vistos até aqui não indicam se os modelos possuem falhas corrigíveis. Por exemplo, se as simulações apresentarem erros significativos comparados aos dados observados, não é possível afirmar se eles ocorreram devido a problemas na estimação de parâmetros ou se a dinâmica que produziu os dados não pôde ser satisfatoriamente representada pelo modelo sendo validado. O procedimento de análise de resíduos informa se os parâmetros do modelo foram ou não estimados corretamente (AGUIRRE, 2007).

A parte dos dados que o modelo não é capaz de reproduzir é chamada de resíduo.

Uma análise destes dados pode revelar informações importantes a respeito da qualidade do modelo estimado, a qual está relacionada com a capacidade do modelo reproduzir bem novos dados de saída. O vetor de resíduos $\varepsilon(t)$ é definido como sendo a diferença entre o vetor de dados observados e de predições um passo à frente do modelo:

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

Pode-se calcular índices de qualidade a partir dos resíduos, conforme indicado a seguir (LJUNG, 1999):

$$S_1 = \max_t |\varepsilon(t)| \quad S_2 = \sqrt{\frac{1}{N} \sum_{t=1}^N \varepsilon^2(t)} \quad (12.24)$$

O emprego destes índices implica nas seguintes afirmações: este modelo nunca gerou um resíduo maior que S_1 ou então apresenta um erro médio de S_2 para todos os dados coletados e é provável que tais afirmações continuarão sendo válidas para dados coletados no futuro. No entanto, o emprego destes índices implica em uma hipótese implícita de invariância: os resíduos não dependem de nada que possa mudar e, em particular, eles não dependem dos sinais de entrada usados em \mathbf{z}^N . Caso eles dependessem dos sinais de entrada, a importância dos Índices (12.24) seria limitada, pois o modelo poderia ser bom apenas para um conjunto de entradas.

O vetor de resíduos depende dos parâmetros estimados bem como dos regressores que geram o modelo. Se o estimador usado for baseado no estimador dos mínimos quadrados, então a soma do quadrado dos resíduos é mínima para o vetor de regressores escolhido, de modo que a variância dos resíduos não é um indicador adequado para validar modelos dinâmicos. Uma das principais condições requeridas para que o estimador dos mínimos quadrados (MQ) não apresente polarização (ou tendência) na determinação de $\hat{\theta}_N$, é que o erro na equação de regressão deva ser ruído branco. Como o modelo identificado $y(t) = \varphi^T \hat{\theta}_N + \varepsilon(t)$ é obtido diretamente da equação de regressão após a estimação dos parâmetros, torna-se claro que para se ter $E[\hat{\theta}_{MQ}] = \theta$ é preciso que os resíduos sejam brancos. Portanto, análise de resíduos designa um conjunto de testes que são efetuados para verificar se os resíduos são aleatórios ou não (AGUIRRE, 2007).

O modelo deve explicar tudo que for explicável nos dados. Se isso ocorrer, então os resíduos conterão apenas aquilo que não é explicável e, conseqüentemente, serão brancos. Portanto, ao se testar o vetor de resíduos se verifique que se trata de uma variável aleatória branca contendo portanto apenas informações aleatórias, isto significa que não há informações úteis nos resíduos, ou seja, o modelo explicou tudo que era possível ser explicado, tendo conseguido representar todas as características contidas no conjunto de dados observados. Infelizmente, isto não quer dizer que a simulação livre do modelo será boa, mas apenas que a simulação de um passo à frente do modelo será boa. Por outro lado, se os resíduos não forem brancos, haverá informações neles indicando que o modelo não conseguiu explicar tudo que era explicável nos dados. Neste caso, o estimador dos mínimos quadrados será polarizado. Pode-se tentar contornar o problema,

escolhendo-se um novo conjunto de regressores e estimando-se os parâmetros novamente. Se os novos regressores forem mais adequados para explicar os dados, o novo vetor de resíduos será mais aleatório (AGUIRRE, 2007).

Um vetor de resíduos $\varepsilon(t)$ é linearmente branco se a sua função autocorrelação for nula para todos os valores de atraso τ maiores ou iguais a um, ou seja, $\hat{c}_\varepsilon(\tau) = 0, \forall \tau \neq 0$. Se os resíduos $\varepsilon(t)$ forem ruído branco, resulta que:

$$\begin{aligned} \hat{c}_\varepsilon(\tau) &\rightarrow 0 && \text{para } \tau \neq 0 \\ \hat{c}_\varepsilon(0) &\rightarrow \lambda^2 = E[\varepsilon^2(t)] && \text{para } \tau = 0 \end{aligned} \quad \text{conforme } N \rightarrow \infty \quad (12.25)$$

Para se trabalhar com valores normalizados, considere:

$$\hat{c}_{\varepsilon, norm}(\tau) = \frac{\hat{c}_\varepsilon(\tau)}{\hat{c}_\varepsilon(0)} \quad (12.26)$$

Se o vetor de resíduos for linearmente branco, isto indica que não há autocorrelação linear no vetor considerado, atestando a sua aleatoriedade. Portanto, uma forma de verificar deficiências no modelo é analisar a função autocorrelação do vetor de resíduos, cuja estimativa pode ser obtida por:

$$\hat{c}_\varepsilon(\tau) = \frac{1}{N} \sum_{t=1}^N \varepsilon(t) \cdot \varepsilon(t-\tau) \quad (12.27)$$

Segundo (12.25), espera-se que a estimativa da autocorrelação normalizada $c_{\varepsilon, norm}(\tau)$ seja pequena para $\tau \neq 0$ e $N \rightarrow \infty$, de modo que $c_{\varepsilon, norm}(\tau)$ deve parecer com uma resposta impulsiva, isto é, tem um pico em $\tau = 0$ e é próxima de zero para os demais valores de τ . Se isto não ocorrer, significa que parte do vetor $\varepsilon(t)$ poderia ter sido predito com dados passados e, conseqüentemente, uma melhor predição da saída $y(t)$ poderia ter sido obtida, o que é um sinal de deficiência do modelo. A quantificação do que pequeno significa é vista na Subseção 12.11.1.

Caso se suponha que o vetor de resíduos seja linearmente não autocorrelacionado, antes de realizar a validação é de suma importância saber quão geral são esses resultados, ou seja, verificar se o modelo teria um desempenho semelhante para outros conjuntos de dados. Uma forma de verificar a generalidade dos resultados é estudar a correlação entre os resíduos e sinais de entrada passados. Se os resíduos e a entrada forem independentes, tem-se que:

$$c_{\varepsilon u}(\tau) = E[\varepsilon(t)u(t-\tau)] = 0$$

Se o modelo não for uma representação adequada do processo, espera-se que $c_{\varepsilon u}(\tau)$ seja diferente de zero para $\tau \geq 0$ e caso se assuma que o modelo seja uma descrição exata do processo, então $c_{\varepsilon u}(\tau) = 0$ para $\tau \geq 0$. Se essa correlação não for nula, então há parte da saída do processo gerada a partir de entradas passadas que foram deixadas nos resíduos e que não foram descritas pelo modelo. Na identificação de processos lineares, este fato normalmente significa que a ordem do modelo está muito baixa (ZHU, 2001). Esta análise permite ainda ter uma ideia se o tempo morto atribuído ao modelo está correto. Considere, por exemplo, que um tempo morto de quatro intervalos de

amostragem tenha sido assumido no modelo, mas que o atraso puro real seja de dois intervalos de amostragem. Então, uma clara correlação entre $u(t-2)$ e $\varepsilon(t)$ surgirá. Para $\tau < 0$, $c_{\varepsilon u}(\tau)$ pode ou não ser nulo, dependendo da autocorrelação de $u(t)$. Por exemplo, se $u(t)$ for ruído branco, então $c_{\varepsilon u}(\tau) = 0$ para $\tau < 0$ mesmo que o modelo seja impreciso. No entanto, pode ocorrer que $c_{\varepsilon u}(\tau) \neq 0$ para $\tau < 0$ se houver realimentação no processo durante a coleta dos dados, de modo que essa correlação é uma indicação de realimentação da saída na entrada e não de deficiência na estrutura do modelo.

Para se estimar a correlação cruzada entre os resíduos e a entrada com base em dados coletados do processo, emprega-se o seguinte cálculo:

$$\hat{c}_{\varepsilon u}(\tau) = \frac{1}{N} \sum_{t=1}^N \varepsilon(t) \cdot u(t-\tau) \quad (12.28)$$

Para se trabalhar com valores normalizados, considere:

$$\hat{c}_{\varepsilon u, norm}(\tau) = \frac{\hat{c}_{\varepsilon u}(\tau)}{\sqrt{\hat{c}_{\varepsilon}(\tau) \cdot \hat{c}_u(\tau)}} \quad (12.29)$$

Espera-se que $\hat{c}_{\varepsilon u, norm}(\tau)$ seja “pequeno” quando $N \rightarrow \infty$ e se assuma que $\varepsilon(t)$ seja independente da entrada passadas, isto é, $E[\varepsilon(t)u(s)] = 0$ para $t > s$ e que $\varepsilon(t)$ seja independente de todas as entradas, isto é, $E[\varepsilon(t)u(s)] = 0 \forall t, s$, que se aplica se o processo opera em malha aberta. Se estes valores forem “pequenos” (o que significa pequeno é definido na Subseção 12.11.1), há razões para acreditar que os Índices (12.24) poderiam ter relevância também quando o modelo fosse aplicado a outras entradas. Se a correlação for nula para diferentes valores de τ , é uma indicação de que o vetor de resíduos $\varepsilon(t)$ não depende do sinal de entrada usado para obter o modelo. Isto implica que as predições um passo à frente do modelo terão características semelhantes, se calculadas para outro conjunto de dados (AGUIRRE, 2007). Uma outra forma de expressar a importância de que os resíduos e o sinal de entrada não tenham correlação ($\hat{c}_{\varepsilon u}(\tau) = 0, \forall \tau$) é que se houver vestígios (correlação) de entradas anteriores $u(t-\tau)$ nos resíduos, então uma parte de $y(t)$, que é originada por tais entradas, não está sendo corretamente representada pelo modelo (LJUNG, 1999). Neste caso, o modelo poderia ser melhorado.

As Equações (12.27) e (12.28) são apenas aproximações das funções de correlação cruzada e de autocorrelação, uma vez que, teoricamente, tais funções deveriam ser calculadas para $N \rightarrow \infty$.

Deve-se enfatizar que a análise de resíduos é importante na medida em que revela se o modelo escolhido está fazendo o que se espera dele, apenas no que diz respeito a seu ajuste aos dados de identificação. Infelizmente, tal análise não garante a qualidade nem a utilidade de um modelo. Por exemplo, é comum obter modelos cujos resíduos sejam brancos, mas cujas predições livres sejam insatisfatórias. Em outras palavras, tais modelos têm boas predições um passo à frente, mas suas predições infinitos passos à frente não são boas. Analogamente, o contrário às vezes ocorre. Modelos cujos resíduos não sejam totalmente brancos representam bem a dinâmica do sistema sendo estudado. Isso serve de

lembrete que na validação de modelos mais de um indicador deve ser usado. Ressalta-se que a análise de resíduos considera o desempenho do modelo utilizando os dados de identificação e não de validação, pois os resíduos são definidos como o erro do preditor um passo à frente aplicado aos dados usados pelo estimador dos mínimos quadrados e, portanto, são gerados dentro do procedimento de estimação de parâmetros (AGUIRRE, 2007).

12.11.1 Definição de intervalo de confiança

Deve-se enfatizar que $\hat{c}_\varepsilon(\tau)$ e $\hat{c}_{\varepsilon U}(\tau)$ são estimativas baseadas em um número finito de amostras, de modo que seus valores desviarão dos teóricos. Desse modo, alguns intervalos devem ser estabelecidos para testar se essas estimativas de correlação são nulas. Assim, na prática, ao analisar as propriedades estatísticas dos resíduos $\varepsilon(t)$, verifica-se se os resultados das Equações (12.27) e (12.28) são suficientemente pequenos, o que significa estar dentro de limites denominados intervalos de confiança (*confidence intervals*), que são determinados de acordo com a precisão que se deseje para as aproximações nos cálculos das funções de correlação.

O intervalo de confiança corresponde à faixa de valores dos resíduos com uma probabilidade específica de ser estatisticamente insignificante para o sistema. Por exemplo, para um intervalo de confiança de 95%, a região em torno do zero representa a faixa de valores dos resíduos que possuem uma probabilidade de 95% de ser estatisticamente insignificante. Pode-se especificar o intervalo de confiança como uma probabilidade (entre 0 e 1) ou como o número de desvios padrão de uma distribuição gaussiana. Por exemplo, uma probabilidade de 0,99 (99%) corresponde a 2,576 desvios padrão.

Assuma que se tomem N amostras de um resíduo. Deve-se estimar o intervalo de confiança a partir dessa amostra, de modo que ele também é uma variável aleatória. Suponha que o valor médio do resíduo $\bar{\varepsilon}$ de uma amostra normalmente distribuída também seja normalmente distribuído, com o mesmo valor esperado μ de uma amostra infinita, mas com desvio padrão σ/\sqrt{N} . Pode-se obter uma variável aleatória, através da seguinte normalização:

$$Z_\alpha = \frac{\bar{\varepsilon} - \mu}{\sigma/\sqrt{N}}$$

A variável aleatória Z_α depende do parâmetro μ , mas com uma distribuição normal que independe de μ . Assim, é possível encontrar números $-z_\alpha$ e z_α , independentes de μ , sendo que Z_α se situa entre esses limites, com probabilidade $1-\alpha$ e α é uma medida de quanta confiança se deseja. z_α denota o nível α da distribuição normal $N(0,1)$ de média nula e covariância unitária. O nível α é o limiar para definir significância estatística. Por exemplo, fazendo-se $\alpha=0,05$ implica que se aceite que haja uma chance de 5% de concluir incorretamente que uma dada intervenção seja efetiva quando, em verdade, ela não tem efeito. Os níveis α são normalmente ajustados em 0,01; 0,05 ou 0,10. Por exemplo, para $\alpha=0,01$ resulta $z_\alpha = 2,576$ e para $\alpha=0,05$ tem-se que $z_\alpha = 1,960$, como se vê na tabela

12.8. Considere inicialmente $\alpha=0,05$, de modo que $1-\alpha=0,95$, fazendo com que:

$$P(-z_\alpha \leq Z_\alpha \leq z_\alpha) = 1 - \alpha = 0,95$$

z_α é calculado com base na função distribuição cumulativa como se segue:

$$\Phi(z_\alpha) = P(Z_\alpha \leq z_\alpha) = 1 - \frac{\alpha}{2} = 0,975$$

Portanto:

$$z_\alpha = \Phi^{-1}(\Phi(z_\alpha)) = \Phi^{-1}(0,975) = 1,960$$

Alguns valores de z_α obtidos de $\Phi(z_\alpha)$ normalmente usados na prática são mostrados na Tabela 12.8.

Tab. 12.8 Números comumente usados em testes de significância.

| | | | | | | |
|------------------|-------|-------|-------|-------|-------|-------|
| α | 0,200 | 0,100 | 0,050 | 0,020 | 0,010 | 0,002 |
| $\Phi(z_\alpha)$ | 0,900 | 0,950 | 0,975 | 0,990 | 0,995 | 0,999 |
| z_α | 1,282 | 1,645 | 1,960 | 2,326 | 2,576 | 3,090 |

Para $\alpha = 0,05$, resulta:

$$\begin{aligned} 0,95 = 1 - \alpha = P(-z_\alpha \leq Z_\alpha \leq z_\alpha) &= P\left(-1,960 \leq \frac{\bar{\varepsilon} - \mu}{\sigma/\sqrt{N}} \leq 1,960\right) = \\ &= P\left(\bar{\varepsilon} - \frac{1,960 \cdot \sigma}{\sqrt{N}} \leq \mu \leq \bar{\varepsilon} + \frac{1,960 \cdot \sigma}{\sqrt{N}}\right) \end{aligned}$$

Isto pode ser interpretado do seguinte modo: existe um intervalo de confiança dentro do qual, com probabilidade de 0,95 (95%), se encontra um parâmetro μ que se situa dentro dos seguintes limites:

$$\bar{\varepsilon} - \frac{1,960 \cdot \sigma}{\sqrt{N}} \quad \text{e} \quad \bar{\varepsilon} + \frac{1,960 \cdot \sigma}{\sqrt{N}}$$

Isto não significa que haja uma probabilidade de 0,95 de encontrar o parâmetro μ no intervalo calculado. Toda a vez que novos dados sejam coletados, haverá um novo valor da média $\bar{\varepsilon}$ da amostra. Em 95% dos casos, μ estará entre os limites calculados a partir dessa média $\bar{\varepsilon}$, mas em 5% dos casos ele não estará. Diz-se então que com nível de confiança $100 \cdot (1 - \alpha)\%$ o parâmetro μ se encontra no intervalo de confiança.

Segundo o texto após a expressão (12.27), o significado de pequeno para $\hat{c}_\varepsilon(\tau)$ é definido a seguir (BILLINGS; VOON, 1986; SÖDERSTRÖM; STOICA, 1989):

$$|\hat{c}_{\varepsilon, norm}(\tau)| \leq \sqrt{\frac{1}{N}} z_\alpha \quad (12.30)$$

Conforme citado no parágrafo seguinte à Equação (12.29), o significado de pequeno para $\hat{c}_{\varepsilon U}(\tau)$ é definido a seguir, sendo que o limite calculado é específico para cada τ (LJUNG, 1999):

$$|\hat{c}_{\varepsilon U}(\tau)| \leq \sqrt{\frac{P_1}{N}} z_\alpha \quad (12.31)$$

$$\text{onde: } P_1 = \sum_{\tau=-\infty}^{\infty} \hat{c}_{\varepsilon,norm}(\tau) \cdot \hat{c}_{u,norm}(\tau) \quad (12.32)$$

$$\hat{c}_{u,norm}(\tau) = \frac{\hat{c}_u(\tau)}{\hat{c}_u(0)}$$

z_{α} denota o nível α da distribuição normal $N(0,1)$ de média nula e covariância unitária. O nível α é o limiar para definir significância estatística. Por exemplo, ajustando-se $\alpha = 0,05$ implica que se aceite que haja uma chance de 5% de concluir incorretamente que uma dada intervenção seja efetiva quando, em verdade, ela não tenha efeito. Os níveis α são normalmente ajustados em 0,01; 0,05 ou 0,10. Por exemplo, para $\alpha = 0,01$ resulta $z_{\alpha} = 2,576$ e para $\alpha = 0,05$ tem-se que $z_{\alpha} = 1,960$, conforme visto na Tabela 12.8.

12.11.2 Exemplo de emprego da análise de resíduos

Para a análise de resíduos é considerado o sistema com estrutura ARX apresentado na Seção 12.8, mas afetado de tempo morto, conforme indicado a seguir:

$$G(s) = \frac{6 \cdot e^{-6 \cdot s}}{960 \cdot s^3 + 296 \cdot s^2 + 30 \cdot s + 1}$$

Assim como foi feito na Seção 12.8, empregam-se 1000 pontos como dados de estimação. Propõe-se inicialmente considerar um modelo ARX com $n_a = 3$, $n_b = 3$ e $n_k = 1$, com variância do ruído branco igual a $\sigma^2 = 10^{-5}$. O sistema é excitado por um sinal PRBS de amplitude unitária e de período $T_b = 8$ s.

O procedimento em Matlab para realizar a análise de resíduos é mostrado a seguir. Expõem-se duas opções: emprego direto do comando de análise de resíduos do Matlab (resid) ou então geração das funções de autoautocorrelação e autocorrelação cruzada e cálculo dos limites da zona de confiança.

```
% Análise de resíduos
Ts = 2;
c1=0; c2=0; c3=0;      % Sistema ARX
Modelo_1a_ordem_validacao_PRBS
sim('Modelo_1a_ordem_validacao_PRBS')
ide = 1000;
ze = iddata(y(1:ide),u(1:ide),Ts);
theta = arx(ze,[3 3 1]);
% Figura com autocorrelação normalizada dos resíduos e com correlação cruzada entre o sinal de
% entrada e os resíduos
[e,r] = resid(theta,ze,'corr',30); % Uso direto do comando de análise de resíduos do Matlab com zona
% de confiança de 99% e 30 atrasos

resid(r)
subplot(2,1,1)
title('Função autocorrelação normalizada dos resíduos')
xlabel('Atrasos \tau')
```

```

axis([0 30 -0.2 1])
set(gca,'XTick',0:3:30,'YTick',-0.2:0.2:1)
subplot(2,1,2)
title('Função correlação cruzada entre o sinal de entrada e os resíduos')
xlabel('Atrasos \tau')
axis([-30 30 -0.12 0.12])
set(gca,'XTick',-30:5:30,'YTick',-0.12:0.04:0.12)
% Duas formas de gerar análise de resíduos
eu = [e.y u(1:ide)];
reux = xcorr(eu,29);      % Usando comando xcorr
reu = covf(eu,30);       % Usando comando covf
% Figura com autocorrelação do resíduo e
atrasos_e = 0:29;
limite_e = 2.576/sqrt(ide);
limite_sup_e = ones(length(atrasos_e),1)*limite_e;
limite_inf_e = ones(length(atrasos_e),1)*(-limite_e);
subplot(2,1,1)
hold
plot(lags,r(1,1:30)/r(1,1),'r*',lags,reu(1,:)/reu(1,1),'g.',lags,reux(30:59,1)/reux(30,1),...
     'k^',lags(1:1:30),limite_sup_e(1:1:30),'k.',lags(1:1:30),limite_inf_e(1:1:30),'k.')
hold
% Figura com correlação cruzada do sinal de entrada u e do resíduo e
somat1 = 0; somat2 = 0;
for i=1:30
    somat1 = somat1 + (reux(i+29,1)/reux(30,1))*(reux(i+29,4)/reux(30,4));
    somat2 = somat2 + (reu(1,i)/reu(1,1))*(reu(4,i)/reu(4,1));
end
limite_ue_1 = 2.576*sqrt(somat1/ide);
limite_ue_2 = 2.576*sqrt(somat2/ide);
limite_ue_sup_1 = ones(2*29+1,1)*limite_ue_1;
limite_ue_inf_1 = ones(2*29+1,1)*(-limite_ue_1);
limite_ue_sup_2 = ones(2*29+1,1)*limite_ue_2;
limite_ue_inf_2 = ones(2*29+1,1)*(-limite_ue_2);
subplot(2,1,2)
hold
K1 = ide/sqrt(reux(30,1)*reux(30,4));
K2 = 1/sqrt(reu(1,1)*reu(4,1));
reu_conc = [reu(3,30:-1:1)'; reu(2,2:30)'];
atrasos_ue = -29:29;
plot(atrasos_ue,K2*reux(:,2)/ide,'r*',atrasos_ue,K3*reu_conc,'g.',atrasos_ue(1:2:59), ...
     limite_ue_sup_1(1:2:59),'g.',atrasos_ue(2:2:58),limite_ue_sup_2(2:2:58),'k.',...
     limite_ue_inf_2(1:2:59),'g.',atrasos_ue(2:2:58),limite_ue_inf_2(2:2:58),'k.')

```

hold

Os gráficos resultantes são mostrados na Figura 12.18, que mostra apenas a primeira figura gerada pelo comando resid, para evitar sobrecarga de pontos no gráfico.

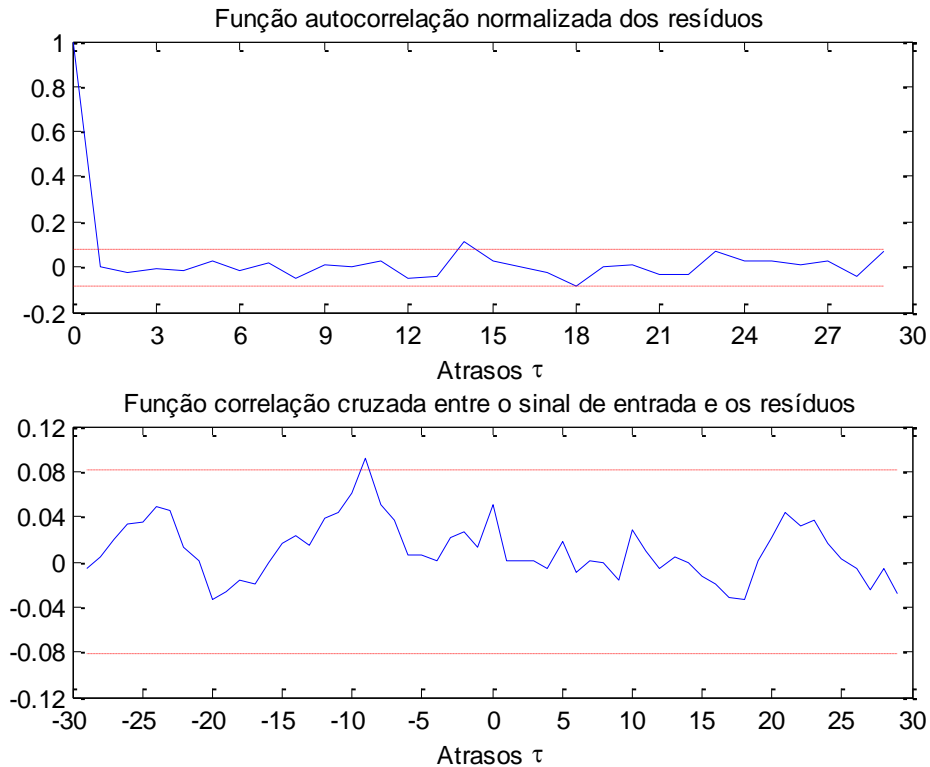


Fig. 12.18 Autocorrelação normalizada dos resíduos e correlação cruzada entre entrada e resíduos para sistema e modelo com estrutura ARX.

Na Figura 12.18 as linhas pontilhadas denotam o intervalo de confiança de 99%. Percebe-se que a autocorrelação dos resíduos viola o intervalo de confiança para o atraso τ de 14 e a correlação cruzada entre a entrada e os resíduos sai do intervalo de confiança para o atraso τ de -9, significando que há informações dinâmicas de entrada e saída que não foram incorporadas pelo modelo.

Suponha agora que se empregue o mesmo modelo ARX com $n_a = 3$, $n_b = 3$ e $n_k = 1$ para identificar o sistema ARMAX da Seção 13.6. O nível de ruído é $\sigma^2 = 10^{-5}$. Resulta a Figura 12.19.

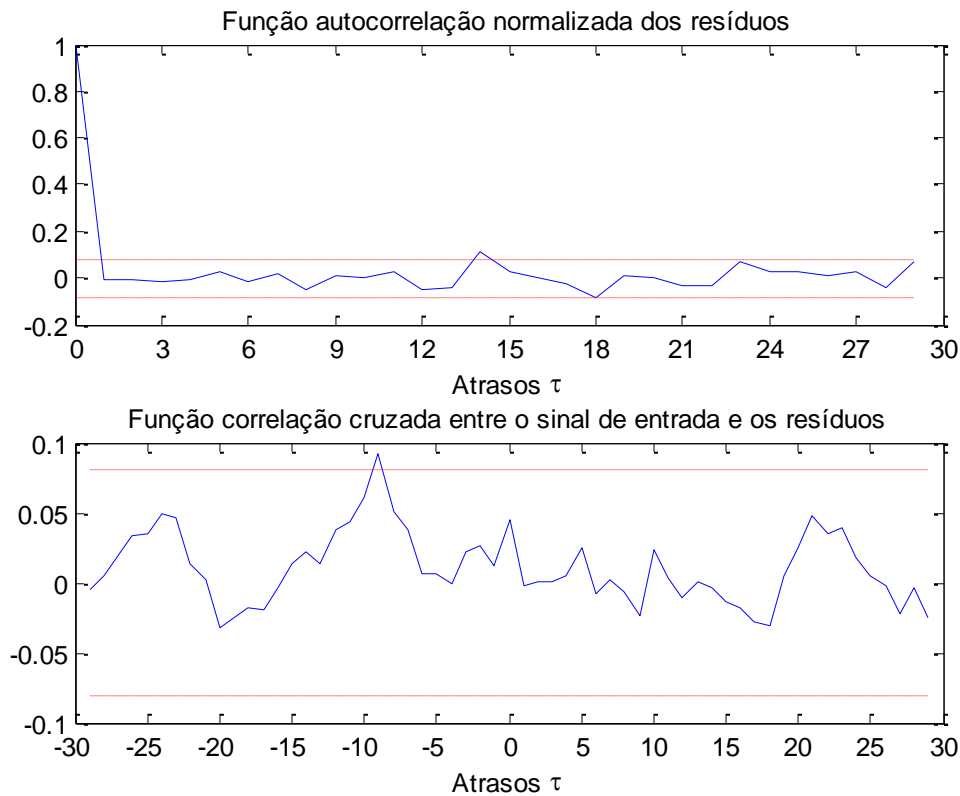


Fig. 12.19 Autocorrelação dos resíduos e correlação cruzada entre entrada e resíduos para sistema com estrutura ARMAX e modelo com estrutura ARX.

Nota-se na Figura 12.19 que a autocorrelação dos resíduos sai do intervalo de confiança para o atraso $\tau = 18$. Já a correlação cruzada entre a entrada e os resíduos não sai do intervalo de confiança. Com isto, pode-se dizer que há informações no conjunto de dados observados que não foram explicadas pelo modelo. Caso a desigualdade em (12.27) não seja obedecida, a hipótese que $\varepsilon(t)$ e $u(t-\tau)$ sejam independentes deve ser descartada. Uma forma interessante de se conduzir o teste é mostrar $\hat{R}_{\varepsilon u}(\tau)$ em função de τ . Como P_1 em (12.28) não depende de τ , os limites de confiança são linhas horizontais. Tal gráfico fornece uma valiosa noção acerca da correção da estrutura do modelo. Se, por exemplo, um tempo morto de dois intervalos de amostragem foi assumido no modelo mas o tempo morto real é um intervalo, então uma clara correlação entre $u(t-1)$ e $\varepsilon(t)$ surgirá.

12.11.3 Teste para verificar presença de realimentação da saída na entrada

É possível detectar se existe realimentação presente na malha sendo identificada através dos gráficos de correlação cruzada entre a entrada e os resíduos.

Assuma que $u(t)$ seja determinado por uma realimentação da saída $y(t)$ e que o resíduo $\varepsilon(t)$ seja uma boa estimativa do ruído branco que gera as perturbações. Então a entrada $u(t)$ no instante t em geral será dependente dos resíduos passados, mas independente de seus futuros valores. Isto significa que (SÖDERSTRÖM; STOICA, 1989):

$$R_{\varepsilon u}(\tau) = 0 \quad \text{para } \tau > 0$$

e, em geral:

$$R_{\varepsilon u}(\tau) \neq 0 \quad \text{para } \tau \leq 0$$

se houver realimentação.

Portanto, ao examinar gráficos de $\hat{R}_{\varepsilon u}(\tau)$ deve-se notar que a correlação entre $\varepsilon(t)$ e $u(t-\tau)$ para valores negativos de τ é uma indicação de realimentação da saída na entrada e não que a estrutura do modelo esteja deficiente (LJUNG, 1999). Portanto, caso haja uma correlação significativa para atrasos negativos, isto indica a presença de realimentação da saída na geração da entrada. Portanto, a análise da presença de realimentação na malha faz parte da validação do modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas** – técnicas lineares e não-lineares aplicadas a sistemas reais. 3.ed. Belo Horizonte, Editora UFMG, 2007.
- BILLINGS, S. A.; VOON, W. S. F. Correlation based model validity tests for non-linear models. **International Journal of Control**, v.44, n.1, p.235-244, 1986.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. San Francisco, Holden-Day, 1976.
- LJUNG, L. **System identification: theory for the user**. 2.ed., Englewood Cliffs, Prentice Hall, 1999.
- RISSANEN, J. Modelling by shortest data description. **Automatica**, v.14, p.465-71, 1978.
- SÖDERSTRÖM, T.; STOICA, P. **System identification**. Hemel Hempstead, U.K., Prentice Hall International, 1989.
- ZHU, Y.; BACKX, T. **Identification of multivariable industrial processes**. London, Springer-Verlag, 1993.