Universidade Federal de Campina Grande

Programa de Pós-Graduação em Engenharia Elétrica

Estimação e Identificação de Sistemas - 21.2

Arthur Dimitri Brito Oliveira

Relatório de Atividade 4

Desenvolvimento

4E.1

Considerando um modelo ARX com estrutura

$$y(t) + a_1 y(t-1) + \ldots + a_{na} y(t-a) = b_1 u(t-1) + \ldots + b_{nb} u(t-n_b) + e(t)$$

$$y(t) = b_1 u(t-1) + \ldots + b_{nb} u(t-n_b) - a_1 y(t-1) + \ldots - a_{na} y(t-a) + e(t)$$

Os parâmetros ajustáveis são representados por:

$$\theta = \begin{bmatrix} a_1 & a_2 \ldots & a_{na} & b_1 & \ldots & b_{nb} \end{bmatrix}$$

Introduzindo o vetor de regressores:

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

A equação do preditor pode ser expressa como:

$$\hat{y}(t|\theta) = \theta^T \varphi(t) = \varphi^T(t)\theta = [-y(t-1) \;\ldots\; -y(t-n_a) \;\; u(t-2) \;\ldots\; u(t-n_b)] \begin{bmatrix} a_1 \\ a_2 \\ \ldots \\ a_{na} \\ b_2 \\ \ldots \\ b_{nb} \end{bmatrix} + b_1 u(t-1)$$

4E.4

Considerando o circuito exibido na Figura 1 e suas entradas como sendo $u_v(t)$ e $u_i(t)$.

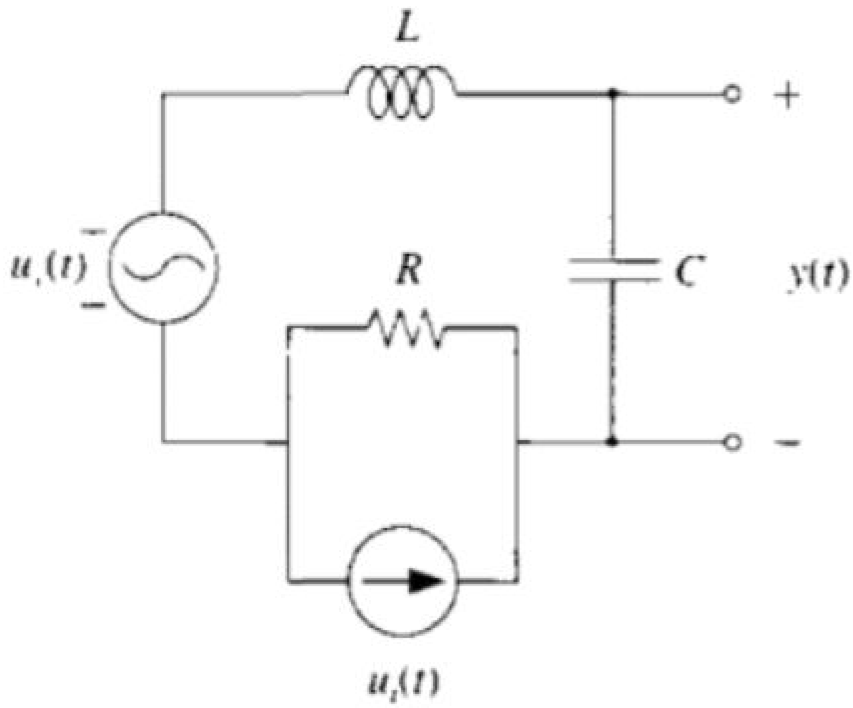

Figura 1: Circuito do problema 4E.1

Aplicando a LKT:

$$-U_v + V_L + V_C + R[i + U_i] = 0 \qquad (1)$$

Sabendo que $i = i_L = i_C = CdV_c/dt = Cdy/dt$ e $\dot{i}_c = Cd^2y/dt^2$, substituindo estas informações na Eq. 1, tem-se:

$$U_v - R[C\frac{dy}{dt} + U_i] - y(t) - LC\frac{d^2y}{dt^2} = 0$$

Dessa forma, na representação de espaço de estados:

$$\begin{bmatrix} \dot{y} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{Lc} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} y \\ \dot{y} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{1}{LC} & -\frac{R}{LC} \end{bmatrix} \begin{bmatrix} U_v \\ U_i \end{bmatrix}$$

Aplicando a transformada de Laplace à expressão $\ddot{y}$,

$$Y(s)\left[s^2 + \frac{R}{L}s + \frac{1}{LC}\right] = \frac{1}{LC}U_v(s) - \frac{R}{LC}U_i(s)$$

Logo, há duas funções de transferência possíveis, já que há duas entradas:

$$G_1(s) = \frac{Y(s)}{Ui(s)} = \frac{\frac{1}{LC}}{s^2 + \frac{R}{L} + \frac{1}{LC}} \text{ e } G_2(s) = \frac{Y(s)}{Uv(s)} = \frac{\frac{-R}{LC}}{s^2 + \frac{R}{L} + \frac{1}{LC}}$$

Ultilizando a transformação bilinear para transformar o sistema de tempo contínuo em tempo discreto, de modo geral pode-se obter uma função de transferência racional própria no tempo discreto:

$$G(q) = \frac{A(z)}{B(z)}$$

Há várias maneiras de parametrizar a função de transferência de um sistema ao considerar os coeficientes do numerador e do denominador como parâmetros. O modelo de um sistema descreve algumas das suas propriedades, sendo adequado a determinados propósitos. A identificação aqui é utilizada para selecionar ou construir modelos de sistemas dinâmicos que se adequem a determinados casos.

Modelos Lineares e Conjuntos de Modelos Lineares

Supondo que o sistema está sujeito a ruídos de medição, de modo a avaliar a influência da perturbação no sistema, considera-se um ruído branco gaussiano a ser adicionado à resposta do sistema:

$$y(t) = G(q)u(t) + H(q)e(t)$$

$$v(t) = H(q)e(t) = \sum_{k=0}^{\infty} h(k)e(t-k) \qquad (2)$$

Uma propriedade crucial da Eq. 2 é que haja invertibilidade, de tal forma que, para $v(t),\ s \le t,\ e(t)$ possa ser determinado. Assumindo que o filtro H é estável, e que $\frac{1}{H(z)}$ é analítica em $|z| \ge 1$, ou seja, que $H(q)$ é um filtro inversível estável:

$$H(z) = \sum_{k=0}^{\infty} h(k)z^{-k}$$

$$H^{-1}(q) = \sum_{k=0}^{\infty} \bar{h}(k)q^{-k} = \frac{1}{H(q)}$$

Predição de $v(t)$ um passo à frente

Dado que há observações de $v(t),\ s \le t-1$, deseja-se predizer o valor de $v(t)$ baseado nessa observação. Tendo em vista que H é mônico, pode-se escrever $v(t)$ como:

$$v(t) = \sum_{k=0}^{\infty} h(k)e(t-k) = e(t) + \sum_{k=1}^{\infty} h(k)e(t-k)$$

O segundo termo do lado direito da equação anterior é conhecido no instante $t-1$. Por facilidade de notação, o denominaremos de $m(t-1)$. A função densidade de probabilidade a posteriori de $v(t)$, dadas observações feitas até o instante $t-1$ é dada por:

$$f_v(x) = f_e(x - m(t-1))$$

Normalmente só provemos um valor que caracteriza essa distribuição de probabilidade e serve como predição para $v(t)$. Este pode ser o valor para o qual a FDP f_v tem o seu máximo. No entanto, devemos trabalhar com o valor médio da distribuição em questão. Assim, a esperança condicional de $v(t)$ é dada por:

$$\hat{v}(t|t-1) = m(t-1) = \sum_{k=1}^{\infty} h(k)e(t-k) = \left[\sum_{k=1}^{\infty} h(k)q^{-k}\right]e(t) = [H(q)-1]e(t)$$

Sabendo que $e(t) = H^{-1}(q)v(t)$, a expressão anterior torna-se:

$$\hat{v}(t|t-1) = [1-H(q)]H^{-1}(q)v(t) = [H(q)-1]v(t) = \textstyle\sum_{k=1}^{\infty} h(k)v(t-k) \qquad (4)$$

Predição de $y(t)$ um passo à frente

Deseja-se obter a predição de:

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

Supondo $y(s)$ e $u(s)$ conhecidos para $s \leq t-1$, tem-se:

$$v(s) = y(s) - G(q)u(s)$$

Conhecendo como obter a estimativa $\hat{v}(t|t-1)$, a esperança condicional de $y(t)$ é dada por:

$$\hat{y}(t|t-1) = G(q)u(t) + \hat{v}(t|t-1)$$

$$\hat{y}(t|t-1) = G(q)u(t) + [1-H^{-1}(q)]v(t) = G(q)u(t) + [1-H^{-1}(q)][y(t) - G(q)u(t)]$$

$$\hat{y}(t|t-1) = H^{-1}(q)G(q)u(t) + y(t)[1-H^{-1}(q)] \qquad (5)$$

O erro de predição

A diferença entre a saída $y(t)$ e $\hat{y}(t|t-1)$, sabendo que $v(t) = y(t) - G(q)u(t)$, é expressa por:

$$y(t) - \hat{y}(t|t-1) = -H^{-1}(q)G(q)u(t) + H^{-1}(q)y(t) = H^{-1}[y(t) - G(q)u(t)] = H^{-1}v(t) = e(t)$$

Assim, $e(t)$ representa a parte da saída $y(t)$ que não pode ser predita a partir dos dados passados.

Modelos lineares e conjuntos de modelos lineares

Um modelo linear invariante no tempo é determinado pela resposta ao imputlso $\{g(k)\}_1^\infty$, pelo espectro do ruído aditivo $\Phi_v(\omega) = \lambda |H(e^{j\omega})|$ e pela FDP do ruído $e(t)$. Assim, o modelo completo é dado por:

$$y(t) = G(q)u(t) + H(q)e(t)$$

$$y(t) = u(t)\sum_{k=1}^{\infty} g(k)q^{-k} + e(t)\left[1 + \sum_{k=1}^{\infty} h(k)q^{-k}\right]$$

$f_e(.)$ é a FDP de e

A determinação das funções de transferência é feita em termos de um número finito de valores, utilizando por exemplo, funções de transferência racionais. Também assume-se que $e(t)$ é gaussiano. Pelo fato de que é quase impossível determinar os coeficientes das funções de transferência a priori, a limitação a uma quantidade finita de termos destas funções permite uma identificação mais fácil, na qual os coeficientes em questão são tratados como parâmetros a serem determinados.

Uma vez especificadas as funções G e H, é possível incorporar à Eq.5, ou seja, a equação do preditor, a dependência do parâmetro θ à formulação do modelo:

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

$f_e(x,\theta)$. a FDP de $e(t)$

$$\theta \in D_M \subset R^d$$

Isso torna a equação de saída de $y(t)$ associada a um conjunto de modelos, e o processo de estimação consistirá em selecionar o conjunto de parâmetros mais adequado. A predição de y um passo à frente, enfatizando a dependência de θ pode ser expressa por:

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + [1 - H^{-1}(q,\theta)]y(t) \qquad (6)$$

Sempre que H e G forem modificados, uma nova equação do preditor será gerada. No entanto, a Eq.6 é o arcabouço básico.

Família de Funções de Transferência

Estrutura do modelo de erro da equação

A relação de entrada-saída mais simples é obtida por meio de uma equação de diferença. A intuição por trás da Eq.7 é que o termo autorregressivo de $y(t)$ e o termo de média móvel de $u(t)$ são iguais, exceto pelo termo de erro $e(t)$.

$$y(t) + a_1 y(t-1) + ... + a_{n_a} y(t-n_a) = b_1 u(t-1) + ... + b_{n_b} u(t-n_b) + e(t) \qquad (7)$$

Este modelo pode ser denominado de **ARX**. As siglas **AR** referem-se à parte autoregressiva $A(q)y(t)$ e **X** à parte exógena $B(q)u(t)$. Os parâmetros ajustáveis são:

$$\theta = [a_1 \ a_2 \ ... \ a_{n_a} \ b_1 \ ... \ b_{n_b}]^T$$

Ao introduzir os polinômios correspondentes à função de transferência do sistema:

$$A(q) = 1 + a_1 q^{-1} + \ \ + a_{n_a} q^{-n_a}$$

$$B(q) = b_1 q^{-1} + \ ... \ + b_n q^{-n_b}$$

$$G(q, \theta) = \frac{B(q)}{A(q)} \qquad (8)$$

$$H(q, \theta) = \frac{1}{A(q)} \qquad (9)$$

O fluxo dos sinais pode ser observado na Figura 2. Assume-se que o ruído branco passa pelo denominador da dinâmica do sistema antes de ser adicionado à saída.

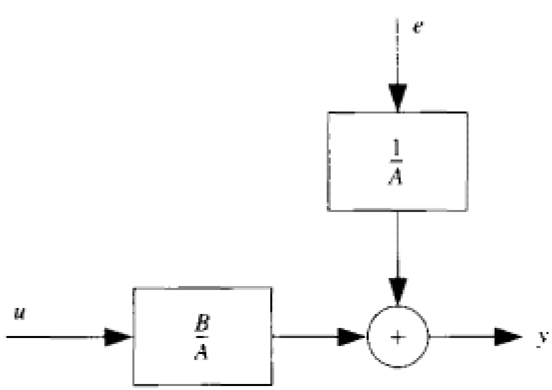

Figura 2: A estrutura do modelo ARX.

A equação de erro do modelo é normalmente escolhida para diversas aplicações, visto que o preditor define uma regressão linear. Ao substituir a Eq. 7 e a Eq.8 na Eq. 6, o **preditor** do modelo ARX pode ser definido como:

$$\hat{y}(t) = A(q)\frac{B(q)}{A(q)}u(t) + [1 - A(q)]y(t) = B(q)u(t) + [1 - A(q)]y(t) \qquad (10)$$

Assim, pode-se introduzir o vetor de regressores:

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

A Eq. 9 pode ser reescrita como:

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

O ruído visto na saída da Fig. 4.2 foi filtrado pela dinâmica do sistema, dada a presença do termo $H(q,\theta) = \frac{1}{A(q)}$. É evidente que a função de transferência do filtro do ruído não possui zeros e está restrita à dinâmica do próprio sistema. Assim, se houver a necessidade de representar ruídos mais complexos, há uma limitação na modelagem da dinâmica do ruído. A **desvantagem** dos modelos ARX, portanto, é que eles modelam a dinâmica determinística, conjuntamente com as dinâmicas não-determinísticas.

Estrutura do modelo ARMAX

De modo a adicionar uma certa flexibilidade à descrição das propriedades do sinal de ruído, pode-se descrever a equação de erro como sendo **MA** (*moving average)* com ruído branco. Essa adição à estrutura do modelo ARX configura o novo modelo como sendo ARMAX. Assim:

$$y(t) + a_1y(t-1) + \ \dots\ a_{n_a}y(t-n_a) = b_1u(t-1) + \ \dots\ b_{b_b}u(t-n_b) + e(t) + c_1e(t-1) + \ \dots\ c_{n_c}e(t-n_c) \qquad (11)$$

Com os seguintes polinômios:

$$A(q) = 1 + a_1q^{-1} + \ + a_{n_a}q^{-n_a}$$

$$B(q) = b_1q^{-1} + \ \dots\ + b_nq^{-n_b}$$

$$C(q) = 1 + c_1q^{-1} + \ \dots\ c_{n_c}q^{-n_c}$$

Assim, a Eq. 10 pode ser reescrita como:

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

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

$$\theta = [a_1\ \dots\ a_{n_a}\ b_1\ \dots\ b_{n_b}\ c_1\ \dots\ c_{n_c}]^T$$

Ao substituir as Equações referentes ao modelo ARMAX na Eq.6, tem-se que o novo predito é dado por:

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

$$C(q)\hat{y}(t|\theta) = B(q)u(t) + [C(q) - A(q)]y(t) \qquad (12)$$

A intuição por trás da Eq.12 é que a predição é obtida ao filtrar u e y por meio de um filtro com dinâmicas do denominador determinadas por $C(q)$. A equação do preditor possui uma memória do próprio $\hat{y}(t|\theta)$. Ao adicionar $[1 - C(q)]\hat{y}(t|\theta)$ a ambos os lados da Eq. 12:

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

$$\hat{y}(t|\theta) = B(q)u(t) + C(q)y(t) - A(q)y(t) + \hat{y}(t|\theta) - C(q)\hat{y}(t|\theta) + y(t) - y(t)$$

A equação do preditor, conforme a Eq. 13, torna-se mais complexa devido à incorporação do modelo de ruído:

$$\hat{y}(t|\theta) = B(q)u(t) + [C(q) - 1][y(t) - \hat{y}(t|\theta)] + [1 - A(q)]y(t) \quad (13)$$

O erro de predição é dado por:

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

E o vetor de regressores agora, além de depender de t, depende do vetor de parâmetros θ, necessário para calcular o erro de estimação $\epsilon(t,\theta)$

$$\varphi(t,\theta) = [-y(t-1)\ \ldots\ -y(t-n_a)\ u(t-1)\ \ldots\ u(t-n_b)\ \varepsilon(t-1.\theta)\ \ldots\ \varepsilon(t-n_c.\theta)]^T$$

A equação do preditor, conforme exibido na Eq. 13, tem uma dependência implícita de θ

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

Oberva-se uma certa similaridade com a regressão linear aplicada anteriormente. No entanto, é importante mencionar que não se trata de uma regressão linear devido ao efeito de θ no vetor de parâmetros $\varphi(t,\theta)$, configurando-se como um regressor pseudolinear.

O que fica claro é que, a cada vez que refina-se o modelo de ruído, incorpora-se uma dependência implícita à equação do preditor.

Outras estruturas de modelos do tipo de equação de erro

Anteriormente, utilizava-se o polinômio $C(q)$ para inserir zeros no filtro de ruído. No caso desse novo tipo de estrutura de modelo, ARARX, acrescenta-se mais pólos no filtro de ruído por meio de $D(q)$, além das dinâmicas determinísticas incorporadas por $A(q)$:

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

$$D(q) = 1 + d_1q^{-1} + \ldots + d_{n_d}q^{-n_d}$$

Utilizando uma descrição ARMA da equação de erro, adiciona-se os polinômios $C(q)$ e $D(q)$ na representação do ruído:

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

Assim, a família dos conjuntos de modelo relacionados à equação de erro pode ser observada na Fig. 3.

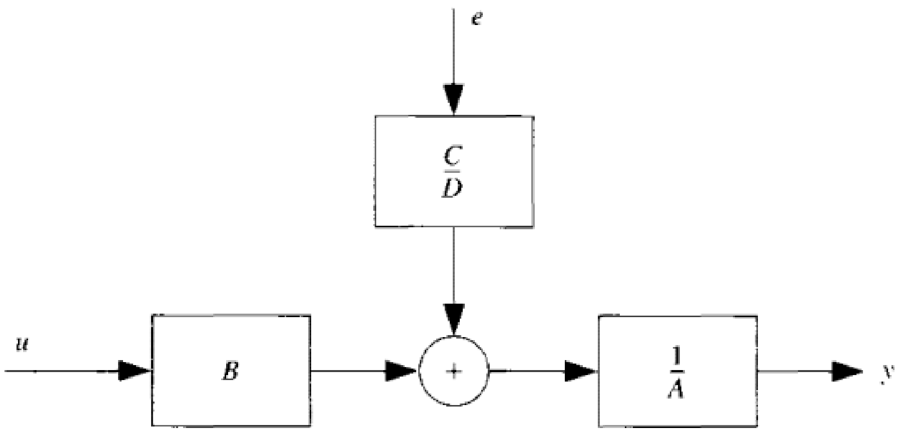

Figura 3: A estrutura geral dos modelos relacionados à equação de erro.

Conforme exibido na Fig.3, o erro $e(t)$ é filtrado por $\frac{C}{D}$ e posteriormente pelas dinâmicas do sistema $\frac{1}{A}$. Isso contribui em um aumento da liberdade de representação do sistema, especificamente em termos do ruído.

Estrutura do modelo de erro de saída

As estruturas anteriores, a parametrização das funções G e H tinha o polinômio $A(q)$ como um fator comum aos denominadores. Levando em consideração o aspecto físico de parametrizar as funções de transferência independentemente, pode se utilizar um outro tipo de estrutura. Supondo que haja um erro de medição associado à saída, e que a relação entre a entrada e a saída intermediária pode ser expressa pela Eq.14, tem-se:

$$w(t) + f_1 w(t-1) + ... + f_{n_f} w(t-n_f) = b_1 u(t-1) + ... + b_{n_b} u(t-n_b)$$

$$F(q) = 1 + f_1 q^{-1} + ... + f_{n_f} q^{-n_f}$$

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

O vetor de parâmetros a ser determinado é, então:

$$\theta = [b_1 \;\; b_2 \;\; ... \; f_1 \;\; f_2 \;\; f_{n_f}]^T$$

Sabendo que $w = y - e$ e incorporando a dependência do sinal $w(t)$ à parametrização dos coeficientes da função de transferência, tem-se:

$$u(t)\frac{B(q)}{F(q)} = y(t) - e(t) = w(t,\theta)$$

$$w(t,\theta) + f_1 w(t-1,\theta) + \ ... \ + f_{n_f} w(t-n_f,\theta) = b_1 u(t-1) + \ + b_{n_b} u)(t-n_b)$$

Ao comparar com a descrição geral do modelo $y(t) = G(q,\theta)u(t) + H(q,\theta)e(t)$, perceb-se que $H(q,\theta) = 1$, o que gera o preditor:

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)}u(t) = w(t,\theta)$$

Sendo assim, $\hat{y}(t|\theta)$ é determinado somente a partir das entradas passadas. Com a definição do vetor de regressores, a predição pode ser expressa por:

$$\varphi(t,\theta) = [u(t-1) \ ... \ u(t-n_b) \ \ -w(t-1,\theta) \ ... \ -w(t-n_f,\theta)]^T$$

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

Referências

[1] L. LJUNG. **System Identification: Theory for the User** Courier Corporation, 2008. ISBN 0486462781.