Relatório de Atividade 04 - Equação de predição

Débora Nunes Pinto de Oliveira

Prof. Antonio Marcus, Estimação e Identificação de Sistemas 21.2

7 de setembro de 2021

Esse documento tem por objetivo descrever a caracterização de um preditor para um sistema, considerando um distúrbio na modelagem.

Table of Contents

Introdução

Considere um sinal genérico de distúrbio $v(t)$ tal que $v(t) = H(q)e(t)$, para $e(t)$ um ruído branco e $H(q)$ a função de transferência de um filtro. Conforme discutido no Capítulo 2 de [1], a função de transferência $H(q)$ pode ser escrita conforme:

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

Logo, a inversa $H^{-1}(q)$ pode ser escrita como:

$$H^{-1}(z) = \sum_{k=0}^{\infty} \tilde{h}(k) z^{-k}$$

Se a $H^{-1}(z)$ existir para $|z| \geq 1$, o filtro $H^{-1}(z)$ é estável. Sendo assim, a função $H(z)$ não deve possuir zeros dentro do círculo unitário.

Considerando que $H(q)$ é um polinômio mônico, é evidente que:

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

Como o segundo termo da soma é conhecido para o instante anterior $t-1$, ele pode ser representado por um sinal $m(t-1)$ tal que:

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

Para $f_v(t)$ e $f_e(t)$ as funções de distribuição de probabilidade de $v(t)$ e $e(t)$, respectivamente, $f_v(t)$ e $f_e(t)$ podem, então, ser relacionadas pela expressão:

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

A função $f_e(t)$ é bem definida para $e(t)$ um ruído branco, uma vez que:

$$E[e(t)] = \int x f_e(x) dx = 0$$

$$E[e^2(t)] = \int x^2 f_e(x) dx = \lambda$$

Logo, é possível utilizar essa função de probabilidade $f_v(t)$ para realizar predições sobre o valor de $v(t)$ a partir do instante $t-1$. Essas predições serão modeladas conforme o equacionamento de uma função preditora, a qual será descrita na próxima seção.

Modelo do preditor

O valor previsto $\hat{v}(t|t-1)$ pode ser definido como o valor para o qual a função de densidade de probabilidade $f_v(t)$ é máxima. Esse valor é a expectativa matemática de $v(t)$:

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

Uma vez que $e(t)$ é um ruído branco de média nula, $E[e(t)] = 0$. Sabendo que $e(t) = H^{-1}(q)v(t)$, tem-se que:

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

Analogamente, para a saída de um sistema $y(t) = G(q)u(t) + v(t)$, a previsão $\hat{y}(t|t-1)$ é dada pela expressão:

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

Substituindo $v(t) = y(t) - G(q)u(q)$, obtém-se que:

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

A equação acima é nomeada equação do preditor. O erro do preditor é, então, dado por:

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

Dessa forma, é evidente que a variável $e(t)$ representa a parte da saída $y(t)$ que não pode ser prevista a partir de dados anteriores. Essa afirmação é confirmada pela característica do ruído $e(t)$ ser branco, já que o valor atual independe dos valores assumidos em instantes anteriores.

Geralmente, não são conhecidos os coeficientes de $G(q)$ e $H(q)$. Esses parâmetros, no entanto, podem ser estimados. Para tal, é importante destacar que ambos os filtros $H^{-1}(q)$ e $H^{-1}(q)G(q)$ devem possuir decaimento rápido no domínio da frequência para minimizar o erro da condição inicial do preditor.

Estimativa dos polinômios G e H

Para esse estudo, considera-se que os coeficientes de $G(q)$ e $H(q)$ compõem um vetor nomeado θ. Dessa forma, o modelo pode ser reescrito como:

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

O regressor para a estimativa dos coeficientes θ pode ser modelado conforme diferentes composições de funções racionais $G(q,\theta)$ e $H(q,\theta)$. De modo generalizado, o modelo de $y(t)$ pode ser dado por:

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

Sendo assim, de forma generalizada, o preditor é dado por:

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

Adaptação para múltiplas variáveis

No caso de multivariáveis, o modelo do processo simulado pode ser dado como:

$$y(t) + A_1 y(t-1) + \ldots + A_{n_a} y(t-n_a) = B_1 u(t-1) + \ldots + B_{n_b} u(t-n_b)$$

para m saídas e p entradas. Logo, A_i e B_i são matrizes de ordem $p \times p$ e $p \times m$, respectivamente, tal que:

$$A(q) = I + A_1 q^{-1} + \ldots + A_{n_a} y(t-n_a)$$
$$B(q) = B_1 q^{-1} + \ldots + B_{n_B} y(t-n_b)$$

Considerando que $\theta = \left[A_1 \ldots A_{n_a}\, B_1 \ldots B_{n_b}\right]$ e $\varphi = \left[-y(t-1)\ \ldots\ -t(t-n_a)\ u(t-1)\ \ldots\ u(t-n_b)\right]$, pode-se reescrever $y(t)$ como:

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

Adaptação para espaço de estados

Um modelo também pode ser representado por um espaço de estados:

$$\dot{x}(t) = F(\theta)x(t) + G(\theta)u(t)$$
$$\eta(t) = Hx(t)$$

Para x o vetor de estados, u o vetor de entradas e η o vetor de saídas. Considerando p o operador derivada, tem-se:

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

Logo, a função de transferência $G_c(p,\theta) = u(t)/\eta(t)$ é escrita como:

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

Se considerarmos que há erro de medição, então:

$$y(t) = Hx(t) + v(t) = G_c(p,\theta)u(t) + v(t).$$

Discretização do modelo em espaço de estados

É importante destacar que as medições do sistema são amostradas em um período T. Logo, a equação diferencial $\dot{x}(t) = F(\theta)x(t) + G(\theta)u(t)$ que modela o sistema é discretizada por:

$$x(kT + T) = A_T(\theta)x(kT) + B_T(\theta) + u(kT)$$

para $A_T(\theta) = e^{F(\theta)T}$ e $B_T(\theta) = \int_{\tau=0}^{T} e^{F(\theta)\tau}G(\theta)d\tau$. Analogamente a $G_c(p,\theta)$, obtém-se em tempo discreto:

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

Dessa forma, de modo semelhante, Se considerarmos que há erro de medição:

$$y(t) = G_T(q,\theta)u(t) + v_T(t)$$

Desenvolvimento

Para o estudo do modelo geral do equacionamento de preditor, serão realizados dois exercícios propostos pela literatura [1]. Na primeira atividade, será analisado um modelo ARMA. Por sua vez, na segunda, serão examinados múltiplos modelos para um circuito RLC. Nesse último item, serão avaliadas as vantagens e desvantagens de cada equacionamento.

Resolva o Problema 4E.1

Considere o modelo ARX:

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

para $b_1 = 0.5$. O modelo do preditor pode ser escrito então como $\widehat{y}(t|\theta) = \theta^T \varphi(t) + \mu(t)$ para $\mu(t)$ um termo conhecido. Inicialmente, considera-se que:

$$\varphi(t) = \left[-y(t-1) \ \ldots \ y(t-n_a) \ u(t-2) \ \ldots \ u(t-n_b)\right] \text{ e } \theta = \left[a_1 \ a_2 \ \ldots \ a_{n_a} \ b_2 \ \ldots \ b_{n_b}\right]$$

já que b_1 é conhecido. Sendo assim, o preditor pode ser escrito como:

$$\widehat{y}(t|\theta) = \theta^T \varphi(t) + 0.5u(t-1)$$

Resolva o Problema 4E.4

Considere o circuito RLC com duas entradas $u_v(t)$ e $u_i(t)$ tal qual ilustrado na Figura 1:

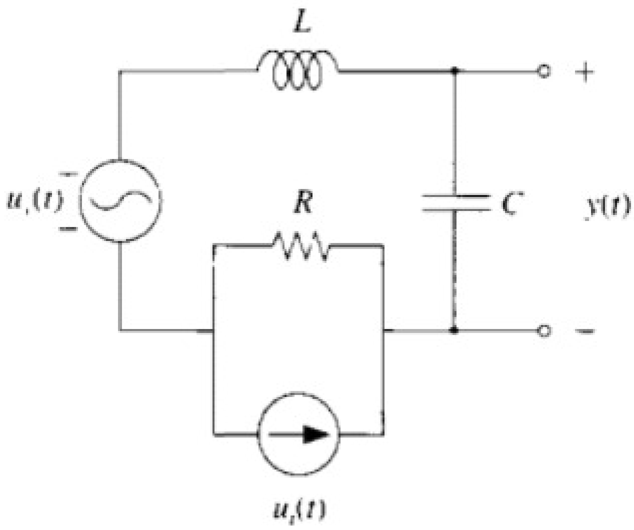

Figura 1: Circuito RLC com duas entradas $u_v(t)$ e $u_i(t)$

Dessa forma, conforme a LKT, tem-se que:

$$u_v(t) = L\frac{di(t)}{dt} + y(t) + R\left[i(t) + u_i(t)\right], \text{ para } y(t) = \frac{1}{C}\int_0^t i(\tau)d\tau$$

Considere $y(t)$ a tensão de saída, e R, L e C constantes desconhecidas. Sabendo que $\frac{di(t)}{dt} = C\frac{d^2y(t)}{dt^2}$, para o vetor de estados $x = [y(t) \ dy(t)/dt]^T$, tem-se o espaço de estados:

$$\begin{bmatrix} \frac{dy(t)}{dt} \\ \frac{d^2y(t)}{dt^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{LC} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} y(t) \\ \frac{dy(t)}{dt} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\frac{R}{LC} & \frac{1}{LC} \end{bmatrix} \begin{bmatrix} u_i(t) \\ u_v(t) \end{bmatrix}$$

Sendo assim, as funções de transferência entre $y(t)$ e as entradas $u_v(t)$ e $u_i(t)$ são dadas como:

$$G_1(s) = \frac{Y(s)}{U_i(s)} = \frac{-R/LC}{s^2 + sR/L + 1/LC} \text{ e } G_2(s) = \frac{Y(s)}{U_v(s)} = \frac{1/LC}{s^2 + sR/L + 1/LC}$$

É possível obter a equivalência em tempo discreto $G_1(z)$ e $G_2(z)$ aplicando, na expressão acima, a aproximação de Tustin ou Euler. Sem perda de generalidade, é possível definir:

$$G_i(z) = \frac{B(q)}{F(q)},\ i = 1, 2$$

Seja $G(z) = [G_1(z)\ G_2(z)]$, pode-se escrever que $y(t) = G(q)u(t)$, para q o operador de atraso (*backshift*). Considerando que a medição não é exata, isto é, que há um ruído $v(t)$, a saída do sistema pode ser escrita como:

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

Conforme definido na Introdução, $v(t) = H(q)e(t)$, para $e(t)$ um ruído branco e $H(q)$ a função de transferência de um filtro. Assim, de modo genérico, é possível reescrever $y(t)$ como:

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

A seguir, serão tratados os principais modelos para esses polinômios $A(q)$, $B(q)$, $F(q)$, $C(q)$ e $D(q)$.

Modelo ARX

Primeiramente, considere que:

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

para $\theta = [a_1\ a_2\ \ldots\ a_{n_a}\ b_1\ b_2\ \ldots\ b_{n_b}]$. Reordenando os termos da expressão acima, é evidente que:

$$G(q, \theta) = \frac{B(q)}{A(q)} = \frac{b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}} \text{ e } H(q, \theta) = \frac{1}{A(q)} = \frac{1}{1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}}$$

Esse modelo é nomeado ARX, uma vez que há uma parte auto regressiva (AR) $A(q)y(t)$ e uma entrada (X) $B(q)u(t)$. Dessa forma, para esse modelo ARX, o preditor pode ser reescrito como:

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

para $\varphi(t) = [-y(t-1)\ \ldots\ y(t - n_a)\ u(t-1)\ \ldots\ u(t - n_b)]$. É importante notar que esse modelo não permite uma flexibilização das características do distúrbio $e(t)$, uma vez que supõe uma dinâmica determinística para um modelo estocástico. Dessa forma, é recomendado utilizar outros modelos, como o ARMAX.

Modelo ARMAX, ARARX e ARARMAX

Por exemplo, $y(t)$ pode ser escrito como:

$$y(t) + a_1 y(t-1) + \ldots + a_{n_a} t(t - n_a) = b_1 u(t-1) + \ldots + b_{n_b} u(t - n_b) + e(t) + c_1 e(t-1) + \ldots + c_{n_c} e(t - n_c)$$

para $\theta = [a_1\ a_2\ \ldots\ a_{n_a}\ b_1\ b_2\ \ldots\ b_{n_b}\ c_1\ c_2\ \ldots\ c_{n_c}]$. Esse modelo é denominado ARMAX, já que, além das propriedades ARX, possuí um média móvel (MA, do inglês *moving average*) $C(q)e(t)$. Considerando $C(q) = 1 + c_1 q^{-1} + \ldots + c_{n_c} q^{-n_c}$, a equação acima pode ser reescrita como:

$$A(q)y(t) = B(q)u(t) + C(q)e(t)$$

Sendo assim, é evidente que:

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

Logo, o preditor é escrito conforme:

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

Adicionando $[1 - C(q)]\hat{y}(t|\theta)$ em ambos os lados da equação, tem-se:

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

Uma vez que o erro de predição é dado por $\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$, o vetor $\varphi(t,\theta)$ é dado por:

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

Como φ é dependente de ambos t e θ, a regressão sobre o vetor θ é nomeada pseudolinear (não linear). Isso não ocorre no modelo ARX, uma vez que φ depende unicamente dos valores de entrada u e y nos instantes anteriores a t.

Também é possível representar o termo o erro $v(t)$ como uma auto regressão. Dessa forma, tem-se que:

$$A(q)y(t) = B(q)u(t) + \frac{1}{D(q)}e(t), \text{ para } D(q) = 1 + d_1q^{-1} + ... + d_{n_d}q^{-n_d}$$

Esse modelo é denominado ARARX, já que possuí duas auto regressões $A(q)y(t)$ e $D^{-1}(q)e(t)$, e uma entrada $B(q)u(t)$. Analogamente, é possível descrever o distúrbio como um modelo ARMA, tal que:

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

Essa formulação é denominada ARARMAX, já que, além das propriedades ARARX, possuí um média móvel $C(q)e(t)$. É importante destacar que, assim como o ARX, o modelo ARARX permite uma regressão linear, uma vez que φ depende unicamente de t. Já o ARARMAX é pseudolinear e semelhante ao modelo ARMAX, já que possuí uma média móvel sobre $e(t)$.

Por fim, é claro que todos os modelos até então apresentados consideram um denominador $A(q)$ comum entre $G(q)$ e $H(q)$. De um ponto de vista físico, é mais natural parametrizar essas funções de transferência independentemente.

Modelo de erro na saída

Para considerar a independência entre $G(q)$ e $H(q)$, pode-se escrever $y(t)$ como:

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t), \text{ para } F(q) = 1 + f_1q^{-1} + ... + f_{n_f}q^{-n_f}$$

Isso significa que há um erro somado diretamente sobre a saída do sistema $w(t) = [B(q)/F(q)]u(t)$, isto é:

$$w(t) + f_1w(t-1) + ... + f_{n_f}w(t-n_f) = b_1u(t-1) + ... + b_{n_b}u(t-n_b)$$

Entretanto, $w(t)$ não é diretamente observado, mas obtido a partir da expressão acima para cada aproximação dos coeficientes f_i e b_i. Dessa forma, $w(t,\theta)$ é alterado a cada iteração do algoritmo de predição.

Seja o vetor θ é composto por:

$$\theta = \left[b_1\ b_2\ ...\ b_{n_b}\ f_1\ f_2\ ...\ f_{n_f}\right]$$

Logo, o vetor $\varphi(t,\theta)$ é dado por:

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

Esse modelo então possuí a mesma restrição do ARMAX e ARARMAX, uma vez que é pseudolinear.

Referências bibliográficas

[1] L. LJUNG. **System Identification: Theory for the User**. Pearson, 1998. 2nd edition, ISBN 9788131744956.