Capítulo 4 Modelagem de Sistemas Dinâmicos Estocásticos

Wagner C. Amaral e João B. R. do Val*

*Depto. de Sistemas e Energia – FEEC – UNICAMP

25 de abril de 2018

FULLSCREEN NEXT

PROCESSOS ESTOCÁSTICOS

Um processo estocástico $\{x_t(\omega), t \in T, \omega \in \Omega\}$ é uma família de variáveis aleatórias indexadas pela variável t e definido no espaço de probabilidades (Ω,A,\mathcal{P}) , onde Ω é o espaço de probabilidades, A é uma σ -algebra e \mathcal{P} uma medida de probabilidade definida em A. Para cada valor de t, $\{x_t(\omega), \omega \in \Omega\}$ é uma variável aleatória. Para cada valor de $\omega \in \Omega$, $\{x_t(\omega), t \in T\}$ é uma t e uma t

PROCESSOS ESTOCÁSTICOS

Seja o processo estocástico contínuo $\{x_t(\mathbf{\omega}), t \in T\}$ para $t = \{t_1, t_2, \dots, t_n\} = \{t_i\} \in T$, nesse caso, o processo estocástico é caracterizado pela distribuição conjunta, ou pela função densidade de probabilidades das variáveis aleatórias, $\{x_t\}$, isso é, $p(x_{t_1}, x_{t_2}, \dots, x_{t_n})$.

Exemplo 1: Seja o processo estocástico $\{x_t, t \geq 0\}$ definido por :

$$x_t = a + bt$$

onde a,b são variáveis aleatórias com distribuição de probabilidade dada. Dessa equação pode-se verificar que as realizações do processo estocástico são retas com declividade e valor constante aleatórios. Essas retas geram o seguinte processo estocástico:

$$\begin{bmatrix} x_{t_1} \\ \vdots \\ x_{t_n} \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Conhecendo-se a distribuição de probabilidade das variáveis aleatórias a,b pode-se calcular a função densidade de probabilidade conjunta $p(x_{t_1},...,x_{t_n})$ que descreve o processo estocástico.

MOMENTOS DE UM PROCESSO ESTOCÁSTICO

A seguir definem-se algumas estatísticas associadas com as densidades de probabilidades que descrevem o processo estocástico. Para $\{\mathbf{x}_t(\mathbf{\omega}), t \in T\}$ um processo estocástico n dimensional definem-se:

Função valor médio

$$\overline{\mathbf{x}} = m_{\mathbf{x}}(t) = \mathcal{E}(\mathbf{x}) = \int \mathbf{x} p(\mathbf{x}, t) d\mathbf{x}$$

onde $p(\mathbf{x},t)$ é a densidade de probabilidade da variável aleatória $\mathbf{x}(t)$

Função de Covariância

$$\mathbf{X}(t_1,t_2) = \mathcal{E}\{(\mathbf{x}(t_1) - \overline{\mathbf{x}}(t_1))(\mathbf{x}(t_2) - \overline{\mathbf{x}}(t_2))^T\}$$

Os elementos da diagonal de $\mathbf{X}(t_1,t_2)$ são as funções de auto-covariância e os elementos fora da diagonal de $\mathbf{X}(t_1,t_2)$ são as funções de covarância cruzada de $\mathbf{X}(t_i,t_j), i \neq j, \{i,j=1,\ldots,n\}$.

MOMENTOS DE UM PROCESSO ESTOCÁSTICO

Exemplo 2: Seja o processo estocástico do exemplo 1, isto é:

$$x_t = a + bt$$

então

$$\bar{x_t} = \mathcal{E}\{x_t\} = \mathcal{E}\{a\} + \mathcal{E}\{b\}t$$

e

$$X(t_1,t_2) = \mathcal{E}\{(x_{t_1} - \bar{x}_{t_1})(x_{t_2} - \bar{x}_{t_2})\} =$$

$$= var(a) + var(b)t_1t_2 + cov\{ab\}(t_1 + t_2)$$

Se a variáveis aleatórias a,b são não correlacionadas com média zero e variância unitária, tem-se que: $\bar{x}_t = 0$ e $X(t_1,t_2) = 1 + t_1t_2$.

PROCESSO ESTOCÁSTICO ESTACIONÁRIO

O processo estocástico $\{x_{t_i}(\mathbf{\omega}), t_i \in T, x \in R\}$, é estacionário se a densidade de probabilidade conjunta de $x_{t_1}, x_{t_2}, x_{t_3}, \dots, x_{t_n}$ é igual à densidade conjunta de $x_{t_1+\tau}, x_{t_2+\tau}, x_{t_3+\tau}, \dots, x_{t_n+\tau}$ para $\forall \tau, n$

Nesse caso o valor médio do processo estocástico é dado por:

$$\overline{x}(t_1) = \int xp(x,t_1)dx = \int xp(x,t_{1+\tau})dx$$
$$= \int xp(x,t_2)dx = \overline{x}(t_2) = \overline{x}$$

Isto é, se a função do valor médio existe, ela é constante.

A função de covariância é dada por:

$$X(t_1,t_2) = X(t_1 + \delta,t_2 + \delta), \forall \delta, \delta = -t_1$$

 $X(t_1,t_2) = X(0,t_2 - t_1) = X(\tau) \ \tau = t_2 - t_1$

Se a função de covariância existe, ela depende somente da diferença entre os instantes de tempo. Propriedades similares poderão ser deduzidas para os momentos de ordem superior.

PROCESSO ESTOCÁSTICO ESTACIONÁRIO NO SENTIDO AMPLO

O processo estocástico $\{x_t(\omega), t \in T\}$ é estacionário no sentido amplo, wide sense, quando ele tem momento finito de segunda ordem, função valor médio constante, isto é:

$$\overline{\mathbf{x}}(t) = \overline{\mathbf{x}}$$

e

$$\mathbf{X}(t_1,t_2) = X(\tau)$$

Nesse caso como, no caso anterior, a matriz de covariância depende somente de $\tau = t_1 - t_2$. Contudo não se impõe nenhuma condição sobre os momentos de ordem superior.

PROCESSO DE MARKOV

Sejam $t_i \in T$, tal que, $t_1 < t_2 ... < t_n < t$ então um processo estocástico $\{x_t(\omega), t \in T\}$ é um processo de Markov quando:

$$\mathcal{P}(\boldsymbol{\omega}: -\infty < x(t) \le x \mid x_{t_1}, x_{t_2}, \dots, x_{t_n}) =$$

$$= \mathcal{P}(\boldsymbol{\omega}: -\infty < x(t) \le x/x_{t_n})$$

onde $\mathcal P$ é uma medida de probabilidade

Para a função densidade de probabilidade tem-se que:

$$p(x_t | x_{t_1}, \dots, x_{t_n}) = p(x_t | x_{t_n})$$

Propriedade: A função densidade de probabilidade conjunta é dada por:

$$p(x_{t_0} x_{t_1} \dots x_{t_n}) = p(x_{t_n} | x_{t_{n-1}}) p(x_{t_{n-1}} | x_{t_{n-2}}) \dots p(x_{t_1} | x_{t_0}) p(x_{t_0})$$

Onde a densidade $p(x_k|x_{x-1})$ é denominada densidade de transição de estado. Isto é, a densidade de probabilidade conjunta pode ser obtida a partir da densidade do instante inicial e da densidade de transição de estado.

PROCESSO DE MARKOV

Exemplo: Seja o sistema descrito pela seguinte equação à diferenças:

$$x_{k+1} = x_k + \omega_k; k = 0, 1, 2, \dots$$

onde $\{\omega_k\}$ é uma sequência de variáveis aleatórias independentes e x_0 é independente de ω_k . A sequência $\{x_k, k=0,1,2,\ldots\}$ é um processo de Markov, pois dado x_k, x_{k+1} depende somente de ω_k que é independente de x_{k-1},\ldots,x_0 , e: $p(x_{k+1}|x_k)=p_{\omega_k}(\omega_k=x_{k+1}-x_k)$.

PROCESSO ESTOCÁSTICO RUÍDO BRANCO

O Processo Ruído Branco é o processo estocástico $\{x_{t_i}(\omega), i=1,2,\ldots\}$ para o qual a densidade transição de estado independe do instante anterior, isto é:

$$p(x_{t_k} | x_{t_{k-1}}) = p(x_{t_k})$$

representando que todos os elementos da sequência $\{x_{t_k}\}$ são mutuamente independentes. Desse modo um processo estocástico ruído branco é totalmente não previsível, isto é, a informação sobre o sistema em instantes anteriores não permite uma melhor previsão sobre o seu comportamento no instante atual.

Se $\{x_{t_i}(\mathbf{\omega}), i=1,2,\ldots\}, x_{t_i} \in R$ um processo estocástico estacionário ruído branco, então a matrix de covariância é dada por :

$$X(t_i,t_l) = \Sigma(t_i)\delta_{t_it_l}$$

onde:

$$\delta_{t_i t_l} = \begin{cases} 1 & i = l \\ 0 & i \neq l \end{cases}$$

é a função delta de Kronecker e $\Sigma(t_i)$ é uma matriz definida não negativa.

DENSIDADE ESPECTRAL

Para um processo estocástico estacionário no sentido amplo $\{x_t(\omega)\}$ com função de covariância $X(\tau)$ define-se a densidade espectral como:

$$\Phi_{X}(\mathbf{\omega}) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \exp(-jk\tau\mathbf{\omega}) X(k\tau)$$

Dado que a função de covariância é simétrica em relação à origem, a função densidade espectral pode ser reeescrita como:

$$\Phi_{X}(\boldsymbol{\omega}) = \frac{1}{2\pi} \{ X(0) + 2\sum_{k=1}^{\infty} \cos(k\tau \boldsymbol{\omega}) X(k\tau) \}$$

Quando o processo estocástico é ruído branco tem-se que:

$$X(k\tau) = \left\{ egin{array}{ll} \Sigma & \mathsf{para} \ k = 0 \\ 0 & \mathsf{para} \ k
eq 0 \end{array}
ight.$$

Portanto

$$\Phi_{x}(\mathbf{\omega}) = \frac{\Sigma}{2\pi}$$

Isto é, a densidade espectral de um ruído branco é constante.

DENSIDADE ESPECTRAL E RUÍDO BRANCO

Exemplo 1
$$y(k) = 1 + u(k) + u(k-1)$$

u(k) é um ruído branco com média zero e covariância dada por:

$$U_k = \left\{ egin{array}{ll} 1 & \mathsf{para} \ k = 0 \ 0 & \mathsf{para} \ k
eq 0 \end{array}
ight.$$

Então $\bar{y} = \mathcal{E}y(k) = 1$ e

$$Y(k-j) = \mathcal{E}\{(y(k) - \bar{y})(y(j) - \bar{y})\}\$$

= $\mathcal{E}\{(u(k) + u(k-1))(u(j) + u(j-1))\}$

Assim, se

$$k = j \rightarrow Y(0) = 2$$

$$k = j + 1 \rightarrow Y(1) = 1$$

$$k > j + 1 \rightarrow Y(k) = 0$$

E a função densidade espectral é dada por:

$$\Phi_{y}(\omega) = \frac{1}{2\pi} \{ 2 + 2\cos \omega T \}$$

DENSIDADE ESPECTRAL E RUÍDO BRANCO

Exemplo 2 $y_k = 1 + u_k - u_{k-1}$

Com a entrada u_k conforme descrita no exemplo anterior. Neste caso tem-se que:

$$Y(0) = 2$$
; $Y(1) = -1$, $Y(l) = 0$, $l \ge 2$

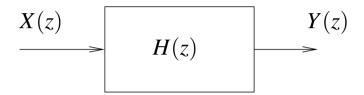
e a função densidade espectral é dada por:

$$\Phi_{y}(\mathbf{\omega}) = \frac{1}{2\pi} \{ 2 - 2\cos \mathbf{\omega}T \}$$

Observar a diferença de comportamento dos sistemas dos exemplos acima.

TRANSMISSÃO DE SÉRIES ESTOCÁSTICAS EM SISTEMAS DINÂMICOS

Seja um sistema linear representado por:



$$Y(z) = H(z)X(z),$$
 $y(k) = \sum_{m=0}^{\infty} h_m x_{k-m}$

Analisar a evolução da função de covariância.

$$Y(k,l) = \mathcal{E}\{y_k y_l^T\} = \mathcal{E}\{(\sum_{m=0}^{\infty} h_m x_{k-m})(\sum_{n=0}^{\infty} h_n x_{l-n})^T\}$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} h_m \mathcal{E}(x_{k-m} x_{l-n}^T) h_n^T$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} h_m X(k-m,l-n) h_n^T$$

Tomando i = k - l, a densidade espectral da saida é dada por:

$$\begin{split} \Phi_{y}(\omega) = & \frac{1}{2\pi} \sum_{i=-\infty}^{\infty} \exp(-jiT\omega)Y(i) \\ = & \frac{1}{2\pi} \sum_{i=-\infty}^{\infty} \exp(-jiT\omega) \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} h_{m}X(i-m+n)h_{n}^{T} \\ = & \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \exp(-jmT\omega)h_{m} \cdot \\ & \left(\frac{1}{2\pi} \sum_{i=-\infty}^{\infty} \exp(-j(i-m+n)T\omega)X(i-m+n)\right) \exp(+jnT\omega)h_{n}^{T} \\ = & \sum_{n=0}^{\infty} H(z)\Phi_{x}(\omega) \exp(jnT\omega)h_{n}^{T} = H(z)\Phi_{x}(\omega)H^{T}(z^{-1}) \end{split}$$

com

$$H(z) = \sum_{m=0}^{\infty} H_m z^{-m}$$

Quando a entrada é um ruído branco tem-se que:

$$\Phi_{x}(\omega) = \Omega \longrightarrow \Phi_{y}(\omega) = H(z)\Omega H^{T}(z^{-1})$$

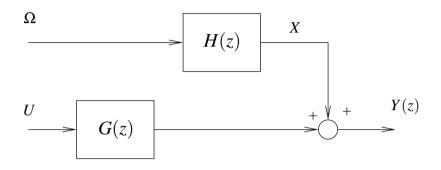
TEOREMA DA REPRESENTAÇÃO

Seja Φ_y uma matriz densidade espectral racional, então existe uma única matriz racional $H(z) \in \mathbb{R}^{n*n}$ e uma única matriz simétrica definida positiva Φ satisfazendo:

a)
$$\Phi_{y} = H(z)\Phi H^{T}(z^{-1})$$

- **b)** H(z) é analítica fora e sobre o circulo unitário.
- c) $H(z^{-1})$ é analítica fora do circulo unitário.
- **d**) $\lim_{n\to\infty} H(z) = I$

Utilizando o teorema da representação são obtidas diversas maneiras de representar os modelos matemáticos de um sistema



- ullet Sistema linear invariante no tempo perturbação X pode ser representada por um processo estocástico estacionário com densidade espectral racional
 - a transformada z aplicada no sistema

$$Y(z) = G(z)U(z) + H(z)\Omega(z)$$
(1)

G(z) é a f.t. da parte determinística do sistema, H(z) é a ft da parte estocástica do sistema, funções racionais e estáveis.

H(z) tem as seguintes propriedades adicionais:

- i) $H(z)^{-1}$ é estavel;
- ii) $\lim_{z\to\infty}H(z)=1$.
- → Caso geral: as f.t.'s da parte determinística e estocástica podem ter alguns polos em comum:

$$Y(z) = \frac{z^{-d}B(z)}{F(z)A(z)}U(z) + \frac{C(z)}{D(z)A(z)}w(z)$$
 (2)

■ A(z),B(z),C(z),D(z),F(z) são polinômios em z, cujas raízes são os polos e zeros da parte deterministica e estocástica do sistema e d é o atraso de transporte do sistema.

* Uma pausa Para um sistema linear invariante no tempo, a transformada z da saída y(k) definida por Y(z) relaciona-se com a transformada z da entrada u(k) definida por U(z) através de A(z)Y(z)=B(z)U(z), em que

$$A(z) = z^{n} + a_{1}z^{n-1} + \dots + a_{n}$$

 $B(z) = z^{m} + b_{1}z^{m-1} + \dots + b_{m}, \quad m \le n$

as raízes dos polinômios A(z) e B(z) são os polos e zeros do sistema. Aplicando a anti-transformada z,

$$y(k+n) + a_1y(k+n-1) + \ldots + a_ny(k) = u(k+m) + b_1u(k+m-1) + \ldots + b_mu(k)$$

* Uma pausa Com o operador avanço qy(k)=y(k+1) e o operador atraso $q^{-1}y(k)=y(k-1)$ tem-se

$$(q^{n} + a_{1}q^{n-1} + \ldots + a_{n})y(k) = (q^{m} + b_{1}q^{m-1} + \ldots + b_{m})u(k)$$

ou ainda

$$(1+a_1q^{-1}+\ldots+a_mq^{-m})y(k)=q^{-(n-m)}(1+b_1q^{-1}+\ldots+b_mq^{-m})u(k)$$

Portanto

$$y(k) = q^{-(n-m)} \frac{B(q^{-1})}{A(q^{-1})} u(k)$$

com

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_m q^{-m}$$
 e $B(q^{-1}) = 1 + b_1 q^{-1} + \ldots + b_m q^{-m}$ Fim *

Aplicando a anti-transformada z e utilizando o operador deslocamento resulta de (2)

$$A(q^{-1})y(k) = q^{-d} \frac{B(q^{-1})}{F(q^{-1})} u(k) + \frac{C(q^{-1})}{D(q^{-1})} \omega(k)$$
(3)

 q^{-1} é o operador atraso tal que $y(k)q^{-1}=y(k-1)$, $\omega(k)$ é o ruído branco.

 $lacksquare A(q^{-1}), B(q^{-1}), C(q^{-1}), D(q^{-1})$ e $F(q^{-1})$ são polinômios definidos na forma:

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

$$B(q^{-1}) = b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d}$$

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

$$(4)$$

 \longrightarrow De acordo com os valores particulares dos polinômios A,B,C,D e F, temos os diferentes modelos para identificação utilizados na literatura.

MODELO DO ERRO PREVISTO

- o modelo obtido da identificação do sistema pode ser utilizado para a previsão da saída do sistema, a partir das saídas e entradas anteriores do processo.
- Como a saída de um processo real não pode ser, em geral, exatamente determinada dada a sua característica estocástica, é importante conhecer no instante atual (k-1), qual é a saída mais provável do sistema no instante seguinte k
- \hookrightarrow Portanto é de interesse determinar o vetor de parâmetros do modelo do processo θ , de modo que o erro previsto, $\varepsilon(k,\theta)$ seja 'pequeno',

$$\varepsilon(k,\theta) = y(k) - \hat{y}(k | k - 1, \theta)$$

 $\hat{y}(k|k-1,\theta)$ é a previsão da saída $y(k) \leadsto$ é função do vetor de parâmetros do modelo θ e do conjunto de medidas até o instante k-1:

$$(y(k-1),u(k-1),y(k-2),u(k-2),...)$$

Essa equação é denominada MODELO DO ERRO PREVISTO.

OBTENDO-SE O PREDITOR PARA O MODELO GERAL

Aplicando a anti-transformada z na equação (1) tem-se que:

$$y(k) = G(q^{-1}; \theta)u(k) + H(q^{-1}; \theta)\omega(k)$$
 (5)

Suponha que $G(0,\theta)=0$, \leadsto o modelo do sistema apresenta pelo menos um instante de atraso de transporte entre a entrada e a saída e denotando

$$G(q^{-1};\theta) = q^{-1}G_1(q^{-1};\theta)$$

pode-se escrever que:

$$y(k) = q^{-1}G_1(q^{-1}; \theta)u(k) + H(q^{-1}; \theta)\omega(k)$$

= $[1 - H^{-1}(q^{-1}; \theta)]y(k) + q^{-1}H^{-1}(q^{-1}; \theta)G_1(q^{-1}; \theta)u(k) + \omega(k)$

pois
$$-H^{-1}y(k) + q^{-1}H^{-1}G_1u(k) + \omega(k) = 0.$$

OBTENDO-SE O PREDITOR PARA O MODELO GERAL

Definindo:

$$L_1(q^{-1};\theta) = q[1 - H^{-1}(q^{-1};\theta)]$$
 e $L_2(q^{-1};\theta) = H^{-1}(q^{-1};\theta)G_1(q^{-1};\theta)$

tem-se que:

$$y(k) = L_1(q^{-1}; \theta)y(k-1) + L_2(q^{-1}; \theta)u(k-1) + \omega(k)$$
(6)

De (6) obtem-se explicitamente, a saída prevista do sistema para um valor particular do estimador, $\theta = \hat{\theta}$, isto é:

$$\hat{y}(k|k-1;\hat{\theta}) = L_1(q^{-1};\hat{\theta})y(k-1) + L_2(q^{-1};\hat{\theta})u(k-1)$$

A seguir apresentam-se os diferentes modelos que são obtidos quando se realiza simplicações na forma geral descrita por (3).

MODELO DE RESPOSTA FINITA AO IMPULSO (FIR)

O somatório de convolução entre a resposta ao impulso, h(k), de um sistema e um sinal de entrada, u(k), resulta na seguinte representação da saída do sistema:

$$y(k) = \sum_{j=0}^{\infty} h(j)u(k-j) + \omega(k)$$
(7)

onde $\omega(k)$ é uma perturbação branca atuando no sistema.

A equação (7) é conhecida como modelo de resposta infinita ao impulso. Se o sistema for estável, então existe $J < \infty$ tal que $|h(k)| < \varepsilon, \forall k > J$. Portanto, truncando a equação (7), obtém-se o modelo de resposta finita ao impulso, FIR¹:

$$y(k) = \sum_{j=0}^{J} h(j)u(k-j) + \omega(k)$$
 (8)

onde J é o numero de elementos da resposta ao impulso.

¹ Finite Impulse Response.

MODELO DE RESPOSTA FINITA AO IMPULSO (FIR)

O modelo FIR pode ser obtido a partir da equação (3) tomando-se

$$A(q^{-1}) = C(q^{-1}) = D(q^{-1}) = F(q^{-1}) = 1$$
 e

 $B(q^{-1})$ um polinômio arbitrário de ordem J,

ou seja, $n_b = J$ (equação (4)). Nesse caso, o modelo FIR (8) pode ser escrito como :

$$y(k) = q^{-d}B(q^{-1})u(k) + \omega(k)$$
(9)

Para essa representação tem-se que a saída prevista $\hat{y}(k | k-1, \theta)$ é dada por:

$$\hat{y}(k \mid k-1, \theta) = q^{-d} B(q^{-1}) u(k) \tag{10}$$

e que também pode ser obtido da equação (5) tomando-se:

$$G(q^{-1}, \theta) = q^{-d}B(q^{-1}) \ H(q^{-1}, \theta) = 1$$
 (11)

MODELO DE RESPOSTA FINITA AO IMPULSO (FIR)

O preditor da saída pode ser representado na forma de regressão com o vetor de regressão, $\varphi(k)$, e o vetor de parâmetros, θ , definidos da seguinte forma:

$$\varphi(k) = [u(k-d) \dots u(k-d-n_b)]^T$$
 (12)

$$\theta = [b_0 \dots b_{n_b}]^T \tag{13}$$

e

$$\hat{\mathbf{y}}(k | k-1, \boldsymbol{\theta}) = \boldsymbol{\varphi}(k)^T \boldsymbol{\theta}$$

Um sistema dinâmico não pode ser descrito exatamente através de um modelo FIR. Contudo, se o sistema for estável e a sua resposta ao impulso decair rapidamente, o sistema normalmente pode ser bem aproximado por um modelo FIR, e a precisão do modelo estará associada com a ordem do polinômio $B(q^{-1})$.

MODELO AUTO-REGRESSIVO COM ENTRADAS EXTERNAS (ARX)

A estrutura do modelo ARX² corresponde a reescrever a equação (3) tomando

$$C(q^{-1})=D(q^{-1})=F(q^{-1})=1$$
 e $A(q^{-1})$ e $B(q^{-1})$ polinômios arbitrários,

resultando em:

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + \omega(k)$$
(14)

O ruído $\omega(k)$ aparece na equação (14) com uma única amostra, e o modelo ARX é classificado como pertencendo à classe de modelos com *erro na equação*.

Como o modelo FIR, o modelo ARX também pode ser reescrito na forma:

$$y(k) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})}u(k) + \frac{1}{A(q^{-1})}\omega(k)$$
(15)

Neste caso, o ruído adicionado à saída, $e(k) = \omega(k)/A(q^{-1})$, não é branco \rightarrow o ruído é modelado como ruído branco filtrado por um modelo auto-regressivo.

²AutoRegressive with eXogenous inputs.

MODELO AUTO-REGRESSIVO COM ENTRADAS EXTERNAS (ARX)

• forma alternativa de representar a estrutura do modelo ARX, baseada na equação do erro previsto:

$$G(q^{-1}, \theta) = q^{-d} \frac{B(q^{-1})}{A(q^{-1})} \qquad H(q^{-1}, \theta) = \frac{1}{A(q^{-1})}$$
(16)

e o preditor toma a seguinte forma:

$$\hat{y}(k \mid k-1, \theta) = q^{-d}B(q^{-1})u(k) + [1 - A(q^{-1})]y(k)
= \varphi^{T}(k)\theta$$
(17)

com

$$\varphi(k) = [y(k-1) \dots y(k-n_a), u(k-d) \dots u(k-d-n_b)]^T$$
 (18)

$$\theta = \begin{bmatrix} -a_1 & \dots & -a_{n_a}, & b_0 & \dots & b_{n_b} \end{bmatrix}^T \tag{19}$$

Desta equação pode-se observar que há somente uma relação algébrica entre a previsão da saída e as entradas passadas e saídas medidas .

MODELO AUTO-REGRESSIVO COM MÉDIA MÓVEL E ENTRADAS (ARMAX)

O modelo $ARMAX^3$ pode ser obtido a partir do modelo geral (3) tomando-se

$$D(q^{-1}) = F(q^{-1}) = 1 \ \mbox{e}$$

$$A(q^{-1}), B(q^{-1}) \ \mbox{e} \ C(q^{-1}) \ \mbox{polinômios arbitrários,}$$

resultando em:

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + C(q^{-1})\omega(k)$$
(20)

ou ainda:

$$y(k) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})}u(k) + \frac{C(q^{-1})}{A(q^{-1})}\omega(k)$$
 (21)

- Assim como o modelo ARX, o modelo ARMAX também é classificado como um modelo de *erro na equação*.
- → Neste caso, a perturbação (não branca), é modelada como a saída de um filtro ARMA com entrada ruído branco.

³AutoRegressive Moving Average with eXogenous inputs.

MODELO AUTO-REGRESSIVO COM MÉDIA MÓVEL E ENTRADAS (ARMAX)

Escolhendo os polinômios G e H conforme segue:

$$G(q^{-1}, \theta) = q^{-d} \frac{B(q^{-1})}{A(q^{-1})} \qquad H(q^{-1}, \theta) = \frac{C(q^{-1})}{A(q^{-1})}$$
(22)

obtém-se o seguinte preditor:

$$\hat{y}(k \mid k-1, \theta) = q^{-d} \frac{B(q^{-1})}{C(q^{-1})} u(k) + \left(1 - \frac{A(q^{-1})}{C(q^{-1})}\right) y(k)$$

então,

$$C(q^{-1})\hat{y}(k \mid k-1, \theta) = q^{-d}B(q^{-1})u(k) + (C(q^{-1}) - A(q^{-1}))y(k)$$
(23)

e assim,

$$\hat{y}(k \mid k-1,\theta) = q^{-d}B(q^{-1})u(k) + (C(q^{-1}) - A(q^{-1}))y(k) - (1 - C(q^{-1}))\hat{y}(k \mid k-1,\theta)
= q^{-d}B(q^{-1})u(k) + [1 - A(q^{-1})]y(k) + [C(q^{-1}) - 1]\varepsilon(k,\theta)
= \varphi^{T}(k,\theta)\theta$$

em que $\varepsilon(k,\theta) = y(k) - \hat{y}(k \mid k-1,\theta)$ representa o $erro\ de\ previsão\$ ou $resíduo\$.

MODELO AUTO-REGRESSIVO COM MÉDIA MÓVEL E ENTRADAS (ARMAX)

Os vetores de regressão e de parâmetros são definidos, respectivamente, por

$$\varphi(k,\theta) = [y(k-1) \dots y(k-n_a), u(k-d) \dots u(k-d-n_b), \\
\varepsilon(k,\theta) \dots \varepsilon(k-n_c,\theta)]^T$$
(24)

$$\theta = [-a_1 \dots -a_{n_a}, b_0 \dots b_{n_b}, c_1 \dots c_{n_c}]^T$$
 (25)

- a dinâmica do modelo do preditor é especificada pelas raízes do polinômio C(z) que devem estar dentro do círculo unitário no plano z para garantir que o preditor seja estável.
- lacktriangle a presença dos pólos no modelo do preditor acarreta que o vetor de regressão depende dos parâmetros do modelo, \leadsto não linearidade em θ na equação do preditor.
- → não linearidade acarretará a necessidade de métodos mais elaborados para a estimação dos parâmetros desconhecidos do modelo.
- a previsão da saída na equação do preditor depende das medidas realizadas e do erro previsto. Na seção PREDITOR ÓTIMO DE SAÍDA reescreve-se a equação do preditor para se obter uma relação dinâmica entre a saída prevista e as saídas medidas.

MODELO DE ERRO NA SAÍDA (OE)

O modelo OE⁴, também conhecido como *modelo com estrutura paralela*, é utilizado quando o ruído que afeta o sistema é do tipo ruído branco . Esta estrutura de modelos é representada por

$$y(k) = q^{-d} \frac{B(q^{-1})}{F(q^{-1})} u(k) + \omega(k)$$
(26)

que resulta na seguinte equação do preditor:

$$\hat{y}(k \mid k-1,\theta) = q^{-d} \frac{B(q^{-1})}{F(q^{-1})} u(k)
= q^{-d} B(q^{-1}) u(k) + [1 - F(q^{-1})] \hat{y}(k \mid k-1,\theta)
= \varphi^{T}(k,\theta) \theta$$
(27)

no qual,

$$\varphi(k,\theta) = [\hat{y}(k-1 \mid \theta) \dots \hat{y}(k-n_f \mid \theta), u(k-d) \dots u(k-d-n_b)]^T$$
 (28)

$$\theta = \begin{bmatrix} -f_1 & \dots & -f_{n_f}, & b_0 & \dots & b_{n_u} \end{bmatrix}^T$$
 (29)

⁴Output Error.

MODELO DE ERRO NA SAÍDA (OE)

- Para que o modelo do preditor seja estável, as raízes de F devem estar dentro do círculo unitário no plano z.
- Dificuldades semelhantes ao do modelo ARMAX

ESTIMADOR DE YULE-WALKER

- No caso geral a função de covariância é calculada utilizando o teorema da representação espectral.
- Para o caso de modelos ARX ($u_k = 0$) no entanto, tem-se o modelo AR que é descrito por:

$$A(q^{-1})y(k) = \mathbf{\omega}_k$$

Supondo que $n_a = n$, resulta

$$(1+a_1q^{-1}+a_2q^{-2}+\ldots+a_nq^{-n})y_k=\omega_k$$

Neste caso a função valor médio da saída é dada por:

$$(1 + a_1q^{-1} + a_2q^{-2} + \ldots + a_nq^{-n})\mathcal{E}y_k = 0$$

A função de Covariância é:

$$Y(k-j) = \mathcal{E}(y_k y_j)$$

$$\mathcal{E}y_k y_j + a_1 \mathcal{E}y_{k-1} y_j + \ldots + a_n \mathcal{E}y_{k-n} y_j = \mathcal{E}(y_j \omega_k)$$

ESTIMADOR DE YULE-WALKER

Equações de Yule-Walker para o modelo AR

$$Y(k-j) + a_1Y(k-j-1) + \dots + a_nY(k-j-n) = 0 \quad j < k$$
$$Y(0) + a_1Y(1) + \dots + a_nY(n) = W(0) \quad j = k$$

Para

$$k - j = 1 \rightarrow Y(1) + a_1 Y(0) + \dots + a_n Y(n - 1) = 0$$

 $k - j = 2 \rightarrow Y(2) + a_1 Y(1) + \dots + a_n Y(n - 2) = 0$
:
 $k - j = n \rightarrow Y(n) + a_1 Y(n - 1) + \dots + a_n Y(0) = 0$

Sistema de n equações e n incógnitas.

ESTIMADOR DE YULE-WALKER

Exercício:

- a) Escolha um sistema de 2a. ordem $y_k + a_1 y_{k-1} + a_2 y_{k-2} = \omega_k$ estável. Considerando que ω_k é um ruído branco com média zero e variância unitária e escolha valores para a_1 a_2 . Gere um conjunto de medidas deste processo $k+1,\ldots,1.000$
- b) Determinar experimentalmente a função de correlação para os modelos AR(1) e AR(2) descritos a seguir:

$$y_k + a_1 y_{k-1} = \mathbf{\omega}_k$$

 $y_k + a_1 y_{k-1} + a_2 y_{k-2} = \mathbf{\omega}_k$

- c) Calculando Y(0),Y(1),Y(2) estime os parâmetros a_i s utilizando as equações de Yule-Walker.
- d) Calcule o resíduo $e_j = y_j \hat{y_j}$ cujo estimador é obtido a partir de \hat{a}_1 e \hat{a}_2 , sobre outro conjunto de dados $y_j, j = 1, \dots, N$. Verifiqe se a correlação de e_j indica ou não ruído branco.

MODELO MÉDIA MÓVEL (MA)

Para o Modelo Média Móvel, MA, isto é, o modelo ARMA com u(k)=0 e $A(q^{-1})=1$ resulta

$$y_k = C(q^{-1})\omega_k$$

$$y_k = (1 + c_1q^{-1} + \dots + c_nq^{-n})\omega_k$$

$$y_k = \omega_k + c_1\omega_{k-1} + \dots + c_n\omega_{k-n}$$

Se
$$\omega_k = 0 \rightarrow \mathcal{E} y_k = 0$$

Neste caso a função de Covariância é dada por:

$$Y(k-j) = \mathcal{E}(y_k y_j)$$

$$Y(k-j) = \mathcal{E}(\boldsymbol{\omega}_k \boldsymbol{\omega}_j) + c_1 \mathcal{E}(\boldsymbol{\omega}_{k-1} \boldsymbol{\omega}_j) + \dots + c_n \mathcal{E}(\boldsymbol{\omega}_{k-n} \boldsymbol{\omega}_j) + \dots + c_n \mathcal{E}(\boldsymbol{\omega}_{k$$

Como

$$\mathcal{E}(\mathbf{\omega}_i\mathbf{\omega}_j) = 0$$
 para $j
eq n o Y(k-j) = 0$ para $k-j > n$

A função de covariância é nula para argumentos maiores do que a ordem do modelo.

PREDITOR ÓTIMO DA SAÍDA.

problema: calcular a previsão da saída de um sistema em um determinado instante de tempo maior do que o instante atual, condicionado à informação disponível até o instante atual.

Seja um sistema cuja saída é denotada por y(k) e deseja-se obter a melhor previsão dessa saída d instantes de tempo a frente,

 \rightarrow calcular a previsão da saída denotada por, $\hat{y}(k+d\mid k)$, que seja ótima no sentido de minimizar um critério especificado como,

$$\min_{z \in \mathbb{R}} J(z) \text{ em que } J(z) = \mathcal{E}(y_{k+d} - z | y_k, y_{k-1}, \dots, y_0)^2$$
 (30)

• o valor que minimiza J é denominado *preditor ótimo* e denotado por $\hat{y}(k+d\mid k)$. \hookrightarrow tem-se que:

$$\hat{y}(k+d \mid k) = \min_{z} \int (y_{k+d} - z)^2 p(y_{k+d}, \mathbf{y}) dy_{k+d} d\mathbf{y}$$
 (31)

em que $\mathbf{y}^T = [y_k \ y_{k-1} \dots y_0]$ é a informação disponível no instante k.

PREDITOR ÓTIMO DA SAÍDA.

De modo análogo ao desenvolvido para o estimador de Bayes a solução desse problema resulta no estimador dado por:

$$\hat{\mathbf{y}}(k+d\mid k) = \mathcal{E}(\mathbf{y}_{k+d}\mid \mathbf{y})$$

Este preditor corresponde à melhor previsão que se pode realizar para instantes futuros da saída a partir da informação disponível no instante atual.

PREVISÃO DA SAÍDA EM MODELOS ARMA

Seja o modelo ARMAX com $u_k = 0 \rightsquigarrow$ denominado modelo ARMA, dado por:

$$y_k = \frac{C(q^{-1})}{A(q^{-1})} \omega_k$$
 e $\varepsilon \omega_k^2 = \lambda^2$

PREVISÃO DA SAÍDA EM MODELOS ARMA

A saída no instante k+d é definida por

$$y_{k+d} = \frac{C(q^{-1})}{A(q^{-1})} \omega_{k+d}$$
(32)

- \hookrightarrow a saída y(k+d) depende de instantes da perturbação superiores ao instante atual k.
- \hookrightarrow para o estimador ótimo, separar a saída y(k+d) em duas parcelas:
- A primeira conterá as perturbações futuras, $\{\omega(k+i)\}_{i=1}^d$. Elas são não correlacionadas com as medidas até o instante k.
- A segunda parcela conterá os termos passados da perturbação $\{\omega(k-i)\}_{i=0}^{\infty}$.

PREVISÃO DA SAÍDA EM MODELOS ARMA

Essas duas parcelas são obtidas utilizando-se a seguinte identidade polinomial.

$$\frac{C(q^{-1})}{A(q^{-1})} = F(q^{-1}) + q^{-d} \frac{G(q^{-1})}{A(q^{-1})}$$

com

$$F(q^{-1}) = 1 + f_1 q^{-1} + \dots + f_{d-1} q^{-d+1}$$

$$G(q^{-1}) = g_0 + g_1 q^{-1} + \dots + g_{n-1} q^{-n+1}$$

logo

$$y_{k+d} = F(q^{-1})\omega_{k+d} + \frac{G(q^{-1})}{A(q^{-1})}\omega_k$$

$$= F(q^{-1})\omega_{k+d} + \frac{G(q^{-1})}{A(q^{-1})}\frac{A(q^{-1})}{C(q^{-1})} = F(q^{-1})\omega_{k+d} + \frac{G(q^{-1})}{C(q^{-1})}y_k$$

$$y_{k+d} = \{ \mathbf{\omega}_{k+d} + f_1 \mathbf{\omega}_{k+d-1} + \dots + f_{d-1} \mathbf{\omega}_{k+1} \} + f(y_k, y_{k-1}, \dots) \leftarrow \mathsf{Passado}$$
 (33)

PREVISÃO DA SAÍDA EM MODELOS ARMA

Assim o preditor ótimo $\hat{y}_{k+d} = \mathcal{E}(y_{k+d} | \mathbf{y})$ é dado por

$$\hat{y}_{k+d} = \frac{G(q^{-1})}{C(q^{-1})} y_k$$

OU

$$C(q^{-1})\hat{y}_{k+d} = G(q^{-1})y_k$$

$$\hat{y}_{k+d|k} + c_1\hat{y}_{k+d-1|k-1} + \dots + c_n\hat{y}_{k+d-n|k-n} = g_0y_k + \dots + g_{n-1}y_{k-n+1}$$

→ Erro de Previsão é dado por:

$$e_{k+d} = y_{k+d} - \hat{y}_{k+d} = F(q^{-1})\omega_{k+d}$$
$$= \{\omega_{k+d} + f_1\omega_{k+d-1} + \dots + f_{d-1}\omega_{k+1}\}$$

→ Variância do Erro de Previsão é :

$$var(e_{k+d}) = \lambda^2 \{1 + f_1^2 + \dots + f_{d-1}^2\}$$

→ a variância do erro de previsão cresce à medida que o horizonte de previsão aumenta.

SOLUÇÃO DA IDENTIDADE POLINOMIAL

Da relação adotada,

$$\frac{C(q^{-1})}{A(q^{-1})} = F(q^{-1}) + q^{-d} \frac{G(q^{-1})}{A(q^{-1})}$$

ou

$$C(q^{-1}) = F(q^{-1})A(q^{-1}) + q^{-d}G(q^{-1})$$
 o que implica nas identidades

$$q^{-1}c_{1} = a_{1} + f_{1}$$

$$q^{-2}c_{2} = a_{2} + a_{1}f_{1} + f_{2}$$

$$\vdots$$

$$q^{-d}c_{d} = a_{d} + a_{d-1}f_{1} + \dots + g_{0}$$

$$\vdots$$

$$q^{-n}c_{n} = a_{n} + a_{n-1}f_{1} + \dots + g_{n-d}$$

$$q^{-n+1}0 = a_{n}f_{d-1} + \dots + g_{n-d+1}$$

$$\vdots$$

$$q^{-n+d+1}0 = a_{n}f_{d-1} + g_{n-1}$$

SOLUÇÃO DA IDENTIDADE POLINOMIAL

As identidades podem ser escritas como:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

em que

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_1 & 1 & 0 & \dots & 0 \\ \vdots & & & & & \\ a_{d-1} & \dots & 1 & \dots & 0 \\ \vdots & & & & & \\ a_{n-1} & \dots & 1 & \dots & 0 \\ \vdots & & & & & \\ 0 & \dots & a_n & \dots & 1 \end{bmatrix}$$

$$\mathbf{x}^T = \begin{bmatrix} f_1 & \dots & f_{d-1} & g_o & \dots & g_{n-1} \end{bmatrix}$$

$$\mathbf{b}^T = \begin{bmatrix} c_1 - a_1 & \dots & c_d - a_n & \dots & c_n - a_n & \dots & 0 \end{bmatrix}$$

SOLUÇÃO DA IDENTIDADE POLINOMIAL - MODELO ARMA

Exercício - Obter o preditor de um e dois passos para o seguinte sistema:

$$y_k + 0.7y_{k-1} + 0.1y_{k-2} = \omega_k + 0.4\omega_{k-1} + 0.03\omega_{k-2}$$

■ É uma aplicação do preditor ótimo, na obtenção de um controlador cujo objetivo é regular a saída de um sistema estocástico modelado por um processo ARMAX.

Seja o sistema descrito por:

$$A(q^{-1})y_k = q^{-d}B(q^{-1})u_k + C(q^{-1})\omega_k$$
(34)

OU

$$y_{k+d} = \frac{B(q^{-1})}{A(q^{-1})} u_k + \frac{C(q^{-1})}{A(q^{-1})} \omega_{k+d}$$

• objetivo é determinar a ação de controle u_k que minimiza a variância da saída d instantes de tempo à frente, isto é:

$$\min_{u_k} \mathcal{E} y_{k+d}^2$$

d é o atraso de transporte do processo.

 \rightarrow A solução é do controle u_k minimiza a variância da melhor previsão da saída no instante k+d.

recai-se no problema de se obter o preditor ótimo agora com o modelo do processo dado por (34). Assim de modo similar ao caso do preditor ótimo utiliza-se a seguinte identidade polinomial:

$$\frac{C(q^{-1})}{A(q^{-1})} = F(q^{-1}) + q^{-d} \frac{G(q^{-1})}{A(q^{-1})}$$

logo

$$y_{k+d} = \frac{B(q^{-1})}{A(q^{-1})} u_k + F(q^{-1}) \omega_{k+d} + q^{-d} \frac{G(q^{-1})}{A(q^{-1})} \omega_{k+d}$$
$$= \frac{B(q^{-1})}{A(q^{-1})} u_k + F(q^{-1}) \omega_{k+d} + \frac{G(q^{-1})}{A(q^{-1})} \omega_k$$

mas do modelo tem-se que:

$$\omega_k = \frac{A(q^{-1})}{C(q^{-1})} y_k - \frac{B(q^{-1})}{C(q^{-1})} q^{-d} u_k$$

Assim

$$y_{k+d} = F(q^{-1})\omega_{k+d} + \left\{\frac{B(q^{-1})}{A(q^{-1})} - q^{-d}\frac{B(q^{-1})}{A(q^{-1})}\frac{G(q^{-1})}{C(q^{-1})}\right\}u_k + \frac{G(q^{-1})}{C(q^{-1})}y_k$$

$$= F(q^{-1})\omega_{k+d} + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u_k + \frac{G(q^{-1})}{C(q^{-1})}y_k$$
(35)

Da equação (35) pode-se obter o preditor ótimo da saída.

Sua variância é

$$\begin{split} \mathcal{E}(y_{k+d}^2) &= \mathcal{E}(\lambda F(q^{-1})\omega_{k+d})^2 + \\ &+ 2\mathcal{E}(\lambda F(q^{-1})\omega_{k+d})(\frac{G(q^{-1})}{C(q^{-1})}y_k + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u_k) + \\ &+ \mathcal{E}\{(\frac{G(q^{-1})}{C(q^{-1})}y_k + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u_k)^2\} \end{split}$$

Como $\{\omega_{k+i}\}_{i=1}^d$ são perturbações não correlacionadas com as variáveis do instante k, tem-se que:

$$\mathcal{E}(y_{k+d}^2) = \mathcal{E}(\lambda F(q^{-1})\omega_{k+d})^2 + \mathcal{E}\{\left(\frac{G(q^{-1})}{C(q^{-1})}y_k + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u_k\right)^2\}$$
(36)

Portanto o controle u_k que minimiza a variância da saída é dado por:

$$\frac{G(q^{-1})}{C(q^{-1})}y_k + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u_k = 0$$

OU

$$u_k = -\frac{G(q^{-1})}{B(q^{-1})F(q^{-1})}y_k$$

Exercício - Considere o exercício do preditor de um e dois passos, agora com uma entrada de controle:

$$y_k + 0.7y_{k-1} + 0.1y_{k-2} = u_{k-d} + \omega_k + 0.4\omega_{k-1} + 0.03\omega_{k-2}$$

- 1) Calcule o controle de variância mínima para d = 2 e d = 3.
- 2) Faça algumas simulações, estime e compare as variâncias do saída y(k) do sistema controlado nos dois casos e para o sistema sem controle.
- 3) Perturbe alguns dos parâmetros numéricos: $0.7 \ 0.1 \ 0.4 \ e \ 0.03$ e verifique se a solução inicial continua adequada.