Chapter 5

Estimação de Parâmetros em Sistemas Dinâmicos Descritos por Equações a Diferenças

5.1 Sistemas Monovariáveis

Modelo ARX:

$$A(q^{-1})y_k = B(q^{-1})u_k + \omega_k \tag{5.1}$$

com

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \end{aligned}$$

em que y_k é a saída do processo, u_k é a entrada e ω_k é a perturbação de natureza estocástica. As dimensões dos vetores da saída e entrada do modelo do sistema são n_a e n_b. Da equação (5.1) tem-se que:

$$y_k = -a_1 y_{k-1} - a_2 y_{k-2} \cdots - a_{n_a} y_{k-n_a} + b_1 u_{k-1} \cdots + b_{nb} u_{k-nb} + \omega_k$$

Definindo:

$$\begin{aligned} \varphi_k &= \begin{bmatrix} -y_{k-1} & \cdots & -y_{k-na} & u_{k-1} & \cdots & u_{k-n_b} \end{bmatrix}^T \\ \theta &= \begin{bmatrix} a_1 & \cdots & a_{n_a} & b_1 & \cdots & b_{n_b} \end{bmatrix}^T \end{aligned}$$

com $\varphi_k \in \mathbb{R}^p$, $\theta \in \mathbb{R}^p$ com $p = n_a + n_b$, sendo p o número de parâmetros a ser estimado. Assim,

$$y_k = \varphi_k^T \theta + \omega_k \tag{5.2}$$

Realizando um conjunto de N medidas tem-se que:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ \vdots \\ \varphi_N^T \end{bmatrix} \theta + \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_N \end{bmatrix} \tag{5.3}$$

ou

$$Y_N = \Phi_n \theta + \omega_N$$

com

$$\begin{aligned} Y_N & \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}^T \\ \Phi &= \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ \vdots \\ \varphi_N^T \end{bmatrix} \in \mathbb{R}^{N \times p} \\ \omega_T &= \begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_N \end{bmatrix}^T \end{aligned}$$

5.2 Comportamento do Estimador dos Mínimos Quadrados

Conforme já desenvolvido na disciplina, o estimador dos mínimos quadrados é obtido minimizando a equação abaixo:

$$J = [Y_N - \Phi_N \theta]^T [Y_N - \Phi_N \theta]^T \tag{5.4}$$

Da minimização resulta:

$$\hat{\theta}_{\mathrm{MQ}} = [\Phi_N^T \Phi_N]^{-1} \Phi_N^T Y_N \tag{5.5}$$

Propriedade: O estimador dos Mínimos Quadrados é assintoticamente não polarizado se o ruído ω_k formar uma sequência de variáveis aleatórias não correlatas.

Prova: Seja o modelo

$$\begin{aligned} A(q^{-1}) y_k &= B(q^{-1}) u_k + \omega_k \\ y_k &= \frac{B(q^{-1})}{A(q^{-1})} u_k + \frac{1}{A(q^{-1})} \omega_k \end{aligned} \tag{5.6}$$

Definindo,

$$y_k^F = \frac{B(q^{-1})}{A(q^{-1})}u_k \quad \text{e} \quad v_k = \frac{1}{A(q^{-1})}\omega_k$$

tem-se que

$$y_k = y_k^F + v_k \tag{5.7}$$

Utilizando a construção matricial com

$$Y_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad V_N = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$$

podemos escrever de (5.7) que para um conjunto de N dados vale

$$Y_N = \Phi_N^F \theta + V_N$$

em que a matriz Φ_N foi decomposta em

$$\begin{aligned} &\Phi_N = \begin{bmatrix} -\mathbf{Y}_N & \big| & \mathbf{U}_N \end{bmatrix} = \\ &\underbrace{\begin{bmatrix} -\mathbf{Y}_N^F & \big| & \mathbf{U}_N \end{bmatrix}}_{\Phi_N^F} + \underbrace{\left[\begin{array}{ccc|ccc} -v_1 & \cdots & -v_{1-n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -v_N & \cdots & -v_{N-n} & 0 & \cdots & 0 \end{array}\right]}_{\Phi_N^v} = \Phi_N^F + \Phi_N^v \end{aligned} \tag{5.8}$$

para quando $n = n_a = n_b$ por facilidade, mas vale para qualquer dimensões n_a e n_b. Da representação acima, podemos escrever também que

$$Y_N = \Phi_N^F \theta + V_N = \left[\Phi_N - \Phi_N^v\right]\theta + V_N$$

Da solução do MQ in (5.5) podemos escrever que

$$\begin{aligned} \hat{\theta}_{\mathrm{MQ}} &= [\Phi_N^T \Phi_N]^{-1}\Phi_N^T Y_N = [\Phi_N^T \Phi_N]^{-1}\Phi_N^T(\Phi_N^F\theta + V_N) \\ &= [\Phi_N^T \Phi_N]^{-1}\Phi_N^T\left[(\Phi_N - \Phi_N^v)\theta + V_N\right] \\ &= \theta - [\Phi_N^T \Phi_N]^{-1}\Phi_N^T\left[\Phi_N^v\theta + V_N\right] \end{aligned} \tag{5.9}$$

Tomando-se o valor esperado,

$$\mathbb{E}[\hat{\theta}_N] = \theta - \mathbb{E}\left\{[\Phi_N^T \Phi_N]^{-1}\Phi_N^T\left[\Phi_N^v\theta + V_N\right]\right\}$$

e a polarização p_N do estimador é facilmente identificada como,

$$p_N = \mathbb{E}\left\{[\Phi_N^T \Phi_N]^{-1}\Phi_N^T\left[\Phi_N^v\theta + V_N\right]\right\}$$

Conceitos

- O limite $\lim \hat{\theta}_N$ quando $N \to \infty$ é entendido como um limite em probabilidade e indicado por
$$\mathbf{P}-\lim_{N\to\infty} \hat{\theta}_N$$
significa que $\limsup_{N\to\infty} \mathbf{P}(|\hat{\theta}_N - \bar{\theta}| > \epsilon) = 0$, em que $\bar{\theta}$ é a v.a. limite. Queremos evidentemente que $\bar{\theta} = \theta$.

- Se as matrizes A_N e B_N convergem em probabilidade,
$$\mathbf{P}-\lim_{N\to\infty} (A_N)^{-1} B_N = \left(\mathbf{P}-\lim_{N\to\infty} A_N\right)^{-1} \mathbf{P}-\lim_{N\to\infty} B_N$$

Utilizando as propriedades acima,

$$\mathbf{P}-\lim_{N\to\infty} p = \left[\mathbf{P}-\lim_{N\to\infty} \frac{\Phi_N^T \Phi_N}{N}\right]^{-1} \underbrace{\left[\mathbf{P}-\lim_{N\to\infty} \frac{\Phi_N^T \Phi_N^v}{N} - \mathbf{P}-\lim_{N\to\infty} \frac{\Phi_N^T V_N}{N}\right]}_{\mathcal{R}}$$

Para que o estimador MQ de θ seja não polarizado é necessário que $\mathcal{R} \equiv 0$.

Analisando os termos de $\mathcal{R}$:

1. Da representação de Φ_N^v em (5.8), temos que
$$\Phi_N^T \Phi_N^v = \left[\begin{array}{c:c} \mathbf{C}_N & 0 \\ \hdashline 0 & 0 \end{array}\right]$$
e portanto,
$$\mathbf{P}-\lim_{N\to\infty} \frac{\Phi_N^T \Phi_N^v}{N} = \left[\begin{array}{c:c} \bar{\mathbf{C}} & 0 \\ \hdashline 0 & 0 \end{array}\right]$$
e $\mathbf{C}_N$ e $\bar{\mathbf{C}}$ são matrizes de covariância com
$$\bar{\mathbf{C}} = \mathrm{Cov}(v_k, v_{k+i})_{i=1,\dots n} = \begin{bmatrix} \mathbf{C}_{vv}(0) & \cdots & \mathbf{C}_{vv}(n-1) \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{vv}(n-1) & \cdots & \mathbf{C}_{vv}(0) \end{bmatrix}, \quad (5.10)$$
supostas aqui definidas positivas.

2. Note que
$$\Phi_N^T V_N = -\begin{bmatrix} \mathbf{C}_N(1,1) \\ \vdots \\ \mathbf{C}_N(n,1) \end{bmatrix}$$

Assim observamos que a polarização p_N no limite é expressa como,

$$\mathbf{P}-\lim_{N\to\infty} p_N = \bar{M}^{-1}\left(\left[\begin{array}{c:c} \bar{\mathbf{C}} & 0 \\ \hdashline 0 & 0 \end{array}\right]\theta - \begin{bmatrix} -\mathbf{C}_{vv}(1,1) \\ \vdots \\ -\mathbf{C}_{vv}(n,1) \end{bmatrix}\right)$$

Assim por inspeção se e somente se,

$$\bar{\mathbf{C}}\,a = \begin{bmatrix} -\mathbf{C}_{vv}(1,1) \\ \vdots \\ -\mathbf{C}_{vv}(n,1) \end{bmatrix} a = -\begin{bmatrix} \mathbf{C}_{vv}(0) \\ \vdots \\ \mathbf{C}_{vv}(n) \end{bmatrix} a, \text{ em que } a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

o estimador será ASSINTOTICAMENTE NÃO POLARIZADO!

Essa condição só ocorrerá quando a perturbação ω_k atuando no processo for branca. Neste caso o modelo de perturbação v_k é dado por

$$v_k = \frac{1}{A(q^{-1})}\omega_k$$

ou

$$v_k = -a_1 v_{k-1} \cdots - a_n + \omega_k$$

Dessa forma, tem-se que

$$\mathbf{C}_{vv}(1) = -a_1\mathbf{C}(0) \cdots - a_n\mathbf{C}_{vv}(n-1) \Longleftrightarrow \bar{\mathbf{C}}\,a == -\begin{bmatrix} \mathbf{C}_{vv}(0) \\ \vdots \\ \mathbf{C}_{vv}(n) \end{bmatrix} a$$

e esta é a única situação em que o estimador de MQ será não polarizado, ou seja somente quando

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

e $\omega_k, k = 1, 2, \ldots$ é uma sequência de ruido branco.

Essa análise confirma e conclui que a única forma possível para que a estimativa paramétrica utilizando o estimador MQ será quando o sistema estocástico admitir um modelo ARX ou AR.

5.5 Método de Variável Instrumental

Um método global que fornece estimadores assintoticamente não polarizados quando a perturbação atuando no processo não pode ser modelada por um ruído branco.

- Idéia - Gerar um novo sinal (variável instrumental) que é correlacionado com os sinais do processo, mas não é correlacionado com a perturbação.

Seja o sistema modelado por:

$$Y_N = \Phi_N \theta + e_N$$

Definindo como Z_N a variável instrumental, tal que:

$$\mathcal{E}\left\{Z_N^T e_N\right\} = 0 \qquad e \qquad \mathcal{E}\left\{Z_N^T \Phi_N\right\} = \psi$$

com ψ uma matriz não singular. Multiplicando-se o modelo do sistema pela variável instrumental resulta:

$$Z_N^T Y_N = Z_N^T \Phi_N \theta + Z_N^T e_N$$

e

$$\theta = \left[Z_N^T \Phi_N\right]^{-1} Z_N^T Y_N - \left[Z_N^T \Phi_N\right]^{-1} Z_N^T e_N$$

Define-se o estimador variável instrumental como:

$$\hat{\theta}_{VI} = \left[Z_N^T \Phi_N\right]^{-1} Z_N^T Y_N$$

Propriedade: O estimador variável instrumental é assintoticamente não polarizado

Da equação do Estimador Variável Instrumental pode-se escrever que:

$$\begin{aligned}
\lim_{N\to\infty} \mathcal{E}\hat{\theta}_{VI} &= P \lim_{N\to\infty} \hat{\theta}_{VI} \\
&= P \lim_{N\to\infty} \left[Z_N^T \Phi_N\right]^{-1} \left[Z_N^T \Phi_N\right]\theta + \left[P \lim_{N\to\infty} Z_N^T \Phi_N\right]^{-1} \left[P \lim_{N\to\infty} Z_N^T e_N\right] \\
\lim_{N\to\infty} \mathcal{E}\hat{\theta}_{VI} &= \theta + \psi .0 = \theta
\end{aligned}$$

Resta o problema prático da escolha de escolha de Z_k ou Z_N. Escolhas comumente utilizada na literatura são:

$$\begin{aligned}
\varphi_k^T &= \begin{bmatrix} -y_{k-1} & \ldots & -y_{k-n_a} & u_{k-1} & \ldots & u_{k-n_b} \end{bmatrix} \\
z_k^T &= \begin{bmatrix} -y^*_{k-1} & \ldots & -y^*_{k-n_a} & u_{k-1} & \ldots & u_{k-n_b} \end{bmatrix}
\end{aligned}$$

onde y_j^* são obtidos utilizando o seguinte modelo auxiliar

$$\hat{A}(q^{-1}) y_k^* = \hat{B}(q^{-1}) u_k$$

Deve-se observar que neste caso não se estima os parâmetros do modelo do ruído.

Algorithm 1 Algoritmo da Variável Instrumental

1: $\hat{\theta}_{VI} = \hat{\theta}_{MQ}$
2: Calcular y_j^* por $\hat{A}(q^{-1})y_k^* = \hat{B}(q^{-1})u_k$
3: Obter z_k e Z_N
4: Calcular $\hat{\theta}_{VI}$
5: Teste de parada

Convergência?

5.5.1 Versão Recursiva

Não é análoga à versão global como no estimador MQ. Neste caso se interrompe o processo iterativo. Dedução análoga ao estimador dos MQR utilizando o lema da inversão.

$$\begin{aligned}
\hat{\theta}_{k+1} &= \hat{\theta}_k + K_{k+1}\Big(y_{k+1} - \varphi_{k+1}^T\hat{\theta}_k\Big) \\
K_{k+1} &= \frac{P_k z_{k+1}}{1 + \varphi_{k+1}^T P_k z_{k+1}} \\
P_{k+1} &= P_k - \frac{P_k z_{k+1}\varphi_{k+1}^T P_k}{1 + \varphi_{k+1}^T P_k z_{k+1}} \\
P_{k+1} &= \Big[Z_{k+1}^T \Phi_{k+1}\Big]^{-1} \\
y_k^* &= z_k^T \hat{\theta}_k \\
z_k^T &= \begin{bmatrix} -y_{k-1}^* & \dots & -y_{k-n_a}^* & u_{k-1} & \dots & u_{k-n_b} \end{bmatrix}
\end{aligned}$$

Não é igual ao VI em lote, como ocorre com o MQ e MQ recursivo

- A matriz P_k não é Definida Positiva. Este problema pode ser contornado utilizando-se a Variável Instrumental Simétrica.

- Variável Instrumental Simétrica

$$\begin{aligned}
P_k &= [Z_k Z_k^T]^{-1} \\
P_{k+1} &= P_k - \frac{P_k z_{k+1} z_{k+1}^T P_k}{1 + z_{k+1}^T P_k z_{k+1}}
\end{aligned}$$

5.6 Estimador dos Mínimos Quadrados Generalizados

Seja o sistema descrito por:

$$A(q^{-1})y_k = B(q^{-1})u_k + v_k$$

Este método utiliza um modelo auto-regressivo para a modelagem do ruído v_k

$$D(q^{-1})v_k = e_k \qquad \text{onde} \quad e_k : \text{ruído branco}$$

Neste caso a função de transferência $D(q^{-1})$ deverá conter a informação sobre a matriz de covariância da perturbação.
Utilizando o modelo da perturbação no modelo do sistema resulta que:

$$A(q^{-1})y_k = B(q^{-1})u_k + D^{-1}(q^{-1})e_k$$

ou

$$D(q^{-1})A(q^{-1})y_k = D(q^{-1})B(q^{-1})u_k + e_k \tag{5.14}$$

Definindo

$$y_k^F = D(q^{-1})y_k \qquad e \qquad u_k^F = D(q^{-1})u_k$$

a equação (5.14) torna-se:

$$A(q^{-1})y_k^F = B(q^{-1})u_k^F + e_k \tag{5.15}$$

Como e_k é um ruído branco pode-se obterum estimador assintóticamente não polarizado para os parâmetros estimados $\hat{A}$ e $\hat{B}$ utilizando-se o estimador dos mínimos quadrados na euqação acima.Resta o problema de se estimar os parâmetros de $D(q^{-1})$ que são desconhecidos.

Problema:$D(q^{-1})$ é desconhecido.

Vamos estimá-lo a partir dos valores estimados $\hat{A}(q^{-1})$ e $\hat{B}(q^{-1})$:

$$\hat{v}_k = \hat{A}(q^{-1})y_k - \hat{B}(q^{-1})u_k$$

Obtém-se o estimador $D(q^{-1})$ utilizando o método dos mínimos quadrados no modelo

$$D(q^{-1})\hat{v}_k = e_k$$

Algorithm 2 Algoritmo do Estimador dos M. Q. Generalizado

1: $\hat{D}(q^{-1}) = 1$
2: Calcular $y_k^F = \hat{D}(q^{-1})y_k$ e $u_k^F = \hat{D}(q^{-1})u_k$
3: Obter estimador MQ de A e B
4: Calcular $\hat{v}_k = \hat{A}(q^{-1})y_k - \hat{B}(q^{-1})u_k$
5: Obter estimador MQ de $D(q^{-1})$
6: Teste de convergência. Fim ou volta ao passo 2

Convergência?

5.7 Algoritmo da Matriz Estendida

Este algoritmo foi desenvolvido, na literatura, utilizando um modelo média-móvel para a perturbação v_k

$$\begin{aligned} A(q^{-1})y_k &= B(q^{-1})u_k + v_k \\ v_k &= C(q^{-1})e_k, \qquad e_k \text{ branco} \end{aligned} \tag{5.16}$$

O algoritmo é denominado matriz estendida porque se constrói um novo vetor φ_k como extensão do vetor φ_k utilizado no estimador dos mínimos quadrados.

Utilizando o modelo (5.16) pode-se escrever que:

$$A(q^{-1})y_k = B(q^{-1})u_k + C(q^{-1})e_k$$

$$y_k = \varphi_k^T\theta + e_k \tag{5.17}$$

com

$$\begin{aligned} \varphi_k^T &= \begin{bmatrix} -y_{k-1} & \dots & -y_{k-n_a} & u_{k-1} & \dots & u_{k-n_b} & e_{k-1} & \dots & e_{k-n_c} \end{bmatrix} \\ \theta^T &= \begin{bmatrix} a_1 & \dots & a_{n_a} & b_1 & \dots & b_{n_b} & c_1 & \dots & c_{n_c} \end{bmatrix} \end{aligned}$$

Assumindo que e_k é mensurável, a equação (5.17) é equivalente à utilizada para se obter os estimador dos mínimos quadrados.

O estimador recursivo é dado por:

$$\begin{aligned} \hat{\theta}_{k+1} &= \hat{\theta}_k + K_{k+1}\Big(y_{k+1} - \varphi_{k+1}^T\hat{\theta}_k\Big) \\ K_{k+1} &= \frac{P_k\varphi_{k+1}}{1 + \varphi_{k+1}^T P_k \varphi_{k+1}} \\ P_{k+1} &= P_k - \frac{P_k\varphi_{k+1}\varphi_{k+1}^T P_k}{1 + \varphi_{k+1}^T P_k \varphi_{k+1}} \end{aligned}$$

Como a seqüência e_k não é mensurável, ela é estimada por

$$\hat{e}_k = y_k - \varphi_k^T \hat{\theta}_{k-1}$$

5.8 Estimador de Máxima Verossimilhança Aproximado

O estimador dos mínimos quadrados pode ser obtido minimizando o seguinte critério:

$$J_k = \frac{1}{2} \sum_{t=1}^{K} (y_t - \hat{y}_t)^T (y_t - \hat{y}_t) \tag{5.18}$$

onde $y_t - \hat{y}_t = \omega_t$ é o erro previsto obtido a partir do modelo do processo.

Seja o modelo dado por:

$$A(q^{-1}) y_k = B(q^{-1}) u_k + C(q^{-1}) e_k$$

Neste caso o erro previsto é dado por:

$$\omega_k = \hat{C}^{-1}(q^{-1}) \left\{ \hat{A}(q^{-1}) y_k - \hat{B}(q^{-1}) u_k \right\}$$

Se $\hat{A} \to A$, $\hat{B} \to B$ e $\hat{C} \to C$, então $\omega_k \to e_k$.

O critério (5.18) torna-se:

$$J_k = \frac{1}{2} \sum_{t=1}^{K} \omega_t^T \omega_t \tag{5.19}$$

O algoritmo de máxima verossimilhança recursivo é obtido a partir das seguintes hipóteses simplificadoras:

a. $\hat{\theta}_k$ é um ponto de mínimo de J_k;

b. Aproximação de segunda ordem do critério J_{k+1}, em torno do último valor estimado $\hat{\theta}_k$.

$$J_{k+1}(\theta) = J_{k+1}(\hat{\theta}_k) + J'_{k+1}(\hat{\theta}_k)(\theta - \hat{\theta}_k) + \frac{1}{2}(\theta - \hat{\theta}_k)^T J''^{T}_{k+1}(\hat{\theta}_k)(\theta - \hat{\theta}_k)$$

onde $J'(.)$ e $J''(.)$ são respectivamente, as derivadas de 1.ª e 2.ª ordem do critério dado pela equação (5.18).

Minimizando o critério $J_{k+1}(\theta)$ obtém-se o estimador no instante $k+1$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + J''^{-1}_{k+1}(\hat{\theta}_k)\Big[-J'_{k+1}(\hat{\theta}_k)\Big]$$

esta equação corresponde à 1.ª iteração do método de Newton-Raphson.

Cálculo recursivo de $J'(.)$ e $J''(.)$.

Da equação (5.19) tem-se que:

$$J_{k+1}(\theta) = J_k(\theta) + \frac{1}{2}\omega_{k+1}(\theta)\omega_{k+1}(\theta) \tag{5.20}$$

e

$$J'_{k+1}(\theta) = J'_k(\theta) + \omega'_{k+1}\omega_{k+1}$$

Como $\theta = \hat{\theta}_k$ é um ponto de mínimo de $J_k(.)$, tem-se que:

linearização do erro previsto com respeito aos parâmetros θ

$$J'_{k+1}(\hat{\theta}_k) = \omega'_{k+1}(\hat{\theta}_k)\omega_{k+1}(\hat{\theta}_k)$$

$$J''_{k+1}(\hat{\theta}_k) = J''_k(\hat{\theta}_k) + \omega'_{k+1}(\hat{\theta}_k)\omega'^T_{k+1}(\hat{\theta}_k) + \text{outros termos}$$

Supondo $J''_k(\hat{\theta}_k) = J''_k(\hat{\theta}_{k+1})$, pode-se escrever que:

$$J''_{k+1}(\hat{\theta}_k) = J''_k(\hat{\theta}_{k-1}) + \omega'_{k+1}(\hat{\theta}_k)\omega'^T_{k+1}(\hat{\theta}_k)$$

Definindo:

$$\begin{aligned} z_{k+1} &= \omega'_{k+1}(\hat{\theta}_k) \\ P_{k+1} &= {J''_{k+1}}^{-1}(\hat{\theta}_k) \end{aligned}$$

tem-se

$$\begin{aligned} \hat{\theta}_{k+1} &= \hat{\theta}_k + K_{k+1}\omega_{k+1}(\hat{\theta}_k) \\ K_{k+1} &= P_{k+1}\Big(-z_{k+1}\Big) \\ P_{k+1} &= P_k - \frac{P_k z_{k+1} z^T_{k+1} P_k}{1 + z^T_{k+1} P_k z_{k+1}} \end{aligned}$$

vide eq. Newton-Raphson

com

$$\begin{array}{rcl} \omega_{k+1} & = & y_k - \varphi_k^T \hat{\theta}_k \\ \varphi_k^T & = & \left[\begin{array}{ccccccccc} -y_{k-1} & \dots & -y_{k-1-n_a} & u_{k-1} & \dots & u_{k-1-n_b} & \hat{\omega}_{k-1} & \dots & \hat{\omega}_{k-1-n_c} \end{array} \right] \\ C(q^{-1})\omega_k & = & A(q^{-1})y_k - B(q^{-1})u_k \\ C(q^{-1})\omega_k & = & y_k - \varphi_k^T \hat{\theta}_k \end{array}$$

logo

$$C(q^{-1})z_k = -\varphi_k$$

e

$$K_{k+1} = \frac{P_{k+1}\varphi_{k+1}}{C(q^{-1})}$$

Obs.: As hipóteses simplificadoras deste algoritmo são válidas após o transitório. Portanto este algoritmo é inicializado com os resultados do estimador da matriz estendida.

Lista no. 4 Identificação Paramétrica

5.9 Algoritmos de Identificação Adaptativa

Quando o algoritmo dos mínimos quadrados converge os elementos da matriz de covariância $P(k)$ e os ganhos de correção tendem a zero. Quando o processo que está sendo identificado é invariante no tempo esta convergência é desejada.

Quando se deseja estimar sistemas com parâmetros variantes no tempo, se os ganhos do estimador decrescem significativamente, o algoritmo perde sua capacidade de adaptação em relação às variações de parâmetros.

Pode-se manter a capacidade de adaptação do estimador limitando-se os valores mínimos dos elementos da matriz de covariância $P(k)$.

A equação de estimador pode ser escrita como:

$$\hat{\theta}_k = \hat{\theta}_{k-1} + K_k \omega_k \tag{5.21}$$

$$\omega_k = y_k - \varphi_k^T \hat{\theta}_{k-1} \tag{5.22}$$

Quando os parâmetros do sistema variam no tempo, a saída do processo desvia-se da saída do modelo, aumentando o erro previsto (eq. 5.22).

Limitando-se os elementos da matriz P_k, o ganho de correção também será limitado ($K_k = P_k \varphi_k$) e o erro previsto corrigirá continuamente $\hat{\theta}_k$ na equação (5.22). Contudo, o erro previsto também inclui perturbações estocásticas que se deseja eliminar. Evitando-se que os elementos da matriz $P(k)$ diminuam, aumentam-se os erros aleatórios no estimador $\hat{\theta}_k$. Portanto tem-se um compromisso entre a qualidade do estimador (P_k "pequeno") e a sua capacidade de adaptação (P_k "grande").

5.9.1 Algoritmos para rastrear variações lentas no tempo

Para processos que variam lentamente no tempo, é necessário que o algoritmo dos mínimos quadrados ponderado recursivo tenha um mínimo de capacidade de adaptação, para impedir que o ganho tenda para zero. Esta capacidade pode ser obtida dando-se uma maior importância às novas medidas (fator de esquecimento) ou através da adição de uma matriz semi-definida positiva à matriz $P(k)$ (busca aleatória).

Regulagem do fator de esquecimento
Traço Limitado

$$\tilde{P}_k = P_{k-1} - \frac{P_{k-1}\varphi_k\varphi_k^T P_{k-1}}{\lambda_{k+1} + \varphi_k^T P_{k-1}\varphi_k} \tag{5.23}$$

Calcular o traço de $\tilde{P}_k$

- Se o traço$(\tilde{P}_k) < tr_0$

$$\lambda_k = \frac{\text{traço}(\tilde{P}_k)}{tr_0}$$

  e

$$P_k = \frac{\tilde{P}_k}{\lambda_k}$$

- Se o traço$(\tilde{P}_k) \geq tr_0$

$$\begin{aligned} \Rightarrow \lambda_k &= 1 \\ P_k &= \tilde{P}_k \end{aligned}$$

Fator de esquecimento direcional

No fator de esquecimento tem-se que:

$$P^{-1}(k) = \lambda P_{k-1}^{-1} + \varphi_k \varphi_k^T$$

Neste caso pondera-se igualmente a informação em todas as direções.
O fator de esquecimento direcional considera

$$P_k^{-1} = P_{k-1}^{-1} + r_{k-1} \varphi_k \varphi_k^T$$

Neste caso pondera-se a informação associada com a direção φ_k

$$P_k = P_{k-1} \left[\ I_m - \frac{\varphi_k \varphi_k^T P_{k-1}}{r_{k+1}^{-1} + \varphi_k^T P_{k-1} \varphi_k} \ \right]$$

com

$$r_{k-1} = \lambda' - \frac{1 - \lambda'}{\varphi_k^T P_{k-1} \varphi_k}$$

se $\lambda' = 1$ $r_{k-1} = 1$

Busca Aleatória

Adiciona-se à matriz P_k uma matriz constante:

$$P_k = \tilde{P}_k + Q_k \tag{5.24}$$

onde $\tilde{P}_k$ é dado pela equação (5.23) com $\lambda_{k-1} = 1$

5.9.2 Detecção de não-estacionariedade

Quando ocorrem variações bruscas nos parâmetros dos sistemas é necessário aumentar a capacidade de adaptação do algoritmo de identificação, para que os novos valores dos estimadores sejam atingidos rapidamente.

Esta condição é obtida aumentando-se bruscamente o ganho de correção K_k.

Problema: Determinação do instante de variação dos parâmetros. Determinação de não-estaciona-riedade.

Critério: Por exemplo compara-se aa variância do erro previsto com a variância da perturbação.

Seja o sistema:

$$y_k = \varphi_k^T \theta + e_k$$

O erro previsto é dado por:

$$\omega_k(\hat{\theta}) = y_k - \varphi_k^T \hat{\theta}_{k-1}$$

$$\omega_k(\hat{\theta}) = \varphi_k^T(\theta - \hat{\theta}_{k-1}) + e_k \tag{5.25}$$

Se $\hat{\theta}_{k-1} \to \theta \Rightarrow \epsilon_k \to v_k$.

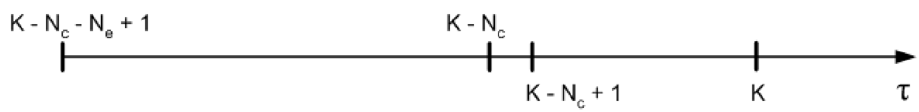
Não-estacionariedade $\Rightarrow \omega_k$ cresce.

Utilizar informação contida na variância (módulo) do erro previsto e da perturbação para detectar uma não-estacionariedade. Contudo, os valores destas variâncias não são conhecidos logo é necessário obter suas estimativas.

Estimativa da variância do erro previsto

$$\sigma_c^2(k) = \frac{1}{N_c} \sum_{\tau=k-N_c+1}^{K} \omega^2(\tau, \hat{\theta})$$

N_c - Horizonte curto



Estimativa da variância da Perturbação

Hipótese: $\tau \leq K - N_c$ sistema em regime

$$\sigma_e^2(k) = \frac{1}{N_e} \sum_{\tau=K-N_c-N_e+1}^{K} \omega^2(\tau, \hat{\theta})$$

N_e - Horizonte longo

Traço Adaptativo

1 - Calcular σ_c^2, σ_e^2

$$J_k = \frac{\sigma_c^2}{\sigma_e^2}$$

2 - Decisão

a. $0 < J_k < J_{min}$

$$tr_0 = tr_{min}$$

b. $J_{min} < J_k < J_{med}$

$$tr_0 = tr_{0int}$$

c. $J_k > J_{max}$

$$tr_0 = tr_{max}$$

3 - Estimador
Algoritmo traço limitado com tr_0 calculado em 2.

5.10 Determinação da ordem do modelo

Seja

$$\varphi^T(k) = \left[\begin{array}{cccccc} u(k-1) & y(k-1) & \dots & u(k-m) & y(k-m) \end{array} \right]$$

e

$$Q(m) = \frac{1}{N} \sum_{k=1}^{N} \varphi(k)\varphi^T(k)$$

Quando a ordem do modelo é maior do que a ordem do sistema $\Rightarrow$ última linha é l. d. $\Rightarrow \det Q = 0$

Analisar

$$DR(m) = \frac{\det Q(m)}{\det Q(m+1)}$$

Quando $DR(m) \uparrow \rightarrow$ ordem é m.
Caso estocástico utiliza variável instrumental.

5.11 Aspectos Computacionais

Estimador dos Mínimos Quadrados é obtido baseado no Lema da Inversão de Matriz. Esta inversão pode ser numericamente mal condicionada devido à subtração de matrizes

$\rightarrow$ Matriz SDN
$\rightarrow$ Divergência do Estimador
Solução: Utilizar a propriedade de que P é simétrica $\Rightarrow$ Fatorar P
$P(k) = S(k)S^T(k)$ - S é a raiz quadrada de P

Fatorização U-D de Bierman → evita o cálculo da raiz quadrada

$$
\begin{aligned}
P(k) &= U(k)D(k)U^T(k) \\
U(k) &= \begin{bmatrix} 1 & u_{12}(k) & \dots & u_{1m}(k) \\ 0 & 1 & \dots & u_{1m}(k) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \\
D(k) &= \begin{bmatrix} d_1(k) & 0 & \dots & 0 \\ 0 & d_1(k) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_m(k) \end{bmatrix}
\end{aligned}
$$

onde m é o número de parâmetros desconhecidos.

Algoritmo U-D
Calcular $D(k)$ e $U(k)$ a partir de $D(k-1)$ e $U(k-1)$.

5.11.1 Algoritmos Rápidos

A fatorização de Bierman tem aproximadamente a mesma eficácia computacional do que o método dos M.Q.

Algoritmo rápido explora o fato de que alguma informação requerida no instante k já está também disponível no instante $k-1$

Processo AR

$$
\varphi(k) = \begin{bmatrix} y(k-1) \\ \vdots \\ y(k-n) \end{bmatrix}
$$

$$
\varphi^*(k) = \begin{bmatrix} y(k) \\ \varphi(k) \end{bmatrix} = \begin{bmatrix} \varphi(k+1) \\ y(k-n) \end{bmatrix}
$$

Utilizando esta condição é possível obter diretamente o ganho do estimador. O número de operações reduz de $O(m^2)$ para $O(m)$.