ID9

ARAR(X) Identification

9.5 PEM ESTIMATION OF ARARX MODELS

The parameters of ARARX models can also be estimated minimizing the cost function $J(\theta)$ by means of the Gauss–Newton algorithm; this PEM estimate is also a maximum likelihood estimate when $w(t)$ is Gaussian. Following this way we have the advantage of a simultaneous estimation of all parameters and of the minimization of a well defined cost function on the set of available sequences; the drawback consists in the necessity of using an iterative implementation, subject to possible convergence problems. Assuming for simplicity $n_\delta = n$, the implementation of the Gauss–Newton algorithmm for the parameter vector

$$\theta^k = \left[\, \alpha_1^k \,\dots\, \alpha_n^k \;\beta_1^k \,\dots\, \beta_n^k \;\delta_1^k \,\dots\, \delta_n^k \,\right]^T , \tag{9.5.1}$$

requires computing the gradient (6.13.3) of $\varepsilon(t,\theta)$. Deriving both members of (9.3.4) we obtain

$$\frac{\partial \varepsilon(t,\theta)}{\partial \alpha_i} = -z^{-(n-i+1)} s(z^{-1})\, y(t) = -s(z^{-1})\, y(t+i-n-1) \tag{9.5.2}$$

$$\frac{\partial \varepsilon(t,\theta)}{\partial \beta_i} = -z^{-(n-i+1)} s(z^{-1})\, u(t) = -s(z^{-1})\, u(t+i-n-1) \tag{9.5.3}$$

$$\begin{aligned} \frac{\partial \varepsilon(t,\theta)}{\partial \delta_i} &= -z^{-(n-i+1)} \big(q(z^{-1})\, y(t) - p(z^{-1})\, u(t)\big) \\ &= -q(z^{-1})\, y(t+i-n-1) + p(z^{-1})\, u(t+i-n-1) \\ &= -e(t+i-n-1). \end{aligned} \tag{9.5.4}$$

By introducing the notations

$$y^F(t) = s(z^{-1})\, y(t) \tag{9.5.5}$$

$$u^F(t) = s(z^{-1})\, u(t)\,, \tag{9.5.6}$$

we obtain the following expression for $\psi(t,\theta)$ (6.13.3)

$$\psi(t,\theta) = \left[\, y^F(t-n) \;\ldots\; y^F(t-1) \;\; u^F(t-n) \;\ldots\; u^F(t-1) \;\; e(t-n) \;\ldots\; e(t-1) \,\right]^T . \tag{9.5.7}$$

Denoting with L the length of the filtered sequences, the iterative algorithm (6.13.16)

$$\theta^{k+1} = \theta^k + \left(H_\psi^T H_\psi\right)^{-1} H_\psi^T \varepsilon^\circ \tag{9.5.8}$$

can be implemented by constructing, at every step, the matrix

$$H_\psi = \tag{9.5.9}$$

$$\begin{bmatrix} y^F(1) & \ldots & y^F(n) & u^F(1) & \ldots & u^F(n) & e(1) & \ldots & e(n) \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ y^F(L-n) & \ldots & y^F(L-1) & u^F(L-n) & \ldots & u^F(L-1) & e(L-n) & \ldots & e(L-1) \end{bmatrix}$$

and the vector ε° (6.13.18)

$$\varepsilon^\circ = \left[\, \varepsilon(n+1) \;\ldots\; \varepsilon(L) \,\right]^T . \tag{9.5.10}$$

It can also be useful to use, instead of (9.5.8), expressions of the type (6.13.19).

Remark 9.5.1 – The extension of (9.5.9) to the general case $n_\delta \neq n$ is straightforward and leads to the expression

$$H_\psi = \tag{9.5.11}$$

$$\begin{bmatrix} y^F(1) & \ldots & y^F(n) & u^F(1) & \ldots & u^F(n) & e(1+n-n_\delta) & \ldots & e(n) \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ y^F(L-n) & \ldots & y^F(L-1) & u^F(L-n) & \ldots & u^F(L-1) & e(L-n_\delta) & \ldots & e(L-1) \end{bmatrix}$$