ID3

ARX Identification

3.7 WEIGHTED LEAST SQUARES

Consider expression (3.6.10) of recursive least squares and relation (3.6.11) linking the gain matrix $K(t)$ to $S(t)^{-1}$; for large values of t it can be expected that $S(t)$ will diverge and, consequently, that its inverse and $K(t)$ will tend to null matrices. In a context of this kind the estimate $\theta^\circ(t)$ is no longer influenced by the innovation $\varepsilon(t)$ even if large. This behavior of the algorithm, due to the same weight attributed to all samples, is disadvantageous when the process to be modeled is nonstationary and the identification algorithm must track parameter variations. This can be obtained by introducing a forgetting factor determining a reduction of the influence of past observations. We can substitute, for this purpose, cost function (3.3.2) with the following

$$J(\theta) = \frac{1}{N} \sum_{t=n+1}^{L} \kappa^{L-t} \big(y(t) - y(t|t-1) \big)^2 = \frac{\varepsilon^T(\theta)\, W \varepsilon(\theta)}{N} = \frac{1}{N} \sum_{t=n+1}^{L} \kappa^{L-t} \varepsilon^2(t) \tag{3.7.1}$$

where

$$W = \mathrm{diag}\big[\, \kappa^{N-1}\ \kappa^{N-2}\ \ldots\ \kappa\ \ 1 \,\big] \tag{3.7.2}$$

and κ is a forgetting factor with values between 0 and 1. The weight of the prevision error at time t, $\varepsilon(t)$ is now equal to 1, the weight of $\varepsilon(t-1)$ is κ, that of $\varepsilon(t-2)$ is κ^2 and so on. The decreasing weights given to less recent data determines, according to the actual value of κ, their influence on the parameter estimate whose expression can be obtained by annihilating the gradient of the cost function, $\partial J(\theta)/\partial\theta$, given by

$$\begin{aligned} \frac{\partial J(\theta)}{\partial\theta} &= \frac{\partial}{\partial\theta} \left[\frac{\varepsilon^T(\theta)\, W \varepsilon(\theta)}{N} \right] = \frac{1}{N} \frac{\partial}{\partial\theta} \big[(H\theta - y^\circ)^T\, W (H\theta - y^\circ) \big] \\ &= \frac{1}{N} \frac{\partial}{\partial\theta} \big[\theta^T H^T W H\theta - \theta^T H^T W y^\circ - y^{\circ T} W H\theta + y^{\circ T} W y^\circ \big] \\ &= \frac{2}{N} \big(H^T W H\theta - H^T W y^\circ \big). \end{aligned} \tag{3.7.3}$$

Equating to zero expression (3.7.3) we obtain, under the assumption of invertibility for $H^T W H$, the formula of weighted least squares

$$\theta^\circ = \left(H^T W H\right)^{-1} H^T W y^\circ. \tag{3.7.4}$$

It is possible to observe that (3.7.4) coincides with (3.3.12) when the weighting factor is equal to 1. The recursive implementation of the algorithm will start from an initial estimate obtained by means of (3.7.4) and will use (3.6.10) and (3.6.14), which remain unchanged, to update the estimate; matrix $R(t)$, instead, is given by

$$R(t) = \frac{H^T(t)\, W H(t)}{N} = \kappa \left(\frac{N-1}{N}\right) R(t-1) + \frac{h^T(t)\, h(t)}{N} \tag{3.7.5}$$

while the update of $R(t)^{-1}$ is given by

$$R(t)^{-1} = \frac{N\, R(t-1)^{-1}}{\kappa\,(N-1)} \left[I - \frac{h^T(t)\, h(t)\, R(t-1)^{-1}}{\kappa\,(N-1) + h(t)\, R(t-1)^{-1} h^T(t)} \right]. \tag{3.7.6}$$

Relation (3.7.6) coincides, obviously, with (3.6.16) when $\kappa = 1$.

SECTIONS | MODULES | QUESTIONS | HOME PAGE
PREV. MODULE | FAQ | TUTOR | NEXT MODULE