The Basic Algorithm

A typical recursive identification algorithm is

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(y(t) - \hat{y}(t)) \qquad \textbf{(3-55)}$$

Here $\hat{\theta}(t)$ is the parameter estimate at time t, and $y(t)$ is the observed output at time *t*. Moreover, $\hat{y}(t)$ is a prediction of the value $y(t)$ based on observations up to time $t-1$ and also based on the current model (and possibly also earlier ones) at time $t-1$. The gain $K(t)$ determines in what way the current prediction error $y(t) - \hat{y}(t)$ affects the update of the parameter estimate. It is typically chosen as

$$K(t) = Q(t)\psi(t) \qquad \textbf{(3-56)}$$

where $\psi(t)$ is (an approximation of) the gradient with respect to θ of $\hat{y}(t|\theta)$. The latter symbol is the prediction of $y(t)$ according the model described by θ. Note that model structures like AR and ARX that correspond to linear regressions can be written as

$$y(t) = \psi^T(t)\theta_0(t) + e(t) \qquad \textbf{(3-57)}$$

where the *regression vector* $\psi(t)$ contains old values of observed inputs and outputs, and $\theta_0(t)$ represents the true description of the system. Moreover, $e(t)$ is the noise source (the innovations). Compare with (3-14). The natural prediction is $\hat{y}(t) = \psi^T(t)\hat{\theta}(t-1)$ and its gradient with respect to θ becomes exactly $\psi(t)$.

For models that cannot be written as linear regressions, you cannot recursively compute the exact prediction and its gradient for the current estimate $\hat{\theta}(t-1)$. Then approximations $\hat{y}(t)$ and $\psi(t)$ must be used instead. Section 11.4 in Ljung (1999) describes suitable ways of computing such approximations for general model structures.

The matrix $Q(t)$ that affects both the adaptation gain and the direction in which the updates are made, can be chosen in several different ways. This is discussed in the following.