ID7

ARMA Identification

7.5 OPTIMAL K–STEP–AHEAD ARMA PREDICTOR

The absence of input allows to deduce optimal predictors for ARMA processes even for more steps ahead; we will now deduce the expression of the optimal k–step–ahead predictor that minimizes

$$\mathrm{E}\left[\left(y(t+k) - y(t+k|t)\right)^2\right]. \tag{7.5.1}$$

Making reference to (7.1.2), consider the polynomials

$$a_k(z^{-1}) = 1 + a_1^k\, z^{-1} + \ldots + a_{k-1}^k\, z^{-(k-1)} \tag{7.5.2}$$

$$b_k(z^{-1}) = b_0^k + b_1^k\, z^{-1} + \ldots + b_{n-1}^k\, z^{-(n-1)} \tag{7.5.3}$$

that are univocally defined, in absence of common factors in $q(z)$ and $r(z)$, by relation

$$r(z^{-1}) = a_k(z^{-1})\, q(z^{-1}) + z^{-k}\, b_k(z^{-1}). \tag{7.5.4}$$

By substituting the expression (7.5.4) to $r(z^{-1})$ in (7.1.2) we obtain

$$\begin{aligned} q(z^{-1})\, y(t) &= a_k(z^{-1})\, q(z^{-1})\; w(t) + z^{-k} b_k(z^{-1})\, w(t) \\ &= a_k(z^{-1})\, q(z^{-1})\, w(t) + z^{-k}\, \frac{b_k(z^{-1})\, q(z^{-1})}{r(z^{-1})}\; y(t). \end{aligned} \tag{7.5.5}$$

The multiplication of both sides of (7.5.5) for $z^k/q(z^{-1})$ leads to relation

$$y(t+k) = a_k(z^{-1})\, w(t+k) + \frac{b_k(z^{-1})}{r(z^{-1})}\; y(t) \tag{7.5.6}$$

which gives the output of the ARMA process (7.1.2) at $t+k$. Consider now the following expression

$$\mathrm{E}\left[\, y(t+k) - y(t+k|t)\,\right]^2 = \mathrm{E}\left[\, a_k(z^{-1})\, w(t+k) + \frac{b_k(z^{-1})}{r(z^{-1})}\; y(t) - y(t+k|t)\right]^2 \tag{7.5.7}$$

where $y(t+k|t)$ denotes the prediction of $y(t+k)$ performed at time t. $a_k(z^{-1})w(t+k)$ is independent of $b_k(z^{-1})y(t)$ and of $y(t+k|t)$; moreover $\mathrm{E}[a_k(z^{-1})w(t+k)] = 0$ so that

$$\mathrm{E}\left[\, y(t+k) - y(t+k|t) \,\right]^2 = \mathrm{E}\left[\, a_k(z^{-1})\, w(t+k) \,\right]^2 + \mathrm{E}\left[\frac{b_k(z^{-1})}{r(z^{-1})}\, y(t) - y(t+k|t) \right]^2 ; \tag{7.5.8}$$

it follows that the optimal k–step–ahead ARMA predictor is defined by

$$y(t+k|t) = \frac{b_k(z^{-1})}{r(z^{-1})}\, y(t). \tag{7.5.9}$$

Note that the corresponding prediction error is described by the MA process

$$\varepsilon_k(t+k) = y(t+k) - y(t+k|t) = a_k(z^{-1})\, w(t+k) \tag{7.5.10}$$

with variance given by

$$\sigma^2_{\varepsilon k} = \left(1 + (a_1^k)^2 + \ldots + (a_{k-1}^k)^2\right) \sigma_w^2 . \tag{7.5.11}$$

$a_k(z^{-1})$ and $b_k(z^{-1})$ can be computed using the following algorithm that allows the determination of their coefficients for $k = 2, 3, \ldots$ starting from the optimal one–step–ahead predictor, used to initialize the coefficients

$$a_0^1 = 1 \tag{7.5.12a}$$

$$b_i^1 = \gamma_{n-i} + \alpha_{n-i} \qquad (i = 0, 1, \ldots, n-1)\,. \tag{7.5.12b}$$

The algorithm consists in the recursive relations

$$a_k(z^{-1}) = a_{k-1}(z^{-1}) + a_{k-1}^k\, z^{-(k-1)} \tag{7.5.13a}$$

$$a_{k-1}^k = b_0^{k-1} \tag{7.5.13b}$$

$$b_i^k = b_{i+1}^{k-1} + \alpha_{n-i}\, a_{k-1}^k \qquad (i = 0, 1, \ldots, n-1). \tag{7.5.13c}$$