ID4

AR Identification

4.5 YULE–WALKER EQUATIONS

Rewrite AR model (4.1.1) in the form

$$y(t) - \alpha_n\, y(t-1) - \alpha_{n-1}\, y(t-2) - \ldots - \alpha_1\, y(t-n) = e(t) \tag{4.5.1}$$

and multiply both members of (4.5.1) by $y(t-k)$; by computing their expected values we obtain the expression

$$\mathrm{E}\left[\, y(t)\, y(t-k) - \alpha_n\, y(t-1)\, y(t-k) - \ldots - \alpha_1\, y(t-n)\, y(t-k)\,\right] = \mathrm{E}\left[\, e(t)\, y(t-k)\,\right]. \tag{4.5.2}$$

Since $e(\cdot)$ is a white and stationary process with null expected value, it follows that

$$\mathrm{E}\left[\, e(t)\, y(t-k)\,\right] = 0 \;\text{ for } k > 0; \tag{4.5.3}$$

defining now the quantities

$$r_k = \mathrm{E}\left[\, y(t)\, y(t-k)\,\right], \tag{4.5.4}$$

which are covariances because $\mathrm{E}[y(t)] = 0$, expression (4.5.2) assumes the form

$$r_k - \alpha_n\, r_{k-1} - \alpha_{n-1}\, r_{k-2} - \ldots - \alpha_1\, r_{k-n} = 0 \quad \text{for } k = 1, 2, \ldots \tag{4.5.5}$$

Relations (4.5.5) are known as Yule–Walker equations. Define now the $(m \times n)$ Toeplitz matrix R_m and the vector ρ_m given by

$$R_m = \begin{bmatrix} r_0 & r_1 & \ldots & r_{n-1} \\ r_1 & r_0 & \ldots & r_{n-2} \\ \vdots & \vdots & \ldots & \vdots \\ r_{m-1} & r_{m-2} & \ldots & r_{m-n} \end{bmatrix} \qquad \rho_m = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} \tag{4.5.6}$$

and denote with θ_R the parameter vector with entries in reverse order with respect to (4.3.1), i.e.

$$\theta_R = \left[\, \alpha_n \; \alpha_{n-1} \; \ldots \; \alpha_2 \; \alpha_1 \,\right]^T . \tag{4.5.7}$$

The Yule–Walker equations that can be obtained for $k = 1, \ldots, m$ can be written, since $r_{-k} = r_k$, in the compact form

$$R_m \, \theta_R = \rho_m \qquad \text{for} \;\; m \geq n \, . \tag{4.5.8}$$

If the quantities r_i are known, the model parameters can be deduced from (4.5.8) taking $m = n$; in fact, since R_n has always full rank, it follows that

$$\theta_R^\circ = R_n^{-1} \rho_n \, . \tag{4.5.9}$$

(4.5.9) is called minimal Yule–Walker estimate. Taking $m > n$ the estimate is given by

$$\theta_R^\circ = R_m^+ \, \rho_m = \left(R_m^T R_m \right)^{-1} R_m^T \, \rho_m \tag{4.5.10}$$

and is called overdetermined Yule–Walker estimate. Note also that

$$\mathrm{E}\left[\, e(t)\, y(t) \,\right] = \mathrm{E}\left[\, \alpha_n \, e(t)\, y(t-1) + \ldots + \alpha_1 \, e(t)\, y(t-n) + e(t)^2 \,\right] = \sigma_e^2 ; \tag{4.5.11}$$

taking thus $k = 0$ in (4.5.5) we obtain the relation

$$r_0 - \alpha_n \, r_1 - \ldots - \alpha_1 \, r_n = \sigma_e^2 \tag{4.5.12}$$

which allows, when the parameters α_i and the covariances r_i are known, to determine the variance of $e(\cdot)$. In the practical implementation of this scheme the covariances r_i will not be known and it will be necessary to substitute R_m and ρ_m with the sample quantities

$$R_m^s = \frac{1}{N} \begin{bmatrix} \sum_t y(t-1)\, y(t-1) & \sum_t y(t-2)\, y(t-1) & \ldots & \sum_t y(t-n)\, y(t-1) \\ \sum_t y(t-1)\, y(t-2) & \sum_t y(t-2)\, y(t-2) & \ldots & \sum_t y(t-n)\, y(t-2) \\ \vdots & \vdots & & \vdots \\ \sum_t y(t-1)\, y(t-m) & \sum_t y(t-2)\, y(t-m) & \ldots & \sum_t y(t-n)\, y(t-m) \end{bmatrix} \tag{4.5.13}$$

$$\rho_m^s = \frac{1}{N} \left[\; \sum_t y(t)\, y(t-1) \quad \sum_t y(t)\, y(t-2) \quad \ldots \quad \sum_t y(t)\, y(t-m) \;\right]^T \tag{4.5.14}$$

where N denotes the number of terms in the sums, where it has been omitted to avoid an unnecessarily heavy notation. Of course it is advantageous to use as many terms as possible so that, assuming $m = n$ and denoting with L the length of the available sequences, N will be taken equal to $L - n$.
It is interesting to observe that taking $m = n$ in relations (4.5.13) and (4.5.14) we obtain the least squares estimate because the entries of R_n^s coincide with those of $H^T H$ and the entries of ρ_n^s with those of $H^T y^\circ$, modulo a reordering linked to the reverse order of the parameters in θ and θ_R. It must, however, be observed that Yule–Walker equations lead to a least squares estimate only taking $k = 1, \ldots, n$ in (4.5.5); using a larger number of equations or shifting k we obtain other estimates that, as will be shown in the following, belong to Instrumental Variable methods. It can also be observed that the coincidence of Yule–Walker and least squares estimates vanishes if we approximate matrix (4.5.13) with a Toeplitz matrix (having equal values on its main diagonal) since in (4.5.13) they will be (slightly) different because they are computed from finite shifted sequences.