THE SUBSAMPLING OF LINEAR STOCHASTIC PROCESSES

Consider the first-order autoregressive AR(1) model

$$y(t) = \alpha y(t-1) = \varepsilon(t). \tag{1}$$

On defining the lag operator L, which has the effect that $Ly(t) = y(t-1)$, this can be written as

$$(1 - \alpha L)y(t) = \varepsilon(t). \tag{2}$$

We wish to consider the effect of sampling every nth point to produce the sequence $x(t) = \{y_t; t = 0, \pm n, \pm 2n \ldots\}$,

Multiplying both sides of (2) by

$$T(L) = \frac{(1 - \{\alpha L\}^n)}{(1 - \alpha L)} = 1 + \alpha L + \cdots + \{\alpha L\}^{n-1}, \tag{3}$$

which is a polynomial of finite degree, we get

$$(1 - \alpha^n L^n)y(t) = (1 + \alpha L + \cdots + \{\alpha L\}^{n-1})\varepsilon(t) = \zeta(t), \tag{4}$$

Here, $\zeta_t = \varepsilon_t + \alpha\varepsilon_{t-1} + \cdots + \alpha^{n-1}\varepsilon_{t-n+1}$ is compounded from elements that postdate y_{t-n}. Therefore, ζ_t and y_{t-n} are statistically independent. Also, ζ_t and ζ_{t-n}, which have no elements in common, are statistically independent; and, therefore, $\eta(t) = \{\zeta_t; t = 0, \pm n, \pm 2n, \ldots\}$ is a white-noise sequence of independently and identically distributed random variables. If follows that a consistent estimate of $\alpha^n = \phi$ can be obtained by applying the usual methods to the equation

$$x(t) = \phi x(t-1) + \eta(t), \tag{5}$$

comprising the subsampled sequence, $x(t)$ and the corresponding subsequence $\eta(t)$ of the disturbances.

Now consider the case of an AR(p) model

$$\alpha(L)y(t) = \prod_{j=1}^{p}(1 - \lambda_j L)y(t) = \varepsilon(t), \tag{6}$$

We may multiply both sides by

$$\begin{aligned} \beta(L) = \frac{\phi^*(L)}{\alpha(L)} &= \frac{\prod_{j=1}^{p}(1 - \{\lambda_j L\}^n)}{\prod_{j=1}^{p}(1 - \lambda_j L)} \\ &= \prod_{j=1}^{p}(1 + \lambda_j L + \cdots + \{\lambda L_j\}^{n-1}) = \sum_{j=0}^{p(n-1)} \beta_j L^j, \end{aligned} \tag{7}$$

Now, there is

$$\sum_{j=0}^{p} \alpha_j L^{nj} y(t) = \sum_{j=0}^{p(n-1)} \beta_j L^j \varepsilon(t) = \zeta(t). \tag{8}$$

and the subsampled data sequence is described by an ARMA model of the form

$$\phi(L)x(t) = \eta(t), \tag{9}$$

where $\phi(L) = 1 + \phi_n^* L + \cdots + \phi_{pn}^* L^p$, which is of degree p, contains the nonzero coefficients of $\phi^*(L)$, which is degree np, and where $\eta(t) = \{\zeta_t; t = 0, \pm n, \pm 2n, \ldots\}$ follows a moving-averge process.

The autocovariances of the disturbances $\eta(t)$ are given by

$$\begin{aligned}\gamma_k = E(\eta_t \eta_{t-nk}) &= E\left\{ \left(\sum_i \beta_i \varepsilon_{t-i} \right) \left(\sum_j \beta_j \varepsilon_{t-j-nk} \right) \right\} \\ &= \sum_i \sum_j \beta_i \beta_j E(\varepsilon_{t-i} \varepsilon_{t-j-nk}).\end{aligned} \tag{9}$$

But

$$E(\varepsilon_{t-i}\varepsilon_{t-j-nk}) = \begin{cases} \sigma^2, & \text{if } i = j + nk; \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

so

$$\gamma_k = \sigma^2 \sum_{j=0}^{p(n-1)} \beta_j \beta_{j+kn}, \tag{11}$$

which becomes zero when $kn > p(n-1)$, which is when $k \geq p$. It follows that the moving-average process describing the subsequence of disturbances $\eta(t)$ has a maximum order of $p-1$. Therefore, the subsampled AR(p) sequence is described by an ARMA($p, p-1$) process

Consider next, the general case of an ARMA(p, q) model which can be denoted by

$$\alpha(L)y(t) = \mu(L)\varepsilon(t), \tag{12}$$

where $\alpha(z)$ is a polynomial of degree p and $\mu(z)$ is a polynomial of degree q. We seek a poynomial $\beta(z)$ of degree h such that

$$\beta(z)\alpha(z) = \phi^*(z), \tag{13}$$

where $\phi^*(z)$ is a polynomial of degree rn of which $1, \phi_n, \ldots, \phi_{rn}$ are the only nonzero coefficients.

The degrees of the products on the LHS and the RHS of (13) must be the same, which imposes the condition that $h + p = rn$. Also, there are $r(n -$

1) coefficients in $\phi(z^n)$, which are subject to the restrictions that they are zero-valued. Since these conditions are imposed by a set of linear restrictions generated by the poynomial $\beta(z)$ of degree h, we must have $h = r(n-1)$. We can see that

$$h + p = rn \quad \text{and} \quad h = r(n-1) \quad \text{implies} \quad r = p, \tag{14}$$

which is the degree of the autoregessive polynomial in the ARMA model that describes the subsampled sequence

To obtain the degree of the moving-average polynomial of sussampled process, we note that the degreee of $\beta(z)\mu(z)$ is $h + q = p(n-1) + q$. It follows that

$$E(\eta_t \eta_{t-kn}) = \begin{cases} \gamma_{kn}, & \text{if } kn \leq p(n-1) + q; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

On defining $b = \text{Trunc}\{p(n-1) + q\}/n$ we can assert that the subsampled sequence $x(t) = \{y_t; t = 0, \pm n, \pm 2n \ldots\}$ follows an ARMA(p, d) process.

References

Brewer, K.R.W., (1973), Some Consequences of Temporal Aggregations and Systematic Sampling for ARMA and ARMAX Models, *Journal of Econometrics,* 1, 133–154.

Telser, L.G., (1967), Discrete Samples and Moving Sums in Stationary Stochastic Processes, *Journal of the American Statistical Association,* 62, 484–499.