FIRST-ORDER AUTOREGRESSIVE DISTURBANCES IN THE CLASSICAL LINEAR REGRESSION MODEL

In the classical linear regression model, it is assumed that the disturbances constitute a sequence $\varepsilon(t) = \{\varepsilon_t; t = 0, \pm 1, \pm 2, \ldots\}$ of independently and identically distributed random variables such that

$$E(\varepsilon_t \varepsilon_s) = \begin{cases} \sigma^2, & \text{if } t = s; \\ 0, & \text{if } t \neq s. \end{cases} \tag{1}$$

The process which generates such disturbances is often called a white-noise process.

Our task is to find models for the disturbance process which are more in accordance with the circumstances of economics where the variables tend to show a high degree of inertia. In econometrics, the traditional means of representing the inertial properties of the disturbance process has been to adopt a simple first-order autoregressive model, or AR(1) model, whose equation takes the form of

$$\eta_t = \phi \eta_{t-1} + \varepsilon_t, \qquad \text{where} \qquad \phi \in (-1, 1). \tag{2}$$

Here it continues to be assumed that ε_t is generated by a white-noise process with $E(\varepsilon_t) = 0$. In many econometric applications, the value of ϕ falls in the more restricted interval $[0, 1)$.

According to this model, the conditional expectation of η_t given η_{t-1} is $E(\eta_t|\eta_{t-1}) = \phi\eta_{t-1}$. That is to say, the expectation of the current disturbance is ϕ times the value of the previous disturbance. This implies that, for a value of ϕ which is closer to unity that to zero, there will be a high degree of correlation amongst successive elements of the sequence $\eta(t) = \{\eta_t; t = 0, \pm 1, \pm 2, \ldots\}$.

We can show that the covariance of two elements of the sequence $\eta(t)$ which are separated by τ time periods is given by

$$C(\eta_{t-\tau}, \eta_t) = \gamma_\tau = \sigma^2 \frac{\phi^\tau}{1 - \phi^2}. \tag{3}$$

It follows that variance of the process, which is formally the autocovariance of lag $\tau = 0$, is given by

$$V(\eta_t) = \gamma_0 = \frac{\sigma^2}{1 - \phi^2}. \tag{4}$$

As ϕ tends to unity, the variance increases without bound.

To find the correlation of two elements from the autoregressive sequence, we note that

$$\text{Corr}(\eta_{t-\tau}, \eta_t) = \frac{C(\eta_{t-\tau}, \eta_t)}{\sqrt{V(\eta_{t-\tau})V(\eta_t)}} = \frac{C(\eta_{t-\tau}, \eta_t)}{V(\eta_t)} = \frac{\gamma_\tau}{\gamma_0}. \tag{5}$$

This implies that the correlation of the two elements separated by τ periods is just ϕ^τ; and thus, as the temporal separation increases, the correlation tends to zero in the manner of a convergent geometric progression.

To demonstrate these results, let us consider substituting for $\eta_{t-1} = \phi\eta_{t-2} + \varepsilon_{t-1}$ in the equation under (2) and then substituting for $\eta_{t-2} = \phi\eta_{t-3} + \varepsilon_{t-2}$, and so on indefinitely. By this process, we find that

$$\begin{aligned}\eta_t &= \phi\eta_{t-1} + \varepsilon_t \\ &= \phi^2\eta_{t-2} + \varepsilon_t + \phi\varepsilon_{t-1} \\ &\vdots \\ &= \left\{\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-1} + \cdots\right\} \\ &= \sum_{i=0}^{\infty}\phi^i\varepsilon_{t-i}.\end{aligned} \tag{6}$$

Here the final expression is justified by the fact that $\phi^n \to 0$ as $n \to \infty$ in consequence of the restriction that $|\phi| < 1$. Thus we see that η_t is formed as a geometrically declining weighted average of all past values of the sequence $\varepsilon(t)$.

Using this result, we can now write

$$\begin{aligned}\gamma_\tau &= C(\eta_{t-\tau}, \eta_t) = E(\eta_{t-\tau}\eta_t) \\ &= E\left(\left\{\sum_{i=0}^{\infty}\phi^i\varepsilon_{t-\tau-i}\right\}\left\{\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}\right\}\right) \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\phi^i\phi^j E(\varepsilon_{t-\tau-i}\varepsilon_{t-j}).\end{aligned} \tag{7}$$

But the assumption that $\varepsilon(t)$ is a white-noise process with zero-valued autocovariances at all nonzero lags implies that

$$E(\varepsilon_{t-\tau-i}\varepsilon_{t-j}) = \begin{cases}\sigma^2, & \text{if } j = \tau + i; \\ 0, & \text{if } j \neq \tau + i.\end{cases} \tag{8}$$

Therefore, on using the above conditions in (7) and on setting $j = \tau + i$, we find that

$$\begin{aligned}\gamma_\tau &= \sigma^2\sum_{i}^{\infty}\phi^i\phi^{i+\tau} = \sigma^2\phi^\tau\sum_{i}^{\infty}\phi^{2i} \\ &= \sigma^2\phi^\tau\left\{1 + \phi^2 + \phi^4 + \phi^6 + \cdots\right\} \\ &= \sigma^2\frac{\phi^\tau}{1-\phi^2}.\end{aligned} \tag{9}$$

This establishes the result under (3).

Now let us imagine a linear regression model in the form of

$$y_t = x_{t1}\beta_1 + x_{t2}\beta_2 + \cdots + x_{tk}\beta_k + \eta_t, \tag{10}$$

where η_t follows a first-order autoregressive process. A set of T instances of the relationship would be written as $y = X\beta + \eta$, where y and η are vectors of T elements and X is a matrix or order $T \times k$. The variance–covariance or dispersion matrix of the vector $\eta = [\eta_1, \eta_2, \eta_3, \ldots \eta_T]'$ takes the form of $[\gamma_{|i-j|}] = \sigma_\varepsilon^2 Q$, where

$$(11) \qquad Q = \frac{1}{1-\phi^2}\begin{bmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{T-1} \\ \phi & 1 & \phi & \ldots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \ldots & \phi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \ldots & 1 \end{bmatrix};$$

and it can be confirmed directly that

$$(12) \qquad Q^{-1} = \begin{bmatrix} 1 & -\phi & 0 & \ldots & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & \ldots & 0 & 0 \\ 0 & -\phi & 1+\phi^2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & \ldots & -\phi & 1 \end{bmatrix}.$$

This is a matrix of three nonzero diagonal bands. The elements of principal diagonal, apart from the first and the last, have the value of $1 + \phi^2$. The first and last elements are units. The elements of the supradiagonal band and of the subdiagonal band have the value of $-\phi$.

Given its sparcity, the matrix Q^{-1} could be used directly in implementing the generalised least-squares estimator for which the formula is

$$(13) \qquad \beta^* = (X'Q^{-1}X)^{-1}X'Q^{-1}y.$$

However, by exploiting the factorisation $Q^{-1} = T'T$, we are able to to implement the estimator by applying an ordinary least-squares procedure to the transformed data $W = TX$ and $g = Ty$. The following equation demonstrates the equivalence of the procedures:

$$(14) \qquad \begin{aligned} \beta^* &= (W'W)^{-1}W'g \\ &= (X'T'TX)^{-1}X'T'Ty \\ &= (X'Q^{-1}X)^{-1}X'Q^{-1}y \end{aligned}$$

The factor T of the matrix $Q^{-1} = T'T$ takes the form of

$$(15) \qquad T = \begin{bmatrix} \sqrt{1-\phi^2} & 0 & 0 & \ldots & 0 \\ -\phi & 1 & 0 & \ldots & 0 \\ 0 & -\phi & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}.$$

This effects a very simple transformation the data. Thus, for example, the element y_1 within the vector $y = [y_1, y_2, y_3, \ldots, y_T]'$ is replaced $y_1\sqrt{1-\phi^2}$ whilst y_t is replaced by $y_t - \phi y_{t-1}$, for all $t > 1$.

Consider, for example, the simple regression model

$$y_t = x_t\beta + \eta_t \qquad \text{with} \quad \eta_t = \phi\eta_{t-1} + \varepsilon_t. \tag{16}$$

For $t > 1$, the transformation gives the equation

$$y_t - \phi y_{t-1} = (x_t - \phi x_{t-1})\beta + \varepsilon_t, \tag{17}$$

which represents a model that fulfils the classical assumptions and for which ordinary least squares regression is the appropriate method of estimation.