REPRESENTATIONS OF COINTEGRATED SYSTEMS

Cointegrated Vector Autoregressive Systems

A closed cointegrated system of N linear dynamic equations is one in which the individual variables follow nonstationary random walks, some of which are subject to linear combinations that result in stationary stochastic processes.

The cointegrating system may be represented as a vector autoregressive system in the form of

$$A(L)y(t) = \varepsilon(t), \tag{1}$$

where $A(z) = A_0 + A_1 z + \cdots + A_p z^p$ is a matrix polynomial, which has its roots on or outside the unit circle and L is the lag operator such that $Lx(t) = x(t-1)$ when $x(t)$ is a time series or a vector of time series.

We may assume that $\varepsilon(t)$ is a vector of mutually correlated white-noise processes with $E(\varepsilon_t) = 0$ and $D(\varepsilon_t) = \Sigma$. Under the additional assumption that the processes are normal, we have $\varepsilon_t \sim N(0, \Sigma)$ for all t.

We may chose, without loss of generality, to set $A_0 = I$ and $A_i = -\Phi_i$ to give $A(L) = I - \Phi_1 Lz - \cdots - \Phi_p L^p$. Then

$$y(t) = \Phi_1 y(t-1) + \cdots + \Phi_p y(t-p) + \varepsilon(t), \tag{2}$$

which is a usual way of representing a vector autoregressive process.

The matrix polynomial $A(z)$ of equation (1) can be factored as

$$\begin{aligned} A(z) = U(z)\Lambda(z)V(z) &= \begin{bmatrix} U_1(z) & U_2(z) \end{bmatrix} \begin{bmatrix} \nabla(z) I_{N-r} & 0 \\ 0 & I_r \end{bmatrix} \begin{bmatrix} V_1(z) \\ V_2(z) \end{bmatrix} \\ &= U_1(z)\{\nabla(z) I_{N-r}\} V_1(z) + U_2(z) V_2(z), \end{aligned} \tag{3}$$

where $U(z)$ and $V(z)$ are full-rank matrix polynomials with all of their roots outside the unit circle, and $\Lambda(z)$ is a diagonal matrix polynomial with its roots on the circle or at the origin. (Here, we should observe that $\nabla = 1 - L$ and that $\nabla(z) = 1 - z$.) In a more general case of a vector autoregressive moving-average system, the operator $A(z)$ constitutes a matrix of rational polynomials.

Since $y(t)$ is a nonstationary process if $r < N$, whereas the forcing function $\varepsilon(t)$ is stationary, the operator $A(L)$ must be effective in reducing $y(t)$ to stationarity. The simplest case is where $A(z)$ has all of its roots on the unit circle, which is to say that $r = 0$ and $\Lambda(L)$ has the difference operator $\nabla = 1 - L$ for each of its diagonal elements. In that case, $A(L) = \nabla U(L)V(L) = A^*(L)(1-L)$ and equation (1) can be written as $A^*(L)\nabla y(t) = \varepsilon(t)$. Then, none of the elements of $y(t)$ are cointegrated and each must be reduced to stationarity by differencing.

Now consider the more general case, represented by (3), where $A(z)$ has $N - r$ of its roots on the unit circle and r stable roots outside. Then, by premultiplying the equation $A(L)y(t) = U(L)\Lambda(L)V(L)y(t) = \varepsilon(t)$ by $U^{-1}(L)$, we get

$$\Lambda(L)V(L)y(t) = U^{-1}(L)\varepsilon(t). \tag{4}$$

Letting $V(L)y(t) = [v'(t), w'(t)]'$, we have

$$\begin{bmatrix} \nabla I_{N-r} & 0 \\ 0 & I_r \end{bmatrix} \begin{bmatrix} v(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} \nabla v(t) \\ w(t) \end{bmatrix} = U^{-1}(L)\varepsilon(t). \tag{5}$$

This shows that $y(t)$ is compounded from $N - r$ nonstationary latent variables in $v(t)$, which constitute random walks, and r latent variables in $w(t)$, which constitute stationary stochastic processes.

Given that there are r latent stationary processes underlying $y(t)$, there are also r independent stationary linear combinations that can be formed from its elements. These correspond to the so-called cointegrating relationships. To find these combinations, consider writing the system matrix of equation (1) as

$$A(L) = A(1)L + D(L)\nabla. \tag{6}$$

Here $A(1) = U(1)\Lambda(1)V(1) = -\Gamma\Delta'$ where $\Gamma = -U_2(1)$ is an $N \times r$ matrix of full column rank $\Delta' = V_2(1)$ and $r \times N$ matrix of full row rank. (This is the result of setting $z = 1$ in $A(z) = U_1(z)\{\nabla(z)I_{N-r}\}V_1(z) + U_2(z)V_2(z)$, since $\nabla(z) = 1 - z$ and $\nabla(1) = 0$.) It follows that equation (1) can be rewritten as

$$D(L)\nabla y(t) - \Gamma\Delta' y(t-1) = \varepsilon(t). \tag{7}$$

The r row vectors of the matrix Δ' are described as cointegrating vectors and the corresponding elements of $\Delta' y(t-1)$ constitute stationary processes. By analogy with the model of factor analysis, Γ may be described as a matrix of factor loadings.

Equation (7) can be represented in a manner that is compatible with equation (2). Given that $A_0 = I$, there is also $D_0 = I$. Therefore, on writing $D(L) = I - B_1 L - \cdots - B_{p-1}L^{p-1}$, we get

$$\nabla y(t) = \Gamma\Delta' y(t-1) + B_1\nabla y(t-1) + \cdots + B_{p-1}\nabla y(t-p+1)\varepsilon(t). \tag{8}$$

Example. To demonstrate more explicitly how equation (2) can be converted to the form of (8), we may consider the case where $p = 3$. Then, subtracting $y(t-1)$ from both sides of

$$y(t) = \Phi_1 y(t-1) + \Phi_2 y(t-2) + \Phi_3 y(t-3) + \varepsilon(t) \tag{9}$$

gives

$$\nabla y(t) = (\Phi_1 - I)y(t-1) + \Phi_2 y(t-2) + \Phi_3 y(t-3) + \varepsilon(t). \tag{10}$$

By virtue of a simple reparametrisation, this becomes

$$\begin{aligned}\nabla y(t) &= \begin{bmatrix} \Phi_1 - I & \Phi_2 & \Phi_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} y(t-1) \\ y(t-2) \\ y(t-3) \end{bmatrix} \\ &= \Pi y(t-1) + B_1 \nabla y(t-1) + B_2 \nabla y(t-2) + \varepsilon(t), \end{aligned} \tag{11}$$

where

$$\Pi = \Gamma\Delta' = (\Phi_1 - I) + \Phi_2 + \Phi_3, \quad B_1 = -(\Phi_2 + \Phi_3) \quad \text{and} \quad B_2 = -\Phi_3. \tag{12}$$

The matrices of zeros and units that are found in midst equation (11) have the identity matrix for their product, which show that (11) and (10) are equivalent.

The $N - r$ unit roots within the diagonal matrix $\Lambda(L)$ of the factorisation of $A(L) = U(L)\Lambda(L)V(L)$ of (3) give rise to the same number of latent or "common" tends that combine to form the trajectories of the elements of $y(t)$.

The trends are revealed in the moving-average representation of the vector process. The inverse of the matrix of (3) is given by

$$A^{-1}(L) = V^{-1}(L)\Lambda^{-1}(L)U^{-1}(L) = \nabla^{-1}C(L), \tag{13}$$

where

$$C(L) = V^{-1}(L) \begin{bmatrix} I_{N-r} & 0 \\ 0 & \nabla I_r \end{bmatrix} U^{-1}(L). \tag{14}$$

The moving-average representation of the system can now be given as

$$\nabla y(t) = C(L)\varepsilon(t). \tag{15}$$

The matrix of equation (15) can be written as $C(L) = C(1) + F(L)\nabla$; and, therefore, the system can be represented by

$$y(t) = \nabla^{-1}C(1)\varepsilon(t) + F(L)\varepsilon(t). \tag{16}$$

This is analogous to the so-called Beveridge–Nelson decomposition of a univariate ARIMA process. The process $\nabla^{-1}C(1)\varepsilon(t)$ is the nonstationary component of $y(t)$, whereas $F(L)\varepsilon(t)$ is its stationary component.

Setting $z = 1$ in the matrix polynomial $C(z)$ gives $C(1) = QR'$, where Q is an $N \times (N - r)$ matrix comprising the leading $N - r$ columns of $V^{-1}(1)$ and R' is an $(N - r) \times N$ matrix comprising the leading rows of $U^{-1}(1)$. Setting $C(1) = QR'$ in (16) and defining $\tau(t) = \nabla^{-1}R'\varepsilon(t)$ gives

$$y(t) = Q\tau(t) + F(L)\varepsilon(t), \tag{17}$$

where $\tau(t)$ represents the set of $N - r$ "common" or latent tends that are combined in the variables of $y(t)$.