Note also that if an $(r \times 1)$ vector $\boldsymbol{\xi}_t$ is premultiplied by a nonstochastic $(n \times r)$ matrix $\mathbf{H}'$, the effect is to premultiply the autocovariance by $\mathbf{H}'$ and postmultiply by $\mathbf{H}$:

$$E[(\mathbf{H}'\boldsymbol{\xi}_t - \mathbf{H}'\boldsymbol{\mu}_\xi)(\mathbf{H}'\boldsymbol{\xi}_{t-j} - \mathbf{H}'\boldsymbol{\mu}_\xi)'] = \mathbf{H}'E[(\boldsymbol{\xi}_t - \boldsymbol{\mu}_\xi)(\boldsymbol{\xi}_{t-j} - \boldsymbol{\mu}_\xi)']\mathbf{H},$$

implying

$$\mathbf{G}_{\mathbf{H}'\xi}(z) = \mathbf{H}'\mathbf{G}_\xi(z)\mathbf{H}.$$

Putting these results together, consider $\boldsymbol{\xi}_t$ the r-dimensional $VAR(1)$ process $\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$ and a new process $\mathbf{u}_t$ given by $\mathbf{u}_t = \mathbf{H}'\boldsymbol{\xi}_t + \mathbf{w}_t$ with $\mathbf{w}_t$ a white noise process that is uncorrelated with $\boldsymbol{\xi}_{t-j}$ for all j. Then

$$\mathbf{G}_\mathbf{U}(z) = \mathbf{H}'\mathbf{G}_\xi(z)\mathbf{H} + \mathbf{G}_\mathbf{W}(z),$$

or, if $\mathbf{R}$ is the variance of $\mathbf{w}_t$,

$$\mathbf{G}_\mathbf{U}(z) = \mathbf{H}'[\mathbf{I}_r - \mathbf{F}z]^{-1}\mathbf{Q}[\mathbf{I}_r - \mathbf{F}'z^{-1}]^{-1}\mathbf{H} + \mathbf{R}. \qquad [10.3.6]$$

More generally, consider an $(n \times 1)$ vector $\mathbf{y}_t$ characterized by

$$\mathbf{y}_t = \boldsymbol{\mu}_\mathbf{Y} + \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t$ is a white noise process with variance-covariance matrix given by $\boldsymbol{\Omega}$ and where $\boldsymbol{\Psi}(L) = \Sigma_{k=0}^{\infty}\boldsymbol{\Psi}_k L^k$ with $\{\boldsymbol{\Psi}_k\}_{k=0}^{\infty}$ absolutely summable. Thus, the autocovariance-generating function for $\mathbf{y}$ is

$$\mathbf{G}_\mathbf{Y}(z) = \boldsymbol{\Psi}(z)\boldsymbol{\Omega}[\boldsymbol{\Psi}(z^{-1})]'. \qquad [10.3.7]$$

Let $\{\mathbf{H}_k\}_{k=-\infty}^{\infty}$ be an absolutely summable sequence of $(r \times n)$ matrices, and suppose that an $(r \times 1)$ vector $\mathbf{x}_t$ is constructed from $\mathbf{y}_t$ according to

$$\mathbf{x}_t = \mathbf{H}(L)\mathbf{y}_t = \sum_{k=-\infty}^{\infty} \mathbf{H}_k\mathbf{y}_{t-k} = \boldsymbol{\mu}_\mathbf{X} + \mathbf{B}(L)\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\mu}_\mathbf{X} = \mathbf{H}(1)\boldsymbol{\mu}_\mathbf{Y}$ and $\mathbf{B}(L) = \mathbf{H}(L)\boldsymbol{\Psi}(L)$ as in [10.2.10] and [10.2.11]. Then the autocovariance-generating function for $\mathbf{x}$ can be found from

$$\mathbf{G}_\mathbf{X}(z) = \mathbf{B}(z)\boldsymbol{\Omega}[\mathbf{B}(z^{-1})]' = [\mathbf{H}(z)\boldsymbol{\Psi}(z)]\boldsymbol{\Omega}[\boldsymbol{\Psi}(z^{-1})]'[\mathbf{H}(z^{-1})]'. \qquad [10.3.8]$$

Comparing [10.3.8] with [10.3.7], the effect of applying the filter $\mathbf{H}(L)$ to $\mathbf{y}_t$ is to premultiply the autocovariance-generating function by $\mathbf{H}(z)$ and to postmultiply by the transpose of $\mathbf{H}(z^{-1})$:

$$\mathbf{G}_\mathbf{X}(z) = [\mathbf{H}(z)]\mathbf{G}_\mathbf{Y}(z)[\mathbf{H}(z^{-1})]'. \qquad [10.3.9]$$

10.4. *The Spectrum for Vector Processes*

Let $\mathbf{y}_t$ be an $(n \times 1)$ vector with mean $E(\mathbf{y}_t) = \boldsymbol{\mu}$ and kth autocovariance matrix

$$E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-k} - \boldsymbol{\mu})'] = \boldsymbol{\Gamma}_k. \qquad [10.4.1]$$

If $\{\boldsymbol{\Gamma}_k\}_{k=-\infty}^{\infty}$ is absolutely summable and if z is a complex scalar, the autocovariance-generating function of $\mathbf{y}$ is given by

$$\mathbf{G}_\mathbf{Y}(z) = \sum_{k=-\infty}^{\infty} \boldsymbol{\Gamma}_k z^k. \qquad [10.4.2]$$

The function $\mathbf{G_Y}(z)$ associates an $(n \times n)$ matrix of complex numbers with the complex scalar z. If [10.4.2] is divided by 2π and evaluated at $z = e^{-i\omega}$, where ω is a real scalar and $i = \sqrt{-1}$, the result is the *population spectrum* of the vector $\mathbf{y}$:

$$\mathbf{s_Y}(\omega) = (2\pi)^{-1}\mathbf{G_Y}(e^{-i\omega}) = (2\pi)^{-1}\sum_{k=-\infty}^{\infty}\mathbf{\Gamma}_k e^{-i\omega k}. \qquad [10.4.3]$$

The population spectrum associates an $(n \times n)$ matrix of complex numbers with the real scalar ω.

Identical calculations to those used to establish Proposition 6.1 indicate that when any element of $\mathbf{s_Y}(\omega)$ is multiplied by $e^{i\omega k}$ and the resulting function of ω is integrated from $-\pi$ to π, the result is the corresponding element of the kth autocovariance matrix of $\mathbf{y}$:

$$\int_{-\pi}^{\pi}\mathbf{s_Y}(\omega)e^{i\omega k}\,d\omega = \mathbf{\Gamma}_k. \qquad [10.4.4]$$

Thus, as in the univariate case, the sequence of autocovariances $\{\mathbf{\Gamma}_k\}_{k=-\infty}^{\infty}$ and the function represented by the population spectrum $\mathbf{s_Y}(\omega)$ contain the identical information.

As a special case, when $k = 0$, equation [10.4.4] implies

$$\int_{-\pi}^{\pi}\mathbf{s_Y}(\omega)\,d\omega = \mathbf{\Gamma}_0. \qquad [10.4.5]$$

In other words, the area under the population spectrum is the unconditional variance-covariance matrix of $\mathbf{y}$.

The jth diagonal element of $\mathbf{\Gamma}_k$ is $E(y_{jt} - \mu_j)(y_{j,t-k} - \mu_j)$, the kth autocovariance of y_{jt}. Thus, the jth diagonal element of the multivariate spectrum $\mathbf{s_Y}(\omega)$ is just the univariate spectrum of the scalar y_{jt}. It follows from the properties of the univariate spectrum discussed in Chapter 6 that the diagonal elements of $\mathbf{s_Y}(\omega)$ are real-valued and nonnegative for all ω. However, the same is not true of the off-diagonal elements of $\mathbf{s_Y}(\omega)$—in general, the off-diagonal elements of $\mathbf{s_Y}(\omega)$ will be complex numbers.

To gain further understanding of the multivariate spectrum, we concentrate on the case of $n = 2$ variables, denoted

$$\mathbf{y}_t = \begin{bmatrix} X_t \\ Y_t \end{bmatrix}.$$

The kth autocovariance matrix is then

$$\begin{aligned}\mathbf{\Gamma}_k &= E\begin{bmatrix} (X_t - \mu_X)(X_{t-k} - \mu_X) & (X_t - \mu_X)(Y_{t-k} - \mu_Y) \\ (Y_t - \mu_Y)(X_{t-k} - \mu_X) & (Y_t - \mu_Y)(Y_{t-k} - \mu_Y)\end{bmatrix} \\ &\equiv \begin{bmatrix} \gamma_{XX}^{(k)} & \gamma_{XY}^{(k)} \\ \gamma_{YX}^{(k)} & \gamma_{YY}^{(k)}\end{bmatrix}.\end{aligned} \qquad [10.4.6]$$

Recall from [10.2.2] that $\mathbf{\Gamma}_k' = \mathbf{\Gamma}_{-k}$. Hence,

$$\gamma_{XX}^{(k)} = \gamma_{XX}^{(-k)} \qquad [10.4.7]$$

$$\gamma_{YY}^{(k)} = \gamma_{YY}^{(-k)} \qquad [10.4.8]$$

$$\gamma_{XY}^{(k)} = \gamma_{YX}^{(-k)}. \qquad [10.4.9]$$

For this $n = 2$ case, the population spectrum [10.4.3] would be

$$\mathbf{s}_{\mathbf{Y}}(\omega) = \frac{1}{2\pi}\begin{bmatrix} \sum_{k=-\infty}^{\infty} \gamma_{XX}^{(k)} e^{-i\omega k} & \sum_{k=-\infty}^{\infty} \gamma_{XY}^{(k)} e^{-i\omega k} \\ \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)} e^{-i\omega k} & \sum_{k=-\infty}^{\infty} \gamma_{YY}^{(k)} e^{-i\omega k} \end{bmatrix}$$

$$= \frac{1}{2\pi}\begin{bmatrix} \sum_{k=-\infty}^{\infty} \gamma_{XX}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} & \sum_{k=-\infty}^{\infty} \gamma_{XY}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} \\ \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} & \sum_{k=-\infty}^{\infty} \gamma_{YY}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} \end{bmatrix}. \quad [10.4.10]$$

Using [10.4.7] and [10.4.8] along with the facts that $\sin(-\omega k) = -\sin(\omega k)$ and $\sin(0) = 0$, the imaginary components disappear from the diagonal terms:

$$\mathbf{s}_{\mathbf{Y}}(\omega) = \frac{1}{2\pi}\begin{bmatrix} \sum_{k=-\infty}^{\infty} \gamma_{XX}^{(k)} \cos(\omega k) & \sum_{k=-\infty}^{\infty} \gamma_{XY}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} \\ \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\} & \sum_{k=-\infty}^{\infty} \gamma_{YY}^{(k)} \cos(\omega k) \end{bmatrix}. \quad [10.4.11]$$

However, since in general $\gamma_{XY}^{(k)} \neq \gamma_{XY}^{(-k)}$, the off-diagonal elements are typically complex numbers.

The Cross Spectrum, Cospectrum, and Quadrature Spectrum

The lower left element of the matrix in [10.4.11] is known as the *population cross spectrum* from X to Y:

$$s_{YX}(\omega) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)}\{\cos(\omega k) - i\cdot\sin(\omega k)\}. \quad [10.4.12]$$

The cross spectrum can be written in terms of its real and imaginary components as

$$s_{YX}(\omega) = c_{YX}(\omega) + i\cdot q_{YX}(\omega). \quad [10.4.13]$$

The real component of the cross spectrum is known as the *cospectrum* between X and Y:

$$c_{YX}(\omega) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)} \cos(\omega k). \quad [10.4.14]$$

One can verify from [10.4.9] and the fact that $\cos(-\omega k) = \cos(\omega k)$ that

$$c_{YX}(\omega) = c_{XY}(\omega). \quad [10.4.15]$$

The imaginary component of the cross spectrum is known as the *quadrature spectrum* from X to Y:

$$q_{YX}(\omega) = -(2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma_{YX}^{(k)} \sin(\omega k). \qquad [10.4.16]$$

One can verify from [10.4.9] and the fact that $\sin(-\omega k) = -\sin(\omega k)$ that the quadrature spectrum from Y to X is the negative of the quadrature spectrum from X to Y:

$$q_{YX}(\omega) = -q_{XY}(\omega).$$

Recalling [10.4.13], these results imply that the off-diagonal elements of $\mathbf{s}_{\mathbf{Y}}(\omega)$ are complex conjugates of each other; in general, the row j, column m element of $\mathbf{s}_{\mathbf{Y}}(\omega)$ is the complex conjugate of the row m, column j element of $\mathbf{s}_{\mathbf{Y}}(\omega)$.

Note that both $c_{YX}(\omega)$ and $q_{YX}(\omega)$ are real-valued periodic functions of ω:

$$c_{YX}(\omega + 2\pi j) = c_{YX}(\omega) \qquad \text{for } j = \pm 1, \pm 2, \ldots$$
$$q_{YX}(\omega + 2\pi j) = q_{YX}(\omega) \qquad \text{for } j = \pm 1, \pm 2, \ldots .$$

It further follows from [10.4.14] that

$$c_{YX}(-\omega) = c_{YX}(\omega),$$

while [10.4.16] implies that

$$q_{YX}(-\omega) = -q_{YX}(\omega). \qquad [10.4.17]$$

Hence, the cospectrum and quadrature spectrum are fully specified by the values they assume as ω ranges between 0 and π.

Result [10.4.5] implies that the cross spectrum integrates to the unconditional covariance between X and Y:

$$\int_{-\pi}^{\pi} s_{YX}(\omega)\, d\omega = E(Y_t - \mu_Y)(X_t - \mu_X).$$

Observe from [10.4.17] that the quadrature spectrum integrates to zero:

$$\int_{-\pi}^{\pi} q_{YX}(\omega)\, d\omega = 0.$$

Hence, the covariance between X and Y can be calculated from the area under the cospectrum between X and Y:

$$\int_{-\pi}^{\pi} c_{YX}(\omega)\, d\omega = E(Y_t - \mu_Y)(X_t - \mu_X). \qquad [10.4.18]$$

The cospectrum between X and Y at frequency ω can thus be interpreted as the portion of the covariance between X and Y that is attributable to cycles with frequency ω. Since the covariance can be positive or negative, the cospectrum can be positive or negative, and indeed, $c_{YX}(\omega)$ may be positive over some frequencies and negative over others.

The Sample Multivariate Periodogram

To gain further understanding of the cospectrum and the quadrature spectrum, let $y_1, y_2, \ldots, y_T$ and $x_1, x_2, \ldots, x_T$ denote samples of T observations on the two variables. If for illustration T is odd, Proposition 6.2 indicates that the value of y_t can be expressed as

$$y_t = \bar{y} + \sum_{j=1}^{M} \{\hat{\alpha}_j \cdot \cos[\omega_j(t-1)] + \hat{\delta}_j \cdot \sin[\omega_j(t-1)]\}, \qquad [10.4.19]$$

where $\bar{y}$ is the sample mean of y, $M = (T-1)/2$, $\omega_j = 2\pi j/T$, and

$$\hat{\alpha}_j = (2/T) \sum_{t=1}^{T} y_t \cdot \cos[\omega_j(t-1)] \qquad [10.4.20]$$

$$\hat{\delta}_j = (2/T) \sum_{t=1}^{T} y_t \cdot \sin[\omega_j(t-1)]. \qquad [10.4.21]$$

An analogous representation for x_t is

$$x_t = \bar{x} + \sum_{j=1}^{M} \{\hat{a}_j \cdot \cos[\omega_j(t-1)] + \hat{d}_j \cdot \sin[\omega_j(t-1)]\} \qquad [10.4.22]$$

$$\hat{a}_j = (2/T) \sum_{t=1}^{T} x_t \cdot \cos[\omega_j(t-1)] \qquad [10.4.23]$$

$$\hat{d}_j = (2/T) \sum_{t=1}^{T} x_t \cdot \sin[\omega_j(t-1)]. \qquad [10.4.24]$$

Recall from [6.2.11] that the periodic regressors in [10.4.19] all have sample mean zero and are mutually orthogonal, while

$$\sum_{t=1}^{T} \cos^2[\omega_j(t-1)] = \sum_{t=1}^{T} \sin^2[\omega_j(t-1)] = T/2. \qquad [10.4.25]$$

Consider the sample covariance between x and y:

$$T^{-1} \sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x}). \qquad [10.4.26]$$

Substituting [10.4.19] and [10.4.22] into [10.4.26] and exploiting the mutual orthogonality of the periodic regressors reveal that

$$\begin{aligned} T^{-1} &\sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x}) \\ &= T^{-1} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{M} \{\hat{\alpha}_j \cdot \cos[\omega_j(t-1)] + \hat{\delta}_j \cdot \sin[\omega_j(t-1)]\} \right. \\ &\quad \left. \times \sum_{j=1}^{M} \{\hat{a}_j \cdot \cos[\omega_j(t-1)] + \hat{d}_j \cdot \sin[\omega_j(t-1)]\} \right\} \qquad [10.4.27] \\ &= T^{-1} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{M} \{\hat{\alpha}_j \hat{a}_j \cdot \cos^2[\omega_j(t-1)] + \hat{\delta}_j \hat{d}_j \cdot \sin^2[\omega_j(t-1)]\} \right\} \\ &= (1/2) \sum_{j=1}^{M} (\hat{\alpha}_j \hat{a}_j + \hat{\delta}_j \hat{d}_j). \end{aligned}$$

Hence, the portion of the sample covariance between x and y that is due to their common dependence on cycles of frequency ω_j is given by

$$(1/2)(\hat{\alpha}_j\hat{a}_j + \hat{\delta}_j\hat{d}_j). \qquad [10.4.28]$$

This magnitude can be related to the sample analog of the cospectrum with calculations similar to those used to establish result (c) of Proposition 6.2. Recall that since

$$\sum_{t=1}^{T} \cos[\omega_j(t-1)] = 0,$$

the magnitude $\hat{\alpha}_j$ in [10.4.20] can alternatively be expressed as

$$\hat{\alpha}_j = (2/T)\sum_{t=1}^{T}(y_t - \bar{y})\cdot\cos[\omega_j(t-1)].$$

Thus,

$$\begin{aligned}
&(\hat{a}_j + i\cdot\hat{d}_j)(\hat{\alpha}_j - i\cdot\hat{\delta}_j)\\
&= (4/T^2)\left\{\sum_{t=1}^{T}(x_t-\bar{x})\cdot\cos[\omega_j(t-1)] + i\cdot\sum_{t=1}^{T}(x_t-\bar{x})\cdot\sin[\omega_j(t-1)]\right\}\\
&\qquad\times\left\{\sum_{\tau=1}^{T}(y_\tau-\bar{y})\cdot\cos[\omega_j(\tau-1)] - i\cdot\sum_{\tau=1}^{T}(y_\tau-\bar{y})\cdot\sin[\omega_j(\tau-1)]\right\}\\
&= (4/T^2)\left\{\sum_{t=1}^{T}(x_t-\bar{x})\cdot\exp[i\cdot\omega_j(t-1)]\right\}\left\{\sum_{\tau=1}^{T}(y_\tau-\bar{y})\cdot\exp[-i\cdot\omega_j(\tau-1)]\right\}\\
&= (4/T^2)\left\{\sum_{t=1}^{T}(x_t-\bar{x})(y_t-\bar{y}) + \sum_{t=1}^{T-1}(x_t-\bar{x})(y_{t+1}-\bar{y})\cdot\exp[-i\omega_j]\right.\\
&\qquad+\sum_{t=2}^{T}(x_t-\bar{x})(y_{t-1}-\bar{y})\cdot\exp[i\omega_j] + \sum_{t=1}^{T-2}(x_t-\bar{x})(y_{t+2}-\bar{y})\cdot\exp[-2i\omega_j]\\
&\qquad+\sum_{t=3}^{T}(x_t-\bar{x})(y_{t-2}-\bar{y})\cdot\exp[2i\omega_j] + \cdots + (x_1-\bar{x})(y_T-\bar{y})\cdot\exp[-(T-1)i\omega_j]\\
&\qquad\left.+(x_T-\bar{x})(y_1-\bar{y})\cdot\exp[(T-1)i\omega_j]\right\}\\
&= (4/T)\left\{\hat{\gamma}_{yx}^{(0)} + \hat{\gamma}_{yx}^{(1)}\cdot\exp[-i\omega_j] + \hat{\gamma}_{yx}^{(-1)}\cdot\exp[i\omega_j]\right.\\
&\qquad+\hat{\gamma}_{yx}^{(2)}\cdot\exp[-2i\omega_j] + \hat{\gamma}_{yx}^{(-2)}\cdot\exp[2i\omega_j] + \cdots\\
&\qquad\left.+\hat{\gamma}_{yx}^{(T-1)}\cdot\exp[-(T-1)i\omega_j] + \hat{\gamma}_{yx}^{(-T+1)}\cdot\exp[(T-1)i\omega_j]\right\},
\end{aligned} \qquad [10.4.29]$$

where $\hat{\gamma}_{yx}^{(k)}$ is the sample covariance between the value of y and the value that x assumed k periods earlier:

$$\hat{\gamma}_{yx}^{(k)} = \begin{cases} (1/T)\displaystyle\sum_{t=1}^{T-k}(x_t-\bar{x})(y_{t+k}-\bar{y}) & \text{for } k = 0, 1, 2, \ldots, T-1\\ (1/T)\displaystyle\sum_{t=-k+1}^{T}(x_t-\bar{x})(y_{t+k}-\bar{y}) & \text{for } k = -1, -2, \ldots, -T+1.\end{cases} \qquad [10.4.30]$$

Result [10.4.29] implies that

$$\tfrac{1}{2}(\hat{a}_j + i\cdot\hat{d}_j)(\hat{\alpha}_j - i\cdot\hat{\delta}_j) = (2/T)\sum_{k=-T+1}^{T-1}\hat{\gamma}_{yx}^{(k)}\cdot\exp[-ki\omega_j] \\ = (4\pi/T)\cdot\hat{s}_{yx}(\omega_j), \qquad [10.4.31]$$

where $\hat{s}_{yx}(\omega_j)$ is the *sample cross periodogram* from x to y at frequency ω_j, or the lower left element of the *sample multivariate periodogram*:

$$\hat{\mathbf{s}}_{\mathbf{y}}(\omega) = (2\pi)^{-1}\begin{bmatrix} \sum_{k=-T+1}^{T-1}\hat{\gamma}_{xx}^{(k)}e^{-i\omega k} & \sum_{k=-T+1}^{T-1}\hat{\gamma}_{xy}^{(k)}e^{-i\omega k} \\ \sum_{k=-T+1}^{T-1}\hat{\gamma}_{yx}^{(k)}e^{-i\omega k} & \sum_{k=-T+1}^{T-1}\hat{\gamma}_{yy}^{(k)}e^{-i\omega k}\end{bmatrix} = \begin{bmatrix}\hat{s}_{xx}(\omega) & \hat{s}_{xy}(\omega) \\ \hat{s}_{yx}(\omega) & \hat{s}_{yy}(\omega)\end{bmatrix}.$$

Expression [10.4.31] states that the sample cross periodogram from x to y at frequency ω_j can be expressed as

$$\hat{s}_{yx}(\omega_j) = [T/(8\pi)]\cdot(\hat{a}_j + i\cdot\hat{d}_j)(\hat{\alpha}_j - i\cdot\hat{\delta}_j) \\ = [T/(8\pi)]\cdot(\hat{a}_j\hat{\alpha}_j + \hat{d}_j\hat{\delta}_j) + i\cdot[T/(8\pi)]\cdot(\hat{d}_j\hat{\alpha}_j - \hat{a}_j\hat{\delta}_j).$$

The real component is the sample analog of the cospectrum, while the imaginary component is the sample analog of the quadrature spectrum:

$$\hat{s}_{yx}(\omega_j) = \hat{c}_{yx}(\omega_j) + i\cdot\hat{q}_{yx}(\omega_j), \qquad [10.4.32]$$

where

$$\hat{c}_{yx}(\omega_j) = [T/(8\pi)]\cdot(\hat{a}_j\hat{\alpha}_j + \hat{d}_j\hat{\delta}_j) \qquad [10.4.33]$$

$$\hat{q}_{yx}(\omega_j) = [T/(8\pi)]\cdot(\hat{d}_j\hat{\alpha}_j - \hat{a}_j\hat{\delta}_j). \qquad [10.4.34]$$

Comparing [10.4.33] with [10.4.28], the sample cospectrum evaluated at ω_j is proportional to the portion of the sample covariance between y and x that is attributable to cycles with frequency ω_j. The population cospectrum admits an analogous interpretation as the portion of the population covariance between Y and X attributable to cycles with frequency ω based on a multivariate version of the spectral representation theorem.

What interpretation are we to attach to the quadrature spectrum? Consider using the weights in [10.4.22] to construct a new series x_t^* by shifting the phase of each of the periodic functions by a quarter cycle:

$$x_t^* = \bar{x} + \sum_{j=1}^{M}\{\hat{a}_j\cdot\cos[\omega_j(t-1) + (\pi/2)] \\ + \hat{d}_j\cdot\sin[\omega_j(t-1) + (\pi/2)]\}. \qquad [10.4.35]$$

The variable x_t^* is driven by the same cycles as x_t, except that at date $t = 1$ each cycle is one-quarter of the way through rather than just beginning as in the case of x_t.

Since $\sin[\theta + (\pi/2)] = \cos(\theta)$ and since $\cos[\theta + (\pi/2)] = -\sin(\theta)$, the variable x_t^* can alternatively be described as

$$x_t^* = \bar{x} + \sum_{j=1}^{M}\{\hat{d}_j\cdot\cos[\omega_j(t-1)] - \hat{a}_j\cdot\sin[\omega_j(t-1)]\}. \qquad [10.4.36]$$

As in [10.4.27], the sample covariance between y_t and x_t^* is found to be

$$T^{-1} \sum_{t=1}^{T} (y_t - \bar{y})(x_t^* - \bar{x}) = (1/2) \sum_{j=1}^{M} (\hat{\alpha}_j \hat{d}_j - \hat{\delta}_j \hat{a}_j).$$

Comparing this with [10.4.34], the sample quadrature spectrum from x to y at frequency ω_j is proportional to the portion of the sample covariance between x^* and y that is due to cycles of frequency ω_j. Cycles of frequency ω_j may be quite important for both x and y individually (as reflected by large values for $\hat{s}_{xx}(\omega)$ and $\hat{s}_{yy}(\omega)$) yet fail to produce much contemporaneous covariance between the variables because at any given date the two series are in a different phase of the cycle. For example, the variable x may respond to an economic recession sooner than y. The quadrature spectrum looks for evidence of such out-of-phase cycles.

Coherence, Phase, and Gain

The *population coherence* between X and Y is a measure of the degree to which X and Y are jointly influenced by cycles of frequency ω. This measure combines the inferences of the cospectrum and the quadrature spectrum, and is defined as[1]

$$h_{YX}(\omega) = \frac{[c_{YX}(\omega)]^2 + [q_{YX}(\omega)]^2}{s_{YY}(\omega)s_{XX}(\omega)},$$

assuming that $s_{YY}(\omega)$ and $s_{XX}(\omega)$ are nonzero. If $s_{YY}(\omega)$ or $s_{XX}(\omega)$ is zero, the coherence is defined to be zero. It can be shown that $0 \leq h_{YX}(\omega) \leq 1$ for all ω as long as X and Y are covariance-stationary with absolutely summable autocovariance matrices.[2] If $h_{YX}(\omega)$ is large, this indicates that Y and X have important cycles of frequency ω in common.

The cospectrum and quadrature spectrum can alternatively be described in polar coordinate form. In this notation, the population cross spectrum from X to Y is written as

$$s_{YX}(\omega) = c_{YX}(\omega) + i \cdot q_{YX}(\omega) = R(\omega) \cdot \exp[i \cdot \theta(\omega)], \qquad [10.4.37]$$

where

$$R(\omega) = \{[c_{YX}(\omega)]^2 + [q_{YX}(\omega)]^2\}^{1/2} \qquad [10.4.38]$$

and $\theta(\omega)$ represents the radian angle satisfying

$$\frac{\sin[\theta(\omega)]}{\cos[\theta(\omega)]} = \frac{q_{YX}(\omega)}{c_{YX}(\omega)}. \qquad [10.4.39]$$

The function $R(\omega)$ is sometimes described as the *gain* while $\theta(\omega)$ is called the *phase*.[3]

[1]The coherence is sometimes alternatively defined as the square root of this magnitude. The sample coherence based on the unsmoothed periodogram is identically equal to 1.

[2]See, for example, Fuller (1976, p. 156).

[3]The gain is sometimes alternatively defined as $R(\omega)/s_{XX}(\omega)$.

The Population Spectrum for Vector MA *and* AR *Processes*

Let $\mathbf{y}_t$ be a vector $MA(\infty)$ process with absolutely summable moving average coefficients:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t,$$

where

$$E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau') = \begin{cases} \boldsymbol{\Omega} & \text{for } t = \tau \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Substituting [10.3.4] into [10.4.3] reveals that the population spectrum for $\mathbf{y}_t$ can be calculated as

$$\mathbf{s}_\mathbf{Y}(\omega) = (2\pi)^{-1}[\boldsymbol{\Psi}(e^{-i\omega})]\boldsymbol{\Omega}[\boldsymbol{\Psi}(e^{i\omega})]'. \qquad [10.4.40]$$

For example, the population spectrum for a stationary $VAR(p)$ as written in [10.1.4] is

$$\begin{aligned} \mathbf{s}_\mathbf{Y}(\omega) = {} & (2\pi)^{-1}\{\mathbf{I}_n - \boldsymbol{\Phi}_1 e^{-i\omega} - \boldsymbol{\Phi}_2 e^{-2i\omega} - \cdots - \boldsymbol{\Phi}_p e^{-pi\omega}\}^{-1}\boldsymbol{\Omega} \\ & \times \{\mathbf{I}_n - \boldsymbol{\Phi}_1' e^{i\omega} - \boldsymbol{\Phi}_2' e^{2i\omega} - \cdots - \boldsymbol{\Phi}_p' e^{pi\omega}\}^{-1}. \end{aligned} \qquad [10.4.41]$$

Estimating the Population Spectrum

If an observed time series $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$ can be reasonably described by a pth-order vector autoregression, one good approach to estimating the population spectrum is to estimate the parameters of the vector autoregression [10.1.4] by *OLS* and then substitute these parameter estimates into equation [10.4.41].

Alternatively, the sample cross periodogram from x to y at frequency $\omega_j = 2\pi j/T$ can be calculated from [10.4.32] to [10.4.34], where $\hat{\alpha}_j$, $\hat{\delta}_j$, $\hat{a}_j$, and $\hat{d}_j$ are as defined in [10.4.20] through [10.4.24]. One would want to smooth these to obtain a more useful estimate of the population cross spectrum. For example, one reasonable estimate of the population cospectrum between X and Y at frequency ω_j would be

$$\hat{c}_{YX}(\omega_j) = \sum_{m=-h}^{h} \left\{\frac{h + 1 - |m|}{(h + 1)^2}\right\} \hat{c}_{yx}(\omega_{j+m}),$$

where $\hat{c}_{yx}(\omega_{j+m})$ denotes the estimate in [10.4.33] evaluated at frequency $\omega_{j+m} = 2\pi(j + m)/T$ and h is a bandwidth parameter reflecting how many different frequencies are to be used in estimating the cospectrum at frequency ω_j.

Another approach is to express the smoothing in terms of weighting coefficients κ_k^* to be applied to $\hat{\boldsymbol{\Gamma}}_k$ when the population autocovariances in expression [10.4.3] are replaced by sample autocovariances. Such an estimate would take the form

$$\hat{\mathbf{s}}_\mathbf{Y}(\omega) = (2\pi)^{-1}\left\{\hat{\boldsymbol{\Gamma}}_0 + \sum_{k=1}^{T-1} \kappa_k^*[\hat{\boldsymbol{\Gamma}}_k e^{-i\omega k} + \hat{\boldsymbol{\Gamma}}_k' e^{i\omega k}]\right\}$$

where

$$\hat{\Gamma}_k = T^{-1}\sum_{t=k+1}^{T}(\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_{t-k} - \bar{\mathbf{y}})'$$

$$\bar{\mathbf{y}} = T^{-1}\sum_{t=1}^{T}\mathbf{y}_t.$$

For example, the modified Bartlett estimate of the multivariate spectrum is

$$\hat{\mathbf{s}}_{\mathbf{Y}}(\omega) = (2\pi)^{-1}\left\{\hat{\Gamma}_0 + \sum_{k=1}^{q}\left[1 - \frac{k}{q+1}\right][\hat{\Gamma}_k e^{-i\omega k} + \hat{\Gamma}'_k e^{i\omega k}]\right\}. \qquad [10.4.42]$$

Filters

Let $\mathbf{x}_t$ be an r-dimensional covariance-stationary process with absolutely summable autocovariances and with $(r \times r)$ population spectrum denoted $\mathbf{s}_{\mathbf{X}}(\omega)$. Let $\{\mathbf{H}_k\}_{k=-\infty}^{\infty}$ be an absolutely summable sequence of $(n \times r)$ matrices, and let $\mathbf{y}_t$ denote the n-dimensional vector process given by

$$\mathbf{y}_t = \mathbf{H}(L)\mathbf{x}_t = \sum_{k=-\infty}^{\infty}\mathbf{H}_k\mathbf{x}_{t-k}.$$

It follows from [10.3.9] that the population spectrum of $\mathbf{y}$ (denoted $\mathbf{s}_{\mathbf{Y}}(\omega)$) is related to that of $\mathbf{x}$ according to

$$\underset{(n\times n)}{\mathbf{s}_{\mathbf{Y}}(\omega)} = \underset{(n\times r)}{[\mathbf{H}(e^{-i\omega})]}\underset{(r\times r)}{\mathbf{s}_{\mathbf{X}}(\omega)}\underset{(r\times n)}{[\mathbf{H}(e^{i\omega})]'}. \qquad [10.4.43]$$

As a special case of this result, let X_t be a univariate stationary stochastic process with continuous spectrum $s_X(\omega)$, and let u_t be a second univariate stationary stochastic process with continuous spectrum $s_U(\omega)$, where X_t and u_τ are uncorrelated for all t and τ. Thus, the population spectrum of the vector $\mathbf{x}_t \equiv (X_t, u_t)'$ is given by

$$\mathbf{s}_{\mathbf{X}}(\omega) = \begin{bmatrix} s_{XX}(\omega) & 0 \\ 0 & s_{UU}(\omega) \end{bmatrix}.$$

Define a new series Y_t according to

$$Y_t = \sum_{k=-\infty}^{\infty} h_k X_{t-k} + u_t \equiv h(L)X_t + u_t, \qquad [10.4.44]$$

where $\{h_k\}_{k=-\infty}^{\infty}$ is absolutely summable. Note that the vector $\mathbf{y}_t \equiv (X_t, Y_t)'$ is obtained from the original vector $\mathbf{x}_t$ by the filter

$$\mathbf{y}_t = \mathbf{H}(L)\mathbf{x}_t,$$

where

$$\mathbf{H}(L) = \begin{bmatrix} 1 & 0 \\ h(L) & 1 \end{bmatrix}.$$

It follows from [10.4.43] that the spectrum of $\mathbf{y}$ is given by

$$\mathbf{s_Y}(\omega) = \begin{bmatrix} 1 & 0 \\ h(e^{-i\omega}) & 1 \end{bmatrix} \begin{bmatrix} s_{XX}(\omega) & 0 \\ 0 & s_{UU}(\omega) \end{bmatrix} \begin{bmatrix} 1 & h(e^{i\omega}) \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} s_{XX}(\omega) & s_{XX}(\omega)h(e^{i\omega}) \\ h(e^{-i\omega})s_{XX}(\omega) & h(e^{-i\omega})s_{XX}(\omega)h(e^{i\omega}) + s_{UU}(\omega) \end{bmatrix}, \qquad [10.4.45]$$

where

$$h(e^{-i\omega}) = \sum_{k=-\infty}^{\infty} h_k e^{-i\omega k}. \qquad [10.4.46]$$

The lower left element of the matrix in [10.4.45] indicates that when Y_t and X_t are related according to [10.4.44], the cross spectrum from X to Y can be calculated by multiplying [10.4.46] by the spectrum of X.

We can also imagine going through these steps in reverse order. Specifically, suppose we are given an observed vector $\mathbf{y}_t = (X_t, Y_t)'$ with absolutely summable autocovariance matrices and with population spectrum given by

$$\mathbf{s_Y}(\omega) = \begin{bmatrix} s_{XX}(\omega) & s_{XY}(\omega) \\ s_{YX}(\omega) & s_{YY}(\omega) \end{bmatrix}. \qquad [10.4.47]$$

Then the linear projection of Y_t on $\{X_{t-k}\}_{k=-\infty}^{\infty}$ exists and is of the form of [10.4.44], where u_t would now be regarded as the population residual associated with the linear projection. The sequence of linear projection coefficients $\{h_k\}_{k=-\infty}^{\infty}$ can be summarized in terms of the function of ω given in [10.4.46]. Comparing the lower left elements of [10.4.47] and [10.4.45], this function must satisfy

$$h(e^{-i\omega})s_{XX}(\omega) = s_{YX}(\omega).$$

In other words, the function $h(e^{-i\omega})$ can be calculated from

$$h(e^{-i\omega}) = \frac{s_{YX}(\omega)}{s_{XX}(\omega)}, \qquad [10.4.48]$$

assuming that $s_{XX}(\omega)$ is not zero. When $s_{XX}(\omega) = 0$, we set $h(e^{-i\omega}) = 0$. This magnitude, the ratio of the cross spectrum from X to Y to the spectrum of X, is known as the *transfer function* from X to Y.

The principles underlying [10.4.4] can further be used to uncover individual transfer function coefficients:

$$h_k = (2\pi)^{-1} \int_{-\pi}^{\pi} h(e^{-i\omega})e^{i\omega k}\, d\omega.$$

In other words, given an observed vector $(X_t, Y_t)'$ with absolutely summable autocovariance matrices and thus with continuous population spectrum of the form of [10.4.47], the coefficient on X_{t-k} in the population linear projection of Y_t on $\{X_{t-k}\}_{k=-\infty}^{\infty}$ can be calculated from

$$h_k = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{s_{YX}(\omega)}{s_{XX}(\omega)} e^{i\omega k}\, d\omega. \qquad [10.4.49]$$

10.5. The Sample Mean of a Vector Process

Variance of the Sample Mean

Suppose we have a sample of size T, $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T\}$, drawn from an n-dimensional covariance-stationary process with

$$E(\mathbf{y}_t) = \boldsymbol{\mu} \qquad [10.5.1]$$

$$E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-j} - \boldsymbol{\mu})'] = \boldsymbol{\Gamma}_j. \qquad [10.5.2]$$

Consider the properties of the sample mean,

$$\bar{\mathbf{y}}_T = (1/T) \sum_{t=1}^{T} \mathbf{y}_t. \qquad [10.5.3]$$

As in the discussion in Section 7.2 of the sample mean of a scalar process, it is clear that $E(\bar{\mathbf{y}}_T) = \boldsymbol{\mu}$ and

$$\begin{aligned}
&E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})'] \\
&\quad = (1/T^2)E\{(\mathbf{y}_1 - \boldsymbol{\mu})[(\mathbf{y}_1 - \boldsymbol{\mu})' + (\mathbf{y}_2 - \boldsymbol{\mu})' + \cdots + (\mathbf{y}_T - \boldsymbol{\mu})'] \\
&\qquad + (\mathbf{y}_2 - \boldsymbol{\mu})[(\mathbf{y}_1 - \boldsymbol{\mu})' + (\mathbf{y}_2 - \boldsymbol{\mu})' + \cdots + (\mathbf{y}_T - \boldsymbol{\mu})'] \\
&\qquad + (\mathbf{y}_3 - \boldsymbol{\mu})[(\mathbf{y}_1 - \boldsymbol{\mu})' + (\mathbf{y}_2 - \boldsymbol{\mu})' + \cdots + (\mathbf{y}_T - \boldsymbol{\mu})'] \\
&\qquad + \cdots + (\mathbf{y}_T - \boldsymbol{\mu})[(\mathbf{y}_1 - \boldsymbol{\mu})' + (\mathbf{y}_2 - \boldsymbol{\mu})' + \cdots + (\mathbf{y}_T - \boldsymbol{\mu})']\} \\
&\quad = (1/T^2)\{[\boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_{-1} + \cdots + \boldsymbol{\Gamma}_{-(T-1)}] \qquad [10.5.4] \\
&\qquad + [\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_{-1} + \cdots + \boldsymbol{\Gamma}_{-(T-2)}] \\
&\qquad + [\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_{-1} + \cdots + \boldsymbol{\Gamma}_{-(T-3)}] \\
&\qquad + \cdots + [\boldsymbol{\Gamma}_{T-1} + \boldsymbol{\Gamma}_{T-2} + \boldsymbol{\Gamma}_{T-3} + \cdots + \boldsymbol{\Gamma}_0]\} \\
&\quad = (1/T^2)\{T\boldsymbol{\Gamma}_0 + (T-1)\boldsymbol{\Gamma}_1 + (T-2)\boldsymbol{\Gamma}_2 + \cdots + \boldsymbol{\Gamma}_{T-1} \\
&\qquad + (T-1)\boldsymbol{\Gamma}_{-1} + (T-2)\boldsymbol{\Gamma}_{-2} + \cdots + \boldsymbol{\Gamma}_{-(T-1)}\}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&T\cdot E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})'] \\
&\quad = \boldsymbol{\Gamma}_0 + [(T-1)/T]\boldsymbol{\Gamma}_1 + [(T-2)/T]\boldsymbol{\Gamma}_2 + \cdots \qquad [10.5.5] \\
&\qquad + [1/T]\boldsymbol{\Gamma}_{T-1} + [(T-1)/T]\boldsymbol{\Gamma}_{-1} + [(T-2)/T]\boldsymbol{\Gamma}_{-2} \\
&\qquad + \cdots + [1/T]\boldsymbol{\Gamma}_{-(T-1)}.
\end{aligned}$$

As in the univariate case, the weights on $\boldsymbol{\Gamma}_k$ for $|k|$ small go to unity as $T \to \infty$, and higher autocovariances go to zero for a covariance-stationary process. Hence, we have the following generalization of Proposition 7.5.

Proposition 10.5: *Let $\mathbf{y}_t$ be a covariance-stationary process with moments given by [10.5.1] and [10.5.2] and with absolutely summable autocovariances. Then the sample mean [10.5.3] satisfies*

(*a*) $\bar{\mathbf{y}}_T \xrightarrow{p} \boldsymbol{\mu}$

(*b*) $\lim_{T\to\infty} \{T\cdot E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})']\} = \sum_{v=-\infty}^{\infty} \boldsymbol{\Gamma}_v.$

The proof of Proposition 10.5 is virtually identical to that of Proposition 7.5. Consider the following $(n \times n)$ matrix:

$$\sum_{v=-\infty}^{\infty} \Gamma_v - T\cdot E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})'] = \sum_{|v|\geq T} \Gamma_v + \sum_{v=-(T-1)}^{T-1} (|v|/T)\Gamma_v, \quad [10.5.6]$$

where the equality follows from [10.5.5]. Let $\gamma_{ij}^{(v)}$ denote the row i, column j element of Γ_v. The row i, column j element of the matrix in [10.5.6] can then be written

$$\sum_{|v|\geq T} \gamma_{ij}^{(v)} + \sum_{v=-(T-1)}^{T-1} (|v|/T)\gamma_{ij}^{(v)}.$$

Absolutely summability of $\{\Gamma_v\}_{v=-\infty}^{\infty}$ implies that for any $\varepsilon > 0$ there exists a q such that

$$\sum_{|v|>q} |\gamma_{ij}^{(v)}| < \varepsilon/2.$$

Thus,

$$\left| \sum_{|v|\geq T} \gamma_{ij}^{(v)} + \sum_{v=-(T-1)}^{T-1} (|v|/T)\gamma_{ij}^{(v)} \right| < \varepsilon/2 + \sum_{v=-q}^{q} (|v|/T)|\gamma_{ij}^{(v)}|.$$

This sum can be made less than ε by choosing T sufficiently large. This establishes claim (b) of Proposition 10.5. From this result, $E(\bar{y}_{i,T} - \mu_i)^2 \to 0$ for each i, implying that $\bar{y}_{i,T} \xrightarrow{p} \mu_i$.

Consistent Estimation of T *Times the Variance of the Sample Mean*

Hypothesis tests about the sample mean require an estimate of the matrix in result (b) of Proposition 10.5. Let $\mathbf{S}$ represent this matrix:

$$\mathbf{S} \equiv \lim_{T\to\infty} T\cdot E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})']. \quad [10.5.7]$$

If the data were generated by a vector $MA(q)$ process, then result (b) would imply

$$\mathbf{S} = \sum_{v=-q}^{q} \Gamma_v. \quad [10.5.8]$$

A natural estimate then is

$$\hat{\mathbf{S}} = \hat{\Gamma}_0 + \sum_{v=1}^{q} (\hat{\Gamma}_v + \hat{\Gamma}_v'), \quad [10.5.9]$$

where

$$\hat{\Gamma}_v = (1/T) \sum_{t=v+1}^{T} (\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_{t-v} - \bar{\mathbf{y}})'. \quad [10.5.10]$$

As long as $\mathbf{y}_t$ is ergodic for second moments, [10.5.9] gives a consistent estimate of [10.5.8]. Indeed, Hansen (1982) and White (1984, Chapter 6) noted that [10.5.9] gives a consistent estimate of the asymptotic variance of the sample mean for a broad class of processes exhibiting time-dependent heteroskedasticity and autocorrelation. To see why, note that for a process satisfying $E(\mathbf{y}_t) = \boldsymbol{\mu}$ with time-varying second moments, the variance of the sample mean is given by

$$\begin{aligned}
&E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})'] \\
&= E\left[(1/T)\sum_{t=1}^{T}(\mathbf{y}_t - \boldsymbol{\mu})\right]\left[(1/T)\sum_{s=1}^{T}(\mathbf{y}_s - \boldsymbol{\mu})\right]' \qquad [10.5.11] \\
&= (1/T^2)\sum_{t=1}^{T}\sum_{s=1}^{T}E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_s - \boldsymbol{\mu})'].
\end{aligned}$$

Suppose, first, that $E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_s - \boldsymbol{\mu})'] = \mathbf{0}$ for $|t - s| > q$, as was the case for the vector $MA(q)$ process, though we generalize from the $MA(q)$ process to allow $E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_s - \boldsymbol{\mu})']$ to be a function of t for $|t - s| \leq q$. Then [10.5.11] implies

$$\begin{aligned}
&T \cdot E[(\bar{\mathbf{y}}_T - \boldsymbol{\mu})(\bar{\mathbf{y}}_T - \boldsymbol{\mu})'] \\
&= (1/T)\sum_{t=1}^{T}E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})'] \\
&\quad + (1/T)\sum_{t=2}^{T}\{E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-1} - \boldsymbol{\mu})'] + E[(\mathbf{y}_{t-1} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})']\} \\
&\quad + (1/T)\sum_{t=3}^{T}\{E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-2} - \boldsymbol{\mu})'] + E[(\mathbf{y}_{t-2} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})']\} + \cdots \\
&\quad + (1/T)\sum_{t=q+1}^{T}\{E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-q} - \boldsymbol{\mu})'] + E[(\mathbf{y}_{t-q} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})']\}. \qquad [10.5.12]
\end{aligned}$$

The estimate [10.5.9] replaces

$$(1/T)\sum_{t=v+1}^{T}E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-v} - \boldsymbol{\mu})'] \qquad [10.5.13]$$

in [10.5.12] with

$$(1/T)\sum_{t=v+1}^{T}(\mathbf{y}_t - \bar{\mathbf{y}}_T)(\mathbf{y}_{t-v} - \bar{\mathbf{y}}_T)', \qquad [10.5.14]$$

and thus [10.5.9] provides a consistent estimate of the limit of [10.5.12] whenever [10.5.14] converges in probability to [10.5.13]. Hence, the estimator proposed in [10.5.9] can give a consistent estimate of T times the variance of the sample mean in the presence of both heteroskedasticity and autocorrelation up through order q.

More generally, even if $E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_s - \boldsymbol{\mu})']$ is nonzero for all t and s, as long as this matrix goes to zero sufficiently quickly as $|t - s| \to \infty$, then there is still a sense in which $\hat{\mathbf{S}}_T$ in [10.5.9] can provide a consistent estimate of $\mathbf{S}$. Specifically, if, as the sample size T grows, a larger number of sample autocovariances q is used to form the estimate, then $\hat{\mathbf{S}}_T \xrightarrow{p} \mathbf{S}$ (see White, 1984, p. 155).

The Newey-West Estimator

Although [10.5.9] gives a consistent estimate of $\mathbf{S}$, it has the drawback that [10.5.9] need not be positive semidefinite in small samples. If $\hat{\mathbf{S}}$ is not positive semidefinite, then some linear combination of the elements of $\bar{\mathbf{y}}$ is asserted to have a negative variance, a considerable handicap in forming a hypothesis test!

Newey and West (1987) suggested the alternative estimate

$$\tilde{\mathbf{S}} = \hat{\boldsymbol{\Gamma}}_0 + \sum_{v=1}^{q}\left[1 - \frac{v}{q+1}\right](\hat{\boldsymbol{\Gamma}}_v + \hat{\boldsymbol{\Gamma}}_v'), \qquad [10.5.15]$$

where $\hat{\Gamma}_v$ is given by [10.5.10]. For example, for $q = 2$,

$$\tilde{S} = \hat{\Gamma}_0 + \tfrac{2}{3}(\hat{\Gamma}_1 + \hat{\Gamma}_1') + \tfrac{1}{3}(\hat{\Gamma}_2 + \hat{\Gamma}_2').$$

Newey and West showed that $\tilde{S}$ is positive semidefinite by construction and has the same consistency properties that were noted for $\hat{S}$, namely, that if q and T both go to infinity with $q/T^{1/4} \to 0$, then $\tilde{S}_T \xrightarrow{p} S$.

Application: Autocorrelation- and Heteroskedasticity-Consistent Standard Errors for Linear Regressions

As an application of using the Newey-West weighting, consider the linear regression model

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$$

for $\mathbf{x}_t$ a $(k \times 1)$ vector of explanatory variables. Recall from equation [8.2.6] that the deviation of the *OLS* estimate $\mathbf{b}_T$ from the true value $\boldsymbol{\beta}$ satisfies

$$\sqrt{T}(\mathbf{b}_T - \boldsymbol{\beta}) = \left[(1/T)\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\left[(1/\sqrt{T})\sum_{t=1}^{T}\mathbf{x}_t u_t\right]. \qquad [10.5.16]$$

In calculating the asymptotic distribution of the *OLS* estimate $\mathbf{b}_T$, we usually assume that the first term in [10.5.16] converges in probability to $\mathbf{Q}^{-1}$:

$$\left[(1/T)\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1} \xrightarrow{p} \mathbf{Q}^{-1}. \qquad [10.5.17]$$

The second term in [10.5.16] can be viewed as $\sqrt{T}$ times the sample mean of the $(k \times 1)$ vector $\mathbf{x}_t u_t$:

$$\begin{aligned}\left[(1/\sqrt{T})\sum_{t=1}^{T}\mathbf{x}_t u_t\right] &= (\sqrt{T})(1/T)\sum_{t=1}^{T}\mathbf{y}_t \\ &= \sqrt{T}\cdot\bar{\mathbf{y}}_T,\end{aligned} \qquad [10.5.18]$$

where $\mathbf{y}_t \equiv \mathbf{x}_t u_t$. Provided that $E(u_t|\mathbf{x}_t) = 0$, the vector $\mathbf{y}_t$ has mean zero. We can allow for conditional heteroskedasticity, autocorrelation, and time variation in the second moments of $\mathbf{y}_t$, as long as

$$\mathbf{S} \equiv \lim_{T\to\infty} T\cdot E(\bar{\mathbf{y}}_T\bar{\mathbf{y}}_T')$$

exists. Under general conditions,[4] it then turns out that

$$\left[(1/\sqrt{T})\sum_{t=1}^{T}\mathbf{x}_t u_t\right] = \sqrt{T}\cdot\bar{\mathbf{y}}_T \xrightarrow{L} N(\mathbf{0}, \mathbf{S}).$$

Substituting this and [10.5.17] into [10.5.16],

$$\sqrt{T}(\mathbf{b}_T - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{Q}^{-1}\mathbf{S}\mathbf{Q}^{-1}). \qquad [10.5.19]$$

In light of the foregoing discussion, we might hope to estimate $\mathbf{S}$ by

$$\hat{\mathbf{S}}_T = \hat{\Gamma}_{0,T} + \sum_{v=1}^{q}\left[1 - \frac{v}{q+1}\right](\hat{\Gamma}_{v,T} + \hat{\Gamma}_{v,T}'). \qquad [10.5.20]$$

[4]See, for example, White (1984, p. 119).

Here,

$$\hat{\Gamma}_{v,T} = (1/T) \sum_{t=v+1}^{T} (\mathbf{x}_t \hat{u}_{t,T} \hat{u}_{t-v,T} \mathbf{x}'_{t-v}),$$

$\hat{u}_{t,T}$ is the *OLS* residual for date t in a sample of size T ($\hat{u}_{t,T} = y_t - \mathbf{x}'_t \mathbf{b}_T$), and q is a lag length beyond which we are willing to assume that the correlation between $\mathbf{x}_t u_t$ and $\mathbf{x}_{t-v} u_{t-v}$ is essentially zero. Clearly, $\mathbf{Q}$ is consistently estimated by $\hat{\mathbf{Q}}_T = (1/T)\Sigma_{t=1}^{T} \mathbf{x}_t \mathbf{x}'_t$. Substituting $\hat{\mathbf{Q}}_T$ and $\hat{\mathbf{S}}_T$ into [10.5.19], the suggestion is to treat the *OLS* estimate $\mathbf{b}_T$ as if

$$\mathbf{b}_T \approx N(\boldsymbol{\beta}, (\hat{\mathbf{V}}_T/T))$$

where

$$\begin{aligned}
\hat{\mathbf{V}}_T &= \hat{\mathbf{Q}}_T^{-1} \hat{\mathbf{S}}_T \hat{\mathbf{Q}}_T^{-1} \\
&= \left[(1/T) \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}'_t\right]^{-1} (1/T) \left[\sum_{t=1}^{T} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}'_t \right. \\
&\quad \left. + \sum_{v=1}^{q} \left[1 - \frac{v}{q+1}\right] \sum_{t=v+1}^{T} (\mathbf{x}_t \hat{u}_t \hat{u}_{t-v} \mathbf{x}'_{t-v} + \mathbf{x}_{t-v} \hat{u}_{t-v} \hat{u}_t \mathbf{x}'_t)\right] \\
&\quad \times \left[(1/T) \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}'_t\right]^{-1},
\end{aligned}$$

that is, the variance of $\mathbf{b}_T$ is approximated by

$$\begin{aligned}
&(\hat{\mathbf{V}}_T/T) \\
&= \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}'_t\right]^{-1} \left[\sum_{t=1}^{T} \hat{u}_t^2 \mathbf{x}_t \mathbf{x}'_t \right. \\
&\quad \left. + \sum_{v=1}^{q} \left[1 - \frac{v}{q+1}\right] \sum_{t=v+1}^{T} (\mathbf{x}_t \hat{u}_t \hat{u}_{t-v} \mathbf{x}'_{t-v} + \mathbf{x}_{t-v} \hat{u}_{t-v} \hat{u}_t \mathbf{x}'_t)\right] \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}'_t\right]^{-1}
\end{aligned} \qquad [10.5.21]$$

where $\hat{u}_t$ is the *OLS* sample residual. The square root of the row i, column i element of $\hat{\mathbf{V}}_T/T$ is known as a *heteroskedasticity- and autocorrelation-consistent standard error* for the ith element of the estimated *OLS* coefficient vector. The hope is that standard errors based on [10.5.21] will be robust to a variety of forms of heteroskedasticity and autocorrelation of the residuals u_t of the regression.

Spectral-Based Estimators

A number of alternative estimates of $\mathbf{S}$ in [10.5.7] have been suggested in the literature. Notice that as in the univariate case discussed in Section 7.2, if $\mathbf{y}_t$ is covariance-stationary, then $\mathbf{S}$ has the interpretation as the autocovariance-generating function $\mathbf{G}_\mathbf{Y}(z) = \Sigma_{v=-\infty}^{\infty} \mathbf{\Gamma}_v z^v$ evaluated at $z = 1$, or, equivalently, as 2π times the population spectrum at frequency zero:

$$\mathbf{S} = \sum_{v=-\infty}^{\infty} \mathbf{\Gamma}_v = 2\pi \mathbf{s}_\mathbf{Y}(0).$$

Indeed, the Newey-West estimator [10.5.15] is numerically identical to 2π times the Bartlett estimate of the multivariate spectrum described in [10.4.42] evaluated at frequency $\omega = 0$. Gallant (1987, p. 533) proposed a similar estimator based on a Parzen kernel,

$$\hat{\mathbf{S}} = \hat{\mathbf{\Gamma}}_0 + \sum_{v=1}^{q} k[v/(q+1)](\hat{\mathbf{\Gamma}}_v + \hat{\mathbf{\Gamma}}'_v),$$

where

$$k(z) = \begin{cases} 1 - 6z^2 + 6z^3 & \text{for } 0 \leq z \leq \frac{1}{2} \\ 2(1 - z)^3 & \text{for } \frac{1}{2} \leq z \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

For example, for $q = 2$, we have

$$\hat{\mathbf{S}} = \hat{\mathbf{\Gamma}}_0 + \tfrac{5}{9}(\hat{\mathbf{\Gamma}}_1 + \hat{\mathbf{\Gamma}}_1') + \tfrac{2}{27}(\hat{\mathbf{\Gamma}}_2 + \hat{\mathbf{\Gamma}}_2').$$

Andrews (1991) examined a number of alternative estimators and found the best results for a quadratic spectral kernel:

$$k(z) = \frac{3}{(6\pi z/5)^2}\left[\frac{\sin(6\pi z/5)}{6\pi z/5} - \cos(6\pi z/5)\right].$$

In contrast to the Newey-West and Gallant estimators, Andrews's suggestion makes use of all $T - 1$ estimated autocovariance estimators:

$$\hat{\mathbf{S}} = \frac{T}{T - k}\left[\hat{\mathbf{\Gamma}}_0 + \sum_{v=1}^{T-1} k\left(\frac{v}{q + 1}\right)(\hat{\mathbf{\Gamma}}_v + \hat{\mathbf{\Gamma}}_v')\right]. \qquad [10.5.22]$$

Even though [10.5.22] makes use of all computed autocovariances, there is still a bandwidth parameter q to be chosen for constructing the kernel. For example, for $q = 2$,

$$\hat{\mathbf{\Gamma}}_0 + \sum_{v=1}^{T-1} k(v/3)(\hat{\mathbf{\Gamma}}_v + \hat{\mathbf{\Gamma}}_v') = \hat{\mathbf{\Gamma}}_0 + 0.85(\hat{\mathbf{\Gamma}}_1 + \hat{\mathbf{\Gamma}}_1') + 0.50(\hat{\mathbf{\Gamma}}_2 + \hat{\mathbf{\Gamma}}_2') + 0.14(\hat{\mathbf{\Gamma}}_3 + \hat{\mathbf{\Gamma}}_3') + \cdots.$$

Andrews recommended multiplying the estimate by $T/(T - k)$, where $\mathbf{y}_t = \mathbf{x}_t\hat{u}_t$ for $\hat{u}_t$ the sample *OLS* residual from a regression with k explanatory variables. Andrews (1991) and Newey and West (1992) also offered some guidance for choosing an optimal value of the lag truncation or bandwidth parameter q for each of the estimators of $\mathbf{S}$ that have been discussed here.

The estimators that have been described will work best when $\mathbf{y}_t$ has a finite moving average representation. Andrews and Monahan (1992) suggested an alternative approach to estimating $\mathbf{S}$ that also takes advantage of any autoregressive structure to the errors. Let $\mathbf{y}_t$ be a zero-mean vector, and let $\mathbf{S}$ be the asymptotic variance of the sample mean of $\mathbf{y}$. For example, if we want to calculate heteroskedasticity- and autocorrelation-consistent standard errors for *OLS* estimation, $\mathbf{y}_t$ would correspond to $\mathbf{x}_t\hat{u}_t$ where $\mathbf{x}_t$ is the vector of explanatory variables for the regression and $\hat{u}_t$ is the *OLS* residual. The first step in estimating $\mathbf{S}$ is to fit a low-order *VAR* for $\mathbf{y}_t$,

$$\mathbf{y}_t = \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \mathbf{v}_t, \qquad [10.5.23]$$

where $\mathbf{v}_t$ is presumed to have some residual autocorrelation not entirely captured by the *VAR*. Note that since $\mathbf{y}_t$ has mean zero, no constant term is included in [10.5.23]. The ith row represented in [10.5.23] can be estimated by an *OLS* regression of the ith element of $\mathbf{y}_t$ on p lags of all the elements of $\mathbf{y}$, though if any eigenvalue of $|\mathbf{I}_n\lambda^p - \hat{\mathbf{\Phi}}_1\lambda^{p-1} - \hat{\mathbf{\Phi}}_2\lambda^{p-2} - \cdots - \hat{\mathbf{\Phi}}_p| = 0$ is too close to the unit circle (say, greater than 0.97 in modulus), Andrews and Monahan (1992, p. 957) recommended altering the *OLS* estimates so as to reduce the largest eigenvalue.

The second step in the Andrews and Monahan procedure is to calculate an estimate $\mathbf{S}^*$ using one of the methods described previously based on the fitted residuals $\hat{\mathbf{v}}_t$ from [10.5.23]. For example,

$$\hat{\mathbf{S}}_T^* = \hat{\mathbf{\Gamma}}_0^* + \sum_{v=1}^{q}\left[1 - \frac{v}{q+1}\right](\hat{\mathbf{\Gamma}}_v^* + \hat{\mathbf{\Gamma}}_v^{*\prime}), \qquad [10.5.24]$$

where

$$\hat{\mathbf{\Gamma}}_v^* = (1/T)\sum_{t=v+1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{v}}_{t-v}'$$

and where q is a parameter representing the maximal order of autocorrelation assumed for $\mathbf{v}_t$. The matrix $\hat{\mathbf{S}}_T^*$ will be recognized as an estimate of $2\pi\cdot\mathbf{s}_{\mathbf{V}}(0)$, where $\mathbf{s}_{\mathbf{V}}(\omega)$ is the spectral density of $\mathbf{v}$:

$$\mathbf{s}_{\mathbf{V}}(\omega) = (2\pi)^{-1}\sum_{v=-\infty}^{\infty}\{E(\mathbf{v}_t\mathbf{v}_{t-v}')\}e^{-i\omega v}.$$

Notice that the original series $\mathbf{y}_t$ can be obtained from $\mathbf{v}_t$ by applying the following filter:

$$\mathbf{y}_t = [\mathbf{I}_n - \mathbf{\Phi}_1 L - \mathbf{\Phi}_2 L^2 - \cdots - \mathbf{\Phi}_p L^p]^{-1}\mathbf{v}_t.$$

Thus, from [10.4.43], the spectral density of $\mathbf{y}$ is related to the spectral density of $\mathbf{v}$ according to

$$\begin{aligned}\mathbf{s}_{\mathbf{Y}}(\omega) = {} & \{[\mathbf{I}_n - \mathbf{\Phi}_1 e^{-i\omega} - \mathbf{\Phi}_2 e^{-2i\omega} - \cdots - \mathbf{\Phi}_p e^{-pi\omega}]\}^{-1}\mathbf{s}_{\mathbf{V}}(\omega)\\ & \times \{[\mathbf{I}_n - \mathbf{\Phi}_1 e^{i\omega} - \mathbf{\Phi}_2 e^{2i\omega} - \cdots - \mathbf{\Phi}_p e^{pi\omega}]'\}^{-1}.\end{aligned}$$

Hence, an estimate of 2π times the spectral density of $\mathbf{y}$ at frequency zero is given by

$$\begin{aligned}\hat{\mathbf{S}}_T = {} & \{[\mathbf{I}_n - \hat{\mathbf{\Phi}}_1 - \hat{\mathbf{\Phi}}_2 - \cdots - \hat{\mathbf{\Phi}}_p]\}^{-1}\hat{\mathbf{S}}_T^*\\ & \times \{[\mathbf{I}_n - \hat{\mathbf{\Phi}}_1 - \hat{\mathbf{\Phi}}_2 - \cdots - \hat{\mathbf{\Phi}}_p]'\}^{-1},\end{aligned} \qquad [10.5.25]$$

where $\hat{\mathbf{S}}_T^*$ is calculated from [10.5.24]. The matrix $\hat{\mathbf{S}}_T$ in [10.5.25] is the Andrews-Monahan (1992) estimate of $\mathbf{S}$, where

$$\mathbf{S} = \lim_{T\to\infty} T\cdot E(\bar{\mathbf{y}}_T\bar{\mathbf{y}}_T').$$

APPENDIX 10.A. *Proofs of Chapter 10 Propositions*

■ **Proof of Proposition 10.1.** The eigenvalues of $\mathbf{F}$ are the values of λ for which the following determinant is zero:

$$\begin{vmatrix} (\mathbf{\Phi}_1 - \lambda\mathbf{I}_n) & \mathbf{\Phi}_2 & \mathbf{\Phi}_3 & \cdots & \mathbf{\Phi}_{p-1} & \mathbf{\Phi}_p \\ \mathbf{I}_n & -\lambda\mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & -\lambda\mathbf{I}_n & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_n & -\lambda\mathbf{I}_n \end{vmatrix}. \qquad [10.A.1]$$

Multiply each of the final block of n columns by $(1/\lambda)$ and add to the previous block. Multiply each of the n columns of this resulting next-to-final block by $(1/\lambda)$ and add the result to the third-to-last block of columns. Proceeding in this manner reveals [10.A.1] to be the same as

$$\begin{vmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{0} & -\lambda\mathbf{I}_{n(p-1)} \end{vmatrix}, \qquad [10.A.2]$$

where $\mathbf{X}_1$ denotes the following $(n \times n)$ matrix:

$$\mathbf{X}_1 \equiv (\mathbf{\Phi}_1 - \lambda\mathbf{I}_n) + (\mathbf{\Phi}_2/\lambda) + (\mathbf{\Phi}_3/\lambda^2) + \cdots + (\mathbf{\Phi}_p/\lambda^{p-1})$$

and $\mathbf{X}_2$ is a related $[n \times n(p-1)]$ matrix. Let $\mathbf{S}$ denote the following $(np \times np)$ matrix:

$$\mathbf{S} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{I}_{n(p-1)} \\ \mathbf{I}_n & \mathbf{0} \end{bmatrix},$$

and note that its inverse is given by

$$\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_{n(p-1)} & \mathbf{0} \end{bmatrix},$$

as may be verified by direct multiplication. Premultiplying a matrix by $\mathbf{S}$ and postmultiplying by $\mathbf{S}^{-1}$ will not change the determinant. Thus, [10.A.2] is equal to

$$\left| \begin{bmatrix} \mathbf{0} & \mathbf{I}_{n(p-1)} \\ \mathbf{I}_n & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{0} & -\lambda\mathbf{I}_{n(p-1)} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_{n(p-1)} & \mathbf{0} \end{bmatrix} \right| = \begin{vmatrix} -\lambda\mathbf{I}_{n(p-1)} & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{X}_1 \end{vmatrix}. \qquad [10.A.3]$$

Applying the formula for calculating a determinant [A.4.5] recursively, [10.A.3] is equal to

$$\begin{aligned} (-\lambda)^{n(p-1)}|\mathbf{X}_1| &= (-\lambda)^{n(p-1)}|\boldsymbol{\Phi}_1 - \lambda\mathbf{I}_n + (\boldsymbol{\Phi}_2/\lambda) + (\boldsymbol{\Phi}_3/\lambda^2) + \cdots + (\boldsymbol{\Phi}_p/\lambda^{p-1})| \\ &= (-1)^{np}|\mathbf{I}_n\lambda^p - \boldsymbol{\Phi}_1\lambda^{p-1} - \boldsymbol{\Phi}_2\lambda^{p-2} - \cdots - \boldsymbol{\Phi}_p|. \end{aligned}$$

Setting this to zero produces equation [10.1.13]. ■

■ **Proof of Proposition 10.2.** It is helpful to define $z_t(i, l)$ to be the component of y_{it} that reflects the cumulative effects of the lth element of $\boldsymbol{\varepsilon}$:

$$z_t(i, l) = \psi_{il}^{(0)}\varepsilon_{lt} + \psi_{il}^{(1)}\varepsilon_{l,t-1} + \psi_{il}^{(2)}\varepsilon_{l,t-2} + \cdots = \sum_{v=0}^{\infty} \psi_{il}^{(v)}\varepsilon_{l,t-v}, \qquad [10.A.4]$$

where $\psi_{il}^{(v)}$ denotes the row i, column l element of the matrix $\boldsymbol{\Psi}_v$. The actual value of the ith variable y_{it} is the sum of the contributions of each of the $l = 1, 2, \ldots, n$ components of $\boldsymbol{\varepsilon}$:

$$y_{it} = \mu_i + \sum_{l=1}^{n} z_t(i, l). \qquad [10.A.5]$$

The results of Proposition 10.2 are all established by first demonstrating absolute summability of the moments of $z_t(i, l)$ and then observing that the moments of $\mathbf{y}_t$ are obtained from finite sums of these expressions based on $z_t(i, l)$.

Proof of (a). Consider the random variable $z_t(i, l)\cdot z_{t-s}(j, m)$, where i, l, j, and m represent arbitrary indices between 1 and n and where s is the order of the autocovariance of $\mathbf{y}$ that is being calculated. Note from [10.A.4] that

$$\begin{aligned} E\{z_t(i, l)\cdot z_{t-s}(j, m)\} &= E\left\{\left[\sum_{r=0}^{\infty} \psi_{il}^{(r)}\varepsilon_{l,t-r}\right] \times \left[\sum_{v=0}^{\infty} \psi_{jm}^{(v)}\varepsilon_{m,t-s-v}\right]\right\} \\ &= \sum_{r=0}^{\infty}\sum_{v=0}^{\infty} \{\psi_{il}^{(r)}\psi_{jm}^{(v)}\}\cdot E\{\varepsilon_{l,t-r}\varepsilon_{m,t-s-v}\}. \end{aligned} \qquad [10.A.6]$$

The expectation operator can be moved inside the summation here because

$$\sum_{r=0}^{\infty}\sum_{v=0}^{\infty} |\psi_{il}^{(r)}\psi_{jm}^{(v)}| = \sum_{r=0}^{\infty}\sum_{v=0}^{\infty} |\psi_{il}^{(r)}|\cdot|\psi_{jm}^{(v)}| = \left\{\sum_{r=0}^{\infty} |\psi_{il}^{(r)}|\right\} \times \left\{\sum_{v=0}^{\infty} |\psi_{jm}^{(v)}|\right\} < \infty.$$

Now, the product of ε's in the final term in [10.A.6] can have nonzero expectation only if the ε's have the same date, that is, if $r = s + v$. Thus, although [10.A.6] involves a sum over an infinite number of values of r, only the value at $r = s + v$ contributes to this sum:

$$E\{z_t(i, l)\cdot z_{t-s}(j, m)\} = \sum_{v=0}^{\infty} \{\psi_{il}^{(s+v)}\psi_{jm}^{(v)}\}\cdot E\{\varepsilon_{l,t-s-v}\varepsilon_{m,t-s-v}\} = \sum_{v=0}^{\infty} \psi_{il}^{(s+v)}\psi_{jm}^{(v)}\sigma_{lm}, \qquad [10.A.7]$$

where σ_{lm} represents the covariance between ε_{lt} and ε_{mt} and is given by the row l, column m element of $\boldsymbol{\Omega}$.

The row i, column j element of $\boldsymbol{\Gamma}_s$ gives the value of

$$\gamma_{ij}^{(s)} = E(y_{it} - \mu_i)(y_{j,t-s} - \mu_j).$$

Using [10.A.5] and [10.A.7], this can be expressed as

$$
\begin{aligned}
E(y_{it} - \mu_i)(y_{j,t-s} - \mu_j) &= E\left\{\left[\sum_{l=1}^{n} z_t(i, l)\right]\left[\sum_{m=1}^{n} z_{t-s}(j, m)\right]\right\} \\
&= \sum_{l=1}^{n} \sum_{m=1}^{n} E\{z_t(i, l) \cdot z_{t-s}(j, m)\} \\
&= \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{v=0}^{\infty} \psi_{il}^{(s+v)} \psi_{jm}^{(v)} \sigma_{lm} \\
&= \sum_{v=0}^{\infty} \sum_{l=1}^{n} \sum_{m=1}^{n} \psi_{il}^{(s+v)} \psi_{jm}^{(v)} \sigma_{lm}.
\end{aligned}
\qquad [10.A.8]
$$

But $\Sigma_{l=1}^{n} \Sigma_{m=1}^{n} \psi_{il}^{(s+v)} \psi_{jm}^{(v)} \sigma_{lm}$ is the row i, column j element of $\mathbf{\Psi}_{v+s} \mathbf{\Omega} \mathbf{\Psi}_v'$. Thus, [10.A.8] states that the row i, column j element of $\mathbf{\Gamma}_s$ is given by the row i, column j element of $\Sigma_{v=0}^{\infty} \mathbf{\Psi}_{v+s} \mathbf{\Omega} \mathbf{\Psi}_v'$, as asserted in part (a).

Proof of (b). Define $h_s(\cdot)$ to be the moment in [10.A.7]:

$$
h_s(i, j, l, m) \equiv E\{z_t(i, l) \cdot z_{t-s}(j, m)\} = \sum_{v=0}^{\infty} \psi_{il}^{(s+v)} \psi_{jm}^{(v)} \sigma_{lm};
$$

and notice that the sequence $\{h_s(\cdot)\}_{s=0}^{\infty}$ is absolutely summable:

$$
\begin{aligned}
\sum_{s=0}^{\infty} |h_s(i, j, l, m)| &\leq \sum_{s=0}^{\infty} \sum_{v=0}^{\infty} |\psi_{il}^{(s+v)}| \cdot |\psi_{jm}^{(v)}| \cdot |\sigma_{lm}| \\
&= |\sigma_{lm}| \sum_{v=0}^{\infty} |\psi_{jm}^{(v)}| \sum_{s=0}^{\infty} |\psi_{il}^{(s+v)}| \\
&\leq |\sigma_{lm}| \sum_{v=0}^{\infty} |\psi_{jm}^{(v)}| \sum_{s=0}^{\infty} |\psi_{il}^{(s)}| \\
&< \infty.
\end{aligned}
\qquad [10.A.9]
$$

Furthermore, the row i, column j element of $\mathbf{\Gamma}_s$ was seen in [10.A.8] to be given by

$$
\gamma_{ij}^{(s)} = \sum_{l=1}^{n} \sum_{m=1}^{n} h_s(i, j, l, m).
$$

Hence,

$$
\sum_{s=0}^{\infty} |\gamma_{ij}^{(s)}| \leq \sum_{s=0}^{\infty} \sum_{l=1}^{n} \sum_{m=1}^{n} |h_s(i, j, l, m)| = \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{s=0}^{\infty} |h_s(i, j, l, m)|. \qquad [10.A.10]
$$

From [10.A.9], there exists an $M < \infty$ such that

$$
\sum_{s=0}^{\infty} |h_s(i, j, l, m)| < M
$$

for any value of i, j, l, or m. Hence, [10.A.10] implies

$$
\sum_{s=0}^{\infty} |\gamma_{ij}^{(s)}| < \sum_{l=1}^{n} \sum_{m=1}^{n} M = n^2 M < \infty,
$$

confirming that the row i, column j element of $\{\mathbf{\Gamma}_s\}_{s=0}^{\infty}$ is absolutely summable, as claimed by part (b).

Proof of (c). Essentially the identical algebra as in the proof of Proposition 7.10 establishes that

$$\begin{aligned}
&E|z_{t_1}(i_1, l_1) \cdot z_{t_2}(i_2, l_2) \cdot z_{t_3}(i_3, l_3) \cdot z_{t_4}(i_4, l_4)| \\
&= E\left|\left\{\sum_{v_1=0}^{\infty} \psi_{i_1,l_1}^{(v_1)} \varepsilon_{l_1,t_1-v_1}\right\} \cdot \left\{\sum_{v_2=0}^{\infty} \psi_{i_2,l_2}^{(v_2)} \varepsilon_{l_2,t_2-v_2}\right\}\right. \\
&\quad \left. \cdot \left\{\sum_{v_3=0}^{\infty} \psi_{i_3,l_3}^{(v_3)} \varepsilon_{l_3,t_3-v_3}\right\} \cdot \left\{\sum_{v_4=0}^{\infty} \psi_{i_4,l_4}^{(v_4)} \varepsilon_{l_4,t_4-v_4}\right\}\right| \qquad [10.A.11] \\
&\le \sum_{v_1=0}^{\infty} \sum_{v_2=0}^{\infty} \sum_{v_3=0}^{\infty} \cdot \sum_{v_4=0}^{\infty} |\psi_{i_1,l_1}^{(v_1)} \psi_{i_2,l_2}^{(v_2)} \psi_{i_3,l_3}^{(v_3)} \psi_{i_4,l_4}^{(v_4)}| \\
&\quad \times E|\varepsilon_{l_1,t_1-v_1} \varepsilon_{l_2,t_2-v_2} \varepsilon_{l_3,t_3-v_3} \varepsilon_{l_4,t_4-v_4}| \\
&< \infty.
\end{aligned}$$

Now,

$$\begin{aligned}
E|y_{i_1,t_1} y_{i_2,t_2} y_{i_3,t_3} y_{i_4,t_4}| &= E\left|\mu_{i_1} + \sum_{l_1=1}^{n} z_{t_1}(i_1, l_1)\right| \cdot \left|\mu_{i_2} + \sum_{l_2=1}^{n} z_{t_2}(i_2, l_2)\right| \\
&\quad \cdot \left|\mu_{i_3} + \sum_{l_3=1}^{n} z_{t_3}(i_3, l_3)\right| \cdot \left|\mu_{i_4} + \sum_{l_4=1}^{n} z_{t_4}(i_4, l_4)\right| \\
&\le E\left\{|\mu_{i_1}| + \sum_{l_1=1}^{n} |z_{t_1}(i_1, l_1)|\right\} \cdot \left\{|\mu_{i_2}| + \sum_{l_2=1}^{n} |z_{t_2}(i_2, l_2)|\right\} \\
&\quad \cdot \left\{|\mu_{i_3}| + \sum_{l_3=1}^{n} |z_{t_3}(i_3, l_3)|\right\} \cdot \left\{|\mu_{i_4}| + \sum_{l_4=1}^{n} |z_{t_4}(i_4, l_4)|\right\}.
\end{aligned}$$

But this is a finite sum involving terms of the form of [10.A.11]—which were seen to be finite—along with terms involving first through third moments of z, which must also be finite.

Proof of (d). Notice that

$$z_t(i, l) \cdot z_{t-s}(j, m) = \sum_{r=0}^{\infty} \sum_{v=0}^{\infty} \psi_{il}^{(r)} \psi_{jm}^{(v)} \varepsilon_{l,t-r} \varepsilon_{m,t-s-v}.$$

The same argument leading to [7.2.14] can be used to establish that

$$(1/T) \sum_{t=1}^{T} z_t(i, l) \cdot z_{t-s}(j, m) \xrightarrow{p} E\{z_t(i, l) \cdot z_{t-s}(j, m)\}. \qquad [10.A.12]$$

To see that [10.A.12] implies ergodicity for the second moments of $\mathbf{y}$, notice from [10.A.5] that

$$\begin{aligned}
(1/T) \sum_{t=1}^{T} y_{it} y_{j,t-s} &= (1/T) \sum_{t=1}^{T} \left[\mu_i + \sum_{l=1}^{n} z_t(i, l)\right]\left[\mu_j + \sum_{m=1}^{n} z_{t-s}(j, m)\right] \\
&= \mu_i \mu_j + \mu_i \sum_{m=1}^{n} \left[(1/T) \sum_{t=1}^{T} z_{t-s}(j, m)\right] + \mu_j \sum_{l=1}^{n} \left[(1/T) \sum_{t=1}^{T} z_t(i, l)\right] \\
&\quad + \sum_{l=1}^{n} \sum_{m=1}^{n} \left[(1/T) \sum_{t=1}^{T} z_t(i, l) z_{t-s}(j, m)\right] \\
&\xrightarrow{p} \mu_i \mu_j + \mu_i \sum_{m=1}^{n} E[z_{t-s}(j, m)] + \mu_j \sum_{l=1}^{n} E[z_t(i, l)] \\
&\quad + \sum_{l=1}^{n} \sum_{m=1}^{n} E[z_t(i, l) z_{t-s}(j, m)] \\
&= E\left\{\left[\mu_i + \sum_{l=1}^{n} z_t(i, l)\right]\left[\mu_j + \sum_{m=1}^{n} z_{t-s}(j, m)\right]\right\} \\
&= E[y_{it} y_{j,t-s}],
\end{aligned}$$

as claimed. ■

■ **Proof of Proposition 10.3.** Writing out [10.2.11] explicitly,

$$\begin{aligned}\mathbf{H}(L)\boldsymbol{\Psi}(L) &= (\cdots + \mathbf{H}_{-1}L^{-1} + \mathbf{H}_0L^0 + \mathbf{H}_1L^1 + \cdots)\\ &\quad \times (\boldsymbol{\Psi}_0L^0 + \boldsymbol{\Psi}_1L^1 + \boldsymbol{\Psi}_2L^2 + \cdots),\end{aligned}$$

from which the coefficient on L^k is

$$\mathbf{B}_k = \mathbf{H}_k\boldsymbol{\Psi}_0 + \mathbf{H}_{k-1}\boldsymbol{\Psi}_1 + \mathbf{H}_{k-2}\boldsymbol{\Psi}_2 + \cdots. \qquad [10.A.13]$$

Let $b_{ij}^{(k)}$ denote the row i, column j element of $\mathbf{B}_k$, and let $h_{ij}^{(k)}$ and $\psi_{ij}^{(k)}$ denote the row i, column j elements of $\mathbf{H}_k$ and $\boldsymbol{\Psi}_k$, respectively. Then the row i, column j element of the matrix equation [10.A.13] states that

$$b_{ij}^{(k)} = \sum_{m=1}^{n} h_{im}^{(k)}\psi_{mj}^{(0)} + \sum_{m=1}^{n} h_{im}^{(k-1)}\psi_{mj}^{(1)} + \sum_{m=1}^{n} h_{im}^{(k-2)}\psi_{mj}^{(2)} + \cdots = \sum_{v=0}^{\infty}\sum_{m=1}^{n} h_{im}^{(k-v)}\psi_{mj}^{(v)}.$$

Thus,

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |b_{ij}^{(k)}| &= \sum_{k=-\infty}^{\infty}\left|\sum_{v=0}^{\infty}\sum_{m=1}^{n} h_{im}^{(k-v)}\psi_{mj}^{(v)}\right|\\ &\le \sum_{k=-\infty}^{\infty}\sum_{v=0}^{\infty}\sum_{m=1}^{n} |h_{im}^{(k-v)}\psi_{mj}^{(v)}| \qquad [10.A.14]\\ &= \sum_{m=1}^{n}\sum_{v=0}^{\infty} |\psi_{mj}^{(v)}| \sum_{k=-\infty}^{\infty} |h_{im}^{(k-v)}|.\end{aligned}$$

But since $\{\mathbf{H}_k\}_{k=-\infty}^{\infty}$ and $\{\boldsymbol{\Psi}_k\}_{k=0}^{\infty}$ are absolutely summable,

$$\sum_{k=-\infty}^{\infty} |h_{im}^{(k-v)}| < M_1 < \infty$$

$$\sum_{v=0}^{\infty} |\psi_{mj}^{(v)}| < M_2 < \infty.$$

Thus, [10.A.14] becomes

$$\sum_{k=-\infty}^{\infty} |b_{ij}^{(k)}| < \sum_{m=1}^{n} M_1M_2 < \infty. \quad ■$$

■ **Proof of Proposition 10.4.** Let $\mathbf{A}$ be $(m \times n)$, $\mathbf{B}$ be $(n \times r)$, and $\mathbf{C}$ be $(r \times q)$. Let the $(n \times 1)$ vector $\mathbf{b}_i$ denote the ith column of $\mathbf{B}$, and let c_{ij} denote the row i, column j element of $\mathbf{C}$. Then

$$\begin{aligned}\mathbf{ABC} &= \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_r]\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1q}\\ c_{21} & c_{22} & \cdots & c_{2q}\\ \vdots & \vdots & \cdots & \vdots\\ c_{r1} & c_{r2} & \cdots & c_{rq}\end{bmatrix}\\ &= [\{\mathbf{Ab}_1c_{11} + \mathbf{Ab}_2c_{21} + \cdots + \mathbf{Ab}_rc_{r1}\}\\ &\qquad \{\mathbf{Ab}_1c_{12} + \mathbf{Ab}_2c_{22} + \cdots + \mathbf{Ab}_rc_{r2}\} \cdots\\ &\qquad \{\mathbf{Ab}_1c_{1q} + \mathbf{Ab}_2c_{2q} + \cdots + \mathbf{Ab}_rc_{rq}\}]\\ &= [\{c_{11}\mathbf{Ab}_1 + c_{21}\mathbf{Ab}_2 + \cdots + c_{r1}\mathbf{Ab}_r\}\\ &\qquad \{c_{12}\mathbf{Ab}_1 + c_{22}\mathbf{Ab}_2 + \cdots + c_{r2}\mathbf{Ab}_r\} \cdots\\ &\qquad \{c_{1q}\mathbf{Ab}_1 + c_{2q}\mathbf{Ab}_2 + \cdots + c_{rq}\mathbf{Ab}_r\}].\end{aligned}$$

Applying the vec operator gives

$$\begin{aligned}\operatorname{vec}(\mathbf{ABC}) &= \begin{bmatrix} c_{11}\mathbf{Ab}_1 + c_{21}\mathbf{Ab}_2 + \cdots + c_{r1}\mathbf{Ab}_r\\ c_{12}\mathbf{Ab}_1 + c_{22}\mathbf{Ab}_2 + \cdots + c_{r2}\mathbf{Ab}_r\\ \vdots\\ c_{1q}\mathbf{Ab}_1 + c_{2q}\mathbf{Ab}_2 + \cdots + c_{rq}\mathbf{Ab}_r\end{bmatrix}\\ &= \begin{bmatrix} c_{11}\mathbf{A} & c_{21}\mathbf{A} & \cdots & c_{r1}\mathbf{A}\\ c_{12}\mathbf{A} & c_{22}\mathbf{A} & \cdots & c_{r2}\mathbf{A}\\ \vdots & \vdots & \cdots & \vdots\\ c_{1q}\mathbf{A} & c_{2q}\mathbf{A} & \cdots & c_{rq}\mathbf{A}\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\\ \vdots\\ \mathbf{b}_r\end{bmatrix}\\ &= (\mathbf{C}' \otimes \mathbf{A})\cdot\operatorname{vec}(\mathbf{B}). \quad ■\end{aligned}$$

Chapter 10 Exercises

10.1. Consider a scalar $AR(p)$ process ($n = 1$). Deduce from equation [10.2.19] that the ($p \times 1$) vector consisting of the variance and first ($p - 1$) autocovariances,

$$\begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{p-1} \end{bmatrix},$$

can be calculated from the first p elements in the first column of the ($p^2 \times p^2$) matrix $\sigma^2[\mathbf{I}_{p^2} - (\mathbf{F} \otimes \mathbf{F})]^{-1}$ for $\mathbf{F}$ the ($p \times p$) matrix defined in equation [1.2.3] in Chapter 1.

10.2. Let $\mathbf{y}_t = (X_t, Y_t)'$ be given by

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}$$
$$Y_t = h_1 X_{t-1} + u_t,$$

where $(\varepsilon_t, u_t)'$ is vector white noise with contemporaneous variance-covariance matrix given by

$$\begin{bmatrix} E(\varepsilon_t^2) & E(\varepsilon_t u_t) \\ E(u_t\varepsilon_t) & E(u_t^2) \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}.$$

(a) Calculate the autocovariance matrices $\{\boldsymbol{\Gamma}_k\}_{k=-\infty}^{\infty}$ for this process.

(b) Use equation [10.4.3] to calculate the population spectrum. Find the cospectrum between X and Y and the quadrature spectrum from X to Y.

(c) Verify that your answer to part (b) could equivalently be calculated from expression [10.4.45].

(d) Verify by integrating your answer to part (b) that [10.4.49] holds; that is, show that

$$(2\pi)^{-1}\int_{-\pi}^{\pi} \frac{s_{YX}(\omega)}{s_{XX}(\omega)} e^{i\omega k}\, d\omega = \begin{cases} h_1 & \text{for } k = 1 \\ 0 & \text{for other integer } k. \end{cases}$$

Chapter 10 References

Andrews, Donald W. K. 1991. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59:817–58.

——— and J. Christopher Monahan. 1992. "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica* 60:953–66.

Fuller, Wayne A. 1976. *Introduction to Statistical Time Series.* New York: Wiley.

Gallant, A. Ronald. 1987. *Nonlinear Statistical Models.* New York: Wiley.

Hansen, Lars P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50:1029–54.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55:703–8.

——— and ———. 1992. "Automatic Lag Selection in Covariance Matrix Estimation." University of Wisconsin, Madison, Mimeo.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica* 48:1–48.

White, Halbert. 1984. *Asymptotic Theory for Econometricians.* Orlando, Fla.: Academic Press.

11

Vector Autoregressions

The previous chapter introduced some basic tools for describing vector time series processes. This chapter looks in greater depth at vector autoregressions, which are particularly convenient for estimation and forecasting. Their popularity for analyzing the dynamics of economic systems is due to Sims's (1980) influential work. The chapter begins with a discussion of maximum likelihood estimation and hypothesis testing. Section 11.2 examines a concept of causation in bivariate systems proposed by Granger (1969). Section 11.3 generalizes the discussion of Granger causality to multivariate systems and examines estimation of restricted vector autoregressions. Sections 11.4 and 11.5 introduce impulse-response functions and variance decompositions, which are used to summarize the dynamic relations between the variables in a vector autoregression. Section 11.6 reviews how such summaries can be used to evaluate structural hypotheses. Section 11.7 develops formulas needed to calculate standard errors for impulse-response functions.

11.1. *Maximum Likelihood Estimation and Hypothesis Testing for an Unrestricted Vector Autoregression*

The Conditional Likelihood Function for a Vector Autoregression

Let $\mathbf{y}_t$ denote an $(n \times 1)$ vector containing the values that n variables assume at date t. The dynamics of $\mathbf{y}_t$ are presumed to be governed by a pth-order Gaussian vector autoregression,

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \qquad [11.1.1]$$

with $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $N(\mathbf{0}, \mathbf{\Omega})$.

Suppose we have observed each of these n variables for $(T + p)$ time periods. As in the scalar autoregression, the simplest approach is to condition on the first p observations (denoted $\mathbf{y}_{-p+1}, \mathbf{y}_{-p+2}, \ldots, \mathbf{y}_0$) and to base estimation on the last T observations (denoted $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$). The objective then is to form the conditional likelihood

$$f_{\mathbf{Y}_T,\mathbf{Y}_{T-1},\ldots,\mathbf{Y}_1|\mathbf{Y}_0,\mathbf{Y}_{-1},\ldots,\mathbf{Y}_{-p+1}}(\mathbf{y}_T, \mathbf{y}_{T-1}, \ldots, \mathbf{y}_1|\mathbf{y}_0, \mathbf{y}_{-1}, \ldots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \qquad [11.1.2]$$

and maximize with respect to $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is a vector that contains the elements of $\mathbf{c}, \mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_p$, and $\mathbf{\Omega}$. Vector autoregressions are invariably estimated on the basis of the conditional likelihood function [11.1.2] rather than the full-sample