APPENDIX 9

Matrices and Polynomials

The Multiplication of Polynomials

Let $\alpha(z) = \alpha_0+\alpha_1 z+\alpha_2 z^2+\cdots \alpha_p z^p$ and $y(z) = y_0+y_1 z+y_2 z^2+\cdots y_n z^n$ be two polynomials of degrees p and n respectively. Then, their product $\gamma(z) = \alpha(z)y(z)$ is a polynomial of degree $p+n$ of which the coefficients comprise combinations of the coefficient of $\alpha(z)$ and $y(z)$.

A simple way of performing the multiplication is via a table of which the margins contain the elements of the two polynomials and in which the cells contain their products. An example of such a table is given below:

$$
(1) \qquad \begin{array}{c|ccc}
 & \alpha_0 & \alpha_1 z & \alpha_2 \\
\hline
y_0 & \alpha_0 y_0 & \alpha_1 y_0 z & \alpha_2 y_0 z^2 \\
y_1 z & \alpha_0 y_1 z & \alpha_1 y_1 z^2 & \alpha_2 y_0 z^3 \\
y_2 z^2 & \alpha_0 y_1 z^2 & \alpha_1 y_2 z^3 & \alpha_2 y_2 z^4
\end{array}
$$

The product is formed by adding all of the elements of the cells. However, if the elements on the SW–NE diagonal are gathered together, then a power of the argument z can be factored from their sum and then the associated coefficient is a coefficient of the product polynomial.

The following is an example from the table above:

$$
(3) \qquad \begin{array}{lcl}
\gamma_0 & & \alpha_0 y_0 \\
+\gamma_1 z & & +(\alpha_0 y_1 + \alpha_1 y_0)z \\
+\gamma_2 z^2 & = & +(\alpha_0 y_2 + \alpha_1 y_1 + \alpha_2 y_0 +)z^2 \\
+\gamma_3 z^3 & & +(\alpha_1 y_2 + \alpha_2 y_1)z^3 \\
+\gamma_4 z^4 & & +\alpha_2 y_4 z^4.
\end{array}
$$

The coefficients of the product polynomial can also be seen as the products of the convolutions of the sequences $\{\alpha_0, \alpha_1, \alpha_2 \ldots \alpha_p\}$ and $\{y_0, y_1, y_2 \ldots y_n\}$.

The coefficients of the product polynomials can also be generated by a simple multiplication of a matrix by a vector. Thus, from the example, we should have

$$
(3)\qquad \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_2 \\ \gamma_v \end{bmatrix} = \begin{bmatrix} y_0 & 0 & 0 \\ y_1 & y_0 & 0 \\ y_2 & y_1 & y_0 \\ 0 & y_2 & y_1 \\ 0 & 0 & y_2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \alpha_0 & 0 & 0 \\ \alpha_1 & \alpha_0 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \\ 0 & \alpha_2 & \alpha_1 \\ 0 & 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}.
$$

To form the elements of the product polynomial $\gamma(z)$, powers of z may be associated with elements of the matrices and the vectors of values indicated by the subscripts.

The argument z is usually described as an algebraic indeterminate. Its place can be taken by any of a wide variety of operators. Examples are provided by the difference operator and the lag operator that are defined in respect of doubly-infinite sequences.

It is also possible to replace z by matrices. However, the fundamental theorem of algebra indicates that all polynomial equations must have solutions that lie in the complex plane. Therefore, it is customary, albeit unnecessary, to regard z as a complex number.

Polynomials with Matrix Arguments

Toeplitz Matrices

There are two matrix arguments of polynomials that are of particular interest in time series analysis. The first is the matrix lag operator. The operator of order T denoted by

$$
(4)\qquad L_T = [e_1, e_2, \ldots, e_{T-1}, 0]
$$

is formed from the identity matrix $I_T = [e_0, e_1, \ldots, e_{T-1}]$ by deleting the leading vector e_0 and by appending a column of zeros to the end of the array. In effect, L_T is the matrix with units on the first subdiagonal band and with zeros elsewhere. Likewise, L_T^2 has units on the second subdiagonal band and with zeros elsewhere, whereas L_T^{T-1} has a single unit in the bottom left (i.e. S–W) corner, and $L_T^{T+r} = 0$ for all $r \geq 0$. In addition, $L_T^0 = I_T$ is identified with the identity matrix of order T. The example of L_4 is given below:

$$L_4^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad L_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$
$$(5)$$
$$L_4^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad L_4^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Putting L_T in place of the argument z of a polynomial in non-negative powers creates a so-called Toeplitz banded lower-triangular matrix of order T. An example is provided by the quadratic polynomials $\alpha(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2$ and $y(z) = y_0 + y_1 z + y_2 z^2$. Then, there is $\alpha(L_3)y(L_3) = y(L_3)\alpha(L_3)$, which is written explicitly as

$$(6) \qquad \begin{bmatrix} \alpha_0 & 0 & 0 \\ \alpha_1 & \alpha_0 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} y_0 & 0 & 0 \\ y_1 & y_0 & 0 \\ y_2 & y_1 & y_0 \end{bmatrix} = \begin{bmatrix} y_0 & 0 & 0 \\ y_1 & y_0 & 0 \\ y_2 & y_1 & y_0 \end{bmatrix} \begin{bmatrix} \alpha_0 & 0 & 0 \\ \alpha_1 & \alpha_0 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix}.$$

The commutativity of the two matrices in multiplication reflects their polynomial nature. Such commutativity is available both for lower-triangular Toeplitz and for upper-triangular Toeplitz matrices, which correspond to polynomials in negative powers of z.

The commutativity is not available for mixtures of upper and lower triangular matrices; and, in this respect, the matrix algebra differs from the corresponding polynomial algebra. An example is provided by the matrix version of the following polynomial identity:

$$(7) \qquad (1 - z)(1 - z^{-1}) = 2z^0 - (z + z^{-1}) = (1 - z^{-1})(1 - z)$$

Putting L_T in place of z in each of these expressions creates three different matrices. This can be illustrated with the case of L_3. Then, $(1 - z)(1 - z^{-1})$ gives rise to

$$(8) \qquad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix},$$

whereas $(1 - z^{-1})(1 - z)$ gives rise to

$$(9) \qquad \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

A Toeplitz matrix, in which each band contains only repetitions of the same element, is obtained from the remaining expression by replacing z by L_3 in $2z^0 - (z + z^{-1})$, wherein z^0 is replaced by the identity matrix:

$$(10) \qquad \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Example. It is straightforward to derive the dispersion matrices that are found within the formulae for the finite-sample estimators from the corresponding autocovariance generating functions. Let $\gamma(z) = \{\gamma_0 + \gamma_1(z + z^{-1}) + \gamma_2(z^2 + z^{-2}) + \cdots\}$ denote the autocovariance generating function of a stationary stochastic process. Then, the corresponding dispersion matrix for a sample of T consecutive elements drawn from the process is

$$(11) \qquad \Gamma = \gamma_0 I_T + \sum_{\tau=1}^{T-1} \gamma_\tau (L_T^\tau + F_T^\tau),$$

where $F_T = L'_T$ is in place of z^{-1}. Since L_T and F_T are nilpotent of degree T, such that $L_T^q, F_T^q = 0$ when $q \geq T$, the index of summation has an upper limit of $T - 1$.

Circulant Matrices

In the second of the matrix representations, which is appropriate to a frequency-domain interpretation of filtering, the argument z is replaced by the full-rank circulant matrix

$$(12) \qquad K_T = [e_1, e_2, \ldots, e_{T-1}, e_0],$$

which is obtained from the identity matrix $I_T = [e_0, e_1, \ldots, e_{T-1}]$ by displacing the leading column to the end of the array. This is an orthonormal matrix of which the transpose is the inverse, such that $K'_T K_T = K_T K'_T = I_T$. The powers of the matrix form a T-periodic sequence such that $K_T^{T+q} = K_T^q$ for all q. The periodicity of these powers is analogous to the periodicity of the powers of the argument $z = \exp\{-\mathrm{i}2\pi/T\}$, which is to be found in the Fourier transform of a sequence of order T.

The example of K_4 is given below

$$
K_4^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad K_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},
$$

$$
(13) \qquad K_4^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad K_4^3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.
$$

The matrices $K_T^0 = I_T, K_T, \ldots, K_T^{T-1}$ form a basis for the set of all circulant matrices of order T—a circulant matrix $X = [x_{ij}]$ of order T being defined as a matrix in which the value of the generic element x_{ij} is determined by the index $\{(i-j) \bmod T\}$. This implies that each column of X is equal to the previous column rotated downwards by one element.

It follows that there exists a one-to-one correspondence between the set of all polynomials of degree less than T and the set of all circulant matrices of order T. Therefore, if $\alpha(z)$ is a polynomial of degree less that T, then there exits a corresponding circulant matrix

$$
(14) \qquad A = \alpha(K_T) = \alpha_0 I_T + \alpha_1 K_T + \cdots + \alpha_{T-1} K_T^{T-1}.
$$

A convergent sequence of an indefinite length can also be mapped into a circulant matrix. Thus, if $\{\gamma_i\}$ is an absolutely summable sequence obeying the condition that $\sum |\gamma_i| < \infty$, then the z-transform of the sequence, which is defined by $\gamma(z) = \sum \gamma_j z^j$, is an analytic function on the unit circle. In that case, replacing z by K_T gives rise to a circulant matrix $\Gamma = \gamma(K_T)$ with finite-valued elements. In consequence of the periodicity of the powers of K_T, it follows that

$$
(15) \qquad \begin{aligned} \Gamma &= \left\{ \sum_{j=0}^{\infty} \gamma_{jT} \right\} I_T + \left\{ \sum_{j=0}^{\infty} \gamma_{(jT+1)} \right\} K_T + \cdots + \left\{ \sum_{j=0}^{\infty} \gamma_{(jT+T-1)} \right\} K^{T-1} \\ &= \varphi_0 I_T + \varphi_1 K_T + \cdots + \varphi_{T-1} K_T^{T-1}. \end{aligned}
$$

Given that $\{\gamma_i\}$ is a convergent sequence, it follows that the sequence of the matrix coefficients $\{\varphi_0, \varphi_1, \ldots, \varphi_{T-1}\}$ converges to $\{\gamma_0, \gamma_1, \ldots, \gamma_{T-1}\}$ as T increases. Notice that the matrix $\varphi(K) = \varphi_0 I_T + \varphi_1 K_T + \cdots + \varphi_{T-1} K_T^{T-1}$, which is derived from a polynomial $\varphi(z)$ of degree $T-1$, is a synonym for the matrix $\gamma(K_T)$, which is derived from the z-transform of an infinite convergent sequence.

The polynomial representation is enough to establish that circulant matrices commute in multiplication and that their product is also a polynomial in K_T. That is to say

(16) If $X = x(K_T)$ and $Y = y(K_T)$ are circulant matrices, then $XY = YX$ is also a circulant matrix.

The matrix operator K_T has a spectral factorisation that is particularly useful in analysing the properties of the discrete Fourier transform. To demonstrate this factorisation, we must first define the so-called Fourier matrix. This is a symmetric matrix

$$U_T = T^{-1/2}[W_T^{jt}; t, j = 0, \ldots, T-1], \tag{17}$$

of which the generic element in the jth row and tth column is

$$\begin{aligned} W_T^{jt} &= \exp(-\mathrm{i}2\pi tj/T) = \cos(\omega_j t) - \mathrm{i}\sin(\omega_j t), \\ &\text{where} \quad \omega_j = 2\pi j/T. \end{aligned} \tag{18}$$

The matrix U_T is a unitary, which is to say that it fulfils the condition

$$\bar{U}_T U_T = U_T \bar{U}_T = I_T, \tag{19}$$

where $\bar{U}_T = T^{-1/2}[W_T^{-jt}; t, j = 0, \ldots, T-1]$ denotes the conjugate matrix.

The operator can be factorised as

$$K_T = \bar{U}_T D_T U_T = U_T \bar{D}_T \bar{U}_T, \tag{20}$$

where

$$D_T = \mathrm{diag}\{1, W, W^2, \ldots, W^{T-1}\} \tag{21}$$

is a diagonal matrix whose elements are the T roots of unity, which are found on the circumference of the unit circle in the complex plane. Observe also that D_T is T-periodic, such that $D_T^{q+T} = D_T^q$, and that $K_T^q = \bar{U}_T D_T^q U_T = U_T \bar{D}_T^q \bar{U}_T$ for any integer q. Since the powers of K_T form the basis for the set of circulant matrices, it follows that such matrices are amenable to a spectral factorisation based on (13).

Example. Consider, in particular, the circulant autocovariances matrix that is obtained by replacing the argument z in the autocovariance generating function $\gamma(z)$ by the matrix K_T. Imagine that the autocovariances form a doubly infinite sequence, as is the case for an autoregressive or an autoregressive moving-average process:

$$\begin{aligned}\Omega^\circ = \gamma(K_T) &= \gamma_0 I_T + \sum_{\tau=1}^{\infty} \gamma_\tau (K_T^\tau + K_T^{-\tau}) \\ &= \varphi_0 I_T + \sum_{\tau=1}^{T-1} \varphi_\tau (K_T^\tau + K_T^{-\tau}).\end{aligned} \tag{22}$$

Here, $\varphi_\tau; \tau = 0, \ldots, T-1$ are the "wrapped" coefficients that are obtained from the original coefficients of the autocovariance generating function in the manner indicated by (15). The spectral factorisation gives

$$\Omega^\circ = \gamma(K_T) = \bar{U}\gamma(D)U. \tag{23}$$

The jth element of the diagonal matrix $\gamma(D)$ is

$$\gamma(\exp\{\mathrm{i}\omega_j\}) = \gamma_0 + 2\sum_{\tau=1}^{\infty} \gamma_\tau \cos(\omega_j \tau). \tag{24}$$

This represents the cosine Fourier transform of the sequence of the ordinary autocovariances; and it corresponds to an ordinate (scaled by 2π) sampled at the point ω_j from the spectral density function of the linear (i.e. non-circular) stationary stochastic process.