LECTURE 3

Models and Methods of Time-Series Analysis

A time-series model is one which postulates a relationship amongst a number of temporal sequences or time series. An example is provided by the simple regression model

$$y(t) = x(t)\beta + \varepsilon(t), \tag{3.1}$$

where $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence, indexed by the time subscript t, which is a combination of an observable signal sequence $x(t) = \{x_t\}$ and an unobservable white-noise sequence $\varepsilon(t) = \{\varepsilon_t\}$ of independently and identically distributed random variables.

A more general model, which we shall call the general temporal regression model, is one which postulates a relationship comprising any number of consecutive elements of $x(t)$, $y(t)$ and $\varepsilon(t)$. The model may be represented by the equation

$$\sum_{i=0}^{p} \alpha_i y(t-i) = \sum_{i=0}^{k} \beta_i x(t-i) + \sum_{i=0}^{q} \mu_i \varepsilon(t-i), \tag{3.2}$$

where it is usually taken for granted that $\alpha_0 = 1$. This normalisation of the leading coefficient on the LHS identifies $y(t)$ as the output sequence. Any of the sums in the equation can be infinite, but if the model is to be viable, the sequences of coefficients $\{\alpha_i\}$, $\{\beta_i\}$ and $\{\mu_i\}$ can depend on only a limited number of parameters.

Although it is convenient to write the general model in the form of (2), it is also common to represent it by the equation

$$y(t) = \sum_{i=1}^{p} \phi_i y(t-i) + \sum_{i=0}^{k} \beta_i x(t-i) + \sum_{i=0}^{q} \mu_i \varepsilon(t-i), \tag{3.3}$$

where $\phi_i = -\alpha_i$ for $i = 1, \ldots, p$. This places the lagged versions of the sequence $y(t)$ on the RHS in the company of the input sequence $x(t)$ and its lags.

Whereas engineers are liable to describe this as a feedback model, economists are more likely to describe it as a model with lagged dependent variables.

The foregoing models are termed regression models by virtue of the inclusion of the observable explanatory sequence $x(t)$. When $x(t)$ is deleted, we obtain a simpler unconditional linear stochastic model:

$$\sum_{i=0}^{p} \alpha_i y(t-i) = \sum_{i=0}^{q} \mu_i \varepsilon(t-i). \tag{3.4}$$

This is the autoregressive moving-average (ARMA) model.

A time-series model can often assume a variety of forms. Consider a simple dynamic regression model of the form

$$y(t) = \phi y(t-1) + x(t)\beta + \varepsilon(t), \tag{3.5}$$

where there is a single lagged dependent variable. By repeated substitution, we obtain

$$\begin{aligned} y(t) &= \phi y(t-1) + \beta x(t) + \varepsilon(t) \\ &= \phi^2 y(t-2) + \beta\big\{x(t) + \phi x(t-1)\big\} + \varepsilon(t) + \phi\varepsilon(t-1) \\ &\;\;\vdots \\ &= \phi^n y(t-n) + \beta\big\{x(t) + \phi x(t-1) + \cdots + \phi^{n-1}x(t-n+1)\big\} \\ &\qquad + \varepsilon(t) + \phi\varepsilon(t-1) + \cdots + \phi^{n-1}\varepsilon(t-n+1). \end{aligned} \tag{3.6}$$

If $|\phi| < 1$, then $\lim(n \to \infty)\phi^n = 0$; and it follows that, if $x(t)$ and $\varepsilon(t)$ are bounded sequences, then, as the number of repeated substitutions increases indefinitely, the equation will tend to the limiting form of

$$y(t) = \beta \sum_{i=0}^{\infty} \phi^i x(t-i) + \sum_{i=0}^{\infty} \phi^i \varepsilon(t-i). \tag{3.7}$$

It is notable that, by this process of repeated substitution, the feedback structure has been eliminated from the model. As a result, it becomes easier to assess the impact upon the output sequence of changes in the values of the input sequence. The direct mapping from the input sequence to the output sequence is described by engineers as a transfer function or as a filter.

For models more complicated than the one above, the method of repeated substitution, if pursued directly, becomes intractable. Thus we are motivated to use more powerful algebraic methods to effect the transformation of the equation. This leads us to consider the use of the so-called lag operator. A proper understanding of the lag operator depends upon a knowledge of the algebra of polynomials and of rational functions.

The Algebra of the Lag Operator

A sequence $x(t) = \{x_t; t = 0, \pm 1, \pm 2, \ldots\}$ is any function mapping from the set of integers $\mathcal{Z} = \{0, \pm 1, \pm 2, \ldots\}$ to the real line. If the set of integers represents a set of dates separated by unit intervals, then $x(t)$ is described as a temporal sequence or a time series.

The set of all time series represents a vector space, and various linear transformations or operators can be defined over the space. The simplest of these is the lag operator L which is defined by

$$Lx(t) = x(t-1). \tag{3.8}$$

Now, $L\{Lx(t)\} = Lx(t-1) = x(t-2)$; so it makes sense to define L^2 by $L^2x(t) = x(t-2)$. More generally, $L^kx(t) = x(t-k)$ and, likewise, $L^{-k}x(t) = x(t+k)$. Other operators are the difference operator $\nabla = I - L$ which has the effect that

$$\nabla x(t) = x(t) - x(t-1), \tag{3.9}$$

the forward-difference operator $\Delta = L^{-1} - I$, and the summation operator $S = (I - L)^{-1} = \{I + L + L^2 + \cdots\}$ which has the effect that

$$Sx(t) = \sum_{i=0}^{\infty} x(t-i). \tag{3.10}$$

In general, we can define polynomials of the lag operator of the form $p(L) = p_0 + p_1L + \cdots + p_nL^n = \sum p_iL^i$ having the effect that

$$\begin{aligned} p(L)x(t) &= p_0x(t) + p_1x(t-1) + \cdots + p_nx(t-n) \\ &= \sum_{i=0}^{n} p_ix(t-i). \end{aligned} \tag{3.11}$$

In these terms, the equation under (2) of the general temporal model becomes

$$\alpha(L)y(t) = \beta(L)x(t) + \mu(L)\varepsilon(t). \tag{3.12}$$

The advantage which comes from defining polynomials in the lag operator stems from the fact that they are isomorphic to the set of ordinary algebraic polynomials. Thus we can rely upon what we know about ordinary polynomials to treat problems concerning lag-operator polynomials.

Algebraic Polynomials

Consider the equation $\phi_0 + \phi_1 z + \phi_2 z^2 = 0$. Once the equation has been divided by ϕ_2, it can be factorised as $(z - \lambda_1)(z - \lambda_2)$ where λ_1, λ_2 are the roots or zeros of the equation which are given by the formula

$$\lambda = \frac{-\phi_1 \pm \sqrt{\phi_1^2 - 4\phi_2\phi_0}}{2\phi_2}. \tag{3.13}$$

If $\phi_1^2 \geq 4\phi_2\phi_0$, then the roots λ_1, λ_2 are real. If $\phi_1^2 = 4\phi_2\phi_0$, then $\lambda_1 = \lambda_2$. If $\phi_1^2 < 4\phi_2\phi_0$, then the roots are the conjugate complex numbers $\lambda = \alpha + i\beta$, $\lambda^* = \alpha - i\beta$, where $i = \sqrt{-1}$.

There are three alternative ways of representing the conjugate complex numbers λ and λ^* :

$$\begin{aligned} \lambda &= \alpha + i\beta = \rho(\cos\theta + i\sin\theta) = \rho e^{i\theta}, \\ \lambda^* &= \alpha - i\beta = \rho(\cos\theta - i\sin\theta) = \rho e^{-i\theta}, \end{aligned} \tag{3.14}$$

where

$$\rho = \sqrt{\alpha^2 + \beta^2} \qquad \text{and} \qquad \theta = \tan^{-1}\left(\frac{\beta}{\alpha}\right). \tag{3.15}$$

These are called, respectively, the Cartesian form, the trigonometrical form and the exponential form.

The Cartesian and trigonometrical representations are understood by considering the Argand diagram:

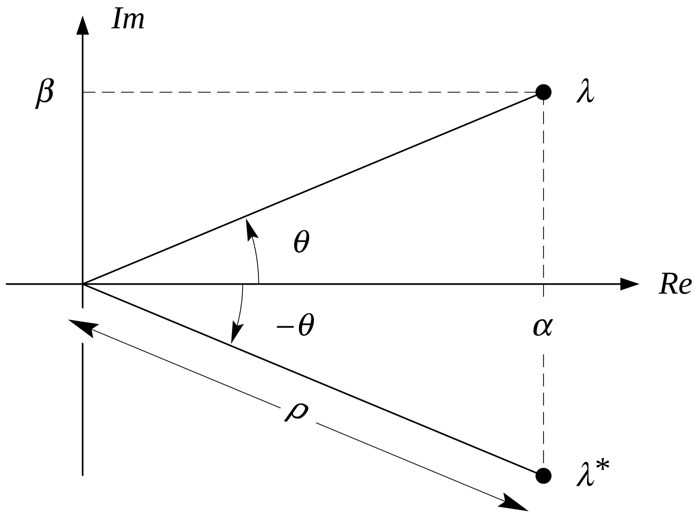

Figure 1. The Argand Diagram showing a complex number $\lambda = \alpha + i\beta$ and its conjugate $\lambda^* = \alpha - i\beta$.

The exponential form is understood by considering the following series expansions of $\cos\theta$ and $i\sin\theta$ about the point $\theta = 0$:

$$(3.16)\qquad \begin{aligned} \cos\theta &= \left\{1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right\}, \\ i\sin\theta &= \left\{i\theta - \frac{i\theta^3}{3!} + \frac{i\theta^5}{5!} - \frac{i\theta^7}{7!} + \cdots\right\}. \end{aligned}$$

Adding these gives

$$(3.17)\qquad \begin{aligned} \cos\theta + i\sin\theta &= \left\{1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \cdots\right\} \\ &= e^{i\theta}. \end{aligned}$$

Likewise, by subtraction, we get

$$(3.18)\qquad \begin{aligned} \cos\theta - i\sin\theta &= \left\{1 - i\theta - \frac{\theta^2}{2!} + \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} - \cdots\right\} \\ &= e^{-i\theta}. \end{aligned}$$

These are Euler's equations. It follows from adding (17) and (18) that

$$(3.19)\qquad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}.$$

Subtracting (18) from (17) gives

$$(3.20)\qquad \begin{aligned} \sin\theta &= \frac{-i}{2}(e^{i\theta} - e^{-i\theta}) \\ &= \frac{1}{2i}(e^{i\theta} - e^{-i\theta}). \end{aligned}$$

Now consider the general equation of the nth order:

$$(3.21)\qquad \phi_0 + \phi_1 z + \phi_2 z^2 + \cdots + \phi_n z^n = 0.$$

On dividing by ϕ_n, we can factorise this as

$$(3.22)\qquad (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n) = 0,$$

where some of the roots may be real and others may be complex. The complex roots come in conjugate pairs, so that, if $\lambda = \alpha + i\beta$ is a complex root, then there is a corresponding root $\lambda^* = \alpha - i\beta$ such that the product $(z-\lambda)(z-\lambda^*) = z^2 - 2\alpha z + (\alpha^2 + \beta^2)$ is real and quadratic. When we multiply the n factors together, we obtain the expansion

$$(3.23)\qquad 0 = z^n - \sum_i \lambda_i z^{n-1} + \sum_i \sum_j \lambda_i \lambda_j z^{n-2} - \cdots (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n.$$

This can be compared with the expression $(\phi_0/\phi_n)+(\phi_1/\phi_n)z+\cdots+z^n = 0$. By equating coefficients of the two expressions, we find that $(\phi_0/\phi_n) = (-1)^n\prod\lambda_i$ or, equivalently,

$$\phi_n = \phi_0\prod_{i=1}^{n}(-\lambda_i)^{-1}. \tag{3.24}$$

Thus we can express the polynomial in any of the following forms:

$$\begin{aligned}\sum\phi_i z^i &= \phi_n\prod(z-\lambda_i)\\ &= \phi_0\prod(-\lambda_i)^{-1}\prod(z-\lambda_i)\\ &= \phi_0\prod\left(1-\frac{z}{\lambda_i}\right).\end{aligned} \tag{3.25}$$

We should also note that, if λ is a root of the primary equation $\sum\phi_i z^i = 0$, where rising powers of z are associated with rising indices on the coefficients, then $\mu = 1/\lambda$ is a root of the equation $\sum\phi_i z^{n-i} = 0$, which has declining powers of z instead. This follows since $\sum\phi_i\lambda^i = \sum\phi_i\mu^{-i} = 0$ implies that $\mu^n\sum\phi_i\mu^{-i} = \sum\phi_i\mu^{n-i} = 0$. Confusion can arise from not knowing which of the two equations one is dealing with.

Rational Functions of Polynomials

If $\delta(z)$ and $\gamma(z)$ are polynomial functions of z of degrees d and g respectively with $d < g$, then the ratio $\delta(z)/\gamma(z)$ is described as a proper rational function. We shall often encounter expressions of the form

$$y(t) = \frac{\delta(L)}{\gamma(L)}x(t). \tag{3.26}$$

For this to have a meaningful interpretation in the context of a time-series model, we normally require that $y(t)$ should be a bounded sequence whenever $x(t)$ is bounded. The necessary and sufficient condition for the boundedness of $y(t)$, in that case, is that the series expansion of $\delta(z)/\gamma(z)$ should be convergent whenever $|z|\leq 1$. We can determine whether or not the sequence will converge by expressing the ratio $\delta(z)/\gamma(z)$ as a sum of partial fractions. The basic result is as follows:

(3.27) If $\delta(z)/\gamma(z) = \delta(z)/\{\gamma_1(z)\gamma_2(z)\}$ is a proper rational function, and if $\gamma_1(z)$ and $\gamma_2(z)$ have no common factor, then the function can be uniquely expressed as

$$\frac{\delta(z)}{\gamma(z)} = \frac{\delta_1(z)}{\gamma_1(z)} + \frac{\delta_2(z)}{\gamma_2(z)},$$

where $\delta_1(z)/\gamma_1(z)$ and $\delta_2(z)/\gamma_2(z)$ are proper rational functions.

Imagine that $\gamma(z) = \prod(1 - z/\lambda_i)$. Then repeated applications of this basic result enables us to write

$$\frac{\delta(z)}{\gamma(z)} = \frac{\kappa_1}{1 - z/\lambda_1} + \frac{\kappa_2}{1 - z/\lambda_2} + \cdots + \frac{\kappa_g}{1 - z/\lambda_g}. \tag{3.28}$$

By adding the terms on the RHS, we find an expression with a numerator of degree $n - 1$. By equating the terms of the numerator with the terms of $\delta(z)$, we can find the values $\kappa_1, \kappa_2, \ldots, \kappa_g$. The convergence of the expansion of $\delta(z)/\gamma(z)$ is a straightforward matter. For the series converges if and only if the expansion of each of the partial fractions converges. For the expansion

$$\frac{\kappa}{1 - z/\lambda} = \kappa\Big\{1 + z/\lambda + (z/\lambda)^2 + \cdots\Big\} \tag{3.29}$$

to converge when $|z| \leq 1$, it is necessary and sufficient that $|\lambda| > 1$.

Example. Consider the function

$$\begin{aligned} \frac{3z}{1 + z - 2z^2} &= \frac{3z}{(1 - z)(1 + 2z)} \\ &= \frac{\kappa_1}{1 - z} + \frac{\kappa_2}{1 + 2z} \\ &= \frac{\kappa_1(1 + 2z) + \kappa_2(1 - z)}{(1 - z)(1 + 2z)}. \end{aligned} \tag{3.30}$$

Equating the terms of the numerator gives

$$3z = (2\kappa_1 - \kappa_2)z + (\kappa_1 + \kappa_2), \tag{3.31}$$

so $\kappa_2 = -\kappa_1$, which gives $3 = (2\kappa_1 - \kappa_2) = 3\kappa_1$; and thus we have $\kappa_1 = 1$, $\kappa_2 = -1$.

Linear Difference Equations

An nth-order linear difference equation is a relationship amongst $n + 1$ consecutive elements of a sequence $x(t)$ of the form

$$\alpha_0 x(t) + \alpha_1 x(t - 1) + \cdots + \alpha_n x(t - n) = u(t), \tag{3.32}$$

where $u(t)$ is some specified sequence which is described as the forcing function. The equation can be written, in a summary notation, as

$$\alpha(L)x(t) = u(t), \tag{3.33}$$

where $\alpha(L) = \alpha_0 + \alpha_1 L + \cdots + \alpha_n L^n$. If n consecutive values of $x(t)$ are given, say $x_1, x_2, \ldots, x_n$, then the relationship can be used to find the succeeding value x_{n+1}. In this way, so long as $u(t)$ is fully specified, it is possible to generate any number of the succeeding elements of the sequence. The values of the sequence prior to $t = 1$ can be generated likewise; and thus, in effect, we can deduce the function $x(t)$ from the difference equation. However, instead of a recursive solution, we often seek an analytic expression for $x(t)$.

The function $x(t; c)$, expressing the analytic solution, will comprise a set of n constants in $c = [c_1, c_2, \ldots, c_n]'$ which can be determined once we are given a set of n consecutive values of $x(t)$ which are called initial conditions. The general analytic solution of the equation $\alpha(L)x(t) = u(t)$ is expressed as $x(t; c) = y(t; c) + z(t)$, where $y(t)$ is the general solution of the homogeneous equation $\alpha(L)y(t) = 0$, and $z(t) = \alpha^{-1}(L)u(t)$ is called a particular solution of the inhomogeneous equation.

We may solve the difference equation in three steps. First, we find the general solution of the homogeneous equation. Next, we find the particular solution $z(t)$ which embodies no unknown quantities. Finally, we use the n initial values of x to determine the constants $c_1, c_2, \ldots, c_n$. We shall discuss in detail only the solution of the homogeneous equation.

Solution of the Homogeneous Difference Equation

If λ_j is a root of the equation $\alpha(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n = 0$ such that $\alpha(\lambda_j) = 0$, then $y_j(t) = (1/\lambda_j)^t$ is a solution of the equation $\alpha(L)y(t) = 0$. This can be see this by considering the expression

$$
\begin{aligned}
\alpha(L)\left(\frac{1}{\lambda_j}\right)^t &= \big(\alpha_0 + \alpha_1 L + \cdots + \alpha_n L^n\big)\left(\frac{1}{\lambda_j}\right)^t \\
&= \alpha_0\left(\frac{1}{\lambda_j}\right)^t + \alpha_1\left(\frac{1}{\lambda_j}\right)^{t-1} + \cdots + \alpha_n\left(\frac{1}{\lambda_j}\right)^{t-n} \\
&= \big(\alpha_0 + \alpha_1\lambda_j + \cdots + \alpha_n\lambda_j^n\big)\left(\frac{1}{\lambda_j}\right)^t \\
&= \alpha(\lambda_j)\left(\frac{1}{\lambda_j}\right)^t .
\end{aligned}
\tag{3.34}
$$

Alternatively, one may consider the factorisation $\alpha(L) = \alpha_0 \prod_i (1 - L/\lambda_i)$. Within this product is the term $1 - L/\lambda_j$; and since

$$\left(1 - \frac{L}{\lambda_j}\right)\left(\frac{1}{\lambda_j}\right)^t = \left(\frac{1}{\lambda_j}\right)^t - \left(\frac{1}{\lambda_j}\right)^t = 0,$$

it follows that $\alpha(L)(1/\lambda_j)^t = 0$.

The general solution, in the case where $\alpha(L) = 0$ has distinct real roots, is given by

$$(3.35)\qquad y(t;c) = c_1\left(\frac{1}{\lambda_1}\right)^t + c_2\left(\frac{1}{\lambda_2}\right)^t + \cdots + c_n\left(\frac{1}{\lambda_n}\right)^t,$$

where $c_1, c_2, \ldots, c_n$ are the constants which are determined by the initial conditions.

In the case where two roots coincide at a value of λ_j, the equation $\alpha(L)y(t) = 0$ has the solutions $y_1(t) = (1/\lambda_j)^t$ and $y_2(t) = t(1/\lambda_j)^t$. To show this, let us extract the term $(1 - L/\lambda_j)^2$ from the factorisation $\alpha(L) = \alpha_0 \prod_i (1 - L/\lambda_i)$. Then, according to the previous argument, we have $(1 - L/\lambda_j)^2(1/\lambda_j)^t = 0$, but, also, we have

$$(3.36)\qquad \begin{aligned}\left(1 - \frac{L}{\lambda_j}\right)^2 t\left(\frac{1}{\lambda_j}\right)^t &= \left(1 - \frac{2L}{\lambda_j} + \frac{L^2}{\lambda_j^2}\right) t\left(\frac{1}{\lambda_j}\right)^t \\ &= t\left(\frac{1}{\lambda_j}\right)^t - 2(t-1)\left(\frac{1}{\lambda_j}\right)^t + (t-2)\left(\frac{1}{\lambda_j}\right)^t = 0.\end{aligned}$$

In general, if there are r repeated roots with a value of λ_j, then all of $(1/\lambda_j)^t$, $t(1/\lambda_j)^t$, $t^2(1/\lambda_j)^t$, $\ldots$, $t^{r-1}(1/\lambda_j)^t$ are solutions to the equation $\alpha(L)y(t) = 0$.

A particularly important special case arises when there are r repeated roots of unit value. Then the functions $1, t, t^2, \ldots, t^{r-1}$ are all solutions to the homogeneous equation. With each solution is associated a coefficient which can be determined in view of the initial conditions. If these coefficients are $d_0, d_1, d_2, \ldots, d_{r-1}$ then, within the general solution of the homogeneous equation, there will be found the term $d_0 + d_1 t + d_2 t^2 + \cdots + d_{r-1}t^{r-1}$ which represents a polynomial in t of degree $r - 1$.

The 2nd-order Difference Equation with Complex Roots

Imagine that the 2nd-order equation $\alpha(L)y(t) = \alpha_0 y(t) + \alpha_1 y(t-1) + \alpha_2 y(t-2) = 0$ is such that $\alpha(z) = 0$ has complex roots $\lambda = 1/\mu$ and $\lambda^* = 1/\mu^*$. If λ, λ^* are conjugate complex numbers, then so too are μ, μ^*. Therefore, let us write

$$(3.37)\qquad \begin{aligned}\mu &= \gamma + i\delta = \kappa(\cos\omega + i\sin\omega) = \kappa e^{i\omega}, \\ \mu^* &= \gamma - i\delta = \kappa(\cos\omega - i\sin\omega) = \kappa e^{-i\omega}.\end{aligned}$$

These will appear in a general solution of the difference equation of the form

$$(3.38)\qquad y(t) = c\mu^t + c^*(\mu^*)^t.$$

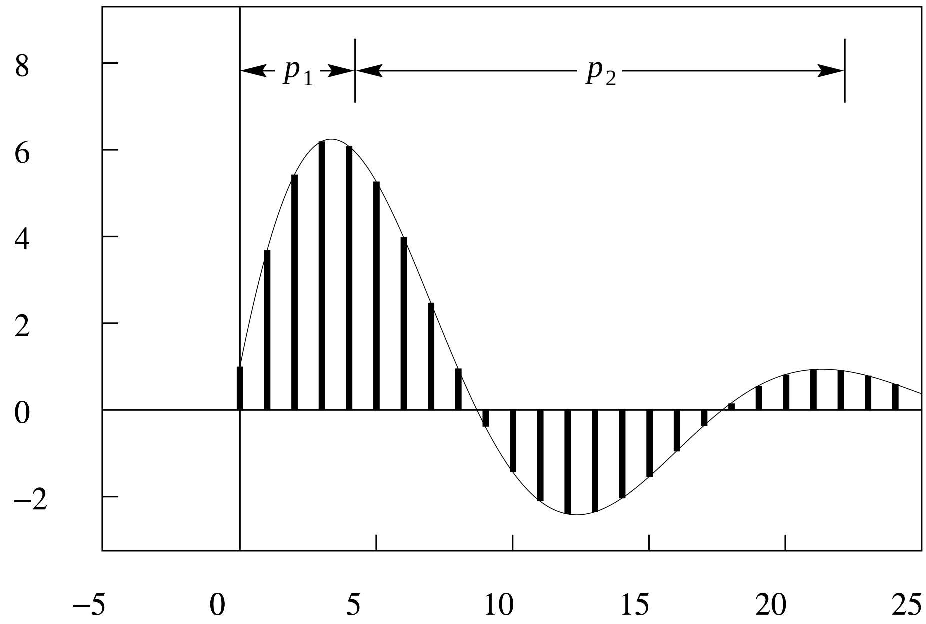

Figure 2. The solution of the homogeneous difference equation $(1 - 1.69L + 0.81L^2)y(t) = 0$ for the initial conditions $y_0 = 1$ and $y_1 = 3.69$. The time lag of the phase displacement p_1 and the duration of the cycle p_2 are also indicated.

This represents a real-valued sequence; and, since a real term must equal its own conjugate, it follows that c and c^* must be conjugate numbers of the form

$$\begin{aligned} c^* &= \rho(\cos\theta + i\sin\theta) = \rho e^{i\theta}, \\ c &= \rho(\cos\theta - i\sin\theta) = \rho e^{-i\theta}. \end{aligned} \tag{3.39}$$

Thus the general solution becomes

$$\begin{aligned} c\mu^t + c^*(\mu^*)^t &= \rho e^{-i\theta}(\kappa e^{i\omega})^t + \rho e^{i\theta}(\kappa e^{-i\omega})^t \\ &= \rho\kappa^t\Big\{e^{i(\omega t-\theta)} + e^{-i(\omega t-\theta)}\Big\} \\ &= 2\rho\kappa^t\cos(\omega t - \theta). \end{aligned} \tag{3.40}$$

To analyse the final expression, consider first the factor $\cos(\omega t - \theta)$. This is a displaced cosine wave. The value ω, which is a number of radians per unit period, is called the angular velocity or the angular frequency of the wave. The value $f = \omega/2\pi$ is its frequency in cycles per unit period. The duration of one cycle, also called the period, is $r = 2\pi/\omega$.

The term θ is called the phase displacement of the cosine wave, and it serves to shift the cosine function along the axis of t so that, in the absence of damping, the peak would occur at the value of $t = \theta/\omega$ instead of at $t = 0$.

Next consider the term κ^t wherein $\kappa = \surd(\gamma^2 + \delta^2)$ is the modulus of the complex roots. When κ has a value of less than unity, it becomes a damping factor which serves to attenuate the cosine wave as t increases. The damping also serves to shift the peaks of the cosine function slightly to the left.

Finally, the factor 2ρ affects the initial amplitude of the cosine wave which is the value which it assumes when $t = 0$. Since ρ is just the modulus of the values c and c^*, this amplitude reflects the initial conditions. The phase angle θ is also a product of the initial conditions.

It is instructive to derive an expression for the second-order difference equation which is in terms of the parameters of the trigonometrical or exponential representations of a pair of complex roots. Consider

$$\begin{aligned}\alpha(z) &= \alpha_0(1 - \mu z)(1 - \mu^* z)\\ &= \alpha_0\big\{1 - (\mu + \mu^*)z + \mu\mu^* z^2\big\},\end{aligned} \tag{3.41}$$

From (37) it follows that

$$\mu + \mu^* = 2\kappa\cos\omega \quad \text{and} \quad \mu\mu^* = \kappa^2. \tag{3.42}$$

Therefore the polynomial operator which is entailed by the difference equation is

$$\alpha_0 + \alpha_1 L + \alpha_2 L^2 = \alpha_0(1 - 2\kappa\cos\omega\ L + \kappa^2 L^2); \tag{3.43}$$

and it is usual to set $\alpha_0 = 1$. This representation indicates that a necessary condition for the roots to be complex, which is not a sufficient condition, is that $\alpha_2/\alpha_0 > 0$.

It is easy to ascertain by inspection whether or not the second-order difference equation is stable. The condition that the roots of $\alpha(z) = 0$ must lie outside the unit circle, which is necessary and sufficient for stability, imposes certain restrictions on the coefficients of $\alpha(z)$ which can be checked easily.

We can reveal these conditions most readily by considering the auxiliary polynomial $\rho(z) = z^2\alpha(z^{-1})$ whose roots, which are the inverses of those of $\alpha(z)$, must lie inside the unit circle. Let the roots of $\rho(z)$, which might be real or complex, be denoted by μ_1, μ_2. Then we can write

$$\begin{aligned}\rho(z) &= \alpha_0 z^2 + \alpha_1 z + \alpha_2\\ &= \alpha_0(z - \mu_1)(z - \mu_2)\\ &= \alpha_0\big\{z^2 - (\mu_1 + \mu_2)z + \mu_1\mu_2\big\},\end{aligned} \tag{3.44}$$

where is is assumed that $\alpha_0 > 0$. This indicates that $\alpha_2/\alpha_0 = \mu_1\mu_2$. Therefore the conditions $|\mu_1|,\ |\mu_2| < 1$ imply that

$$-\alpha_0 < \alpha_2 < \alpha_0. \tag{3.45}$$

If the roots are complex conjugate numbers $\mu, \mu^* = \gamma \pm i\delta$, then this condition will ensure that $\mu^*\mu = \alpha_2/\alpha_0 < 1$, which is the condition that they are within the unit circle.

Now consider the fact that, if $\alpha_0 > 0$, then the function $\rho(z)$ will have a minimum value over the real line which is greater than zero if the roots are complex and no greater than zero if they are real. If the roots are real, then they will be found in the interval $(-1, 1)$ if and only if

$$\begin{aligned} \rho(-1) &= \alpha_0 - \alpha_1 + \alpha_2 > 0 \quad \text{and} \\ \rho(1) &= \alpha_0 + \alpha_1 + \alpha_2 > 0. \end{aligned} \tag{3.46}$$

If the roots are complex then these conditions are bound to be satisfied.

From these arguments, it follows that the conditions under (45) and (46) in combination are necessary and sufficient to ensure that the roots of $\rho(z) = 0$ are within the unit circle and that the roots of $\alpha(z) = 0$ are outside.

State-Space Models

An nth-order difference equation in a single variable can be transformed into a first-order system in n variables which are the elements of a so-called state vector.

There is a wide variety of alternative forms which can be assumed by a first-order vector difference equation corresponding to the nth-order scalar equation. However, certain of these are described as canonical forms by virtue of special structures in the matrix.

In demonstrating one of the more common canonical forms, let us consider again the nth-order difference equation of (32), in reference to which we may define the following variables:

$$\begin{aligned} \xi_1(t) &= x(t), \\ \xi_2(t) &= \xi_1(t-1) = x(t-1), \\ &\vdots \\ \xi_n(t) &= \xi_{n-1}(t-1) = x(t-n+1). \end{aligned} \tag{3.47}$$

On the basis of these definitions, a first-order vector equation may be constructed in the form of

$$\begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \vdots \\ \xi_n(t) \end{bmatrix} = \begin{bmatrix} -\alpha_1 & \ldots & -\alpha_{n-1} & -\alpha_n \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} \xi_1(t-1) \\ \xi_2(t-1) \\ \vdots \\ \xi_n(t-1) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon(t). \tag{3.48}$$

The matrix in this structure is sometimes described as the companion form. Here it is manifest, in view of the definitions under (47), that the leading equation of the system, which is

$$\xi_1(t) = -\alpha_1\xi_1(t-1) + \cdots + \alpha_n\xi_n(t-1) + \varepsilon(t), \tag{3.49}$$

is precisely the equation under (32).

Example. An example of a system which is not in a canonical form is provided by the following matrix equation:

$$\begin{bmatrix} y(t) \\ z(t) \end{bmatrix} = \kappa \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} y(t-1) \\ z(t-1) \end{bmatrix} + \begin{bmatrix} \upsilon(t) \\ \zeta(t) \end{bmatrix}. \tag{3.50}$$

With the use of the lag operator, the equation can also be written as

$$\begin{bmatrix} 1-\kappa\cos\omega L & \kappa\sin\omega L \\ -\kappa\sin\omega L & 1-\kappa\cos\omega L \end{bmatrix} \begin{bmatrix} y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \upsilon(t) \\ \zeta(t) \end{bmatrix}. \tag{3.51}$$

On premultiplying the equation by the inverse of the matrix on the LHS, we get

(3.52)
$$\begin{bmatrix} y(t) \\ z(t) \end{bmatrix} = \frac{1}{1-2\kappa\cos\omega L + \kappa^2L^2} \begin{bmatrix} 1-\kappa\cos\omega L & -\kappa\sin\omega L \\ \kappa\sin\omega L & 1-\kappa\cos\omega L \end{bmatrix} \begin{bmatrix} \upsilon(t) \\ \zeta(t) \end{bmatrix}.$$

A special case arises when

$$\begin{bmatrix} \upsilon(t) \\ \zeta(t) \end{bmatrix} = \begin{bmatrix} -\sin\omega \\ \cos\omega \end{bmatrix} \eta(t), \tag{3.53}$$

where $\eta(t)$ is a white-noise sequence. Then the equation becomes

$$\begin{bmatrix} y(t) \\ z(t) \end{bmatrix} = \frac{1}{1-2\kappa\cos\omega L + \kappa^2L^2} \begin{bmatrix} -\sin\omega \\ \cos\omega \end{bmatrix} \eta(t). \tag{3.54}$$

On defining $\varepsilon(t) = -\sin\omega\eta(t)$ we may write the first of these equations as

$$(1-2\kappa\cos\omega L + \kappa^2L^2)y(t) = \varepsilon(t). \tag{3.55}$$

This is just a second-order difference equation with a white-noise forcing function; and, by virtue of the inclusion of the damping factor $\kappa \in [0,1)$, it represents a generalisation of the equation to be found under (2.24).

Transfer Functions

Consider again the simple dynamic model of equation (5):

$$y(t) = \phi y(t-1) + x(t)\beta + \varepsilon(t). \tag{3.56}$$

With the use of the lag operator, this can be rewritten as

$$(1-\phi L)y(t) = \beta x(t) + \varepsilon(t) \tag{3.57}$$

or, equivalently, as

$$y(t) = \frac{\beta}{1-\phi L}x(t) + \frac{1}{1-\phi L}\varepsilon(t). \tag{3.58}$$

The latter is the so-called rational transfer-function form of the equation. The operator L within the transfer functions or filters can be replaced by a complex number z. Then the transfer function which is associated with the signal $x(t)$ becomes

$$\frac{\beta}{1-\phi z} = \beta\big\{1 + \phi z + \phi^2 z^2 + \cdots\big\}, \tag{3.59}$$

where the RHS comes from a familiar power-series expansion.

The sequence $\{\beta, \beta\phi, \beta\phi^2, \ldots\}$ of the coefficients of the expansion constitutes the impulse response of the transfer function. That is to to say, if we imagine that, on the input side, the signal is a unit-impulse sequence of the form

$$x(t) = \{\ldots, 0, 1, 0, 0, \ldots\}, \tag{3.60}$$

which has zero values at all but one instant, then its mapping through the transfer function would result in an output sequence of

$$r(t) = \{\ldots, 0, \beta, \beta\phi, \beta\phi^2, \ldots\}. \tag{3.61}$$

Another important concept is the step response of the filter. We may imagine that the input sequence is zero-valued up to a point in time when it assumes a constant unit value:

$$x(t) = \{\ldots, 0, 1, 1, 1, \ldots\}. \tag{3.62}$$

The mapping of this sequence through the transfer function would result in an output sequence of

$$s(t) = \{\ldots, 0, \beta, \beta + \beta\phi, \beta + \beta\phi + \beta\phi^2, \ldots\} \tag{3.63}$$

whose elements, from the point when the step occurs in $x(t)$, are simply the partial sums of the impulse-response sequence. This sequence of partial sums $\{\beta, \beta + \beta\phi, \beta + \beta\phi + \beta\phi^2, \ldots\}$ is described as the step response. Given that $|\phi| < 1$, the step response converges to a value

$$\gamma = \frac{\beta}{1-\phi} \tag{3.64}$$

which is described as the steady-state gain or the long-term multiplier of the transfer function.

These various concepts apply to models of any order. Consider the equation

$$\alpha(L)y(t) = \beta(L)x(t) + \varepsilon(t), \tag{3.65}$$

where

$$\begin{aligned} \alpha(L) &= 1 + \alpha_1 L + \cdots + \alpha_p L^p \\ &= 1 - \phi_1 L - \cdots - \phi_p L^p, \\ \beta(L) &= \beta_0 + \beta_1 L + \cdots + \beta_k L^k \end{aligned} \tag{3.66}$$

are polynomials of the lag operator. The transfer-function form of the model is simply

$$y(t) = \frac{\beta(L)}{\alpha(L)}x(t) + \frac{1}{\alpha(L)}\varepsilon(t), \tag{3.67}$$

The rational function associated with $x(t)$ has a series expansion

$$\begin{aligned} \frac{\beta(z)}{\alpha(z)} &= \omega(z) \\ &= \{\omega_0 + \omega_1 z + \omega_2 z^2 + \cdots\}; \end{aligned} \tag{3.68}$$

and the sequence of the coefficients of this expansion constitutes the impulse-response function. The partial sums of the coefficients constitute the step-response function. The gain of the transfer function is defined by

$$\gamma = \frac{\beta(1)}{\alpha(1)} = \frac{\beta_0 + \beta_1 + \cdots + \beta_k}{1 + \alpha_1 + \cdots + \alpha_p}. \tag{3.69}$$

The method of finding the coefficients of the series expansion of the transfer function in the general case can be illustrated by the second-order case:

$$\frac{\beta_0 + \beta_1 z}{1 - \phi_1 z - \phi_2 z^2} = \{\omega_0 + \omega_1 z + \omega_2 z^2 + \cdots\}. \tag{3.70}$$

We rewrite this equation as

$$(3.71)\qquad \beta_0+\beta_1 z=\big\{1-\phi_1 z-\phi_2 z^2\big\}\big\{\omega_0+\omega_1 z+\omega_2 z^2+\cdots\big\}.$$

Then, by performing the multiplication on the RHS, and by equating the coefficients of the same powers of z on the two sides of the equation, we find that

$$(3.72)\qquad \begin{aligned} \beta_0&=\omega_0, & \omega_0&=\beta_0,\\ \beta_1&=\omega_1-\phi_1\omega_0, & \omega_1&=\beta_1+\phi_1\omega_0,\\ 0&=\omega_2-\phi_1\omega_1-\phi_2\omega_0, & \omega_2&=\phi_1\omega_1+\phi_2\omega_0,\\ &\ \vdots & &\ \vdots\\ 0&=\omega_n-\phi_1\omega_{n-1}-\phi_2\omega_{n-2}, & \omega_n&=\phi_1\omega_{n-1}+\phi_2\omega_{n-2}. \end{aligned}$$

The necessary and sufficient condition for the convergence of the sequence $\{\omega_i\}$ is that the roots of the primary polynomial equation $1-\phi_1 z-\phi_2 z^2=0$ should lie outside the unit circle or, equivalently, that the roots of the auxiliary equation $z^2-\phi_1 z-\phi_2=0$—which are the inverses of the former roots—should lie inside the unit circle. If the roots of these equations are real, then the sequence will converge monotonically to zero whereas, if the roots are complex-valued, then the sequence will converge in the manner of a damped sinusoid.

It is clear that the equation

$$(3.73)\qquad \omega(n)=\phi_1\omega(n-1)+\phi_2\omega(n-2),$$

which serves to generate the elements of the impulse response, is nothing but a second-order homogeneous difference equation. In fact, Figure 2, which has been presented as the solution to a homogeneous difference equation, represents the impulse response of the transfer function $(1+2L)/(1-1.69L+0.81L^2)$.

In the light of this result, it is apparent that the coefficients of the denominator polynomial $1-\phi_1 z-\phi_2 z^2$ serve to determine the period and the damping factor of a complex impulse response. The coefficients in the numerator polynomial $\beta_0+\beta_1 z$ serve to determine the initial amplitude of the response and its phase lag. It seems that all four coefficients must be present if a second-order transfer function is to have complete flexibility in modelling a dynamic response.

The Frequency Response

In many applications within forecasting and time-series analysis, it is of interest to consider the response of a transfer function to a signal which is a simple sinusoid. As we have indicated in a previous lecture, it is possible

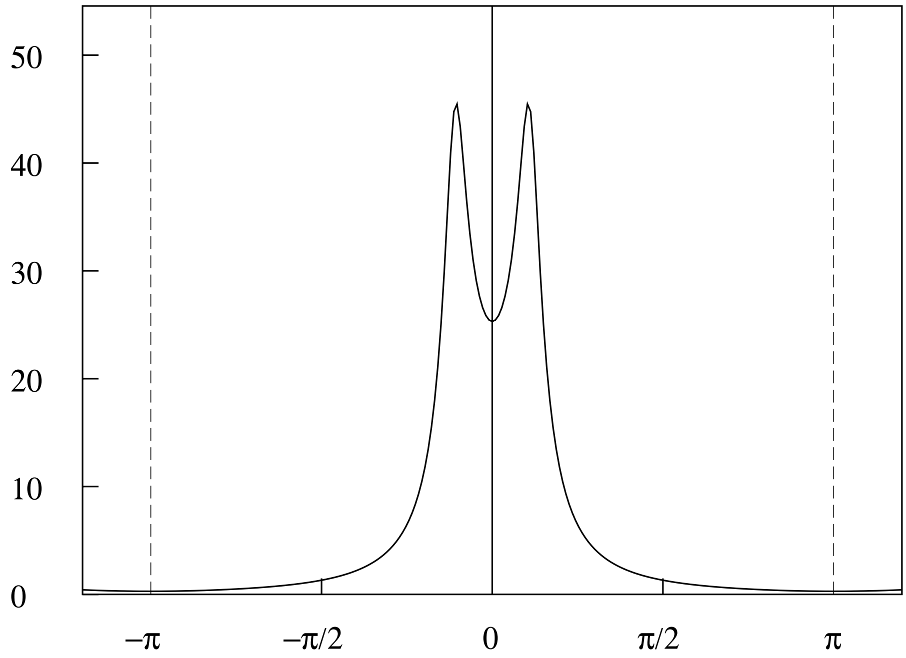

Figure 3. The gain of the transfer function $(1 + 2L^2)/(1 - 1.69L + 0.81L^2)$.

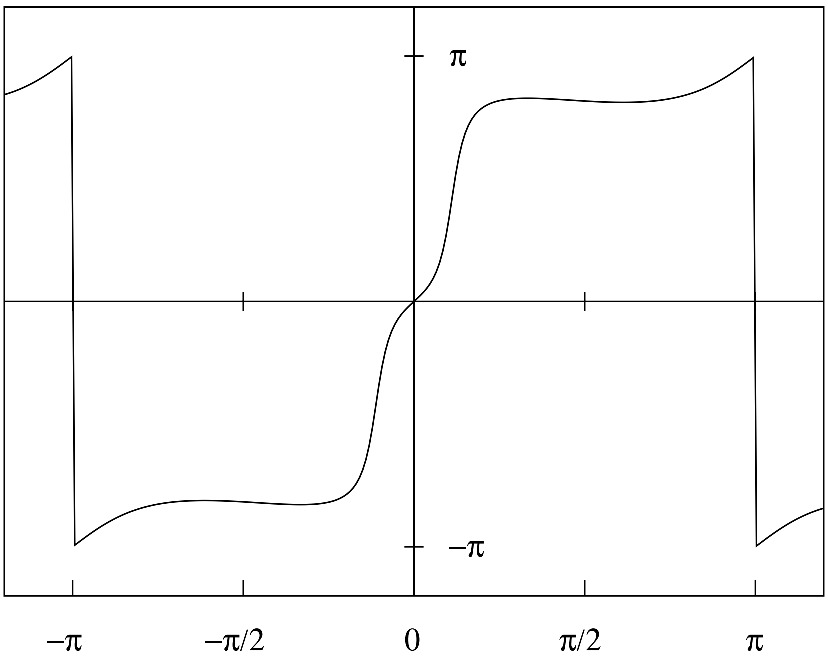

Figure 4. The phase diagram of the transfer function $(1 + 2L^2)/(1 - 1.69L + 0.81L^2)$.

to represent a finite sequence as a sum of sine and cosine functions whose frequencies are integer multiples of a fundamental frequency. More generally, it is possible, as we shall see later, to represent an arbitrary stationary stochastic process as a combination of an infinite number of sine and cosine functions whose frequencies range continuously in the interval $[0, \pi]$. It follows that the effect of a transfer function upon stationary signals can be characterised in terms of its effect upon the sinusoidal functions.

Consider therefore the consequences of mapping the signal $x(t) = \cos(\omega t)$ through the transfer function $\gamma(L) = \gamma_0 + \gamma_1 L + \cdots + \gamma_g L^g$. The output is

$$\begin{aligned} y(t) &= \gamma(L)\cos(\omega t) \\ &= \sum_{j=0}^{g} \gamma_j \cos\big(\omega[t-j]\big). \end{aligned} \tag{3.74}$$

The trigonometrical identity $\cos(A - B) = \cos A \cos B + \sin A \sin B$ enables us to write this as

$$\begin{aligned} y(t) &= \Big\{\sum_j \gamma_j \cos(\omega j)\Big\}\cos(\omega t) + \Big\{\sum_j \gamma_j \sin(\omega j)\Big\}\sin(\omega t) \\ &= \alpha\cos(\omega t) + \beta\sin(\omega t) = \rho\cos(\omega t - \theta). \end{aligned} \tag{3.75}$$

Here we have defined

$$\begin{aligned} &\alpha = \sum_{j=0}^{g} \gamma_j \cos(\omega j), \qquad \beta = \sum_{j=0}^{g} \gamma_j \sin(\omega j), \\ &\rho = \sqrt{\alpha^2 + \beta^2} \qquad \text{and} \qquad \theta = \tan^{-1}\Big(\frac{\beta}{\alpha}\Big). \end{aligned} \tag{3.76}$$

It can be seen from (75) that the effect of the filter upon the signal is twofold. First there is a *gain effect* whereby the amplitude of the sinusoid has been increased or diminished by a factor of ρ. Also there is a *phase effect* whereby the peak of the sinusoid is displaced by a time delay of θ/ω periods. Figures 3 and 4 represent the two effects of a simple rational transfer function on the set of sinusoids whose frequencies range from 0 to π.