# 1 Difference Equations

## 1.1. *First-Order Difference Equations*

This book is concerned with the dynamic consequences of events over time. Let's say we are studying a variable whose value at date $t$ is denoted $y_t$. Suppose we are given a dynamic equation relating the value $y$ takes on at date $t$ to another variable $w_t$ and to the value $y$ took on in the previous period:

$$y_t = \phi y_{t-1} + w_t. \qquad [1.1.1]$$

Equation [1.1.1] is a *linear first-order difference equation.* A *difference equation* is an expression relating a variable $y_t$ to its previous values. This is a *first-order* difference equation because only the first lag of the variable ($y_{t-1}$) appears in the equation. Note that it expresses $y_t$ as a linear function of $y_{t-1}$ and $w_t$.

An example of [1.1.1] is Goldfeld's (1973) estimated money demand function for the United States. Goldfeld's model related the log of the real money holdings of the public ($m_t$) to the log of aggregate real income ($I_t$), the log of the interest rate on bank accounts ($r_{bt}$), and the log of the interest rate on commercial paper ($r_{ct}$):

$$m_t = 0.27 + 0.72m_{t-1} + 0.19I_t - 0.045r_{bt} - 0.019r_{ct}. \qquad [1.1.2]$$

This is a special case of [1.1.1] with $y_t = m_t$, $\phi = 0.72$, and

$$w_t = 0.27 + 0.19I_t - 0.045r_{bt} - 0.019r_{ct}.$$

For purposes of analyzing the dynamics of such a system, it simplifies the algebra a little to summarize the effects of all the input variables ($I_t$, $r_{bt}$, and $r_{ct}$) in terms of a scalar $w_t$ as here.

In Chapter 3 the input variable $w_t$ will be regarded as a random variable, and the implications of [1.1.1] for the statistical properties of the output series $y_t$ will be explored. In preparation for this discussion, it is necessary first to understand the mechanics of difference equations. For the discussion in Chapters 1 and 2, the values for the input variable $\{w_1, w_2, \ldots\}$ will simply be regarded as a sequence of deterministic numbers. Our goal is to answer the following question: If a dynamic system is described by [1.1.1], what are the effects on $y$ of changes in the value of $w$?

### *Solving a Difference Equation by Recursive Substitution*

The presumption is that the dynamic equation [1.1.1] governs the behavior of $y$ for all dates $t$. Thus, for each date we have an equation relating the value of

$y$ for that date to its previous value and the current value of $w$:

| Date | Equation | |
|---|---|---|
| 0 | $y_0 = \phi y_{-1} + w_0$ | [1.1.3] |
| 1 | $y_1 = \phi y_0 + w_1$ | [1.1.4] |
| 2 | $y_2 = \phi y_1 + w_2$ | [1.1.5] |
| $\vdots$ | $\vdots$ | |
| $t$ | $y_t = \phi y_{t-1} + w_t.$ | [1.1.6] |

If we know the starting value of $y$ for date $t = -1$ and the value of $w$ for dates $t = 0, 1, 2, \ldots$, then it is possible to simulate this dynamic system to find the value of $y$ for any date. For example, if we know the value of $y$ for $t = -1$ and the value of $w$ for $t = 0$, we can calculate the value of $y$ for $t = 0$ directly from [1.1.3]. Given this value of $y_0$ and the value of $w$ for $t = 1$, we can calculate the value of $y$ for $t = 1$ from [1.1.4]:

$$y_1 = \phi y_0 + w_1 = \phi(\phi y_{-1} + w_0) + w_1,$$

or

$$y_1 = \phi^2 y_{-1} + \phi w_0 + w_1.$$

Given this value of $y_1$ and the value of $w$ for $t = 2$, we can calculate the value of $y$ for $t = 2$ from [1.1.5]:

$$y_2 = \phi y_1 + w_2 = \phi(\phi^2 y_{-1} + \phi w_0 + w_1) + w_2,$$

or

$$y_2 = \phi^3 y_{-1} + \phi^2 w_0 + \phi w_1 + w_2.$$

Continuing recursively in this fashion, the value that $y$ takes on at date $t$ can be described as a function of its initial value $y_{-1}$ and the history of $w$ between date 0 and date $t$:

$$y_t = \phi^{t+1} y_{-1} + \phi^t w_0 + \phi^{t-1} w_1 + \phi^{t-2} w_2 + \cdots + \phi w_{t-1} + w_t. \qquad [1.1.7]$$

This procedure is known as solving the difference equation [1.1.1] by *recursive substitution.*

## *Dynamic Multipliers*

Note that [1.1.7] expresses $y_t$ as a linear function of the initial value $y_{-1}$ and the historical values of $w$. This makes it very easy to calculate the effect of $w_0$ on $y_t$. If $w_0$ were to change with $y_{-1}$ and $w_1, w_2, \ldots, w_t$ taken as unaffected, the effect on $y_t$ would be given by

$$\frac{\partial y_t}{\partial w_0} = \phi^t. \qquad [1.1.8]$$

Note that the calculations would be exactly the same if the dynamic simulation were started at date $t$ (taking $y_{t-1}$ as given); then $y_{t+j}$ could be described as a



function of $y_{t-1}$ and $w_t, w_{t+1}, \ldots, w_{t+j}$:

$$y_{t+j} = \phi^{j+1} y_{t-1} + \phi^j w_t + \phi^{j-1} w_{t+1} + \phi^{j-2} w_{t+2} + \cdots + \phi w_{t+j-1} + w_{t+j}. \qquad [1.1.9]$$

The effect of $w_t$ on $y_{t+j}$ is given by

$$\frac{\partial y_{t+j}}{\partial w_t} = \phi^j. \qquad [1.1.10]$$

Thus the *dynamic multiplier* [1.1.10] depends only on $j$, the length of time separating the disturbance to the input ($w_t$) and the observed value of the output ($y_{t+j}$). The multiplier does not depend on $t$; that is, it does not depend on the dates of the observations themselves. This is true of any linear difference equation.

As an example of calculating a dynamic multiplier, consider again Goldfeld's money demand specification [1.1.2]. Suppose we want to know what will happen to money demand two quarters from now if current income $I_t$ were to increase by one unit today with future income $I_{t+1}$ and $I_{t+2}$ unaffected:

$$\frac{\partial m_{t+2}}{\partial I_t} = \frac{\partial m_{t+2}}{\partial w_t} \times \frac{\partial w_t}{\partial I_t} = \phi^2 \times \frac{\partial w_t}{\partial I_t}.$$

From [1.1.2], a one-unit increase in $I_t$ will increase $w_t$ by 0.19 units, meaning that $\partial w_t/\partial I_t = 0.19$. Since $\phi = 0.72$, we calculate

$$\frac{\partial m_{t+2}}{\partial I_t} = (0.72)^2(0.19) = 0.098.$$

Because $I_t$ is the log of income, an increase in $I_t$ of 0.01 units corresponds to a 1% increase in income. An increase in $m_t$ of $(0.01)\cdot(0.098) \cong 0.001$ corresponds to a 0.1% increase in money holdings. Thus the public would be expected to increase its money holdings by a little less than 0.1% two quarters following a 1% increase in income.

Different values of $\phi$ in [1.1.1] can produce a variety of dynamic responses of $y$ to $w$. If $0 < \phi < 1$, the multiplier $\partial y_{t+j}/\partial w_t$ in [1.1.10] decays geometrically toward zero. Panel (a) of Figure 1.1 plots $\phi^j$ as a function of $j$ for $\phi = 0.8$. If $-1 < \phi < 0$, the multiplier $\partial y_{t+j}/\partial w_t$ will alternate in sign as in panel (b). In this case an increase in $w_t$ will cause $y_t$ to be higher, $y_{t+1}$ to be lower, $y_{t+2}$ to be higher, and so on. Again the absolute value of the effect decays geometrically toward zero. If $\phi > 1$, the dynamic multiplier increases exponentially over time as in panel (c). A given increase in $w_t$ has a larger effect the farther into the future one goes. For $\phi < -1$, the system [1.1.1] exhibits explosive oscillation as in panel (d).

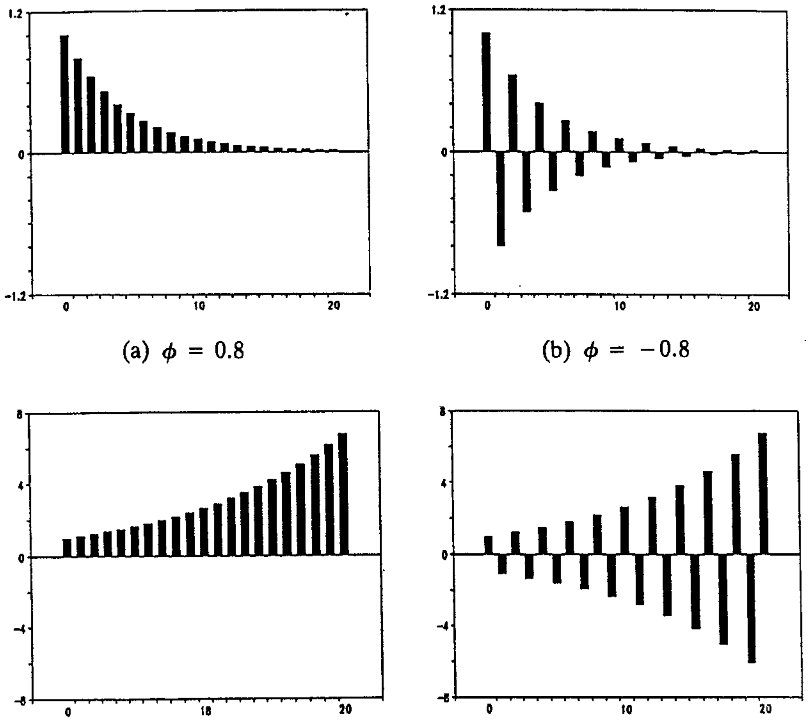
Thus, if $|\phi| < 1$, the system is stable; the consequences of a given change in $w_t$ will eventually die out. If $|\phi| > 1$, the system is explosive. An interesting possibility is the borderline case, $\phi = 1$. In this case, the solution [1.1.9] becomes

$$y_{t+j} = y_{t-1} + w_t + w_{t+1} + w_{t+2} + \cdots + w_{t+j-1} + w_{t+j}. \qquad [1.1.11]$$

Here the output variable $y$ is the sum of the historical inputs $w$. A one-unit increase in $w$ will cause a permanent one-unit increase in $y$:

$$\frac{\partial y_{t+j}}{\partial w_t} = 1 \qquad \text{for } j = 0, 1, \ldots.$$

We might also be interested in the effect of $w$ on the present value of the stream of future realizations of $y$. For a given stream of future values $y_t, y_{t+1},$



(a) $\phi = 0.8$ (b) $\phi = -0.8$

(c) $\phi = 1.1$ (d) $\phi = -1.1$

**FIGURE 1.1** Dynamic multiplier for first-order difference equation for different values of $\phi$ (plot of $\partial y_{t+j}/\partial w_t = \phi^j$ as a function of the lag $j$).

$y_{t+2}, \ldots$ and a constant interest rate[1] $r > 0$, the *present value* of the stream at time $t$ is given by

$$y_t + \frac{y_{t+1}}{1+r} + \frac{y_{t+2}}{(1+r)^2} + \frac{y_{t+3}}{(1+r)^3} + \cdots. \qquad [1.1.12]$$

Let $\beta$ denote the discount factor:

$$\beta \equiv 1/(1+r).$$

Note that $0 < \beta < 1$. Then the present value [1.1.12] can be written as

$$\sum_{j=0}^{\infty} \beta^j y_{t+j}. \qquad [1.1.13]$$

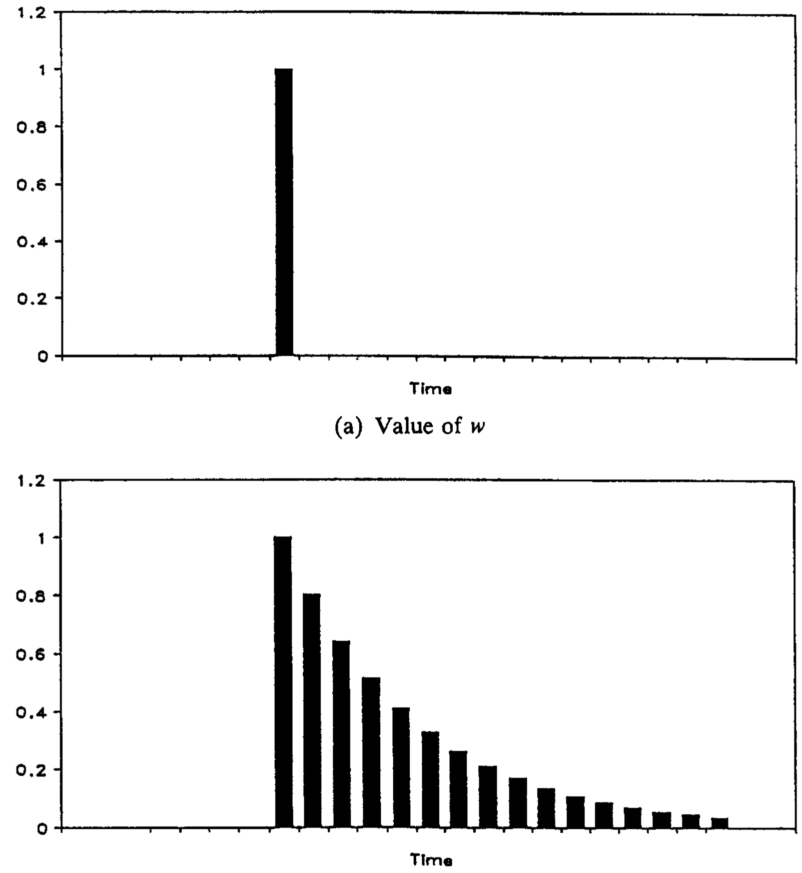
Consider what would happen if there were a one-unit increase in $w_t$ with $w_{t+1}, w_{t+2}, \ldots$ unaffected. The consequences of this change for the present value of $y$ are found by differentiating [1.1.13] with respect to $w_t$ and then using [1.1.10] to evaluate each derivative:

$$\sum_{j=0}^{\infty} \beta^j \frac{\partial y_{t+j}}{\partial w_t} = \sum_{j=0}^{\infty} \beta^j \phi^j = 1/(1 - \beta\phi), \qquad [1.1.14]$$

provided that $|\beta\phi| < 1$.

In calculating the dynamic multipliers [1.1.10] or [1.1.14], we were asking what would happen if $w_t$ were to increase by one unit with $w_{t+1}, w_{t+2}, \ldots, w_{t+j}$ unaffected. We were thus finding the effect of a purely transitory change in $w$. Panel (a) of Figure 1.2 shows the time path of $w$ associated with this question, and panel (b) shows the implied path for $y$. Because the dynamic multiplier [1.1.10] calculates the response of $y$ to a single impulse in $w$, it is also referred to as the *impulse-response function.*

[1]The interest rate is measured here as a fraction of 1; thus $r = 0.1$ corresponds to a 10% interest rate.




(a) Value of $w$

(b) Value of $y$

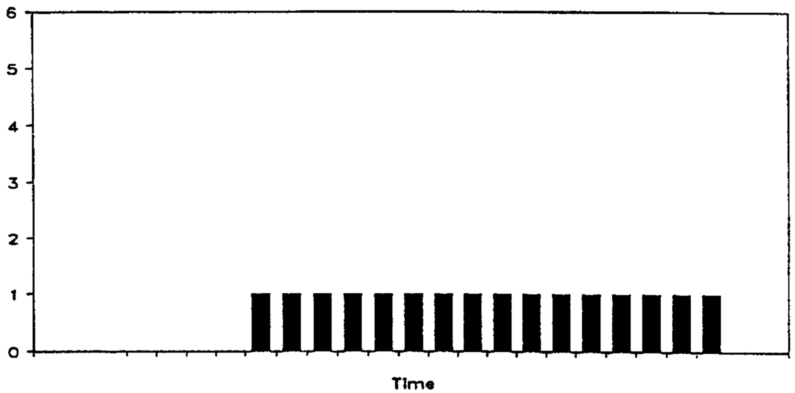
**FIGURE 1.2** Paths of input variable ($w_t$) and output variable ($y_t$) assumed for dynamic multiplier and present-value calculations.

Sometimes we might instead be interested in the consequences of a permanent change in $w$. A permanent change in $w$ means that $w_t, w_{t+1}, \ldots,$ and $w_{t+j}$ would all increase by one unit, as in Figure 1.3. From formula [1.1.10], the effect on $y_{t+j}$ of a permanent change in $w$ beginning in period $t$ is given by

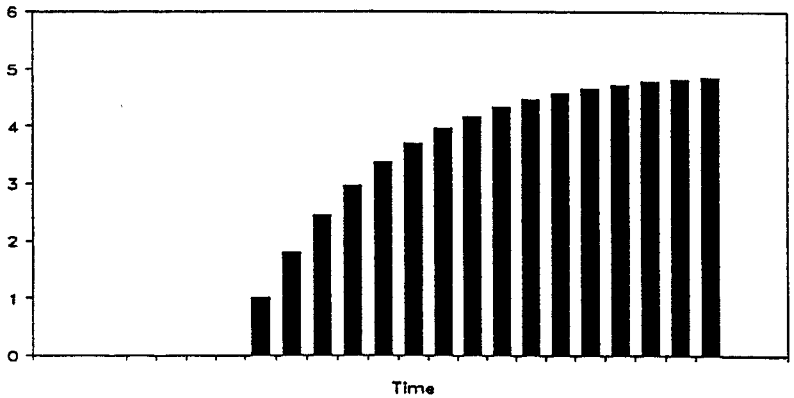
$$\frac{\partial y_{t+j}}{\partial w_t} + \frac{\partial y_{t+j}}{\partial w_{t+1}} + \frac{\partial y_{t+j}}{\partial w_{t+2}} + \cdots + \frac{\partial y_{t+j}}{\partial w_{t+j}} = \phi^j + \phi^{j-1} + \phi^{j-2} + \cdots + \phi + 1.$$

When $|\phi| < 1$, the limit of this expression as $j$ goes to infinity is sometimes described as the "long-run" effect of $w$ on $y$:

$$\lim_{j\to\infty}\left[\frac{\partial y_{t+j}}{\partial w_t} + \frac{\partial y_{t+j}}{\partial w_{t+1}} + \frac{\partial y_{t+j}}{\partial w_{t+2}} + \cdots + \frac{\partial y_{t+j}}{\partial w_{t+j}}\right] = 1 + \phi + \phi^2 + \cdots = 1/(1-\phi). \qquad [1.1.15]$$




(a) Value of $w$




(b) Value of $y$

**FIGURE 1.3** Paths of input variable ($w_t$) and output variable ($y_t$) assumed for long-run effect calculations.



For example, the long-run income elasticity of money demand in the system [1.1.2] is given by

$$\frac{0.19}{1 - 0.72} = 0.68.$$

A permanent 1% increase in income will eventually lead to a 0.68% increase in money demand.

Another related question concerns the cumulative consequences for $y$ of a one-time change in $w$. Here we consider a transitory disturbance to $w$ as in panel (a) of Figure 1.2, but wish to calculate the sum of the consequences for all future values of $y$. Another way to think of this is as the effect on the present value of $y$ [1.1.13] with the discount rate $\beta = 1$. Setting $\beta = 1$ in [1.1.14] shows this cumulative effect to be equal to

$$\sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial w_t} = 1/(1-\phi), \qquad [1.1.16]$$

provided that $|\phi| < 1$. Note that the cumulative effect on $y$ of a transitory change in $w$ (expression [1.1.16]) is the same as the long-run effect on $y$ of a permanent change in $w$ (expression [1.1.15]).

## 1.2. *p*th-Order Difference Equations

Let us now generalize the dynamic system [1.1.1] by allowing the value of $y$ at date $t$ to depend on $p$ of its own lags along with the current value of the input variable $w_t$:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + w_t. \qquad [1.2.1]$$

Equation [1.2.1] is a linear $p$th-order difference equation.

It is often convenient to rewrite the $p$th-order difference equation [1.2.1] in the scalar $y_t$ as a first-order difference equation in a vector $\boldsymbol{\xi}_t$. Define the $(p \times 1)$ vector $\boldsymbol{\xi}_t$ by

$$\boldsymbol{\xi}_t \equiv \begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{bmatrix}. \qquad [1.2.2]$$

That is, the first element of the vector $\boldsymbol{\xi}$ at date $t$ is the value $y$ took on at date $t$. The second element of $\boldsymbol{\xi}_t$ is the value $y$ took on at date $t-1$, and so on. Define the $(p \times p)$ matrix $\mathbf{F}$ by

$$\mathbf{F} \equiv \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \qquad [1.2.3]$$

For example, for $p = 4$, $\mathbf{F}$ refers to the following $4 \times 4$ matrix:

$$\mathbf{F} = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

For $p = 1$ (the first-order difference equation [1.1.1]), $\mathbf{F}$ is just the scalar $\phi$. Finally, define the $(p \times 1)$ vector $\mathbf{v}_t$ by

$$\mathbf{v}_t \equiv \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \qquad [1.2.4]$$

Consider the following first-order vector difference equation:

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t, \qquad [1.2.5]$$

or

$$\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This is a system of $p$ equations. The first equation in this system is identical to equation [1.2.1]. The second equation is simply the identity

$$y_{t-1} = y_{t-1},$$

owing to the fact that the second element of $\boldsymbol{\xi}_t$ is the same as the first element of $\boldsymbol{\xi}_{t-1}$. The third equation in [1.2.5] states that $y_{t-2} = y_{t-2}$; the $p$th equation states that $y_{t-p+1} = y_{t-p+1}$.

Thus, the first-order vector system [1.2.5] is simply an alternative representation of the $p$th-order scalar system [1.2.1]. The advantage of rewriting the $p$th-order system [1.2.1] in the form of a first-order system [1.2.5] is that first-order systems are often easier to work with than $p$th-order systems.

A dynamic multiplier for [1.2.5] can be found in exactly the same way as was done for the first-order scalar system of Section 1.1. If we knew the value of the vector $\boldsymbol{\xi}$ for date $t = -1$ and of $\mathbf{v}$ for date $t = 0$, we could find the value of $\boldsymbol{\xi}$ for date 0 from

$$\boldsymbol{\xi}_0 = \mathbf{F}\boldsymbol{\xi}_{-1} + \mathbf{v}_0.$$

The value of $\boldsymbol{\xi}$ for date 1 is

$$\boldsymbol{\xi}_1 = \mathbf{F}\boldsymbol{\xi}_0 + \mathbf{v}_1 = \mathbf{F}(\mathbf{F}\boldsymbol{\xi}_{-1} + \mathbf{v}_0) + \mathbf{v}_1 = \mathbf{F}^2\boldsymbol{\xi}_{-1} + \mathbf{F}\mathbf{v}_0 + \mathbf{v}_1.$$

Proceeding recursively in this fashion produces a generalization of [1.1.7]:

$$\boldsymbol{\xi}_t = \mathbf{F}^{t+1}\boldsymbol{\xi}_{-1} + \mathbf{F}^t\mathbf{v}_0 + \mathbf{F}^{t-1}\mathbf{v}_1 + \mathbf{F}^{t-2}\mathbf{v}_2 + \cdots + \mathbf{F}\mathbf{v}_{t-1} + \mathbf{v}_t. \qquad [1.2.6]$$



Writing this out in terms of the definitions of $\boldsymbol{\xi}$ and $\mathbf{v}$,

$$\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \mathbf{F}^{t+1} \begin{bmatrix} y_{-1} \\ y_{-2} \\ y_{-3} \\ \vdots \\ y_{-p} \end{bmatrix} + \mathbf{F}^t \begin{bmatrix} w_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \mathbf{F}^{t-1} \begin{bmatrix} w_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \mathbf{F}^1 \begin{bmatrix} w_{t-1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \qquad [1.2.7]$$

Consider the first equation of this system, which characterizes the value of $y_t$. Let $f_{11}^{(t)}$ denote the (1, 1) element of $\mathbf{F}^t$, $f_{12}^{(t)}$ the (1, 2) element of $\mathbf{F}^t$, and so on. Then the first equation of [1.2.7] states that

$$\begin{aligned} y_t = {} & f_{11}^{(t+1)}y_{-1} + f_{12}^{(t+1)}y_{-2} + \cdots + f_{1p}^{(t+1)}y_{-p} + f_{11}^{(t)}w_0 \\ & + f_{11}^{(t-1)}w_1 + \cdots + f_{11}^{(1)}w_{t-1} + w_t. \end{aligned} \qquad [1.2.8]$$

This describes the value of $y$ at date $t$ as a linear function of $p$ initial values of $y$ $(y_{-1}, y_{-2}, \ldots, y_{-p})$ and the history of the input variable $w$ since time 0 $(w_0, w_1, \ldots, w_t)$. Note that whereas only one initial value for $y$ (the value $y_{-1}$) was needed in the case of a first-order difference equation, $p$ initial values for $y$ (the values $y_{-1}, y_{-2}, \ldots, y_{-p}$) are needed in the case of a $p$th-order difference equation.

The obvious generalization of [1.1.9] is

$$\begin{aligned} \boldsymbol{\xi}_{t+j} = {} & \mathbf{F}^{j+1}\boldsymbol{\xi}_{t-1} + \mathbf{F}^j\mathbf{v}_t + \mathbf{F}^{j-1}\mathbf{v}_{t+1} + \mathbf{F}^{j-2}\mathbf{v}_{t+2} + \cdots \\ & + \mathbf{F}\mathbf{v}_{t+j-1} + \mathbf{v}_{t+j} \end{aligned} \qquad [1.2.9]$$

from which

$$\begin{aligned} y_{t+j} = {} & f_{11}^{(j+1)}y_{t-1} + f_{12}^{(j+1)}y_{t-2} + \cdots + f_{1p}^{(j+1)}y_{t-p} + f_{11}^{(j)}w_t \\ & + f_{11}^{(j-1)}w_{t+1} + f_{11}^{(j-2)}w_{t+2} + \cdots + f_{11}^{(1)}w_{t+j-1} + w_{t+j}. \end{aligned} \qquad [1.2.10]$$

Thus, for a $p$th-order difference equation, the dynamic multiplier is given by

$$\frac{\partial y_{t+j}}{\partial w_t} = f_{11}^{(j)} \qquad [1.2.11]$$

where $f_{11}^{(j)}$ denotes the (1, 1) element of $\mathbf{F}^j$. For $j = 1$, this is simply the (1, 1) element of $\mathbf{F}$, or the parameter $\phi_1$. Thus, for any $p$th-order system, the effect on $y_{t+1}$ of a one-unit increase in $w_t$ is given by the coefficient relating $y_t$ to $y_{t-1}$ in equation [1.2.1]:

$$\frac{\partial y_{t+1}}{\partial w_t} = \phi_1.$$

Direct multiplication of [1.2.3] reveals that the (1, 1) element of $\mathbf{F}^2$ is $(\phi_1^2 + \phi_2)$, so

$$\frac{\partial y_{t+2}}{\partial w_t} = \phi_1^2 + \phi_2$$

in a $p$th-order system.

For larger values of $j$, an easy way to obtain a numerical value for the dynamic multiplier $\partial y_{t+j}/\partial w_t$ is to simulate the system. This is done as follows. Set $y_{-1} = y_{-2} = \cdots = y_{-p} = 0$, $w_0 = 1$, and set the value of $w$ for all other dates to 0. Then use [1.2.1] to calculate the value of $y_t$ for $t = 0$ (namely, $y_0 = 1$). Next substitute this value along with $y_{t-1}, y_{t-2}, \ldots, y_{t-p+1}$ back into [1.2.1] to calculate $y_{t+1}$, and continue recursively in this fashion. The value of $y$ at step $t$ gives the effect of a one-unit change in $w_0$ on $y_t$.

Although numerical simulation may be adequate for many circumstances, it is also useful to have a simple analytical characterization of $\partial y_{t+j}/\partial w_t$, which, we know from [1.2.11], is given by the (1, 1) element of $\mathbf{F}^j$. This is fairly easy to obtain in terms of the eigenvalues of the matrix $\mathbf{F}$. Recall that the eigenvalues of a matrix $\mathbf{F}$ are those numbers $\lambda$ for which

$$|\mathbf{F} - \lambda \mathbf{I}_p| = 0. \qquad [1.2.12]$$

For example, for $p = 2$ the eigenvalues are the solutions to

$$\left| \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

or

$$\begin{vmatrix} (\phi_1 - \lambda) & \phi_2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - \phi_1\lambda - \phi_2 = 0. \qquad [1.2.13]$$

The two eigenvalues of $\mathbf{F}$ for a second-order difference equation are thus given by

$$\lambda_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \qquad [1.2.14]$$

$$\lambda_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}. \qquad [1.2.15]$$

For a general $p$th-order system, the determinant in [1.2.12] is a $p$th-order polynomial in $\lambda$ whose $p$ solutions characterize the $p$ eigenvalues of $\mathbf{F}$. This polynomial turns out to take a very similar form to [1.2.13]. The following result is proved in Appendix 1.A at the end of this chapter.

***Proposition 1.1:*** *The eigenvalues of the matrix* $\mathbf{F}$ *defined in equation [1.2.3] are the values of* $\lambda$ *that satisfy*

$$\lambda^p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_{p-1}\lambda - \phi_p = 0. \qquad [1.2.16]$$

Once we know the eigenvalues, it is straightforward to characterize the dynamic behavior of the system. First we consider the case when the eigenvalues of $\mathbf{F}$ are distinct; for example, we require that $\lambda_1$ and $\lambda_2$ in [1.2.14] and [1.2.15] be different numbers.



### *General Solution of a pth-Order Difference Equation with Distinct Eigenvalues*

Recall[2] that if the eigenvalues of a $(p \times p)$ matrix $\mathbf{F}$ are distinct, there exists a nonsingular $(p \times p)$ matrix $\mathbf{T}$ such that

$$\mathbf{F} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \qquad [1.2.17]$$

where $\mathbf{\Lambda}$ is a $(p \times p)$ matrix with the eigenvalues of $\mathbf{F}$ along the principal diagonal and zeros elsewhere:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p \end{bmatrix}. \qquad [1.2.18]$$

This enables us to characterize the dynamic multiplier (the (1, 1) element of $\mathbf{F}^j$ in [1.2.11]) very easily. For example, from [1.2.17] we can write $\mathbf{F}^2$ as

$$\begin{aligned} \mathbf{F}^2 &= \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \times \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \\ &= \mathbf{T} \times \mathbf{\Lambda} \times (\mathbf{T}^{-1}\mathbf{T}) \times \mathbf{\Lambda} \times \mathbf{T}^{-1} \\ &= \mathbf{T} \times \mathbf{\Lambda} \times \mathbf{I}_p \times \mathbf{\Lambda} \times \mathbf{T}^{-1} \\ &= \mathbf{T}\mathbf{\Lambda}^2\mathbf{T}^{-1}. \end{aligned}$$

The diagonal structure of $\mathbf{\Lambda}$ implies that $\mathbf{\Lambda}^2$ is also a diagonal matrix whose elements are the squares of the eigenvalues of $\mathbf{F}$:

$$\mathbf{\Lambda}^2 = \begin{bmatrix} \lambda_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p^2 \end{bmatrix}.$$

More generally, we can characterize $\mathbf{F}^j$ in terms of the eigenvalues of $\mathbf{F}$ as

$$\begin{aligned} \mathbf{F}^j &= \underbrace{\mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \times \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \times \cdots \times \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}}_{j \text{ terms}} \\ &= \mathbf{T} \times \mathbf{\Lambda} \times (\mathbf{T}^{-1}\mathbf{T}) \times \mathbf{\Lambda} \times (\mathbf{T}^{-1}\mathbf{T}) \times \cdots \times \mathbf{\Lambda} \times \mathbf{T}^{-1}, \end{aligned}$$

which simplifies to

$$\mathbf{F}^j = \mathbf{T}\mathbf{\Lambda}^j\mathbf{T}^{-1} \qquad [1.2.19]$$

where

$$\mathbf{\Lambda}^j = \begin{bmatrix} \lambda_1^j & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^j & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p^j \end{bmatrix}.$$

[2] See equation [A.4.24] in the Mathematical Review (Appendix A) at the end of the book.

Let $t_{ij}$ denote the row $i$, column $j$ element of $\mathbf{T}$ and let $t^{ij}$ denote the row $i$, column $j$ element of $\mathbf{T}^{-1}$. Equation [1.2.19] written out explicitly becomes

$$\mathbf{F}^j = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ t_{21} & t_{22} & \cdots & t_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ t_{p1} & t_{p2} & \cdots & t_{pp} \end{bmatrix} \begin{bmatrix} \lambda_1^j & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^j & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p^j \end{bmatrix} \begin{bmatrix} t^{11} & t^{12} & \cdots & t^{1p} \\ t^{21} & t^{22} & \cdots & t^{2p} \\ \vdots & \vdots & \cdots & \vdots \\ t^{p1} & t^{p2} & \cdots & t^{pp} \end{bmatrix}$$

$$= \begin{bmatrix} t_{11}\lambda_1^j & t_{12}\lambda_2^j & \cdots & t_{1p}\lambda_p^j \\ t_{21}\lambda_1^j & t_{22}\lambda_2^j & \cdots & t_{2p}\lambda_p^j \\ \vdots & \vdots & \cdots & \vdots \\ t_{p1}\lambda_1^j & t_{p2}\lambda_2^j & \cdots & t_{pp}\lambda_p^j \end{bmatrix} \begin{bmatrix} t^{11} & t^{12} & \cdots & t^{1p} \\ t^{21} & t^{22} & \cdots & t^{2p} \\ \vdots & \vdots & \cdots & \vdots \\ t^{p1} & t^{p2} & \cdots & t^{pp} \end{bmatrix}$$

from which the (1, 1) element of $\mathbf{F}^j$ is given by

$$f_{11}^{(j)} = [t_{11}t^{11}]\lambda_1^j + [t_{12}t^{21}]\lambda_2^j + \cdots + [t_{1p}t^{p1}]\lambda_p^j$$

or

$$f_{11}^{(j)} = c_1\lambda_1^j + c_2\lambda_2^j + \cdots + c_p\lambda_p^j \qquad [1.2.20]$$

where

$$c_i = [t_{1i}t^{i1}]. \qquad [1.2.21]$$

Note that the sum of the $c_i$ terms has the following interpretation:

$$c_1 + c_2 + \cdots + c_p = [t_{11}t^{11}] + [t_{12}t^{21}] + \cdots + [t_{1p}t^{p1}], \qquad [1.2.22]$$

which is the (1, 1) element of $\mathbf{T}\cdot\mathbf{T}^{-1}$. Since $\mathbf{T}\cdot\mathbf{T}^{-1}$ is just the $(p \times p)$ identity matrix, [1.2.22] implies that the $c_i$ terms sum to unity:

$$c_1 + c_2 + \cdots + c_p = 1. \qquad [1.2.23]$$

Substituting [1.2.20] into [1.2.11] gives the form of the dynamic multiplier for a $p$th-order difference equation:

$$\frac{\partial y_{t+j}}{\partial w_t} = c_1\lambda_1^j + c_2\lambda_2^j + \cdots + c_p\lambda_p^j. \qquad [1.2.24]$$

Equation [1.2.24] characterizes the dynamic multiplier as a weighted average of each of the $p$ eigenvalues raised to the $j$th power.

The following result provides a closed-form expression for the constants $(c_1, c_2, \ldots, c_p)$.

*Proposition 1.2: If the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ of the matrix* $\mathbf{F}$ *in [1.2.3] are distinct, then the magnitude $c_i$ in [1.2.21] can be written*

$$c_i = \frac{\lambda_i^{p-1}}{\prod_{\substack{k=1 \\ k \neq i}}^{p} (\lambda_i - \lambda_k)}. \qquad [1.2.25]$$

To summarize, the $p$th-order difference equation [1.2.1] implies that

$$y_{t+j} = f_{11}^{(j+1)}y_{t-1} + f_{12}^{(j+1)}y_{t-2} + \cdots + f_{1p}^{(j+1)}y_{t-p} \qquad [1.2.26]$$
$$+ w_{t+j} + \psi_1 w_{t+j-1} + \psi_2 w_{t+j-2} + \cdots + \psi_{j-1}w_{t+1} + \psi_j w_t.$$



The dynamic multiplier

$$\frac{\partial y_{t+j}}{\partial w_t} = \psi_j \qquad [1.2.27]$$

is given by the (1, 1) element of $\mathbf{F}^j$:

$$\psi_j = f_{11}^{(j)}. \qquad [1.2.28]$$

A closed-form expression for $\psi_j$ can be obtained by finding the eigenvalues of $\mathbf{F}$, or the values of $\lambda$ satisfying [1.2.16]. Denoting these $p$ values by $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and assuming them to be distinct, the dynamic multiplier is given by

$$\psi_j = c_1\lambda_1^j + c_2\lambda_2^j + \cdots + c_p\lambda_p^j \qquad [1.2.29]$$

where $(c_1, c_2, \ldots, c_p)$ is a set of constants summing to unity given by expression [1.2.25].

For a first-order system ($p = 1$), this rule would have us solve [1.2.16],

$$\lambda - \phi_1 = 0,$$

which has the single solution

$$\lambda_1 = \phi_1. \qquad [1.2.30]$$

According to [1.2.29], the dynamic multiplier is given by

$$\frac{\partial y_{t+j}}{\partial w_t} = c_1\lambda_1^j. \qquad [1.2.31]$$

From [1.2.23], $c_1 = 1$. Substituting this and [1.2.30] into [1.2.31] gives

$$\frac{\partial y_{t+j}}{\partial w_t} = \phi_1^j,$$

or the same result found in Section 1.1.

For higher-order systems, [1.2.29] allows a variety of more complicated dynamics. Suppose first that all the eigenvalues of $\mathbf{F}$ (or solutions to [1.2.16]) are real. This would be the case, for example, if $p = 2$ and $\phi_1^2 + 4\phi_2 > 0$ in the solutions [1.2.14] and [1.2.15] for the second-order system. If, furthermore, all of the eigenvalues are less than 1 in absolute value, then the system is stable, and its dynamics are represented as a weighted average of decaying exponentials or decaying exponentials oscillating in sign. For example, consider the following second-order difference equation:

$$y_t = 0.6y_{t-1} + 0.2y_{t-2} + w_t.$$

From equations [1.2.14] and [1.2.15], the eigenvalues of this system are given by

$$\lambda_1 = \frac{0.6 + \sqrt{(0.6)^2 + 4(0.2)}}{2} = 0.84$$

$$\lambda_2 = \frac{0.6 - \sqrt{(0.6)^2 + 4(0.2)}}{2} = -0.24.$$

From [1.2.25], we have

$$c_1 = \lambda_1/(\lambda_1 - \lambda_2) = 0.778$$
$$c_2 = \lambda_2/(\lambda_2 - \lambda_1) = 0.222.$$

The dynamic multiplier for this system,

$$\frac{\partial y_{t+j}}{\partial w_t} = c_1\lambda_1^j + c_2\lambda_2^j,$$

is plotted as a function of $j$ in panel (a) of Figure 1.4.[3] Note that as $j$ becomes larger, the pattern is dominated by the larger eigenvalue ($\lambda_1$), approximating a simple geometric decay at rate $\lambda_1$.

If the eigenvalues (the solutions to [1.2.16]) are real but at least one is greater than unity in absolute value, the system is explosive. If $\lambda_1$ denotes the eigenvalue that is largest in absolute value, the dynamic multiplier is eventually dominated by an exponential function of that eigenvalue:

$$\lim_{j\to\infty} \frac{\partial y_{t+j}}{\partial w_t} \cdot \frac{1}{\lambda_1^j} = c_1.$$

Other interesting possibilities arise if some of the eigenvalues are complex. Whenever this is the case, they appear as complex conjugates. For example, if $p = 2$ and $\phi_1^2 + 4\phi_2 < 0$, then the solutions $\lambda_1$ and $\lambda_2$ in [1.2.14] and [1.2.15] are complex conjugates. Suppose that $\lambda_1$ and $\lambda_2$ are complex conjugates, written as

$$\lambda_1 = a + bi \qquad [1.2.32]$$

$$\lambda_2 = a - bi. \qquad [1.2.33]$$

For the $p = 2$ case of [1.2.14] and [1.2.15], we would have

$$a = \phi_1/2 \qquad [1.2.34]$$

$$b = (1/2)\sqrt{-\phi_1^2 - 4\phi_2}. \qquad [1.2.35]$$

Our goal is to characterize the contribution to the dynamic multiplier $c_1\lambda_1^j$ when $\lambda_1$ is a complex number as in [1.2.32]. Recall that to raise a complex number to a power, we rewrite [1.2.32] in polar coordinate form:

$$\lambda_1 = R\cdot[\cos(\theta) + i\cdot\sin(\theta)], \qquad [1.2.36]$$

where $\theta$ and $R$ are defined in terms of $a$ and $b$ by the following equations:

$$R = \sqrt{a^2 + b^2}$$

$$\cos(\theta) = a/R$$

$$\sin(\theta) = b/R.$$

Note that $R$ is equal to the modulus of the complex number $\lambda_1$.

The eigenvalue $\lambda_1$ in [1.2.36] can be written as[4]

$$\lambda_1 = R[e^{i\theta}],$$

and so

$$\lambda_1^j = R^j[e^{i\theta j}] = R^j[\cos(\theta j) + i\cdot\sin(\theta j)]. \qquad [1.2.37]$$

Analogously, if $\lambda_2$ is the complex conjugate of $\lambda_1$, then

$$\lambda_2 = R[\cos(\theta) - i\cdot\sin(\theta)],$$

which can be written[5]

$$\lambda_2 = R[e^{-i\theta}].$$

Thus

$$\lambda_2^j = R^j[e^{-i\theta j}] = R^j[\cos(\theta j) - i\cdot\sin(\theta j)]. \qquad [1.2.38]$$

[3] Again, if one's purpose is solely to generate a numerical plot as in Figure 1.4, the easiest approach is numerical simulation of the system.

[4] See equation [A.3.25] in the Mathematical Review (Appendix A) at the end of the book.

[5] See equation [A.3.26].



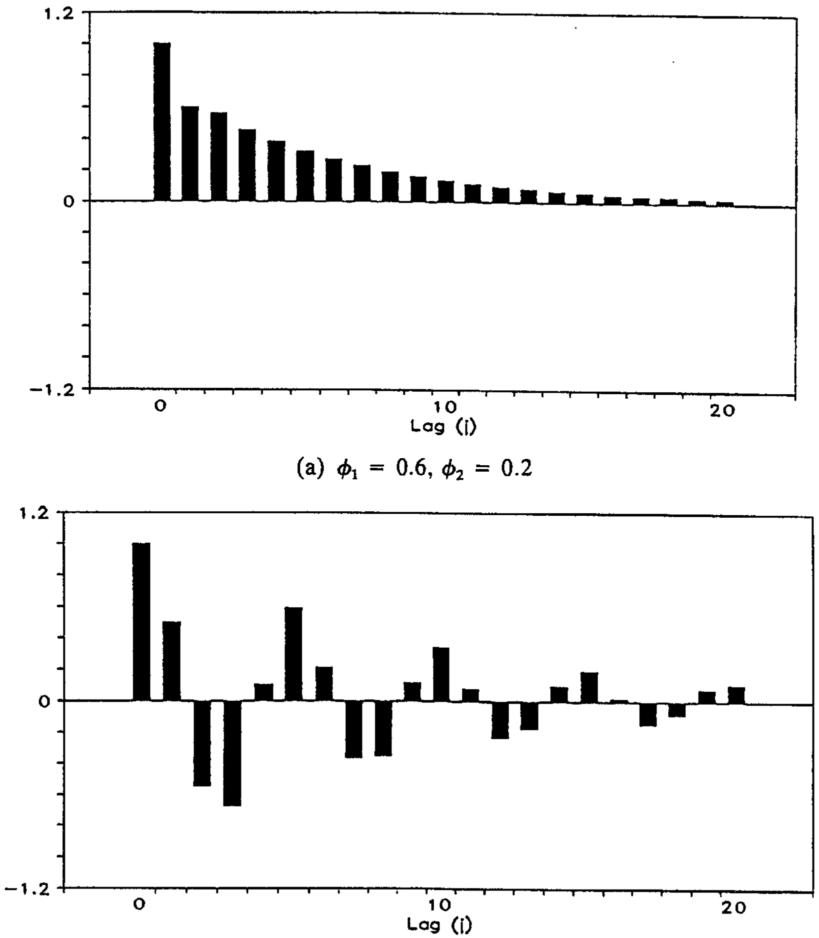



**FIGURE 1.4** Dynamic multiplier for second-order difference equation for different values of $\phi_1$ and $\phi_2$ (plot of $\partial y_{t+j}/\partial w_t$ as a function of the lag $j$).

Substituting [1.2.37] and [1.2.38] into [1.2.29] gives the contribution of the complex conjugates to the dynamic multiplier $\partial y_{t+j}/\partial w_t$:

$$\begin{aligned} c_1\lambda_1^j + c_2\lambda_2^j &= c_1R^j[\cos(\theta j) + i\cdot\sin(\theta j)] + c_2R^j[\cos(\theta j) - i\cdot\sin(\theta j)] \\ &= [c_1 + c_2]\cdot R^j\cdot\cos(\theta j) + i\cdot[c_1 - c_2]\cdot R^j\cdot\sin(\theta j). \end{aligned} \qquad [1.2.39]$$

The appearance of the imaginary number $i$ in [1.2.39] may seem a little troubling. After all, this calculation was intended to give the effect of a change in the real-valued variable $w_t$ on the real-valued variable $y_{t+j}$ as predicted by the real-valued system [1.2.1], and it would be odd indeed if the correct answer involved the imaginary number $i$! Fortunately, it turns out from [1.2.25] that if $\lambda_1$ and $\lambda_2$ are complex conjugates, then $c_1$ and $c_2$ are complex conjugates; that is, they can be written as

$$c_1 = \alpha + \beta i$$
$$c_2 = \alpha - \beta i$$

for some real numbers $\alpha$ and $\beta$. Substituting these expressions into [1.2.39] yields

$$\begin{aligned} c_1\lambda_1^j + c_2\lambda_2^j &= [(\alpha + \beta i) + (\alpha - \beta i)]\cdot R^j\cos(\theta j) + i\cdot[(\alpha + \beta i) - (\alpha - \beta i)]\cdot R^j\sin(\theta j) \\ &= [2\alpha]\cdot R^j\cos(\theta j) + i\cdot[2\beta i]\cdot R^j\sin(\theta j) \\ &= 2\alpha R^j\cos(\theta j) - 2\beta R^j\sin(\theta j), \end{aligned}$$

which is strictly real.

Thus, when some of the eigenvalues are complex, they contribute terms proportional to $R^j\cos(\theta j)$ and $R^j\sin(\theta j)$ to the dynamic multiplier $\partial y_{t+j}/\partial w_t$. Note that if $R = 1$—that is, if the complex eigenvalues have unit modulus—the multipliers are periodic sine and cosine functions of $j$. A given increase in $w_t$ increases $y_{t+j}$ for some ranges of $j$ and decreases $y_{t+j}$ over other ranges, with the impulse never dying out as $j \to \infty$. If the complex eigenvalues are less than 1 in modulus ($R < 1$), the impulse again follows a sinusoidal pattern though its amplitude decays at the rate $R^j$. If the complex eigenvalues are greater than 1 in modulus ($R > 1$), the amplitude of the sinusoids explodes at the rate $R^j$.

For an example of dynamic behavior characterized by decaying sinusoids, consider the second-order system

$$y_t = 0.5y_{t-1} - 0.8y_{t-2} + w_t.$$

The eigenvalues for this system are given from [1.2.14] and [1.2.15]:

$$\lambda_1 = \frac{0.5 + \sqrt{(0.5)^2 - 4(0.8)}}{2} = 0.25 + 0.86i$$

$$\lambda_2 = \frac{0.5 - \sqrt{(0.5)^2 - 4(0.8)}}{2} = 0.25 - 0.86i,$$

with modulus

$$R = \sqrt{(0.25)^2 + (0.86)^2} = 0.9.$$

Since $R < 1$, the dynamic multiplier follows a pattern of damped oscillation plotted in panel (b) of Figure 1.4. The frequency[6] of these oscillations is given by the parameter $\theta$ in [1.2.39], which was defined implicitly by

$$\cos(\theta) = a/R = (0.25)/(0.9) = 0.28$$

or

$$\theta = 1.29.$$

The cycles associated with the dynamic multiplier function [1.2.39] thus have a period of

$$\frac{2\pi}{\theta} = \frac{(2)(3.14159)}{1.29} = 4.9;$$

that is, the peaks in the pattern in panel (b) of Figure 1.4 appear about five periods apart.

[6]See Section A.1 of the Mathematical Review (Appendix A) at the end of the book for a discussion of the frequency and period of a sinusoidal function.



### *Solution of a Second-Order Difference Equation with Distinct Eigenvalues*

The second-order difference equation ($p = 2$) comes up sufficiently often that it is useful to summarize the properties of the solution as a general function of $\phi_1$ and $\phi_2$, which we now do.[7]

The eigenvalues $\lambda_1$ and $\lambda_2$ in [1.2.14] and [1.2.15] are complex whenever

$$\phi_1^2 + 4\phi_2 < 0,$$

or whenever ($\phi_1$, $\phi_2$) lies below the parabola indicated in Figure 1.5. For the case of complex eigenvalues, the modulus $R$ satisfies

$$R^2 = a^2 + b^2,$$

or, from [1.2.34] and [1.2.35],

$$R^2 = (\phi_1/2)^2 - (\phi_1^2 + 4\phi_2)/4 = -\phi_2.$$

Thus, a system with complex eigenvalues is explosive whenever $\phi_2 < -1$. Also, when the eigenvalues are complex, the frequency of oscillations is given by

$$\theta = \cos^{-1}(a/R) = \cos^{-1}[\phi_1/(2\sqrt{-\phi_2})],$$

where "$\cos^{-1}(x)$" denotes the inverse of the cosine function, or the radian measure of an angle whose cosine is $x$.

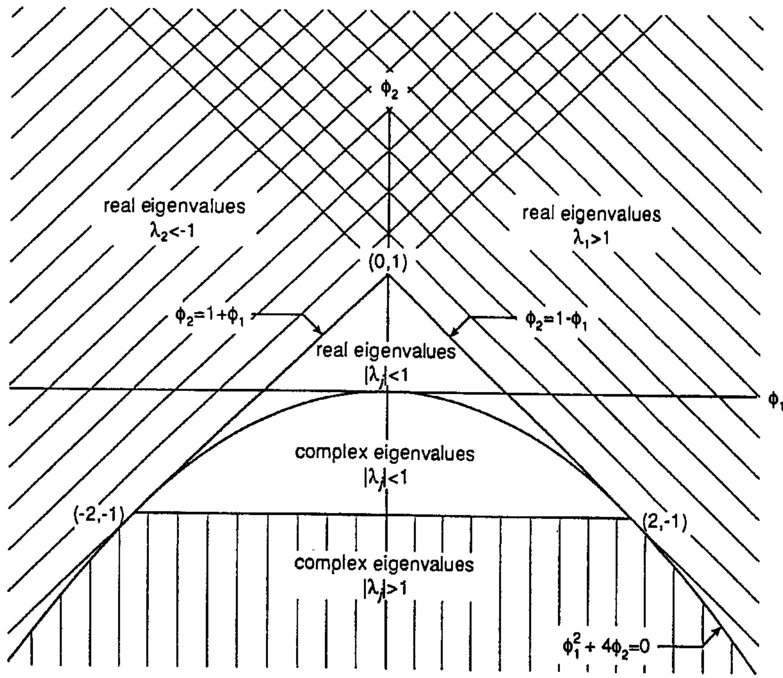



**FIGURE 1.5** Summary of dynamics for a second-order difference equation.

[7]This discussion closely follows Sargent (1987, pp. 188–89).

For the case of real eigenvalues, the arithmetically larger eigenvalue ($\lambda_1$) will be greater than unity whenever

$$\frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} > 1$$

or

$$\sqrt{\phi_1^2 + 4\phi_2} > 2 - \phi_1.$$

Assuming that $\lambda_1$ is real, the left side of this expression is a positive number and the inequality would be satisfied for any value of $\phi_1 > 2$. If, on the other hand, $\phi_1 < 2$, we can square both sides to conclude that $\lambda_1$ will exceed unity whenever

$$\phi_1^2 + 4\phi_2 > 4 - 4\phi_1 + \phi_1^2$$

or

$$\phi_2 > 1 - \phi_1.$$

Thus, in the real region, $\lambda_1$ will be greater than unity either if $\phi_1 > 2$ or if $(\phi_1, \phi_2)$ lies northeast of the line $\phi_2 = 1 - \phi_1$ in Figure 1.5. Similarly, with real eigenvalues, the arithmetically smaller eigenvalue ($\lambda_2$) will be less than $-1$ whenever

$$\frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} < -1$$

$$-\sqrt{\phi_1^2 + 4\phi_2} < -2 - \phi_1$$

$$\sqrt{\phi_1^2 + 4\phi_2} > 2 + \phi_1.$$

Again, if $\phi_1 < -2$, this must be satisfied, and in the case when $\phi_1 > -2$, we can square both sides:

$$\phi_1^2 + 4\phi_2 > 4 + 4\phi_1 + \phi_1^2$$

$$\phi_2 > 1 + \phi_1.$$

Thus, in the real region, $\lambda_2$ will be less than $-1$ if either $\phi_1 < -2$ or $(\phi_1, \phi_2)$ lies to the northwest of the line $\phi_2 = 1 + \phi_1$ in Figure 1.5.

The system is thus stable whenever $(\phi_1, \phi_2)$ lies within the triangular region of Figure 1.5.

## *General Solution of a pth-Order Difference Equation with Repeated Eigenvalues*

In the more general case of a difference equation for which $\mathbf{F}$ has repeated eigenvalues and $s < p$ linearly independent eigenvectors, result [1.2.17] is generalized by using the Jordan decomposition,

$$\mathbf{F} = \mathbf{M}\mathbf{J}\mathbf{M}^{-1} \qquad [1.2.40]$$

where $\mathbf{M}$ is a $(p \times p)$ matrix and $\mathbf{J}$ takes the form

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_s \end{bmatrix}$$



with

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} \qquad [1.2.41]$$

for $\lambda_i$ an eigenvalue of $\mathbf{F}$. If [1.2.17] is replaced by [1.2.40], then equation [1.2.19] generalizes to

$$\mathbf{F}^j = \mathbf{M}\mathbf{J}^j\mathbf{M}^{-1} \qquad [1.2.42]$$

where

$$\mathbf{J}^j = \begin{bmatrix} \mathbf{J}_1^j & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2^j & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_s^j \end{bmatrix}.$$

Moreover, from [1.2.41], if $\mathbf{J}_i$ is of dimension $(n_i \times n_i)$, then[8]

$$\mathbf{J}_i^j = \begin{bmatrix} \lambda_i^j & \binom{j}{1}\lambda_i^{j-1} & \binom{j}{2}\lambda_i^{j-2} & \cdots & \binom{j}{n_i-1}\lambda_i^{j-n_i+1} \\ 0 & \lambda_i^j & \binom{j}{1}\lambda_i^{j-1} & \cdots & \binom{j}{n_i-2}\lambda_i^{j-n_i+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i^j \end{bmatrix} \qquad [1.2.43]$$

where

$$\binom{j}{n} \equiv \begin{cases} \dfrac{j(j-1)(j-2)\cdots(j-n+1)}{n(n-1)\cdots 3\cdot 2\cdot 1} & \text{for } j \geq n \\ 0 & \text{otherwise.} \end{cases}$$

Equation [1.2.43] may be verified by induction by multiplying [1.2.41] by [1.2.43] and noticing that $\binom{j}{n} + \binom{j}{n-1} = \binom{j+1}{n}$.

For example, consider again the second-order difference equation, this time with repeated roots. Then

$$\mathbf{F}^j = \mathbf{M}\begin{bmatrix} \lambda^j & j\lambda^{j-1} \\ 0 & \lambda^j \end{bmatrix}\mathbf{M}^{-1},$$

so that the dynamic multiplier takes the form

$$\frac{\partial y_{t+j}}{\partial w_t} = f_{11}^{(j)} = k_1\lambda^j + k_2 j\lambda^{j-1}.$$

## *Long-Run and Present-Value Calculations*

If the eigenvalues are all less than 1 in modulus, then $\mathbf{F}^j$ in [1.2.9] goes to zero as $j$ becomes large. If all values of $w$ and $y$ are taken to be bounded, we can

[8]This expression is taken from Chiang (1980, p. 444).

think of a "solution" of $y_t$ in terms of the infinite history of $w$,

$$y_t = w_t + \psi_1 w_{t-1} + \psi_2 w_{t-2} + \psi_3 w_{t-3} + \cdots, \qquad [1.2.44]$$

where $\psi_j$ is given by the (1, 1) element of $\mathbf{F}^j$ and takes the particular form of [1.2.29] in the case of distinct eigenvalues.

It is also straightforward to calculate the effect on the present value of $y$ of a transitory increase in $w$. This is simplest to find if we first consider the slightly more general problem of the hypothetical consequences of a change in any element of the vector $\mathbf{v}_t$ on any element of $\boldsymbol{\xi}_{t+j}$ in a general system of the form of [1.2.5]. The answer to this more general problem can be inferred immediately from [1.2.9]:

$$\frac{\partial \boldsymbol{\xi}_{t+j}}{\partial \mathbf{v}_t'} = \mathbf{F}^j. \qquad [1.2.45]$$

The true dynamic multiplier of interest, $\partial y_{t+j}/\partial w_t$, is just the (1, 1) element of the $(p \times p)$ matrix in [1.2.45]. The effect on the present value of $\boldsymbol{\xi}$ of a change in $\mathbf{v}$ is given by

$$\frac{\partial \sum_{j=0}^{\infty} \beta^j \boldsymbol{\xi}_{t+j}}{\partial \mathbf{v}_t'} = \sum_{j=0}^{\infty} \beta^j \mathbf{F}^j = (\mathbf{I}_p - \beta\mathbf{F})^{-1}, \qquad [1.2.46]$$

provided that the eigenvalues of $\mathbf{F}$ are all less than $\beta^{-1}$ in modulus. The effect on the present value of $y$ of a change in $w$,

$$\frac{\partial \sum_{j=0}^{\infty} \beta^j y_{t+j}}{\partial w_t},$$

is thus the (1, 1) element of the $(p \times p)$ matrix in [1.2.46]. This value is given by the following proposition.

*Proposition 1.3: If the eigenvalues of the $(p \times p)$ matrix $\mathbf{F}$ defined in [1.2.3] are all less than $\beta^{-1}$ in modulus, then the matrix $(\mathbf{I}_p - \beta\mathbf{F})^{-1}$ exists and the effect of $w$ on the present value of $y$ is given by its (1, 1) element:*

$$1/(1 - \phi_1\beta - \phi_2\beta^2 - \cdots - \phi_{p-1}\beta^{p-1} - \phi_p\beta^p).$$

Note that Proposition 1.3 includes the earlier result for a first-order system (equation [1.1.14]) as a special case.

The cumulative effect of a one-time change in $w_t$ on $y_t, y_{t+1}, \ldots$ can be considered a special case of Proposition 1.3 with no discounting. Setting $\beta = 1$ in Proposition 1.3 shows that, provided the eigenvalues of $\mathbf{F}$ are all less than 1 in modulus, the cumulative effect of a one-time change in $w$ on $y$ is given by

$$\sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial w_t} = 1/(1 - \phi_1 - \phi_2 - \cdots - \phi_p). \qquad [1.2.47]$$

Notice again that [1.2.47] can alternatively be interpreted as giving the eventual long-run effect on $y$ of a permanent change in $w$:

$$\lim_{j\to\infty} \frac{\partial y_{t+j}}{\partial w_t} + \frac{\partial y_{t+j}}{\partial w_{t+1}} + \frac{\partial y_{t+j}}{\partial w_{t+2}} + \cdots + \frac{\partial y_{t+j}}{\partial w_{t+j}} = 1/(1 - \phi_1 - \phi_2 - \cdots - \phi_p).$$



## APPENDIX 1.A. *Proofs of Chapter 1 Propositions*

■ **Proof of Proposition 1.1.** The eigenvalues of $\mathbf{F}$ satisfy

$$|\mathbf{F} - \lambda\mathbf{I}_p| = 0. \qquad [1.A.1]$$

For the matrix $\mathbf{F}$ defined in equation [1.2.3], this determinant would be

$$\left| \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \right|$$

$$= \begin{vmatrix} (\phi_1 - \lambda) & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & -\lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\lambda \end{vmatrix}. \qquad [1.A.2]$$

Recall that if we multiply a column of a matrix by a constant and add the result to another column, the determinant of the matrix is unchanged. If we multiply the $p$th column of the matrix in [1.A.2] by $(1/\lambda)$ and add the result to the $(p-1)$th column, the result is a matrix with the same determinant as that in [1.A.2]:

$$|\mathbf{F} - \lambda\mathbf{I}_p| = \begin{vmatrix} \phi_1 - \lambda & \phi_2 & \phi_3 & \cdots & \phi_{p-2} & \phi_{p-1} + (\phi_p/\lambda) & \phi_p \\ 1 & -\lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -\lambda \end{vmatrix}.$$

Next, multiply the $(p-1)$th column by $(1/\lambda)$ and add the result to the $(p-2)$th column:

$$|\mathbf{F} - \lambda\mathbf{I}_p| = \begin{vmatrix} \phi_1 - \lambda & \phi_2 & \phi_3 & \cdots & \phi_{p-2} + \phi_{p-1}/\lambda + \phi_p/\lambda^2 & \phi_{p-1} + \phi_p/\lambda & \phi_p \\ 1 & -\lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -\lambda \end{vmatrix}.$$

Continuing in this fashion shows [1.A.1] to be equivalent to the determinant of the following upper triangular matrix:

$$|\mathbf{F} - \lambda\mathbf{I}_p| = \begin{vmatrix} \phi_1 - \lambda + \phi_2/\lambda + \phi_3/\lambda^2 + \cdots + \phi_p/\lambda^{p-1} & \phi_2 + \phi_3/\lambda + \phi_4/\lambda^2 + \cdots + \phi_p/\lambda^{p-2} & \cdots & \phi_{p-1} + \phi_p/\lambda & \phi_p \\ 0 & -\lambda & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & -\lambda & 0 \\ 0 & 0 & \cdots & 0 & -\lambda \end{vmatrix}.$$

But the determinant of an upper triangular matrix is simply the product of the terms along the principal diagonal:

$$\begin{aligned} |\mathbf{F} - \lambda\mathbf{I}_p| &= [\phi_1 - \lambda + \phi_2/\lambda + \phi_3/\lambda^2 + \cdots + \phi_p/\lambda^{p-1}] \cdot [-\lambda]^{p-1} \\ &= (-1)^p \cdot [\lambda^p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_p]. \end{aligned} \qquad [1.A.3]$$

The eigenvalues of $\mathbf{F}$ are thus the values of $\lambda$ for which [1.A.3] is zero, or for which

$$\lambda^p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_p = 0,$$

as asserted in Proposition 1.1. ■

■ **Proof of Proposition 1.2.** Assuming that the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ are distinct, the matrix $\mathbf{T}$ in equation [1.2.17] can be constructed from the eigenvectors of $\mathbf{F}$. Let $\mathbf{t}_i$ denote the following $(p \times 1)$ vector,

$$\mathbf{t}_i = \begin{bmatrix} \lambda_i^{p-1} \\ \lambda_i^{p-2} \\ \lambda_i^{p-3} \\ \vdots \\ \lambda_i^1 \\ 1 \end{bmatrix}, \qquad [1.A.4]$$

where $\lambda_i$ denotes the $i$th eigenvalue of $\mathbf{F}$. Notice

$$\mathbf{Ft}_i = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_i^{p-1} \\ \lambda_i^{p-2} \\ \lambda_i^{p-3} \\ \vdots \\ \lambda_i^1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \phi_1\lambda_i^{p-1} + \phi_2\lambda_i^{p-2} + \phi_3\lambda_i^{p-3} + \cdots + \phi_{p-1}\lambda_i + \phi_p \\ \lambda_i^{p-1} \\ \lambda_i^{p-2} \\ \vdots \\ \lambda_i^2 \\ \lambda_i \end{bmatrix}. \qquad [1.A.5]$$

Since $\lambda_i$ is an eigenvalue of $\mathbf{F}$, it satisfies [1.2.16]:

$$\lambda_i^p - \phi_1\lambda_i^{p-1} - \phi_2\lambda_i^{p-2} - \cdots - \phi_{p-1}\lambda_i - \phi_p = 0. \qquad [1.A.6]$$

Substituting [1.A.6] into [1.A.5] reveals

$$\mathbf{Ft}_i = \begin{bmatrix} \lambda_i^p \\ \lambda_i^{p-1} \\ \lambda_i^{p-2} \\ \vdots \\ \lambda_i^2 \\ \lambda_i \end{bmatrix} = \lambda_i \begin{bmatrix} \lambda_i^{p-1} \\ \lambda_i^{p-2} \\ \lambda_i^{p-3} \\ \vdots \\ \lambda_i^1 \\ 1 \end{bmatrix}$$

or

$$\mathbf{Ft}_i = \lambda_i\mathbf{t}_i. \qquad [1.A.7]$$

Thus $\mathbf{t}_i$ is an eigenvector of $\mathbf{F}$ associated with the eigenvalue $\lambda_i$.

We can calculate the matrix $\mathbf{T}$ by combining the eigenvectors $(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_p)$ into a $(p \times p)$ matrix

$$\mathbf{T} = [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \cdots \quad \mathbf{t}_p]. \qquad [1.A.8]$$

To calculate the particular values for $c_i$ in equation [1.2.21], recall that $\mathbf{T}^{-1}$ is characterized by

$$\mathbf{TT}^{-1} = \mathbf{I}_p, \qquad [1.A.9]$$



where $\mathbf{T}$ is given by [1.A.4] and [1.A.8]. Writing out the first column of the matrix system of equations [1.A.9] explicitly, we have

$$\begin{bmatrix} \lambda_1^{p-1} & \lambda_2^{p-1} & \cdots & \lambda_p^{p-1} \\ \lambda_1^{p-2} & \lambda_2^{p-2} & \cdots & \lambda_p^{p-2} \\ \lambda_1^{p-3} & \lambda_2^{p-3} & \cdots & \lambda_p^{p-3} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_p^1 \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} t^{11} \\ t^{21} \\ t^{31} \\ \vdots \\ t^{p-1,1} \\ t^{p1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

This gives a system of $p$ linear equations in the $p$ unknowns $(t^{11}, t^{21}, \ldots, t^{p1})$. Provided that the $\lambda_i$ are all distinct, the solution can be shown to be[9]

$$t^{11} = \frac{1}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \cdots (\lambda_1 - \lambda_p)}$$

$$t^{21} = \frac{1}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3) \cdots (\lambda_2 - \lambda_p)}$$

$$\vdots$$

$$t^{p1} = \frac{1}{(\lambda_p - \lambda_1)(\lambda_p - \lambda_2) \cdots (\lambda_p - \lambda_{p-1})}.$$

Substituting these values into [1.2.21] gives equation [1.2.25]. ■

■ **Proof of Proposition 1.3.** The first claim in this proposition is that if the eigenvalues of $\mathbf{F}$ are less than $\beta^{-1}$ in modulus, then the inverse of $(\mathbf{I}_p - \beta\mathbf{F})$ exists. Suppose the inverse of $(\mathbf{I}_p - \beta\mathbf{F})$ did not exist. Then the determinant $|\mathbf{I}_p - \beta\mathbf{F}|$ would have to be zero. But

$$|\mathbf{I}_p - \beta\mathbf{F}| = |-\beta \cdot [\mathbf{F} - \beta^{-1}\mathbf{I}_p]| = (-\beta)^p|\mathbf{F} - \beta^{-1}\mathbf{I}_p|,$$

so that $|\mathbf{F} - \beta^{-1}\mathbf{I}_p|$ would have to be zero whenever the inverse of $(\mathbf{I}_p - \beta\mathbf{F})$ fails to exist. But this would mean that $\beta^{-1}$ is an eigenvalue of $\mathbf{F}$, which is ruled out by the assumption that all eigenvalues of $\mathbf{F}$ are strictly less than $\beta^{-1}$ in modulus. Thus, the matrix $\mathbf{I}_p - \beta\mathbf{F}$ must be nonsingular.

Since $[\mathbf{I}_p - \beta\mathbf{F}]^{-1}$ exists, it satisfies the equation

$$[\mathbf{I}_p - \beta\mathbf{F}]^{-1}[\mathbf{I}_p - \beta\mathbf{F}] = \mathbf{I}_p. \qquad [1.A.10]$$

Let $x_{ij}$ denote the row $i$, column $j$ element of $[\mathbf{I}_p - \beta\mathbf{F}]^{-1}$, and write [1.A.10] as

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pp} \end{bmatrix} \begin{bmatrix} 1 - \beta\phi_1 & -\beta\phi_2 & \cdots & -\beta\phi_{p-1} & -\beta\phi_p \\ -\beta & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & -\beta & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \qquad [1.A.11]$$

The task is then to find the (1, 1) element of $[\mathbf{I}_p - \beta\mathbf{F}]^{-1}$, that is, to find the value of $x_{11}$. To do this we need only consider the first row of equations in [1.A.11]:

$$[x_{11} \quad x_{12} \quad \cdots \quad x_{1p}] \begin{bmatrix} 1 - \beta\phi_1 & -\beta\phi_2 & \cdots & -\beta\phi_{p-1} & -\beta\phi_p \\ -\beta & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & -\beta & 1 \end{bmatrix}$$

$$= [1 \quad 0 \quad \cdots \quad 0 \quad 0]. \qquad [1.A.12]$$

[9]See Lemma 2 of Chiang (1980, p. 144).

Consider postmultiplying this system of equations by a matrix with 1s along the principal diagonal, $\beta$ in the row $p$, column $p - 1$ position, and 0s elsewhere:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \beta & 1 \end{bmatrix}.$$

The effect of this operation is to multiply the $p$th column of a matrix by $\beta$ and add the result to the $(p - 1)$th column:

$$[x_{11} \quad x_{12} \cdots x_{1p}] \begin{bmatrix} 1 - \beta\phi_1 & -\beta\phi_2 & \cdots & -\beta\phi_{p-1} - \beta^2\phi_p & -\beta\phi_p \\ -\beta & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = [1 \quad 0 \cdots 0 \quad 0].$$

Next multiply the $(p - 1)$th column by $\beta$ and add the result to the $(p - 2)$th column. Proceeding in this fashion, we arrive at

$$[x_{11} \quad x_{12} \quad \cdots \quad x_{1p}] \times$$

$$\begin{bmatrix} 1 - \beta\phi_1 - \beta^2\phi_2 - \cdots - \beta^{p-1}\phi_{p-1} - \beta^p\phi_p & -\beta\phi_2 - \beta^2\phi_3 - \cdots - \beta^{p-1}\phi_p & \cdots & -\beta\phi_{p-1} - \beta^2\phi_p & -\beta\phi_p \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$= [1 \quad 0 \quad \cdots \quad 0 \quad 0]. \qquad [1.A.13]$$

The first equation in [1.A.13] states that

$$x_{11} \cdot (1 - \beta\phi_1 - \beta^2\phi_2 - \cdots - \beta^{p-1}\phi_{p-1} - \beta^p\phi_p) = 1$$

or

$$x_{11} = 1/(1 - \beta\phi_1 - \beta^2\phi_2 - \cdots - \beta^p\phi_p),$$

as claimed in Proposition 1.3. ■

## *Chapter 1 References*


Chiang, Chin Long. 1980. *An Introduction to Stochastic Processes and Their Applications.* Huntington, N.Y.: Krieger.

Goldfeld, Stephen M. 1973. "The Demand for Money Revisited," *Brookings Papers on Economic Activity* 3:577–638.

Sargent, Thomas J. 1987. *Macroeconomic Theory*, 2d ed. Boston: Academic Press.




# 2 *Lag Operators*

## 2.1. *Introduction*

The previous chapter analyzed the dynamics of linear difference equations using matrix algebra. This chapter develops some of the same results using time series operators. We begin with some introductory remarks on some useful time series operators.

A time series is a collection of observations indexed by the date of each observation. Usually we have collected data beginning at some particular date (say, $t = 1$) and ending at another (say, $t = T$):

$$(y_1, y_2, \ldots, y_T).$$

We often imagine that we could have obtained earlier observations $(y_0, y_{-1}, y_{-2}, \ldots)$ or later observations $(y_{T+1}, y_{T+2}, \ldots)$ had the process been observed for more time. The observed sample $(y_1, y_2, \ldots, y_T)$ could then be viewed as a finite segment of a doubly infinite sequence, denoted $\{y_t\}_{t=-\infty}^{\infty}$:

$$\{y_t\}_{t=-\infty}^{\infty} = \{\ldots, y_{-1}, y_0, \underbrace{y_1, y_2, \ldots, y_T}_{\text{observed sample}}, y_{T+1}, y_{T+2}, \ldots\}.$$

Typically, a time series $\{y_t\}_{t=-\infty}^{\infty}$ is identified by describing the $t$th element. For example, a *time trend* is a series whose value at date $t$ is simply the date of the observation:

$$y_t = t.$$

We could also consider a time series in which each element is equal to a constant $c$, regardless of the date of the observation $t$:

$$y_t = c.$$

Another important time series is a *Gaussian white noise process*, denoted

$$y_t = \varepsilon_t,$$

where $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is a sequence of independent random variables each of which has a $N(0, \sigma^2)$ distribution.

We are used to thinking of a function such as $y = f(x)$ or $y = g(x, w)$ as an operation that accepts as input a number $(x)$ or group of numbers $(x, w)$ and produces the output $(y)$. A time series *operator* transforms one time series or group of time series into a new time series. It accepts as input a sequence such as $\{x_t\}_{t=-\infty}^{\infty}$ or a group of sequences such as $(\{x_t\}_{t=-\infty}^{\infty}, \{w_t\}_{t=-\infty}^{\infty})$ and has as output a new sequence $\{y_t\}_{t=-\infty}^{\infty}$. Again, the operator is summarized by describing the value of a typical element of $\{y_t\}_{t=-\infty}^{\infty}$ in terms of the corresponding elements of $\{x_t\}_{t=-\infty}^{\infty}$.

An example of a time series operator is the multiplication operator, represented as

$$y_t = \beta x_t. \qquad [2.1.1]$$

Although it is written exactly the same way as simple scalar multiplication, equation [2.1.1] is actually shorthand for an infinite sequence of multiplications, one for each date $t$. The operator multiplies the value $x$ takes on at any date $t$ by some constant $\beta$ to generate the value of $y$ for that date.

Another example of a time series operator is the addition operator:

$$y_t = x_t + w_t.$$

Here the value of $y$ at any date $t$ is the sum of the values that $x$ and $w$ take on for that date.

Since the multiplication or addition operators amount to element-by-element multiplication or addition, they obey all the standard rules of algebra. For example, if we multiply each observation of $\{x_t\}_{t=-\infty}^{\infty}$ by $\beta$ and each observation of $\{w_t\}_{t=-\infty}^{\infty}$ by $\beta$ and add the results,

$$\beta x_t + \beta w_t,$$

the outcome is the same as if we had first added $\{x_t\}_{t=-\infty}^{\infty}$ to $\{w_t\}_{t=-\infty}^{\infty}$ and then multiplied each element of the resulting series by $\beta$:

$$\beta(x_t + w_t).$$

A highly useful operator is the lag operator. Suppose that we start with a sequence $\{x_t\}_{t=-\infty}^{\infty}$ and generate a new sequence $\{y_t\}_{t=-\infty}^{\infty}$, where the value of $y$ for date $t$ is equal to the value $x$ took on at date $t - 1$:

$$y_t = x_{t-1}. \qquad [2.1.2]$$

This is described as applying the *lag operator* to $\{x_t\}_{t=-\infty}^{\infty}$. The operation is represented by the symbol $L$:

$$Lx_t \equiv x_{t-1}. \qquad [2.1.3]$$

Consider the result of applying the lag operator twice to a series:

$$L(Lx_t) = L(x_{t-1}) = x_{t-2}.$$

Such a double application of the lag operator is indicated by "$L^2$":

$$L^2 x_t = x_{t-2}.$$

In general, for any integer $k$,

$$L^k x_t = x_{t-k}. \qquad [2.1.4]$$

Notice that if we first apply the multiplication operator and then the lag operator, as in

$$x_t \to \beta x_t \to \beta x_{t-1},$$

the result will be exactly the same as if we had applied the lag operator first and then the multiplication operator:

$$x_t \to x_{t-1} \to \beta x_{t-1}.$$



Thus the lag operator and multiplication operator are commutative:

$$L(\beta x_t) = \beta \cdot Lx_t.$$

Similarly, if we first add two series and then apply the lag operator to the result,

$$(x_t, w_t) \to x_t + w_t \to x_{t-1} + w_{t-1},$$

the result is the same as if we had applied the lag operator before adding:

$$(x_t, w_t) \to (x_{t-1}, w_{t-1}) \to x_{t-1} + w_{t-1}.$$

Thus, the lag operator is distributive over the addition operator:

$$L(x_t + w_t) = Lx_t + Lw_t.$$

We thus see that the lag operator follows exactly the same algebraic rules as the multiplication operator. For this reason, it is tempting to use the expression "multiply $y_t$ by $L$" rather than "operate on $\{y_t\}_{t=-\infty}^{\infty}$ by $L$." Although the latter expression is technically more correct, this text will often use the former shorthand expression to facilitate the exposition.

Faced with a time series defined in terms of compound operators, we are free to use the standard commutative, associative, and distributive algebraic laws for multiplication and addition to express the compound operator in an alternative form. For example, the process defined by

$$y_t = (a + bL)Lx_t$$

is exactly the same as

$$y_t = (aL + bL^2)x_t = ax_{t-1} + bx_{t-2}.$$

To take another example,

$$\begin{aligned}(1 - \lambda_1 L)(1 - \lambda_2 L)x_t &= (1 - \lambda_1 L - \lambda_2 L + \lambda_1\lambda_2 L^2)x_t \\ &= (1 - [\lambda_1 + \lambda_2]L + \lambda_1\lambda_2 L^2)x_t \\ &= x_t - (\lambda_1 + \lambda_2)x_{t-1} + (\lambda_1\lambda_2)x_{t-2}.\end{aligned} \qquad [2.1.5]$$

An expression such as $(aL + bL^2)$ is referred to as a *polynomial in the lag operator*. It is algebraically similar to a simple polynomial $(az + bz^2)$ where $z$ is a scalar. The difference is that the simple polynomial $(az + bz^2)$ refers to a particular number, whereas a polynomial in the lag operator $(aL + bL^2)$ refers to an operator that would be applied to one time series $\{x_t\}_{t=-\infty}^{\infty}$ to produce a new time series $\{y_t\}_{t=-\infty}^{\infty}$.

Notice that if $\{x_t\}_{t=-\infty}^{\infty}$ is just a series of constants,

$$x_t = c \quad \text{for all } t,$$

then the lag operator applied to $x_t$ produces the same series of constants:

$$Lx_t = x_{t-1} = c.$$

Thus, for example,

$$(\alpha L + \beta L^2 + \gamma L^3)c = (\alpha + \beta + \gamma) \cdot c. \qquad [2.1.6]$$

## 2.2. *First-Order Difference Equations*

Let us now return to the first-order difference equation analyzed in Section 1.1:

$$y_t = \phi y_{t-1} + w_t. \qquad [2.2.1]$$

Equation [2.2.1] can be rewritten using the lag operator [2.1.3] as

$$y_t = \phi L y_t + w_t.$$

This equation, in turn, can be rearranged using standard algebra,

$$y_t - \phi L y_t = w_t,$$

or

$$(1 - \phi L)y_t = w_t. \qquad [2.2.2]$$

Next consider "multiplying" both sides of [2.2.2] by the following operator:

$$(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t). \qquad [2.2.3]$$

The result would be

$$\begin{aligned}(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)(1 - \phi L)y_t \\ = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)w_t.\end{aligned} \qquad [2.2.4]$$

Expanding out the compound operator on the left side of [2.2.4] results in

$$\begin{aligned}(1 &+ \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)(1 - \phi L) \\ &= (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t) \\ &\quad - (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)\phi L \\ &= (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t) \\ &\quad - (\phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t + \phi^{t+1} L^{t+1}) \\ &= (1 - \phi^{t+1} L^{t+1}).\end{aligned} \qquad [2.2.5]$$

Substituting [2.2.5] into [2.2.4] yields

$$(1 - \phi^{t+1} L^{t+1})y_t = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)w_t. \qquad [2.2.6]$$

Writing [2.2.6] out explicitly using [2.1.4] produces

$$y_t - \phi^{t+1} y_{t-(t+1)} = w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + \cdots + \phi^t w_{t-t}$$

or

$$y_t = \phi^{t+1} y_{-1} + w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + \cdots + \phi^t w_0. \qquad [2.2.7]$$

Notice that equation [2.2.7] is identical to equation [1.1.7]. Applying the operator [2.2.3] is performing exactly the same set of recursive substitutions that were employed in the previous chapter to arrive at [1.1.7].

It is interesting to reflect on the nature of the operator [2.2.3] as $t$ becomes large. We saw in [2.2.5] that

$$(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)(1 - \phi L)y_t = y_t - \phi^{t+1} y_{-1}.$$

That is, $(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)(1 - \phi L)y_t$ differs from $y_t$ by the term $\phi^{t+1} y_{-1}$. If $|\phi| < 1$ and if $y_{-1}$ is a finite number, this residual $\phi^{t+1} y_{-1}$ will become negligible as $t$ becomes large:

$$(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^t L^t)(1 - \phi L)y_t \cong y_t \qquad \text{for } t \text{ large.}$$

A sequence $\{y_t\}_{t=-\infty}^{\infty}$ is said to be *bounded* if there exists a finite number $\bar{y}$ such that

$$|y_t| < \bar{y} \qquad \text{for all } t.$$

Thus, when $|\phi| < 1$ and when we are considering applying an operator to a bounded sequence, we can think of

$$(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^j L^j)$$



as approximating the inverse of the operator $(1 - \phi L)$, with this approximation made arbitrarily accurate by choosing $j$ sufficiently large:

$$(1 - \phi L)^{-1} = \lim_{j\to\infty} (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \cdots + \phi^j L^j). \qquad [2.2.8]$$

This operator $(1 - \phi L)^{-1}$ has the property

$$(1 - \phi L)^{-1}(1 - \phi L) = 1,$$

where "1" denotes the identity operator:

$$1y_t = y_t.$$

The following chapter discusses stochastic sequences rather than the deterministic sequences studied here. There we will speak of mean square convergence and stationary stochastic processes in place of limits of bounded deterministic sequences, though the practical meaning of [2.2.8] will be little changed.

Provided that $|\phi| < 1$ and we restrict ourselves to bounded sequences or stationary stochastic processes, both sides of [2.2.2] can be "divided" by $(1 - \phi L)$ to obtain

$$y_t = (1 - \phi L)^{-1} w_t$$

or

$$y_t = w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + \cdots. \qquad [2.2.9]$$

It should be emphasized that if we were not restricted to considering bounded sequences or stationary stochastic processes $\{w_t\}_{t=-\infty}^{\infty}$ and $\{y_t\}_{t=-\infty}^{\infty}$, then expression [2.2.9] would not be a necessary implication of [2.2.1]. Equation [2.2.9] is consistent with [2.2.1], but adding a term $a_0\phi^t$,

$$y_t = a_0\phi^t + w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + \cdots, \qquad [2.2.10]$$

produces another series consistent with [2.2.1] for any constant $a_0$. To verify that [2.2.10] is consistent with [2.2.1], multiply [2.2.10] by $(1 - \phi L)$:

$$\begin{aligned}(1 - \phi L)y_t &= (1 - \phi L)a_0\phi^t + (1 - \phi L)(1 - \phi L)^{-1} w_t \\ &= a_0\phi^t - \phi \cdot a_0\phi^{t-1} + w_t \\ &= w_t,\end{aligned}$$

so that [2.2.10] is consistent with [2.2.1] for any constant $a_0$.

Although any process of the form of [2.2.10] is consistent with the difference equation [2.2.1], notice that since $|\phi| < 1$,

$$|a_0\phi^t| \to \infty \qquad \text{as} \qquad t \to -\infty.$$

Thus, even if $\{w_t\}_{t=-\infty}^{\infty}$ is a bounded sequence, the solution $\{y_t\}_{t=-\infty}^{\infty}$ given by [2.2.10] is unbounded unless $a_0 = 0$ in [2.2.10]. Thus, there was a particular reason for defining the operator [2.2.8] to be the inverse of $(1 - \phi L)$—namely, $(1 - \phi L)^{-1}$ defined in [2.2.8] is the unique operator satisfying

$$(1 - \phi L)^{-1}(1 - \phi L) = 1$$

that maps a bounded sequence $\{w_t\}_{t=-\infty}^{\infty}$ into a bounded sequence $\{y_t\}_{t=-\infty}^{\infty}$.

The nature of $(1 - \phi L)^{-1}$ when $|\phi| \geq 1$ will be discussed in Section 2.5.

## 2.3. *Second-Order Difference Equations*

Consider next a second-order difference equation:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + w_t. \qquad [2.3.1]$$

Rewriting this in lag operator form produces

$$(1 - \phi_1 L - \phi_2 L^2)y_t = w_t. \qquad [2.3.2]$$

The left side of [2.3.2] contains a second-order polynomial in the lag operator $L$. Suppose we factor this polynomial, that is, find numbers $\lambda_1$ and $\lambda_2$ such that

$$(1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L) = (1 - [\lambda_1 + \lambda_2]L + \lambda_1\lambda_2 L^2). \qquad [2.3.3]$$

This is just the operation in [2.1.5] in reverse. Given values for $\phi_1$ and $\phi_2$, we seek numbers $\lambda_1$ and $\lambda_2$ with the properties that

$$\lambda_1 + \lambda_2 = \phi_1$$

and

$$\lambda_1\lambda_2 = -\phi_2.$$

For example, if $\phi_1 = 0.6$ and $\phi_2 = -0.08$, then we should choose $\lambda_1 = 0.4$ and $\lambda_2 = 0.2$:

$$(1 - 0.6L + 0.08L^2) = (1 - 0.4L)(1 - 0.2L). \qquad [2.3.4]$$

It is easy enough to see that these values of $\lambda_1$ and $\lambda_2$ work for this numerical example, but how are $\lambda_1$ and $\lambda_2$ found in general? The task is to choose $\lambda_1$ and $\lambda_2$ so as to make sure that the operator on the right side of [2.3.3] is identical to that on the left side. This will be true whenever the following represent the identical functions of $z$:

$$(1 - \phi_1 z - \phi_2 z^2) = (1 - \lambda_1 z)(1 - \lambda_2 z). \qquad [2.3.5]$$

This equation simply replaces the lag operator $L$ in [2.3.3] with a scalar $z$. What is the point of doing so? With [2.3.5], we can now ask, For what values of $z$ is the right side of [2.3.5] equal to zero? The answer is, if either $z = \lambda_1^{-1}$ or $z = \lambda_2^{-1}$, then the right side of [2.3.5] would be zero. It would not have made sense to ask an analogous question of [2.3.3]—$L$ denotes a particular operator, not a number, and $L = \lambda_1^{-1}$ is not a sensible statement.

Why should we care that the right side of [2.3.5] is zero if $z = \lambda_1^{-1}$ or if $z = \lambda_2^{-1}$? Recall that the goal was to choose $\lambda_1$ and $\lambda_2$ so that the two sides of [2.3.5] represented the identical polynomial in $z$. This means that for any particular value $z$ the two functions must produce the same number. If we find a value of $z$ that sets the right side to zero, that same value of $z$ must set the left side to zero as well. But the values of $z$ that set the left side to zero,

$$(1 - \phi_1 z - \phi_2 z^2) = 0, \qquad [2.3.6]$$

are given by the quadratic formula:

$$z_1 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} \qquad [2.3.7]$$

$$z_2 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}. \qquad [2.3.8]$$

Setting $z = z_1$ or $z_2$ makes the left side of [2.3.5] zero, while $z = \lambda_1^{-1}$ or $\lambda_2^{-1}$ sets the right side of [2.3.5] to zero. Thus

$$\lambda_1^{-1} = z_1 \qquad [2.3.9]$$

$$\lambda_2^{-1} = z_2. \qquad [2.3.10]$$



Returning to the numerical example [2.3.4] in which $\phi_1 = 0.6$ and $\phi_2 = -0.08$, we would calculate

$$z_1 = \frac{0.6 - \sqrt{(0.6)^2 - 4(0.08)}}{2(0.08)} = 2.5$$

$$z_2 = \frac{0.6 + \sqrt{(0.6)^2 - 4(0.08)}}{2(0.08)} = 5.0,$$

and so

$$\lambda_1 = 1/(2.5) = 0.4$$

$$\lambda_2 = 1/(5.0) = 0.2,$$

as was found in [2.3.4].

When $\phi_1^2 + 4\phi_2 < 0$, the values $z_1$ and $z_2$ are complex conjugates, and their reciprocals $\lambda_1$ and $\lambda_2$ can be found by first writing the complex number in polar coordinate form. Specifically, write

$$z_1 = a + bi$$

as

$$z_1 = R\cdot[\cos(\theta) + i\cdot\sin(\theta)] = R\cdot e^{i\theta}.$$

Then

$$z_1^{-1} = R^{-1}\cdot e^{-i\theta} = R^{-1}\cdot[\cos(\theta) - i\cdot\sin(\theta)].$$

Actually, there is a more direct method for calculating the values of $\lambda_1$ and $\lambda_2$ from $\phi_1$ and $\phi_2$. Divide both sides of [2.3.5] by $z^2$:

$$(z^{-2} - \phi_1 z^{-1} - \phi_2) = (z^{-1} - \lambda_1)(z^{-1} - \lambda_2) \qquad [2.3.11]$$

and define $\lambda$ to be the variable $z^{-1}$:

$$\lambda \equiv z^{-1}. \qquad [2.3.12]$$

Substituting [2.3.12] into [2.3.11] produces

$$(\lambda^2 - \phi_1\lambda - \phi_2) = (\lambda - \lambda_1)(\lambda - \lambda_2). \qquad [2.3.13]$$

Again, [2.3.13] must hold for all values of $\lambda$ in order for the two sides of [2.3.5] to represent the same polynomial. The values of $\lambda$ that set the right side to zero are $\lambda = \lambda_1$ and $\lambda = \lambda_2$. These same values must set the left side of [2.3.13] to zero as well:

$$(\lambda^2 - \phi_1\lambda - \phi_2) = 0. \qquad [2.3.14]$$

Thus, to calculate the values of $\lambda_1$ and $\lambda_2$ that factor the polynomial in [2.3.3], we can find the roots of [2.3.14] directly from the quadratic formula:

$$\lambda_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \qquad [2.3.15]$$

$$\lambda_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}. \qquad [2.3.16]$$

For the example of [2.3.4], we would thus calculate

$$\lambda_1 = \frac{0.6 + \sqrt{(0.6)^2 - 4(0.08)}}{2} = 0.4$$

$$\lambda_2 = \frac{0.6 - \sqrt{(0.6)^2 - 4(0.08)}}{2} = 0.2.$$

It is instructive to compare these results with those in Chapter 1. There the dynamics of the second-order difference equation [2.3.1] were summarized by calculating the eigenvalues of the matrix $\mathbf{F}$ given by

$$\mathbf{F} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}. \qquad [2.3.17]$$

The eigenvalues of $\mathbf{F}$ were seen to be the two values of $\lambda$ that satisfy equation [1.2.13]:

$$(\lambda^2 - \phi_1\lambda - \phi_2) = 0.$$

But this is the same calculation as in [2.3.14]. This finding is summarized in the following proposition.

***Proposition 2.1:*** *Factoring the polynomial* $(1 - \phi_1 L - \phi_2 L^2)$ *as*

$$(1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L) \qquad [2.3.18]$$

*is the same calculation as finding the eigenvalues of the matrix* $\mathbf{F}$ *in [2.3.17]. The eigenvalues* $\lambda_1$ *and* $\lambda_2$ *of* $\mathbf{F}$ *are the same as the parameters* $\lambda_1$ *and* $\lambda_2$ *in [2.3.18], and are given by equations [2.3.15] and [2.3.16].*

The correspondence between calculating the eigenvalues of a matrix and factoring a polynomial in the lag operator is very instructive. However, it introduces one minor source of possible semantic confusion about which we have to be careful. Recall from Chapter 1 that the system [2.3.1] is stable if both $\lambda_1$ and $\lambda_2$ are less than 1 in modulus and explosive if either $\lambda_1$ or $\lambda_2$ is greater than 1 in modulus. Sometimes this is described as the requirement that the roots of

$$(\lambda^2 - \phi_1\lambda - \phi_2) = 0 \qquad [2.3.19]$$

lie inside the unit circle. The possible confusion is that it is often convenient to work directly with the polynomial in the form in which it appears in [2.3.2],

$$(1 - \phi_1 z - \phi_2 z^2) = 0, \qquad [2.3.20]$$

whose roots, we have seen, are the reciprocals of those of [2.3.19]. Thus, we could say with equal accuracy that "the difference equation [2.3.1] is stable whenever the roots of [2.3.19] lie *inside* the unit circle" or that "the difference equation [2.3.1] is stable whenever the roots of [2.3.20] lie *outside* the unit circle." The two statements mean exactly the same thing. Some scholars refer simply to the "roots of the difference equation [2.3.1]," though this raises the possibility of confusion between [2.3.19] and [2.3.20]. This book will follow the convention of using the term "eigenvalues" to refer to the roots of [2.3.19]. Wherever the term "roots" is used, we will indicate explicitly the equation whose roots are being described.

From here on in this section, it is assumed that the second-order difference equation is stable, with the eigenvalues $\lambda_1$ and $\lambda_2$ distinct and both inside the unit circle. Where this is the case, the inverses

$$(1 - \lambda_1 L)^{-1} = 1 + \lambda_1^1 L + \lambda_1^2 L^2 + \lambda_1^3 L^3 + \cdots$$
$$(1 - \lambda_2 L)^{-1} = 1 + \lambda_2^1 L + \lambda_2^2 L^2 + \lambda_2^3 L^3 + \cdots$$

are well defined for bounded sequences. Write [2.3.2] in factored form:

$$(1 - \lambda_1 L)(1 - \lambda_2 L)y_t = w_t$$

and operate on both sides by $(1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1}$:

$$y_t = (1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1} w_t. \qquad [2.3.21]$$



Following Sargent (1987, p. 184), when $\lambda_1 \neq \lambda_2$, we can use the following operator:

$$(\lambda_1 - \lambda_2)^{-1}\left\{\frac{\lambda_1}{1 - \lambda_1 L} - \frac{\lambda_2}{1 - \lambda_2 L}\right\}. \qquad [2.3.22]$$

Notice that this is simply another way of writing the operator in [2.3.21]:

$$\begin{aligned}(\lambda_1 - \lambda_2)^{-1}&\left\{\frac{\lambda_1}{1 - \lambda_1 L} - \frac{\lambda_2}{1 - \lambda_2 L}\right\} \\ &= (\lambda_1 - \lambda_2)^{-1}\left\{\frac{\lambda_1(1 - \lambda_2 L) - \lambda_2(1 - \lambda_1 L)}{(1 - \lambda_1 L)\cdot(1 - \lambda_2 L)}\right\} \\ &= \frac{1}{(1 - \lambda_1 L)\cdot(1 - \lambda_2 L)}.\end{aligned}$$

Thus, [2.3.21] can be written as

$$\begin{aligned} y_t &= (\lambda_1 - \lambda_2)^{-1}\left\{\frac{\lambda_1}{1 - \lambda_1 L} - \frac{\lambda_2}{1 - \lambda_2 L}\right\} w_t \\ &= \left\{\frac{\lambda_1}{\lambda_1 - \lambda_2}[1 + \lambda_1 L + \lambda_1^2 L^2 + \lambda_1^3 L^3 + \cdots] \right. \\ &\qquad \left. - \frac{\lambda_2}{\lambda_1 - \lambda_2}[1 + \lambda_2 L + \lambda_2^2 L^2 + \lambda_2^3 L^3 + \cdots]\right\} w_t \end{aligned}$$

or

$$\begin{aligned} y_t = [c_1 + c_2]w_t + [c_1\lambda_1 + c_2\lambda_2]w_{t-1} + [c_1\lambda_1^2 + c_2\lambda_2^2]w_{t-2} \\ + [c_1\lambda_1^3 + c_2\lambda_2^3]w_{t-3} + \cdots, \end{aligned} \qquad [2.3.23]$$

where

$$c_1 = \lambda_1/(\lambda_1 - \lambda_2) \qquad [2.3.24]$$
$$c_2 = -\lambda_2/(\lambda_1 - \lambda_2). \qquad [2.3.25]$$

From [2.3.23] the dynamic multiplier can be read off directly as

$$\frac{\partial y_{t+j}}{\partial w_t} = c_1\lambda_1^j + c_2\lambda_2^j,$$

the same result arrived at in equations [1.2.24] and [1.2.25].

## 2.4. *p*th-Order Difference Equations

These techniques generalize in a straightforward way to a $p$th-order difference equation of the form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + w_t. \qquad [2.4.1]$$

Write [2.4.1] in terms of lag operators as

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)y_t = w_t. \qquad [2.4.2]$$

Factor the operator on the left side of [2.4.2] as

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) = (1 - \lambda_1 L)(1 - \lambda_2 L)\cdots(1 - \lambda_p L). \qquad [2.4.3]$$

This is the same as finding the values of $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ such that the following polynomials are the same for all $z$:

$$(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p) = (1 - \lambda_1 z)(1 - \lambda_2 z)\cdots(1 - \lambda_p z).$$

As in the second-order system, we multiply both sides of this equation by $z^{-p}$ and define $\lambda \equiv z^{-1}$:

$$\begin{aligned}(\lambda^p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_{p-1}\lambda - \phi_p)& \\ = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_p).&\end{aligned} \qquad [2.4.4]$$

Clearly, setting $\lambda = \lambda_i$ for $i = 1, 2, \ldots,$ or $p$ causes the right side of [2.4.4] to equal zero. Thus the values $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ must be the numbers that set the left side of expression [2.4.4] to zero as well:

$$\lambda_p - \phi_1\lambda^{p-1} - \phi_2\lambda^{p-2} - \cdots - \phi_{p-1}\lambda - \phi_p = 0. \qquad [2.4.5]$$

This expression again is identical to that given in Proposition 1.1, which characterized the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ of the matrix $\mathbf{F}$ defined in equation [1.2.3]. Thus, Proposition 2.1 readily generalizes.

***Proposition 2.2:*** *Factoring a pth-order polynomial in the lag operator,*

$$(1 - \phi_1L - \phi_2L^2 - \cdots - \phi_pL^p) = (1 - \lambda_1L)(1 - \lambda_2L)\cdots(1 - \lambda_pL),$$

*is the same calculation as finding the eigenvalues of the matrix* $\mathbf{F}$ *defined in [1.2.3]. The eigenvalues* $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ *of* $\mathbf{F}$ *are the same as the parameters* $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ *in [2.4.3] and are given by the solutions to equation [2.4.5].*

The difference equation [2.4.1] is stable if the eigenvalues (the roots of [2.4.5]) lie inside the unit circle, or equivalently if the roots of

$$1 - \phi_1z - \phi_2z^2 - \cdots - \phi_pz^p = 0 \qquad [2.4.6]$$

lie outside the unit circle.

Assuming that the eigenvalues are inside the unit circle and that we are restricting ourselves to considering bounded sequences, the inverses $(1 - \lambda_1L)^{-1}$, $(1 - \lambda_2L)^{-1}, \ldots, (1 - \lambda_pL)^{-1}$ all exist, permitting the difference equation

$$(1 - \lambda_1L)(1 - \lambda_2L)\cdots(1 - \lambda_pL)y_t = w_t$$

to be written as

$$y_t = (1 - \lambda_1L)^{-1}(1 - \lambda_2L)^{-1}\cdots(1 - \lambda_pL)^{-1}w_t. \qquad [2.4.7]$$

Provided further that the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ are all distinct, the polynomial associated with the operator on the right side of [2.4.7] can again be expanded with partial fractions:

$$\begin{aligned}&\frac{1}{(1 - \lambda_1z)(1 - \lambda_2z)\cdots(1 - \lambda_pz)} \\ &\qquad = \frac{c_1}{(1 - \lambda_1z)} + \frac{c_2}{(1 - \lambda_2z)} + \cdots + \frac{c_p}{(1 - \lambda_pz)}.\end{aligned} \qquad [2.4.8]$$

Following Sargent (1987, pp. 192–93), the values of $(c_1, c_2, \ldots, c_p)$ that make [2.4.8] true can be found by multiplying both sides by $(1 - \lambda_1z)(1 - \lambda_2z)\cdots(1 - \lambda_pz)$:

$$\begin{aligned}1 = {}& c_1(1 - \lambda_2z)(1 - \lambda_3z)\cdots(1 - \lambda_pz) \\ &+ c_2(1 - \lambda_1z)(1 - \lambda_3z)\cdots(1 - \lambda_pz) + \cdots \\ &+ c_p(1 - \lambda_1z)(1 - \lambda_2z)\cdots(1 - \lambda_{p-1}z).\end{aligned} \qquad [2.4.9]$$

Equation [2.4.9] has to hold for all values of $z$. Since it is a $(p - 1)$th-order polynomial, if $(c_1, c_2, \ldots, c_p)$ are chosen so that [2.4.9] holds for $p$ particular



distinct values of $z$, then [2.4.9] must hold for all $z$. To ensure that [2.4.9] holds at $z = \lambda_1^{-1}$ requires that

$$1 = c_1(1 - \lambda_2\lambda_1^{-1})(1 - \lambda_3\lambda_1^{-1})\cdots(1 - \lambda_p\lambda_1^{-1})$$

or

$$c_1 = \frac{\lambda_1^{p-1}}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)\cdots(\lambda_1 - \lambda_p)}. \qquad [2.4.10]$$

For [2.4.9] to hold for $z = \lambda_2^{-1}, \lambda_3^{-1}, \ldots, \lambda_p^{-1}$ requires

$$c_2 = \frac{\lambda_2^{p-1}}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)\cdots(\lambda_2 - \lambda_p)} \qquad [2.4.11]$$

$$\vdots$$

$$c_p = \frac{\lambda_p^{p-1}}{(\lambda_p - \lambda_1)(\lambda_p - \lambda_2)\cdots(\lambda_p - \lambda_{p-1})}. \qquad [2.4.12]$$

Note again that these are identical to expression [1.2.25] in Chapter 1. Recall from the discussion there that $c_1 + c_2 + \cdots + c_p = 1$.

To conclude, [2.4.7] can be written

$$\begin{aligned}y_t &= \frac{c_1}{(1 - \lambda_1L)}w_t + \frac{c_2}{(1 - \lambda_2L)}w_t + \cdots + \frac{c_p}{(1 - \lambda_pL)}w_t \\ &= c_1(1 + \lambda_1L + \lambda_1^2L^2 + \lambda_1^3L^3 + \cdots)w_t + c_2(1 + \lambda_2L + \lambda_2^2L^2 + \lambda_2^3L^3 + \cdots)w_t \\ &\quad + \cdots + c_p(1 + \lambda_pL + \lambda_p^2L^2 + \lambda_p^3L^3 + \cdots)w_t\end{aligned}$$

or

$$\begin{aligned}y_t = {}& [c_1 + c_2 + \cdots + c_p]w_t + [c_1\lambda_1 + c_2\lambda_2 + \cdots + c_p\lambda_p]w_{t-1} \\ &+ [c_1\lambda_1^2 + c_2\lambda_2^2 + \cdots + c_p\lambda_p^2]w_{t-2} \\ &+ [c_1\lambda_1^3 + c_2\lambda_2^3 + \cdots + c_p\lambda_p^3]w_{t-3} + \cdots\end{aligned} \qquad [2.4.13]$$

where $(c_1, c_2, \ldots, c_p)$ are given by equations [2.4.10] through [2.4.12]. Again, the dynamic multiplier can be read directly off [2.4.13]:

$$\frac{\partial y_{t+j}}{\partial w_t} = [c_1\lambda_1^j + c_2\lambda_2^j + \cdots + c_p\lambda_p^j], \qquad [2.4.14]$$

reproducing the result from Chapter 1.

There is a very convenient way to calculate the effect of $w$ on the present value of $y$ using the lag operator representation. Write [2.4.13] as

$$y_t = \psi_0w_t + \psi_1w_{t-1} + \psi_2w_{t-2} + \psi_3w_{t-3} + \cdots \qquad [2.4.15]$$

where

$$\psi_j = [c_1\lambda_1^j + c_2\lambda_2^j + \cdots + c_p\lambda_p^j]. \qquad [2.4.16]$$

Next rewrite [2.4.15] in lag operator notation as

$$y_t = \psi(L)w_t, \qquad [2.4.17]$$

where $\psi(L)$ denotes an infinite-order polynomial in the lag operator:

$$\psi(L) = \psi_0 + \psi_1L + \psi_2L^2 + \psi_3L^3 + \cdots.$$

Notice that $\psi_j$ is the dynamic multiplier [2.4.14]. The effect of $w_t$ on the present value of $y$ is given by

$$\frac{\partial \sum_{j=0}^{\infty} \beta^j y_{t+j}}{\partial w_t} = \sum_{j=0}^{\infty} \beta^j \frac{\partial y_{t+j}}{\partial w_t} = \sum_{j=0}^{\infty} \beta^j \psi_j. \qquad [2.4.18]$$

Thinking of $\psi(z)$ as a polynomial in a real number $z$,

$$\psi(z) = \psi_0 + \psi_1 z + \psi_2 z^2 + \psi_3 z^3 + \cdots,$$

it appears that the multiplier [2.4.18] is simply this polynomial evaluated at $z = \beta$:

$$\frac{\partial \sum_{j=0}^{\infty} \beta^j y_{t+j}}{\partial w_t} = \psi(\beta) = \psi_0 + \psi_1 \beta + \psi_2 \beta^2 + \psi_3 \beta^3 + \cdots. \qquad [2.4.19]$$

But comparing [2.4.17] with [2.4.7], it is apparent that

$$\psi(L) = [(1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_p L)]^{-1},$$

and from [2.4.3] this means that

$$\psi(L) = [1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p]^{-1}.$$

We conclude that

$$\psi(z) = [1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p]^{-1}$$

for any value of $z$, so, in particular,

$$\psi(\beta) = [1 - \phi_1 \beta - \phi_2 \beta^2 - \cdots - \phi_p \beta^p]^{-1}. \qquad [2.4.20]$$

Substituting [2.4.20] into [2.4.19] reveals that

$$\frac{\partial \sum_{j=0}^{\infty} \beta^j y_{t+j}}{\partial w_t} = \frac{1}{1 - \phi_1 \beta - \phi_2 \beta^2 - \cdots - \phi_p \beta^p}, \qquad [2.4.21]$$

reproducing the claim in Proposition 1.3. Again, the long-run multiplier obtains as the special case of [2.4.21] with $\beta = 1$:

$$\lim_{j\to\infty} \left[\frac{\partial y_{t+j}}{\partial w_t} + \frac{\partial y_{t+j}}{\partial w_{t+1}} + \cdots + \frac{\partial y_{t+j}}{\partial w_{t+j}}\right] = \frac{1}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}.$$

## 2.5. *Initial Conditions and Unbounded Sequences*

Section 1.2 analyzed the following problem. Given a $p$th-order difference equation

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + w_t, \qquad [2.5.1]$$

$p$ initial values of $y$,

$$y_{-1}, y_{-2}, \ldots, y_{-p}, \qquad [2.5.2]$$

and a sequence of values for the input variable $w$,

$$\{w_0, w_1, \ldots, w_t\}, \qquad [2.5.3]$$



we sought to calculate the sequence of values for the output variable $y$:

$$\{y_0, y_1, \ldots, y_t\}.$$

Certainly there are systems where the question is posed in precisely this form. We may know the equation of motion for the system [2.5.1] and its current state [2.5.2] and wish to characterize the values that $\{y_0, y_1, \ldots, y_t\}$ might take on for different specifications of $\{w_0, w_1, \ldots, w_t\}$.

However, there are many examples in economics and finance in which a theory specifies just the equation of motion [2.5.1] and a sequence of driving variables [2.5.3]. Clearly, these two pieces of information alone are insufficient to determine the sequence $\{y_0, y_1, \ldots, y_t\}$, and some additional theory beyond that contained in the difference equation [2.5.1] is needed to describe fully the dependence of $y$ on $w$. These additional restrictions can be of interest in their own right and also help give some insight into some of the technical details of manipulating difference equations. For these reasons, this section discusses in some depth an example of the role of initial conditions and their implications for solving difference equations.

Let $P_t$ denote the price of a stock and $D_t$ its dividend payment. If an investor buys the stock at date $t$ and sells it at $t + 1$, the investor will earn a yield of $D_t/P_t$ from the dividend and a yield of $(P_{t+1} - P_t)/P_t$ in capital gains. The investor's total return $(r_{t+1})$ is thus

$$r_{t+1} = (P_{t+1} - P_t)/P_t + D_t/P_t.$$

A very simple model of the stock market posits that the return investors earn on stocks is constant across time periods:

$$r = (P_{t+1} - P_t)/P_t + D_t/P_t \qquad r > 0. \qquad [2.5.4]$$

Equation [2.5.4] may seem too simplistic to be of much practical interest; it assumes among other things that investors have perfect foresight about future stock prices and dividends. However, a slightly more realistic model in which *expected* stock returns are constant involves a very similar set of technical issues. The advantage of the perfect-foresight model [2.5.4] is that it can be discussed using the tools already in hand to gain some further insight into using lag operators to solve difference equations.

Multiply [2.5.4] by $P_t$ to arrive at

$$rP_t = P_{t+1} - P_t + D_t$$

or

$$P_{t+1} = (1 + r)P_t - D_t. \qquad [2.5.5]$$

Equation [2.5.5] will be recognized as a first-order difference equation of the form of [1.1.1] with $y_t = P_{t+1}$, $\phi = (1 + r)$, and $w_t = -D_t$. From [1.1.7], we know that [2.5.5] implies that

$$P_{t+1} = (1 + r)^{t+1}P_0 - (1 + r)^t D_0 - (1 + r)^{t-1} D_1 - (1 + r)^{t-2} D_2 - \cdots - (1 + r)D_{t-1} - D_t. \qquad [2.5.6]$$

If the sequence $\{D_0, D_1, \ldots, D_t\}$ *and* the value of $P_0$ were given, then [2.5.6] could determine the values of $\{P_1, P_2, \ldots, P_{t+1}\}$. But if only the values $\{D_0, D_1, \ldots, D_t\}$ are given, then equation [2.5.6] would not be enough to pin down $\{P_1, P_2, \ldots, P_{t+1}\}$. There are an infinite number of possible sequences $\{P_1, P_2, \ldots, P_{t+1}\}$ consistent with [2.5.5] and with a given $\{D_0, D_1, \ldots, D_t\}$. This infinite number of possibilities is indexed by the initial value $P_0$.

A further simplifying assumption helps clarify the nature of these different paths for $\{P_1, P_2, \ldots, P_{t+1}\}$. Suppose that dividends are constant over time:

$$D_t = D \qquad \text{for all } t.$$

Then [2.5.6] becomes

$$\begin{aligned} P_{t+1} &= (1+r)^{t+1}P_0 - [(1+r)^t + (1+r)^{t-1} \\ &\quad + \cdots + (1+r) + 1]D \\ &= (1+r)^{t+1}P_0 - \frac{1-(1+r)^{t+1}}{1-(1+r)}D \\ &= (1+r)^{t+1}[P_0 - (D/r)] + (D/r). \end{aligned} \qquad [2.5.7]$$

Consider first the solution in which $P_0 = D/r$. If the initial stock price should happen to take this value, then [2.5.7] implies that

$$P_t = D/r \qquad [2.5.8]$$

for all $t$. In this solution, dividends are constant at $D$ and the stock price is constant at $D/r$. With no change in stock prices, investors never have any capital gains or losses, and their return is solely the dividend yield $D/P = r$. In a world with no changes in dividends this seems to be a sensible expression of the theory represented by [2.5.4]. Equation [2.5.8] is sometimes described as the "market fundamentals" solution to [2.5.4] for the case of constant dividends.

However, even with constant dividends, equation [2.5.8] is not the only result consistent with [2.5.4]. Suppose that the initial price exceeded $D/r$:

$$P_0 > D/r.$$

Investors seem to be valuing the stock beyond the potential of its constant dividend stream. From [2.5.7] this could be consistent with the asset pricing theory [2.5.4] provided that $P_1$ exceeds $D/r$ by an even larger amount. As long as investors all believe that prices will continue to rise over time, each will earn the required return $r$ from the realized capital gain and [2.5.4] will be satisfied. This scenario has reminded many economists of a speculative bubble in stock prices.

If such bubbles are to be ruled out, additional knowledge about the process for $\{P_t\}_{t=-\infty}^{\infty}$ is required beyond that contained in the theory of [2.5.4]. For example, we might argue that finite world resources put an upper limit on feasible stock prices, as in

$$|P_t| < \bar{P} \qquad \text{for all } t. \qquad [2.5.9]$$

Then the only sequence for $\{P_t\}_{t=-\infty}^{\infty}$ consistent with both [2.5.4] and [2.5.9] would be the market fundamentals solution [2.5.8].

Let us now relax the assumption that dividends are constant and replace it with the assumption that $\{D_t\}_{t=-\infty}^{\infty}$ is a bounded sequence. What path for $\{P_t\}_{t=-\infty}^{\infty}$ in [2.5.6] is consistent with [2.5.9] in this case? The answer can be found by returning to the difference equation [2.5.5]. We arrived at the form [2.5.6] by recursively substituting this equation backward. That is, we used the fact that [2.5.5] held for dates $t, t-1, t-2, \ldots, 0$ and recursively substituted to arrive at [2.5.6] as a logical implication of [2.5.5]. Equation [2.5.5] could equally well be solved recursively *forward*. To do so, equation [2.5.5] is written as

$$P_t = \frac{1}{1+r}[P_{t+1} + D_t]. \qquad [2.5.10]$$



An analogous equation must hold for date $t+1$:

$$P_{t+1} = \frac{1}{1+r}[P_{t+2} + D_{t+1}]. \qquad [2.5.11]$$

Substitute [2.5.11] into [2.5.10] to deduce

$$\begin{aligned} P_t &= \frac{1}{1+r}\left[\frac{1}{1+r}[P_{t+2} + D_{t+1}] + D_t\right] \\ &= \left[\frac{1}{1+r}\right]^2 P_{t+2} + \left[\frac{1}{1+r}\right]^2 D_{t+1} + \left[\frac{1}{1+r}\right] D_t. \end{aligned} \qquad [2.5.12]$$

Using [2.5.10] for date $t+2$,

$$P_{t+2} = \frac{1}{1+r}[P_{t+3} + D_{t+2}],$$

and substituting into [2.5.12] gives

$$P_t = \left[\frac{1}{1+r}\right]^3 P_{t+3} + \left[\frac{1}{1+r}\right]^3 D_{t+2} + \left[\frac{1}{1+r}\right]^2 D_{t+1} + \left[\frac{1}{1+r}\right] D_t.$$

Continuing in this fashion $T$ periods into the future produces

$$\begin{aligned} P_t &= \left[\frac{1}{1+r}\right]^T P_{t+T} + \left[\frac{1}{1+r}\right]^T D_{t+T-1} + \left[\frac{1}{1+r}\right]^{T-1} D_{t+T-2} \\ &\quad + \cdots + \left[\frac{1}{1+r}\right]^2 D_{t+1} + \left[\frac{1}{1+r}\right] D_t. \end{aligned} \qquad [2.5.13]$$

If the sequence $\{P_t\}_{t=-\infty}^{\infty}$ is to satisfy [2.5.9], then

$$\lim_{T\to\infty} \left[\frac{1}{1+r}\right]^T P_{t+T} = 0.$$

If $\{D_t\}_{t=-\infty}^{\infty}$ is likewise a bounded sequence, then the following limit exists:

$$\lim_{T\to\infty} \sum_{j=0}^{T} \left[\frac{1}{1+r}\right]^{j+1} D_{t+j}.$$

Thus, if $\{P_t\}_{t=-\infty}^{\infty}$ is to be a bounded sequence, then we can take the limit of [2.5.13] as $T \to \infty$ to conclude

$$P_t = \sum_{j=0}^{\infty} \left[\frac{1}{1+r}\right]^{j+1} D_{t+j}, \qquad [2.5.14]$$

which is referred to as the "market fundamentals" solution of [2.5.5] for the general case of time-varying dividends. Notice that [2.5.14] produces [2.5.8] as a special case when $D_t = D$ for all $t$.

Describing the value of a variable at time $t$ as a function of future realizations of another variable as in [2.5.14] may seem an artifact of assuming a perfect-foresight model of stock prices. However, an analogous set of operations turns out to be appropriate in a system similar to [2.5.4] in which expected returns are constant.[1] In such systems [2.5.14] generalizes to

$$P_t = \sum_{j=0}^{\infty} \left[\frac{1}{1+r}\right]^{j+1} E_t D_{t+j},$$

[1]See Sargent (1987) and Whiteman (1983) for an introduction to the manipulation of difference equations involving expectations.

where $E_t$ denotes an expectation of an unknown future quantity based on information available to investors at date $t$.

Expression [2.5.14] determines the particular value for the initial price $P_0$ that is consistent with the boundedness condition [2.5.9]. Setting $t = 0$ in [2.5.14] and substituting into [2.5.6] produces

$$\begin{aligned} P_{t+1} = (1 + r)^{t+1}\biggl\{\left[\frac{1}{1 + r}\right]D_0 &+ \left[\frac{1}{1 + r}\right]^2 D_1 + \left[\frac{1}{1 + r}\right]^3 D_2 \\ &+ \cdots + \left[\frac{1}{1 + r}\right]^{t+1} D_t + \left[\frac{1}{1 + r}\right]^{t+2} D_{t+1} + \cdots\biggr\} - (1 + r)^t D_0 \\ &- (1 + r)^{t-1}D_1 - (1 + r)^{t-2}D_2 - \cdots - (1 + r)D_{t-1} - D_t \\ &= \left[\frac{1}{1 + r}\right]D_{t+1} + \left[\frac{1}{1 + r}\right]^2 D_{t+2} + \left[\frac{1}{1 + r}\right]^3 D_{t+3} + \cdots. \end{aligned}$$

Thus, setting the initial condition $P_0$ to satisfy [2.5.14] is sufficient to ensure that it holds for all $t$. Choosing $P_0$ equal to any other value would cause the consequences of each period's dividends to accumulate over time so as to lead to a violation of [2.5.9] eventually.

It is useful to discuss these same calculations from the perspective of lag operators. In Section 2.2 the recursive substitution backward that led from [2.5.5] to [2.5.6] was represented by writing [2.5.5] in terms of lag operators as

$$[1 - (1 + r)L]P_{t+1} = -D_t \qquad [2.5.15]$$

and multiplying both sides of [2.5.15] by the following operator:

$$[1 + (1 + r)L + (1 + r)^2L^2 + \cdots + (1 + r)^tL^t]. \qquad [2.5.16]$$

If $(1 + r)$ were less than unity, it would be natural to consider the limit of [2.5.16] as $t \to \infty$:

$$[1 - (1 + r)L]^{-1} = 1 + (1 + r)L + (1 + r)^2L^2 + \cdots.$$

In the case of the theory of stock returns discussed here, however, $r > 0$ and this operator is not defined. In this case, a lag operator representation can be sought for the recursive substitution forward that led from [2.5.5] to [2.5.13]. This is accomplished using the inverse of the lag operator,

$$L^{-1}w_t = w_{t+1},$$

which extends result [2.1.4] to negative values of $k$. Note that $L^{-1}$ is indeed the inverse of the operator $L$:

$$L^{-1}(Lw_t) = L^{-1}w_{t-1} = w_t.$$

In general,

$$L^{-k}L^j = L^{j-k},$$

with $L^0$ defined as the identity operator:

$$L^0w_t \equiv w_t.$$



Now consider multiplying [2.5.15] by

$$\begin{aligned} [1 + (1 + r)^{-1}L^{-1} + (1 + r)^{-2}L^{-2} + \cdots + (1 + r)^{-(T-1)}L^{-(T-1)}] \\ \times [-(1 + r)^{-1}L^{-1}] \end{aligned} \qquad [2.5.17]$$

to obtain

$$\begin{aligned} &[1 + (1 + r)^{-1}L^{-1} + (1 + r)^{-2}L^{-2} + \cdots + (1 + r)^{-(T-1)}L^{-(T-1)}] \\ &\times [1 - (1 + r)^{-1}L^{-1}]P_{t+1} \\ &= [1 + (1 + r)^{-1}L^{-1} + (1 + r)^{-2}L^{-2} + \cdots \\ &\quad + (1 + r)^{-(T-1)}L^{-(T-1)}] \times (1 + r)^{-1}D_{t+1} \end{aligned}$$

or

$$\begin{aligned} [1 - (1 + r)^{-T}L^{-T}]P_{t+1} &= \left[\frac{1}{1 + r}\right]D_{t+1} + \left[\frac{1}{1 + r}\right]^2 D_{t+2} \\ &+ \left[\frac{1}{1 + r}\right]^3 D_{t+3} + \cdots + \left[\frac{1}{1 + r}\right]^T D_{t+T}, \end{aligned}$$

which is identical to [2.5.13] with $t$ in [2.5.13] replaced with $t + 1$.

When $r > 0$ and $\{P_t\}_{t=-\infty}^{\infty}$ is a bounded sequence, the left side of the preceding equation will approach $P_{t+1}$ as $T$ becomes large. Thus, when $r > 0$ and $\{P_t\}_{t=-\infty}^{\infty}$ and $\{D_t\}_{t=-\infty}^{\infty}$ are bounded sequences, the limit of the operator in [2.5.17] exists and could be viewed as the inverse of the operator on the left side of [2.5.15]:

$$\begin{aligned} [1 - (1 + r)L]^{-1} &= -(1 + r)^{-1}L^{-1} \\ &\times [1 + (1 + r)^{-1}L^{-1} + (1 + r)^{-2}L^{-2} + \cdots]. \end{aligned}$$

Applying this limiting operator to [2.5.15] amounts to solving the difference equation forward as in [2.5.14] and selecting the market fundamentals solution among the set of possible time paths for $\{P_t\}_{t=-\infty}^{\infty}$ given a particular time path for dividends $\{D_t\}_{t=-\infty}^{\infty}$.

Thus, given a first-order difference equation of the form

$$(1 - \phi L)y_t = w_t, \qquad [2.5.18]$$

Sargent's (1987) advice was to solve the equation "backward" when $|\phi| < 1$ by multiplying by

$$[1 - \phi L]^{-1} = [1 + \phi L + \phi^2L^2 + \phi^3L^3 + \cdots] \qquad [2.5.19]$$

and to solve the equation "forward" when $|\phi| > 1$ by multiplying by

$$\begin{aligned} [1 - \phi L]^{-1} &= \frac{-\phi^{-1}L^{-1}}{1 - \phi^{-1}L^{-1}} \\ &= -\phi^{-1}L^{-1}[1 + \phi^{-1}L^{-1} + \phi^{-2}L^{-2} + \phi^{-3}L^{-3} + \cdots]. \end{aligned} \qquad [2.5.20]$$

Defining the inverse of $[1 - \phi L]$ in this way amounts to selecting an operator $[1 - \phi L]^{-1}$ with the properties that

$$[1 - \phi L]^{-1} \times [1 - \phi L] = 1 \qquad \text{(the identity operator)}$$

and that, when it is applied to a bounded sequence $\{w_t\}_{t=-\infty}^{\infty}$,

$$[1 - \phi L]^{-1}w_t,$$

the result is another bounded sequence.

The conclusion from this discussion is that in applying an operator such as $[1 - \phi L]^{-1}$, we are implicitly imposing a boundedness assumption that rules out phenomena such as the speculative bubbles of equation [2.5.7] a priori. Where that is our intention, so much the better, though we should not apply the rules [2.5.19] or [2.5.20] without some reflection on their economic content.

## Chapter 2 References

Sargent, Thomas J. 1987. *Macroeconomic Theory*, 2d ed. Boston: Academic Press.
Whiteman, Charles H. 1983. *Linear Rational Expectations Models: A User's Guide*. Minneapolis: University of Minnesota Press.



# 3 Stationary ARMA Processes

This chapter introduces univariate *ARMA* processes, which provide a very useful class of models for describing the dynamics of an individual time series. The chapter begins with definitions of some of the key concepts used in time series analysis. Sections 3.2 through 3.5 then investigate the properties of various *ARMA* processes. Section 3.6 introduces the autocovariance-generating function, which is useful for analyzing the consequences of combining different time series and for an understanding of the population spectrum. The chapter concludes with a discussion of invertibility (Section 3.7), which can be important for selecting the *ARMA* representation of an observed time series that is appropriate given the uses to be made of the model.

## 3.1. *Expectations, Stationarity, and Ergodicity*

### *Expectations and Stochastic Processes*

Suppose we have observed a sample of size $T$ of some random variable $Y_t$:

$$\{y_1, y_2, \ldots, y_T\}. \qquad [3.1.1]$$

For example, consider a collection of $T$ independent and identically distributed (i.i.d.) variables $\varepsilon_t$,

$$\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T\}, \qquad [3.1.2]$$

with

$$\varepsilon_t \sim N(0, \sigma^2).$$

This is referred to as a sample of size $T$ from a *Gaussian white noise* process.

The observed sample [3.1.1] represents $T$ particular numbers, but this set of $T$ numbers is only one possible outcome of the underlying stochastic process that generated the data. Indeed, even if we were to imagine having observed the process for an infinite period of time, arriving at the sequence

$$\{y_t\}_{t=-\infty}^{\infty} = \{\ldots, y_{-1}, y_0, y_1, y_2, \ldots, y_T, y_{T+1}, y_{T+2}, \ldots\},$$

the infinite sequence $\{y_t\}_{t=-\infty}^{\infty}$ would still be viewed as a single realization from a time series process. For example, we might set one computer to work generating an infinite sequence of i.i.d. $N(0, \sigma^2)$ variates, $\{\varepsilon_t^{(1)}\}_{t=-\infty}^{\infty}$, and a second computer generating a separate sequence, $\{\varepsilon_t^{(2)}\}_{t=-\infty}^{\infty}$. We would then view these as two independent realizations of a Gaussian white noise process.