CHAPTER NINE

Seasonal Models

In Chapters 3 to 8 we have considered the properties of a class of linear stochastic models, which are of value in representing stationary and nonstationary time series, and we have seen how these models may be used for forecasting. We then considered the practical problems of identification, fitting, and diagnostic checking that arise when relating these models to actual data. In the present chapter we apply these methods to analyzing and forecasting seasonal series and also provide an opportunity to show how the ideas of the previous chapters fit together.

9.1 PARSIMONIOUS MODELS FOR SEASONAL TIME SERIES

Figure 9.1 shows the totals of international airline passengers for 1952, 1953, and 1954. It is part of a longer series (12 years of data) quoted by Brown [79] and listed as Series G in the Collection of Time Series in Part Five. The series shows a marked seasonal pattern since travel is at its highest in the late summer months, while a secondary peak occurs in the spring. Many other series, particularly sales data, show similar seasonal characteristics.

In general, we say that a series exhibits periodic behavior with period s, when similarities in the series occur after s basic time intervals. In the example above, the basic time interval is 1 month and the period is $s = 12$ months. However, examples occur when s can take on other values. For example, $s = 4$ for quarterly data showing seasonal effects within years. It sometimes happens that there is more than one periodicity. Thus, because bills tend to be paid monthly, we would expect weekly business done by a bank to show a periodicity of about 4 within months, while monthly business shows a periodicity of 12.

9.1.1 Fitting versus Forecasting

One of the deficiencies in the analysis of time series in the past has been the confusion between *fitting* a series and *forecasting* it. For example, suppose that a

Time Series Analysis, Fourth Edition. By G. E. Box, G. M. Jenkins, and G. C. Reinsel

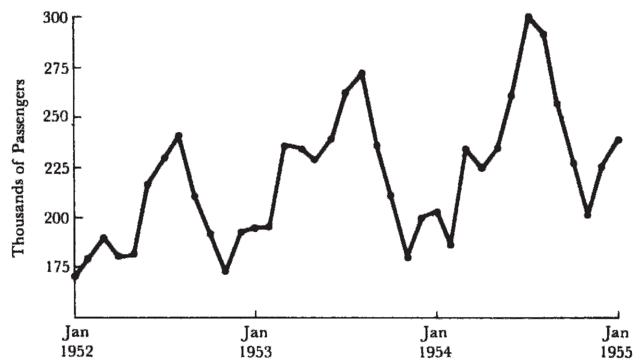

FIGURE 9.1 Totals of international airline passengers in thousands (part of Series G).

time series has shown a tendency to increase over a particular period and also to follow a seasonal pattern. A common method of analysis is to decompose the series arbitrarily into three components: a *trend*, a *seasonal component*, and a *random component*. The trend might be fitted by a polynomial and the seasonal component by a Fourier series. A forecast was then made by projecting these fitted functions.

Such methods can give extremely misleading results. For example, we have already seen that the behavior of IBM stock prices (Series B) is closely approximated by the random walk model $\nabla z_t = a_t$, that is,

$$z_t = z_0 + \sum_{j=0}^{t-1} a_{t-j} \tag{9.1.1}$$

which implies that $\hat{z}_t(l) = z_t$. In other words, the best forecast of future values of the stock is very nearly today's price. Now it is true that short lengths of Series B do look as if they might be fitted by quadratic curves. This simply reflects the fact that a sum of random deviates can sometimes have this appearance. However, there is no basis for the use of a quadratic forecast function, which produces very poor forecasts. Of course, genuine systematic effects that can be explained physically should be taken into account by the inclusion of a suitable deterministic component in the model. For example, if it is known that heat is being steadily added to a system, it would be sensible to explain the resulting increase in temperature by means of a suitable deterministic function of time, in addition to the stochastic component.

9.1.2 Seasonal Models Involving Adaptive Sines and Cosines

The general linear model

$$\tilde{z}_t = \sum_{j=1}^{\infty} \pi_j \tilde{z}_{t-j} + a_t = \sum_{j=1}^{\infty} \psi_j a_{t-j} + a_t \tag{9.1.2}$$

with suitable values for the coefficients π_j and ψ_j, is entirely adequate to describe many seasonal time series. The problem is to choose a suitable system of *parsimonious parameterization* for such models. As we have said before, this is not a mathematical problem but a question of finding out how the world tends to behave. To do this one can only proceed by trying out ideas on actual time series and developing those concepts that seem fruitful.

We have seen that for nonseasonal series, it is usually possible to obtain a useful and parsimonious representation in the form

$$\varphi(B)\tilde{z}_t = \theta(B)a_t \tag{9.1.3}$$

Moreover, the generalized autoregressive operator $\varphi(B)$ determines the eventual forecast function, which is the solution of the difference equation

$$\varphi(B)\hat{z}_t(l) = 0$$

where B is understood to operate on l. In representing seasonal behavior, we shall want the forecast function to trace out a periodic pattern. Our first thought might be that $\varphi(B)$ should produce a forecast function consisting of a mixture of sines and cosines, and possibly mixed with polynomial terms, to allow for changes in the level of the series and changes in the seasonal pattern. Such a forecast function could arise perfectly naturally within the structure of the general model (9.1.3). For example, with monthly data, a forecast function that is a sine wave with a 12-month period, adaptive in phase and amplitude, will satisfy the difference equation

$$(1 - \sqrt{3}B + B^2)\hat{z}_t(l) = 0$$

where B is understood to operate on l. However, it is not true that periodic behavior is necessarily represented *economically* by mixtures of sines and cosines. Many sine–cosine components would, for example, be needed to represent sales data affected by Christmas, Easter, and other seasonal buying. To take an extreme case, sales of fireworks in Britain are largely confined to the weeks immediately prior to November 5, when the abortive attempt of Guy Fawkes to blow up the Houses of Parliament is celebrated. An attempt to represent the "single spike" of fireworks sales data directly by sines and cosines might be unprofitable. It is clear that a more careful consideration of the problem is needed.

Now, in our previous analysis, we have not necessarily estimated *all* the components of $\varphi(B)$. Where differencing d times was needed to induce stationarity, we have written $\varphi(B) = \phi(B)(1 - B)^d$, which is equivalent to setting d roots of the equation $\varphi(B) = 0$ equal to unity. When such a representation proved adequate, we could proceed with the simpler analysis of $w_t = \nabla^d z_t$. Thus, we have used $\nabla = 1 - B$ as a simplifying operator. In other problems, different types of simplifying operators might be appropriate. For example, the consumption of fuel oil for heat is highly dependent on ambient temperature, which, because the Earth rotates around the sun, is known to follow approximately a sine wave with period 12 months. In analyzing sales of fuel oil, it might then be sensible to introduce $1 - \sqrt{3}B + B^2$ as a simplifying operator, constituting one of the contributing components of the generalized autoregressive operator $\varphi(B)$. If such a representation proved useful, we could then proceed with the simpler analysis of $w_t = (1 - \sqrt{3}B + B^2)z_t$. This operator, it may be noted, is of the homogeneous nonstationary variety, having zeros $e^{\pm(i2\pi/12)}$ on the unit circle.

9.1.3 General Multiplicative Seasonal Model

***Simplifying Operator* $\mathbf{1 - B^s}$**

The fundamental fact about seasonal time series with period s is that observations that are s intervals apart are similar. Therefore, one might expect that the operation $B^s z_t = z_{t-s}$ would play a particularly important role in the analysis of seasonal series and, furthermore, since nonstationarity is to be expected in the series $z_t, z_{t-s}, z_{t-2s}, \ldots$, the simplifying operation

$$\nabla_s z_t = (1 - B^s)z_t = z_t - z_{t-s}$$

might be useful. This nonstationary operator $1 - B^s$ has s zeros $e^{i(2\pi k/s)}$ ($k = 0, 1, \ldots, s-1$) evenly spaced on the unit circle. Furthermore, the eventual forecast function satisfies $(1 - B^s)\hat{z}_t(l) = 0$ and so may (but need not) be represented by a full complement of sines and cosines:

$$\hat{z}_t(l) = b_0^{(t)} + \sum_{j=1}^{[s/2]} \left[b_{1j}^{(t)} \quad \cos\left(\frac{2\pi jl}{s}\right) + b_{2j}^{(t)} \sin\left(\frac{2\pi jl}{s}\right) \right]$$

where the b's are adaptive coefficients, and where $[s/2] = \frac{1}{2}s$ if s is even and $[s/2] = \frac{1}{2}(s-1)$ if s is odd.

Multiplicative Model When we have a series exhibiting seasonal behavior with known periodicity s, it is of value to set down the data in the form of a table containing s columns, such as Table 9.1, which shows the logarithms of the airline data. For seasonal data special care is needed in selecting an appropriate transformation. In this example (see Section 9.3.5) data analysis supports the use of the logarithm.

The arrangement of Table 9.1 emphasizes the fact that in periodic data; there are not one but two time intervals of importance. For this example, these intervals correspond to months and to years. Specifically, we expect relationships to occur (a) between the observations for successive months in a particular year and (b) between the observations for the same month in successive years. The situation is somewhat like that in a two-way analysis of variance model, where similarities can be expected between observations in the same column and between observations in the same row.

Referring to the airline data of Table 9.1, the seasonal effect implies that an observation for a particular month, say April, is related to the observations for previous Aprils. Suppose that the tth observation z_t is for the month of April. We

Table 9.1 Natural Logarithms of Monthly Passenger Totals (Measured in Thousands) in International Air Travel (Series G)

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1949	4.718	4.771	4.883	4.860	4.796	4.905	4.997	4.997	4.913	4.779	4.644	4.771
1950	4.745	4.836	4.949	4.905	4.828	5.004	5.136	5.136	5.063	4.890	4.736	4.942
1951	4.977	5.011	5.182	5.094	5.147	5.182	5.293	5.293	5.215	5.088	4.984	5.112
1952	5.142	5.193	5.263	5.199	5.209	5.384	5.438	5.489	5.342	5.252	5.147	5.268
1953	5.278	5.278	5.464	5.460	5.434	5.493	5.576	5.606	5.468	5.352	5.193	5.303
1954	5.318	5.236	5.460	5.245	5.455	5.576	5.710	5.680	5.557	5.434	5.313	5.434
1955	5.489	5.451	5.587	5.595	5.598	5.753	5.897	5.849	5.743	5.613	5.648	5.628
1956	5.649	5.624	5.759	5.746	5.762	5.924	6.023	6.004	5.872	5.724	5.602	5.724
1957	5.753	5.707	5.875	5.852	5.872	6.045	6.142	6.146	6.001	5.849	5.720	5.817
1958	5.829	5.762	5.892	5.852	5.894	6.075	6.196	6.225	6.001	5.883	5.737	5.820
1959	5.886	5.835	6.006	5.981	6.040	6.157	6.306	6.326	6.138	6.009	5.892	6.004
1960	6.033	5.969	6.038	6.133	6.157	6.282	6.433	6.407	6.230	6.133	5.966	6.068

might be able to link this observation z_t to observations in previous Aprils by a model of the form

$$\Phi(B^s)\nabla_s^D z_t = \Theta(B^s)\alpha_t \tag{9.1.4}$$

where $s = 12$, $\nabla_s = 1 - B^s$, and $\Phi(B^s)$, $\Theta(B^s)$ are polynomials in B^s of degrees P and Q, respectively, and satisfying stationarity and invertibility conditions. Similarly, a model

$$\Phi(B^s)\nabla_s^D z_{t-1} = \Theta(B^s)\alpha_{t-1} \tag{9.1.5}$$

might be used to link the current behavior for March with previous March observations, and so on, for each of the 12 months. Moreover, it would usually be reasonable to assume that the parameters $\mathbf{\Phi}$ and $\mathbf{\Theta}$ contained in these monthly models would be approximately the same for each month.

Now the error components $\alpha_t, \alpha_{t-1}, \ldots,$ in these models would not in general be uncorrelated. For example, the total of airline passengers in April 1960, while related to previous April totals, would also be related to totals in March 1960, February 1960, January 1960, and so on. Thus, we would expect that α_t in (9.1.4) would be related to α_{t-1} in (9.1.5) and to α_{t-2}, and so on. Therefore, to take care of such relationships, we introduce a second model

$$\phi(B)\nabla^d \alpha_t = \theta(B)a_t \tag{9.1.6}$$

where now a_t is a white noise process and $\phi(B)$ and $\theta(B)$ are polynomials in B of degrees p and q, respectively, and satisfying stationarity and invertibility conditions, and $\nabla = \nabla_1 = 1 - B$.

Substituting (9.1.6) in (9.1.4), we finally obtain a general multiplicative model

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t = \theta_q(B)\Theta_Q(B^s)a_t \tag{9.1.7}$$

where for this particular example, $s = 12$. Also, in (9.1.7) the subscripts p, P, q, Q have been added to remind the reader of the orders of the various operators. The resulting multiplicative process will be said to be of order $(p, d, q) \times (P, D, Q)_s$. A similar argument can be used to obtain models with three or more periodic components to take care of multiple seasonalities.

In the remainder of this chapter we consider the basic forms for seasonal models of the kind just introduced and their potential for forecasting. We also consider the problems of identification, estimation, and diagnostic checking that arise in relating such models to data. No new principles are needed to do this, merely an application of the procedures and ideas we have already discussed in detail in Chapters 6 to 8. We proceed in Section 9.2 by discussing a particular example in considerable detail. In Section 9.3 we discuss those aspects of the general seasonal ARIMA model that call for special mention. Section 9.4 discusses an alternate structural component model approach to representing stochastic seasonal and trend behavior, including consideration of the possibility for components to be deterministic. In Section 9.5 we consider the topic, which is related to the presence of deterministic components, of regression models for time series data with autocorrelated errors.

9.2 REPRESENTATION OF THE AIRLINE DATA BY A MULTIPLICATIVE $(0, 1, 1) \times (0, 1, 1)_{12}$ MODEL

The detailed illustration in this section will consist of relating a seasonal ARIMA model of order $(0, 1, 1) \times (0, 1, 1)_{12}$ to the airline data of Series G. In Section 9.2.1 we consider the model itself; in Section 9.2.2, its forecasting; in Section 9.2.3, its identification; in Section 9.2.4, its fitting; and finally in Section 9.2.5, its diagnostic checking.

9.2.1 Multiplicative $(0, 1, 1) \times (0, 1, 1)_{12}$ Model

We have already seen that a simple and widely applicable stochastic model for the analysis of nonstationary time series, which contains no seasonal component, is the IMA(0, 1, 1) process. Suppose, following the argument of Section 9.1.3, that we employed such a model:

$$\nabla_{12} z_t = (1 - \Theta B^{12})\alpha_t$$

for linking z's 1 year apart. Suppose further that we employed a similar model

$$\nabla \alpha_t = (1 - \theta B)a_t$$

for linking α's 1 month apart, where in general θ and Θ will have different values. Then, on combining these expressions, we would obtain the seasonal multiplicative model

$$\nabla \nabla_{12} z_t = (1 - \theta B)(1 - \Theta B^{12})a_t \tag{9.2.1}$$

of order $(0, 1, 1) \times (0, 1, 1)_{12}$. The model written explicitly is

$$z_t - z_{t-1} - z_{t-12} + z_{t-13} = a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta\Theta a_{t-13} \tag{9.2.2}$$

The invertibility region for this model, required by the condition that the roots of $(1 - \theta B)(1 - \Theta B^{12}) = 0$ lie outside the unit circle, is defined by the inequalities $-1 < \theta < 1$ and $-1 < \Theta < 1$. Note that the moving average operator $(1 - \theta B)(1 - \Theta B^{12}) = 1 - \theta B - \Theta B^{12} + \theta\Theta B^{13}$, on the right of (9.2.1), is of order $q + sQ = 1 + 12(1) = 13$.

We shall show in Sections 9.2.3, 9.2.4, and 9.2.5 that the logged airline data is well fitted by a model of the form (9.2.1), where to a sufficient approximation, $\hat{\theta} = 0.4$, $\hat{\Theta} = 0.6$, and $\hat{\sigma}_a^2 = 1.34 \times 10^{-3}$. It is convenient as a preliminary to consider how, using this model and with these parameter values inserted, future values of the series may be forecast.

9.2.2 Forecasting

In Chapter 4 we saw that there are three basically different ways of considering the general model, each giving rise in Chapter 5 to a different way of viewing the forecast. We consider now these three approaches for the forecasting of the seasonal model (9.2.1).

Difference Equation Approach Forecasts are best *computed* directly from the difference equation itself. Thus, since

$$z_{t+l} = z_{t+l-1} + z_{t+l-12} - z_{t+l-13} + a_{t+l} - \theta a_{t+l-1} - \Theta a_{t+l-12} + \theta\Theta a_{t+l-13} \tag{9.2.3}$$

after setting $\theta = 0.4$, $\Theta = 0.6$, the minimum mean square error forecast at lead time l and origin t is given immediately by

$$\begin{aligned} \hat{z}_t(l) = [&z_{t+l-1} + z_{t+l-12} - z_{t+l-13} + a_{t+l} - 0.4a_{t+l-1} \\ &- 0.6a_{t+l-12} + 0.24a_{t+l-13}] \end{aligned} \tag{9.2.4}$$

As in Chapter 5, we refer to

$$[z_{t+l}] = E[z_{t+l}|z_t, z_{t-1}, \ldots; \theta, \Theta]$$

as the conditional expectation of z_{t+l} taken at origin t. In this expression the parameters are supposed exactly known, and knowledge of the series z_t, $z_{t-1}, \ldots$ is supposed to extend into the remote past.

Practical application depends upon the following facts:

1. Invertible models fitted to actual data usually yield forecasts that depend appreciably only on recent values of the series.
2. The forecasts are insensitive to small changes in parameter values such as are introduced by estimation errors.

Now

$$[z_{t+j}] = \begin{cases} z_{t+j} & j \le 0 \\ \hat{z}_t(j) & j > 0 \end{cases} \tag{9.2.5}$$

$$[a_{t+j}] = \begin{cases} a_{t+j} & j \le 0 \\ 0 & j > 0 \end{cases} \tag{9.2.6}$$

Thus, to obtain the forecasts, as in Chapter 5, we simply replace unknown z's by forecasts and unknown a's by zeros. The known a's are, of course, the one-step-ahead forecast errors already computed, that is, $a_t = z_t - \hat{z}_{t-1}(1)$.

For example, to obtain the 3-months-ahead forecast, we have

$$z_{t+3} = z_{t+2} + z_{t-9} - z_{t-10} + a_{t+3} - 0.4a_{t+2} - 0.6a_{t-9} + 0.24a_{t-10}$$

Taking conditional expectations at the origin t gives

$$\hat{z}_t(3) = \hat{z}_t(2) + z_{t-9} - z_{t-10} - 0.6a_{t-9} + 0.24a_{t-10}$$

Substituting for $a_{t-9} = z_{t-9} - \hat{z}_{t-10}(1)$ and $a_{t-10} = z_{t-10} - \hat{z}_{t-11}(1)$ on the right also yields

$$\hat{z}_t(3) = \hat{z}_t(2) + 0.4z_{t-9} - 0.76z_{t-10} + 0.6\hat{z}_{t-10}(1) - 0.24\hat{z}_{t-11}(1) \tag{9.2.7}$$

which expresses the forecast in terms of previous z's and previous forecasts of z's. Although separate expressions for each lead time may readily be written down, computation of the forecasts is best carried out by using the single expression (9.2.4) directly, the elements of its right-hand side being defined by (9.2.5) and (9.2.6).

Figure 9.2 shows the forecasts for lead times up to 36 months, all made at the arbitrarily selected origin, July 1957. We see that the simple model, containing only two parameters, faithfully reproduces the seasonal pattern and supplies excellent forecasts. It is to be remembered, of course, that like all predictions obtained from the general linear stochastic model, the forecast function is adaptive. When changes occur in the seasonal pattern, these will be appropriately projected into the forecast. It will be noticed that when the 1-month-ahead forecast is too high, there is a tendency for all future forecasts from the point to be high. This is to be expected because, as has been noted in Appendix A5.1, forecast errors from the same origin, but for different lead times, are highly correlated. Of course, a forecast for a long lead time, such as 36 months, may necessarily contain a fairly large error.

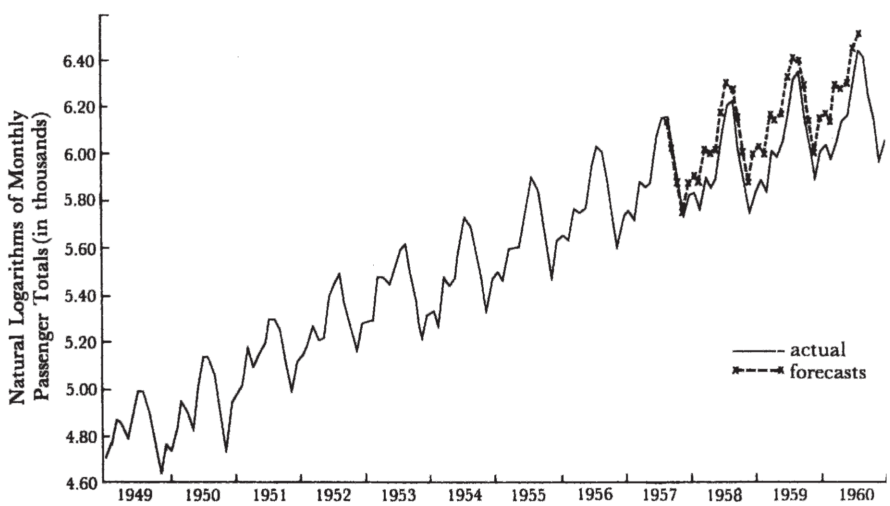

FIGURE 9.2 Series G, with forecasts for 1, 2, 3, ..., 36 months ahead, all made from an arbitrary selected origin, July 1957.

However, in practice, an initially remote forecast will be updated continually, and as the lead shortens, greater accuracy will be possible.

The preceding forecasting procedure is robust to moderate changes in the values of the parameters. Thus, if we used $\theta = 0.5$ and $\Theta = 0.5$, instead of the values $\theta = 0.4$ and $\Theta = 0.6$, the forecasts would not be greatly affected. This is true even for forecasts made several steps ahead (e.g., 12 months). The approximate effect on the one-step-ahead forecasts of modifying the values of the parameters can be seen by studying the sum-of-squares surface. Thus, we know that the approximate confidence region for the k parameters $\boldsymbol{\beta}$ is bounded, in general, by the contour $S(\boldsymbol{\beta}) = S(\hat{\boldsymbol{\beta}})[1 + \chi_\varepsilon^2(k)/n]$, which includes the true parameter point with probability $1 - \varepsilon$. Therefore, we know that, had the *true* parameter values been employed, with this same probability the mean square of the one-step-ahead forecast errors could not have been increased by a factor greater that $1 + \chi_\varepsilon^2(k)/n$.

Forecast Function, Its Updating, and the Forecast Error Variance As we said in Chapter 5, in practice, the difference equation procedure is by far the simplest and most convenient way for actually *computing* forecasts and updating them. However, the difference equation itself does not reveal very much about the *nature* of the forecasts so computed and about their updating. It is to cast light on these aspects (and not to provide alternative computational procedures) that we now consider the forecasts from other points of view.

Forecast Function Using (5.1.12) yields $z_{t+l} = \hat{z}_t(l) + e_t(l)$, where

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1} \tag{9.2.8}$$

Now, the moving average operator on the right of (9.2.1) is of order 13. Hence, from (5.3.2) and for $l > 13$, the forecasts satisfy the difference equation

$$(1 - B)(1 - B^{12})\hat{z}_t(l) = 0 \qquad l > 13 \tag{9.2.9}$$

where in this equation B operates on the lead time l.

We now write $l = (r, m) = 12r + m$, $r = 0, 1, 2, \ldots$ and $m = 1, 2, \ldots, 12$, to represent a lead time of r years and m months, so that, for example, $l = 15 = (1, 3)$. Then, the forecast function, which is the solution of (9.2.9), with starting conditions given by the first 13 forecasts, is of the form

$$\hat{z}_t(l) = \hat{z}_t(r, m) = b_{0,m}^{(t)} + r b_1^{(t)} \qquad l > 0 \tag{9.2.10}$$

This forecast function contains 13 adjustable coefficients $b_{0,1}^{(t)}, b_{0,2}^{(t)}, \ldots, b_{0,12}^{(t)}, b_1^{(t)}$. These represent 12 monthly contributions and one yearly contribution and are determined by the first 13 forecasts. The nature of this function is more clearly understood from Figure 9.3, which shows a forecast function of this kind, but with period $s = 5$, so that there are six adjustable coefficients $b_{0,1}^{(t)}, b_{0,2}^{(t)}, \ldots, b_{0,5}^{(t)}, b_1^{(t)}$.

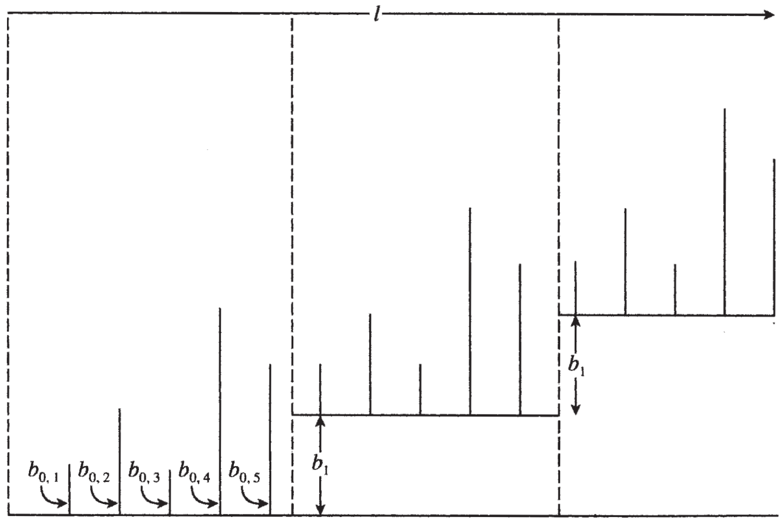

FIGURE 9.3 Seasonal forecast function generated by the model $\nabla\nabla_{12}z_t = (1-\theta B)(1-\Theta B^s)a_t$, with $s = 5$.

Equivalently, since $\hat{z}_t(l)$ satisfies (9.2.9) and the roots of $(1-B)(1-B^{12}) = 0$ are $1, 1, -1, e^{\pm(i2\pi k/12)}$, $k = 1, \ldots, 5$, on the unit circle, the forecast function, as in (5.3.3), can be represented as

$$\hat{z}_t(l) = \sum_{j=1}^{5}\left[b_{1j}^{(t)} \cos\left(\frac{2\pi jl}{12}\right) + b_{2j}^{(t)} \sin\left(\frac{2\pi jl}{12}\right)\right] + b_{16}^{(t)}(-1)^l + b_0^{(t)} + b_1^{*(t)}l$$

This shows that $\hat{z}_t(l)$ consists of a mixture of sinusoids at the seasonal frequencies $2\pi j/12$, $j = 1, \ldots, 6$, plus a linear trend with slope $b_1^{*(t)}$. The coefficients $b_{1j}^{(t)}$, $b_{2j}^{(t)}$, $b_0^{(t)}$, and $b_1^{*(t)}$ in the expression above are all adaptive with regard to the forecast origin t, being determined by the first 13 forecasts.In comparison to (9.2.10), it is clear, for example, that $b_1^{(t)} = 12b_1^{*(t)}$, and represents the *annual* rate of change in the forecasts $\hat{z}_t(l)$, whereas $b_1^{*(t)}$ is the *monthly* rate of change.

The ψ Weights To determine updating formulas and to obtain the variance of the forecast error $e_t(l)$ in (9.2.8), we need the ψ weights in the form $z_t = \sum_{j=0}^{\infty}\psi_j a_{t-j}$ of the model. We can write the moving average operator in (9.2.1) in the form

$$(1-\theta B)(1-\Theta B^{12}) = (\nabla + \lambda B)(\nabla_{12} + \Lambda B^{12})$$

where $\lambda = 1-\theta$, $\Lambda = 1-\Theta$, $\nabla_{12} = 1-B^{12}$. Hence the model (9.2.1) may be written

$$\nabla\nabla_{12}z_t = (\nabla + \lambda B)(\nabla_{12} + \Lambda B^{12})a_t$$

By equating coefficients in $\nabla\nabla_{12}\psi(B) = (\nabla + \lambda B)(\nabla_{12} + \Lambda B^{12})$, it can be seen that the ψ weights satisfy $\psi_0 = 1$, $\psi_1 - \psi_0 = \lambda - 1$, $\psi_{12} - \psi_{11} - \psi_0 = \Lambda - 1$, $\psi_{13} - \psi_{12} - \psi_1 + \psi_0 = (\lambda - 1)(\Lambda - 1)$, and $\psi_j - \psi_{j-1} - \psi_{j-12} + \psi_{j-13} = 0$ otherwise. Thus, the ψ weights for this process are

$$\begin{aligned}
&\psi_1 = \psi_2 = \cdots = \psi_{11} = \lambda && \psi_{12} = \lambda + \Lambda \\
&\psi_{13} = \psi_{14} = \cdots = \psi_{23} = \lambda(1 + \Lambda) && \psi_{24} = \lambda(1 + \Lambda) + \Lambda \\
&\psi_{25} = \psi_{26} = \cdots = \psi_{35} = \lambda(1 + 2\Lambda) && \psi_{36} = \lambda(1 + 2\Lambda) + \Lambda
\end{aligned}$$

and so on. Writing ψ_j as $\psi_{r,m} = \psi_{12r+m}$, where $r = 0, 1, 2, \ldots$ and $m = 1, 2, \ldots, 12$, refer, respectively, to years and months, we obtain

$$\psi_{r,m} = \lambda(1 + r\Lambda) + \delta\Lambda \tag{9.2.11}$$

where

$$\delta = \begin{cases} 1 & \text{when } m = 12 \\ 0 & \text{when } m \neq 12 \end{cases}$$

Updating The general updating formula (5.2.5) is

$$\hat{z}_{t+1}(l) = \hat{z}_t(l + 1) + \psi_l a_{t+1}$$

Thus, if $m \neq s = 12$,

$$b_{0,m}^{(t+1)} + rb_1^{(t+1)} = b_{0,m+1}^{(t)} + rb_1^{(t)} + (\lambda + r\lambda\Lambda)a_{t+1}$$

and on equating coefficients of r, the updating formulas are

$$\begin{aligned}
b_{0,m}^{(t+1)} &= b_{0,m+1}^{(t)} + \lambda a_{t+1} \\
b_1^{(t+1)} &= b_1^{(t)} + \lambda\Lambda a_{t+1}
\end{aligned} \tag{9.2.12}$$

Alternatively, if $m = s = 12$,

$$b_{0,12}^{(t+1)} + rb_1^{(t+1)} = b_{0,1}^{(t)} + (r + 1)b_1^{(t)} + (\lambda + \Lambda + r\lambda\Lambda)a_{t+1}$$

and in this case,

$$\begin{aligned}
b_{0,12}^{(t+1)} &= b_{0,1}^{(t)} + b_1^{(t)} + (\lambda + \Lambda)a_{t+1} \\
b_1^{(t+1)} &= b_1^{(t)} + \lambda\Lambda a_{t+1}
\end{aligned} \tag{9.2.13}$$

In studying these relations, it should be remembered that $b_{0,m}^{(t+1)}$ will be the updated version of $b_{0,m+1}^{(t)}$. Thus, if the origin t was January of a particular year, $b_{0,2}^{(t)}$ would be the coefficient for March. After a month had elapsed, we should move the forecast origin to February and the updated version for the March coefficient would now be $b_{0,1}^{(t+1)}$.

Forecast Error Variance Knowledge of the ψ weights enables us to calculate the variance of the forecast errors at any lead time l, using the result (5.1.16), namely

$$V(l) = (1 + \psi_1^2 + \cdots + \psi_{l-1}^2)\sigma_a^2 \tag{9.2.14}$$

Setting $\lambda = 0.6$, $\Lambda = 0.4$, $\sigma_a^2 = 1.34 \times 10^{-3}$ in (9.2.11) and (9.2.14), the estimated standard deviations $\hat{\sigma}(l)$ of the forecast errors of the log airline data for lead times 1 to 36 are shown in Table 9.2.

Forecasts as a Weighted Average of Previous Observations If we write the model in the form

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t$$

the one-step-ahead forecast is

$$\hat{z}_t(1) = \sum_{j=1}^{\infty} \pi_j z_{t+1-j}$$

The π weights may be obtained by equating coefficients in

$$(1 - B)(1 - B^{12}) = (1 - \theta B)(1 - \Theta B^{12})(1 - \pi_1 B - \pi_2 B^2 - \cdots)$$

Table 9.2 Estimated Standard Deviations of Forecast Errors for Logarithms of Airline Series at Various Lead Times

Forecast Lead Times	$\hat{\sigma}(l) \times 10^{-2}$	Forecast Lead Times	$\hat{\sigma}(l) \times 10^{-2}$	Forecast Lead Times	$\hat{\sigma}(l) \times 10^{-2}$
1	3.7	13	9.0	25	14.4
2	4.3	14	9.5	26	15.0
3	4.8	15	10.0	27	15.5
4	5.3	16	10.5	28	16.0
5	5.8	17	10.9	29	16.4
6	6.2	18	11.4	30	17.0
7	6.6	19	11.7	31	17.4
8	6.9	20	12.1	32	17.8
9	7.2	21	12.6	33	18.3
10	7.6	22	13.0	34	18.7
11	8.0	23	13.3	35	19.2
12	8.2	24	13.6	36	19.6

Thus,

$$\begin{aligned}
\pi_j &= \theta^{j-1}(1-\theta) \qquad j = 1, 2, \ldots, 11 \\
\pi_{12} &= \theta^{11}(1-\theta) + (1-\Theta) \\
\pi_{13} &= \theta^{12}(1-\theta) - (1-\theta)(1-\Theta) \\
\pi_j - \theta\pi_{j-1} &- \Theta\pi_{j-12} + \theta\Theta\pi_{j-13} = 0 \qquad j \geq 14
\end{aligned} \tag{9.2.15}$$

These weights are plotted in Figure 9.4 for the parameter values $\theta = 0.4$ and $\Theta = 0.6$.

The reason that the weight function takes the particular form shown in the figure may be understood as follows: The process (9.2.1) may be written

$$a_{t+1} = \left(1 - \frac{\lambda B}{1-\theta B}\right)\left(1 - \frac{\Lambda B^{12}}{1-\Theta B^{12}}\right) z_{t+1} \tag{9.2.16}$$

We now use the notation $\text{EWMA}_\lambda(z_t)$ to mean an exponentially weighted moving average, with parameter $\lambda = 1 - \theta$, of values z_t, z_{t-1}, z_{t-2}, ..., so that

$$\text{EWMA}_\lambda(z_t) = \frac{\lambda}{1-\theta B} z_t = \lambda z_t + \lambda\theta z_{t-1} + \lambda\theta^2 z_{t-2} + \cdots$$

Similarly, we use $\text{EWMA}_\Lambda(z_t)$ to mean an exponentially weighted moving average, with parameter $\Lambda = 1 - \Theta$, of values z_t, z_{t-12}, z_{t-24}, ..., so that

$$\text{EWMA}_\Lambda(z_t) = \frac{\Lambda}{1-\Theta B^{12}} z_t = \Lambda z_t + \Lambda\Theta z_{t-12} + \Lambda\Theta^2 z_{t-24} + \cdots$$

Substituting $\hat{z}_t(1) = z_{t+1} - a_{t+1}$ in (9.2.16), we obtain

$$\hat{z}_t(1) = \text{EWMA}_\lambda(z_t) + \text{EWMA}_\Lambda\big(z_{t-11} - \text{EWMA}_\lambda(z_{t-12})\big) \tag{9.2.17}$$

Thus, the forecast is an EWMA taken over previous months, modified by a second EWMA of discrepancies found between similar monthly EWMAs and actual performance in previous years. As a particular case, if $\theta = 0$ ($\lambda = 1$), (9.2.17) would

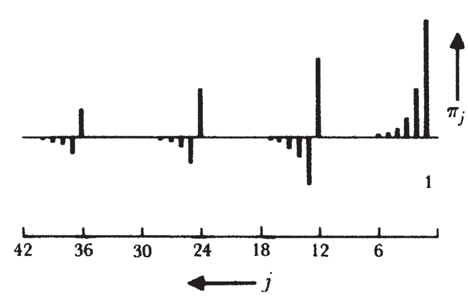

FIGURE 9.4 The π weights for $(0, 1, 1) \times (0, 1, 1)_{12}$ process fitted to Series G ($\theta = 0.4$, $\Theta = 0.6$).

reduce to

$$\begin{aligned}\hat{z}_t(1) &= z_t + \text{EWMA}_\Lambda(z_{t-11} - z_{t-12}) \\ &= z_t + \Lambda[(z_{t-11} - z_{t-12}) + \Theta(z_{t-23} - z_{t-24}) + \cdots]\end{aligned}$$

which shows that first differences are forecast as the seasonal EWMA of first differences for similar months from previous years.

For example, suppose that we were attempting to predict December sales for a department store. These sales would include a heavy component from Christmas buying. The first term on the right of (9.2.17) would be an EWMA taken over previous months up to November. However, we know this will be an underestimate, so we correct it by taking a second EWMA over previous years of the *discrepancies* between actual December sales and the corresponding monthly EWMAs taken over previous months in those years.

The forecasts for lead times $l > 1$ can be generated from the π weights by substituting forecasts of shorter lead time for unknown values, as displayed in the general expression (5.3.6) of Section 5.3.3. Alternatively, explicit values for the weights applied directly to $z_t, z_{t-1}, z_{t-2}, \ldots$ may be computed, for example, from (5.3.9) or from (A5.2.3).

9.2.3 Identification

The identification of the nonseasonal IMA(0, 1, 1) process depends upon the fact that, after taking first differences, the autocorrelations for all lags beyond the first are zero. For the multiplicative $(0, 1, 1) \times (0, 1, 1)_{12}$ process (9.2.1), the only nonzero autocorrelations of $\nabla\nabla_{12}z_t$ are those at lags 1, 11, 12, and 13. In fact, from (9.2.2) the model is viewed as

$$w_t = a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta\Theta a_{t-13}$$

an MA model of order 13 for $w_t = \nabla\nabla_{12}z_t$. So the autocovariances of w_t are obtained directly from results in Section 3.3, and hence they are

$$\begin{aligned}\gamma_0 &= [1 + \theta^2 + \Theta^2 + (\theta\Theta)^2]\sigma_a^2 = (1 + \theta^2)(1 + \Theta^2)\sigma_a^2 \\ \gamma_1 &= [-\theta - \Theta(\theta\Theta)]\sigma_a^2 = -\theta(1 + \Theta^2)\sigma_a^2 \\ \gamma_{11} &= \theta\Theta\sigma_a^2 \\ \gamma_{12} &= [-\Theta - \theta(\theta\Theta)]\sigma_a^2 = -\Theta(1 + \theta^2)\sigma_a^2 \\ \gamma_{13} &= \theta\Theta\sigma_a^2\end{aligned} \tag{9.2.18}$$

In particular, these expressions imply that

$$\rho_1 = \frac{-\theta}{1 + \theta^2} \qquad \text{and} \qquad \rho_{12} = \frac{-\Theta}{1 + \Theta^2}$$

so that the value ρ_1 is unaffected by the presence of the seasonal MA factor $(1 - \Theta B^{12})$ in the model (9.2.1) while the value of ρ_{12} is unaffected by the nonseasonal or regular MA factor $(1 - \theta B)$ in (9.2.1).

Table 9.3 shows the estimated autocorrelations of the logged airline data for (a) the original series, z_t, (b) the logged series differenced with respect to months only, ∇z_t, (c) the logged series differenced with respect to years only, $\nabla_{12} z_t$, and (d) the logged series differenced with respect to months and years, $\nabla\nabla_{12} z_t$. The autocorrelations for z_t are large and fail to die out at higher lags. While simple differencing reduces the correlations in general, a very heavy periodic component remains. This is evidenced particularly by very large correlations at lags 12, 24, 36, and 48. Simple differencing with respect to period 12 results in correlations which are first persistently positive and then persistently negative. By contrast, the differencing $\nabla\nabla_{12}$ markedly reduces correlations throughout.

The autocorrelations of $\nabla\nabla_{12} z_t$ are seen to exhibit spikes at lags 1 and 12, compatible with the features in (9.2.18) corresponding to the model (9.2.1). As an alternative, however, the autocorrelations for $\nabla_{12} z_t$ might be viewed as dying out at a (slow) exponential rate, for low lags, after lag one, so there is also the possibility that $\nabla_{12} z_t$ may follow a model with nonseasonal operator of the form of an ARMA(1, 1), with the value of ϕ relatively close to one, rather than the nonstationary IMA(1, 1) model as in (9.2.1). The distinction between these two model possibilities may not be substantial, in practice, as has been discussed in Chapter 6, and the latter model possibility will not be explored further here. The choice between the nonstationary and stationary AR(1) factor could, in fact, be tested by methods similar to those described in Section 7.5.

On the assumption that the model is of the form (9.2.1), the variances for the estimated higher lag autocorrelations are approximated by Bartlett's formula (2.1.15), which in this case becomes

$$\text{var}[r_k] \simeq \frac{1 + 2(\rho_1^2 + \rho_{11}^2 + \rho_{12}^2 + \rho_{13}^2)}{n} \qquad k > 13 \tag{9.2.19}$$

Substituting estimated correlations for the ρ's and setting $n = 144 - 13 = 131$ in (9.2.19), where $n = 131$ is the number of differences $\nabla\nabla_{12} z_t$, we obtain a standard error $\hat{\sigma}(r) \simeq 0.11$.

In Table 9.4 observed frequencies of the 35 sample autocorrelations r_k, $k > 13$, are compared with those for a normal distribution having mean zero and standard deviation 0.11. This rough check suggests that the model is worthy of further investigation.

Preliminary Estimates As with the nonseasonal model, by equating appropriate observed sample correlations to their expected values, approximate values can be obtained for the parameters θ and Θ. On substituting the sample estimates $r_1 = -0.34$ and $r_{12} = -0.39$ in the expressions

$$\rho_1 = \frac{-\theta}{1 + \theta^2} \qquad \rho_{12} = \frac{-\Theta}{1 + \Theta^2}$$

Table 9.3 Estimated Autocorrelations of Various Differences of the Logged Airline Data

	Lag	Autocorrelation											
(a) z	1–12	0.95	0.90	0.85	0.81	0.78	0.76	0.74	0.73	0.73	0.74	0.76	0.76
	13–24	0.72	0.66	0.62	0.58	0.54	0.52	0.50	0.49	0.50	0.50	0.52	0.52
	25–36	0.48	0.44	0.40	0.36	0.34	0.31	0.30	0.29	0.30	0.30	0.31	0.32
	37–48	0.29	0.24	0.21	0.17	0.15	0.12	0.11	0.10	0.10	0.11	0.12	0.13
(b) ∇z	1–12	0.20	−0.12	−0.15	−0.32	−0.08	0.03	−0.11	−0.34	−0.12	−0.11	0.21	0.84
	13–24	0.22	−0.14	−0.12	−0.28	−0.05	0.01	−0.11	−0.34	−0.11	−0.08	0.20	0.74
	25–36	0.20	−0.12	−0.10	−0.21	−0.06	0.02	−0.12	−0.29	−0.13	−0.04	0.15	0.66
	37–48	0.19	−0.13	−0.06	−0.16	−0.06	0.01	−0.11	−0.28	−0.11	−0.03	0.12	0.59
(c) $\nabla_{12} z$	1–12	0.71	0.62	0.48	0.44	0.39	0.32	0.24	0.19	0.15	−0.01	−0.12	−0.24
	13–24	−0.14	−0.14	−0.10	−0.15	−0.10	−0.11	−0.14	−0.16	−0.11	−0.08	0.00	−0.05
	25–36	−0.10	−0.09	−0.13	−0.15	−0.19	−0.20	−0.19	−0.15	−0.22	−0.23	−0.27	−0.22
	37–48	−0.18	−0.16	−0.14	−0.10	−0.05	0.02	0.04	0.10	0.15	0.22	0.29	0.30
(d) $\nabla\nabla_{12} z$	1–12	−0.34	0.11	−0.20	0.02	0.06	0.03	−0.06	0.00	0.18	−0.08	0.06	−0.39
	13–24	0.15	−0.06	0.15	−0.14	0.07	0.02	−0.01	−0.12	0.04	−0.09	0.22	−0.02
	25–36	−0.10	0.05	−0.03	0.05	−0.02	−0.05	−0.05	0.20	−0.12	0.08	−0.15	−0.01
	37–48	0.05	0.03	−0.02	−0.03	−0.07	0.10	−0.09	0.03	−0.04	−0.04	0.11	−0.05

Table 9.4 Comparison of Observed and Expected Frequencies for Sample Autocorrelations of $\nabla\nabla_{12}z_t$ at Lags Greater Than 13

	Expected from Normal Distribution Mean Zero and Std. Dev. 0.11	Observed [a]
$0 < \|r_k\| < 0.11$	23.9	27.5
$0.11 < \|r_k\| < 0.22$	9.5	7.0
$0.22 < \|r_k\|$	1.6	0.5
	35.0	35.0

[a] Observations on the cell boundary are allocated half to each adjacent cell.

we obtain rough estimates $\hat{\theta} \simeq 0.39$ and $\hat{\Theta} \simeq 0.48$. A table summarizing the behavior of the autocorrelation function for some specimen seasonal models, useful in identification and in obtaining preliminary estimates of the parameters, is given in Appendix A9.1.

9.2.4 Estimation

Contours of the sum-of-squares function $S(\theta, \Theta)$ for the airline data fitted to the model (9.2.1) are shown in Figure 9.5, together with the appropriate 95% confidence region. The least squares estimates are seen to be very nearly $\hat{\theta} = 0.4$ and $\hat{\Theta} = 0.6$. The grid of values for $S(\theta, \Theta)$ was computed using the technique described in Chapter 7. It was shown there that given n observations $\mathbf{w}$ from a linear process defined by

$$\phi(B)w_t = \theta(B)a_t$$

the quadratic form $\mathbf{w}'\mathbf{M}_n\mathbf{w}$, which appears in the exponent of the likelihood, can always be expressed in terms of a sum of squares of the conditional expectation of a's and a quadratic function of the conditional expectation of the $p + q$ initial values $\mathbf{e}_* = (w_{1-p}, \ldots, w_0, a_{1-q}, \ldots, a_0)'$, that is,

$$\mathbf{w}'\mathbf{M}_n\mathbf{w} = S(\boldsymbol{\phi}, \boldsymbol{\theta}) = \sum_{t=-\infty}^{n} [a_t]^2 = \sum_{t=1}^{n} [a_t]^2 + [\mathbf{e}_*]'\boldsymbol{\Omega}^{-1}[\mathbf{e}_*]$$

where $[a_t] = [a_t|\mathbf{w}, \boldsymbol{\phi}, \boldsymbol{\theta}]$, $[\mathbf{e}_*] = [\mathbf{e}_*|\mathbf{w}, \boldsymbol{\phi}, \boldsymbol{\theta}]$, and $\text{cov}[\mathbf{e}_*] = \sigma_a^2\boldsymbol{\Omega}$. Furthermore, it was shown that $S(\boldsymbol{\phi}, \boldsymbol{\theta})$ plays a central role in the estimation of the parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ from both a sampling theory and a likelihood or Bayesian point of view.

The computation for seasonal models follows precisely the same course as described in Section 7.1.5 for nonseasonal models. We illustrate by considering the computation of $S(\theta, \Theta)$ for the airline data in relation to the model

$$\nabla\nabla_{12}z_t = w_t = (1 - \theta B)(1 - \Theta B^{12})a_t$$

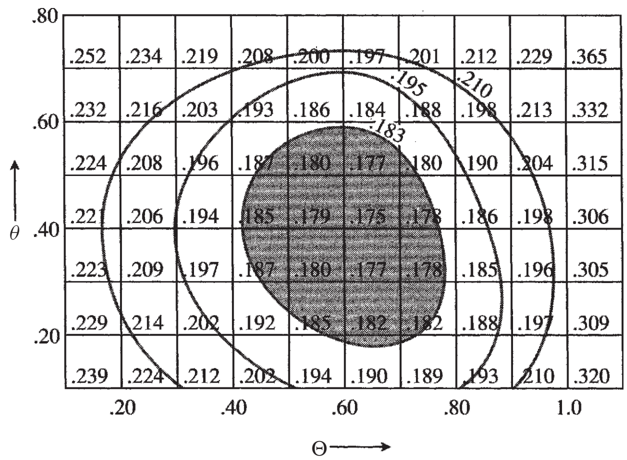

FIGURE 9.5 Series G fitted by the model $\nabla\nabla_{12}z_t = (1 - \theta B)(1 - \Theta B^{12})a_t$; contours of $S(\theta, \Theta)$ with shaded 95% confidence region.

Table 9.5 Airline Data: Computation Table for the [a]'s and Hence for $S(\theta, \Theta)$

z_t	t	$[a_t]$	$[w_t]$	$[e_t]$	a_t^0	u_t
		0	0	0		
z_{-12}	-12	$[a_{-12}]$	$[w_{-12}]$	0	0	
z_{-11}	-11	$[a_{-11}]$	$[w_{-11}]$	0	0	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
z_0	0	$[a_0]$	$[w_0]$	0	0	
z_1	1	$[a_1]$	w_1	$[e_1]$	a_1^0	u_1
z_2	2	$[a_2]$	w_2	$[e_2]$	a_2^0	u_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
z_{131}	131	$[a_{131}]$	w_{131}	$[e_{131}]$	a_{131}^0	u_{131}

A format for the computation of the [a]'s is shown in Table 9.5. If there are N observations of z_t, then in general, with a difference operator $\nabla^d\nabla_s^D$, we can compute $n = N - d - sD$ values of $w_t = \nabla^d\nabla_s^D z_t$. Therefore, it is convenient to use a numbering system so that the first observation in the z series has a subscript $1 - d - sD$. The first observation in the w series then has a subscript 1 and the last has subscript n. There are $N = 144$ observations in the airline series. Accordingly, in Table 9.5 these are designated as $z_{-12}, z_{-11}, \ldots, z_{131}$. The w's, obtained by differencing, then form the series $w_1, w_2, \ldots, w_n$, where $n = 131$. These values are set out in the center of the table.

As noted in Section 7.1.5, an approximate method to obtain the $[a_t]$ using the corresponding backward model can be used for computational convenience. The fundamental formulas, on which backward and forward recursions are based in this method, may be obtained, as before, by taking conditional expectations in the backward and forward forms of the model. In this instance they yield

$$[e_t] = [w_t] + \theta[e_{t+1}] + \Theta[e_{t+12}] - \theta\Theta[e_{t+13}] \tag{9.2.20}$$

$$[a_t] = [w_t] + \theta[a_{t-1}] + \Theta[a_{t-12}] - \theta\Theta[a_{t-13}] \tag{9.2.21}$$

In general, for seasonal models, we might have a stationary autoregressive operator $\phi(B)\Phi(B^s)$ of degree $(p + sP)$. If we wished the back computation of Section˜7.1.5 to begin as far back in the series as possible, the recursion would begin with the calculation of an approximate value for $[e_{n-p-sP}]$, obtained by setting unknown $[e]$'s equal to zero. In the present example, $p = P = 0$ and hence, using (9.2.20), we can begin with

$$[e_{131}] = w_{131} + (\theta \times 0) + (\Theta \times 0) - (\theta\Theta \times 0)$$

$$[e_{130}] = w_{131} + (\theta \times [e_{131}]) + (\Theta \times 0) - (\theta\Theta \times 0)$$

and so on, until $[e_1]$ is obtained. Recalling that $[e_{-j}] = 0$ when $j \geq 0$, we can now use (9.2.20) to compute the back-forecasts $[w_0]$, $[w_{-1}], \ldots, [w_{-12}]$. Furthermore, values of $[w_{-j}]$ for $j > 12$ are all zero, and since each $[a_t]$ is a function of previously occurring $[w]$'s, it follows (and otherwise is obvious directly from the form of the model) that $[a_{-j}] = 0$, $j > 12$. Thus, (9.2.21) may now be used directly to compute the $[a]$'s, and hence to calculate $S(\theta, \Theta) = \sum_{t=-12}^{131} [a_t]^2$. In almost all cases of interest, the transients introduced by the approximation at the beginning of the back recursion will have negligible effect on the calculation of the preliminary $[w]$'s, so that $S(\theta, \Theta)$ computed in this way will be virtually exact. However, it is possible, as indicated in Section 7.1.5, to continue the "up and down" iteration. The next iteration would involve recomputing the $[e]$'s, starting off the iteration using forecasts $[w_{n+1}]$, $[w_{n+2}], \ldots, [w_{n+13}]$ obtained from the $[a]$'s already calculated.

Alternatively, the exact method to conveniently compute the $[a_t]$, as discussed in Section 7.1.5, can be employed. For the present model, this involves first computing the conditional estimates of the a_t, using zero initial values $a^0_{-12} = a^0_{-11} = \cdots = a^0_0 = 0$, through recursive calculations similar to (9.2.21), as

$$a^0_t = w_t + \theta a^0_{t-1} + \Theta a^0_{t-12} - \theta\Theta a^0_{t-13} \qquad t = 1, \ldots, n$$

Then a backward recursion is used to obtain values of u_t as

$$u_t = a^0_t + \theta u_{t+1} + \Theta u_{t+12} - \theta\Theta u_{t+13} \qquad t = n, \ldots, 1$$

using zero initial values $u_{n+1} = \cdots = u_{n+13} = 0$. Finally, the exact back-forecasts for the vector of initial values $\mathbf{a}'_* = (a_{-12}, \ldots, a_0)$ is obtained by solving the equations $\mathbf{D}[\mathbf{a}_*] = \mathbf{F}'\mathbf{u}$, as described in general in (A7.3.12) of Appendix A7.3.

We note that if the vector $\mathbf{F}'\mathbf{u}$ is denoted as $\mathbf{h} = \mathbf{F}'\mathbf{u} = (h_{-12}, h_{-11}, \ldots, h_0)'$, then the values h_{-j} are computed as

$$h_{-j} = -(\theta u_{-j+1} + \Theta u_{-j+12} - \theta\Theta u_{-j+13})$$

where the convention $u_{-j} = 0,\ j \geq 0$, must be used. Once the back-forecasted values are obtained, the remaining $[a_t]$ values for $t = 1, 2, \ldots, n$ are obtained recursively from (9.2.21) exactly as in the previous method, and hence the exact sum of squares $S(\theta, \Theta) = \sum_{t=-12}^{131}[a_t]^2$ is obtained. The format for computation of the $[a_t]$ by this method is also shown in the last two columns of Table 9.5.

Iterative Calculation of Least Squares Estimates $\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Theta}}$ As discussed in Section 7.2, while it is essential to plot sums-of-squares surfaces in a new situation, or whenever difficulties arise, an iterative linearization technique may be used in straightforward situations to supply the least squares estimates and their approximate standard errors. The procedure has been set out in Section 7.2.1, and no new difficulties arise in estimating the parameters of seasonal models.

For the present example, we can write approximately

$$a_{t,0} = (\theta - \theta_0)x_{t,1} + (\Theta - \Theta_0)x_{t,2} + a_t$$

where

$$x_{t,1} = -\left.\frac{\partial a_t}{\partial \theta}\right|_{\theta_0,\Theta_0} \qquad x_{t,2} = -\left.\frac{\partial a_t}{\partial \Theta}\right|_{\theta_0,\Theta_0}$$

and where θ_0 and Θ_0 are guessed values and $a_{t,0} = [a_t|\theta_0, \Theta_0]$. As explained and illustrated in Section 7.2.2, the derivatives are most easily computed numerically. Proceeding in this way and using as starting values the preliminary estimates $\hat{\theta} = 0.39$, $\hat{\Theta} = 0.48$ obtained in Section 9.2.3 from the estimated autocorrelations, the iteration proceeded as in Table 9.6. Alternatively, proceeding as in Section 7.2.3, the derivatives could be obtained to any degree of required accuracy by recursive calculation.

Thus, values of the parameters correct to two decimals, which is the most that would be needed in practice, are available in three iterations. The estimated variance

Table 9.6 Iterative Estimation of θ and Θ for the Logged Airline Data

Iteration	θ	Θ
Starting values	0.390	0.480
1	0.404	0.640
2	0.395	0.612
3	0.396	0.614
4	0.396	0.614

of the residuals is $\hat{\sigma}_a^2 = 1.34 \times 10^{-3}$. From the inverse of the matrix of sums of squares and products of the x's on the last iteration, the standard errors of the estimates may now be calculated. The least squares estimates followed by their standard errors are then

$$\hat{\theta} = 0.40 \pm 0.08$$
$$\hat{\Theta} = 0.61 \pm 0.07$$

agreeing closely with the values obtained from the sum-of-squares plot.

Large-Sample Variances and Covariances for the Estimates As in Section 7.2.6, large-sample formulas for the variances and covariances of the parameter estimates may be obtained. In this case, from the model equation $w_t = a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta\Theta a_{t-13}$, the derivatives $x_{t,1} = -\partial a_t/\partial\theta$ are seen to satisfy

$$x_{t,1} - \theta x_{t-1,1} - \Theta x_{t-12,1} + \theta\Theta x_{t-13,1} + a_{t-1} - \Theta a_{t-13} = 0$$

hence $(1 - \theta B)(1 - \Theta B^{12})x_{t,1} = -(1 - \Theta B^{12})a_{t-1}$, or simply $(1 - \theta B)x_{t,1} = -a_{t-1}$. Thus, using a similar derivation for $x_{t,2} = -\partial a_t/\partial\Theta$, we obtain that

$$x_{t,1} \simeq -(1 - \theta B)^{-1}a_{t-1} = -\sum_{j=0}^{\infty} \theta^j B^j a_{t-1}$$

$$x_{t,2} \simeq -(1 - \Theta B^{12})^{-1}a_{t-12} = -\sum_{i=0}^{\infty} \Theta^i B^{12i} a_{t-12}$$

Therefore, for large samples, the information matrix is

$$\mathbf{I}(\theta, \Theta) = n \begin{bmatrix} (1 - \theta^2)^{-1} & \theta^{11}(1 - \theta^{12}\Theta)^{-1} \\ \theta^{11}(1 - \theta^{12}\Theta)^{-1} & (1 - \Theta^2)^{-1} \end{bmatrix}$$

Provided that $|\theta|$ is not close to unity, the off-diagonal term is negligible and approximate values for the variances and covariances of $\hat{\theta}$ and $\hat{\Theta}$ are

$$V(\hat{\theta}) \simeq n^{-1}(1 - \theta^2) \qquad V(\hat{\Theta}) \simeq n^{-1}(1 - \Theta^2) \qquad \text{cov}[\hat{\theta}, \hat{\Theta}] \simeq 0 \tag{9.2.22}$$

In the present example, substituting the values $\hat{\theta} = 0.40$, $\hat{\Theta} = 0.61$, and $n = 131$, we obtain

$$V(\hat{\theta}) \simeq 0.0064 \qquad V(\hat{\Theta}) \simeq 0.0048$$

and

$$\sigma(\hat{\theta}) \simeq 0.08 \qquad \sigma(\hat{\Theta}) \simeq 0.07$$

which, to this accuracy, are identical with the values obtained directly from the iteration. It is also interesting to note that the parameter estimates $\hat{\theta}$ and $\hat{\Theta}$, associated with months and years, respectively, are virtually uncorrelated.

9.2.5 Diagnostic Checking

Before proceeding further, we check the adequacy of fit of the model by examining the residuals from the fitted process.

Autocorrelation Check The estimated autocorrelations of the residuals

$$\hat{a}_t = \nabla\nabla_{12}z_t + 0.40\hat{a}_{t-1} + 0.61\hat{a}_{t-12} - 0.24\hat{a}_{t-13}$$

are shown in Table 9.7. A number of individual correlations appear rather large compared with the upper bound 0.09 of their standard error, and the value $r_{23} = 0.22$, which is about 2.5 times this upper bound, is particularly discrepant. However, among 48 random deviates one would expect some large deviations.

An overall check is provided by the quantity $\tilde{Q} = n(n+2)\sum_{k=1}^{24} r_k^2(\hat{a})/(n-k)$, which (see Section 8.2.2) is approximately distributed as χ^2 with 22 degrees of freedom, since two parameters have been fitted. The observed value of $\tilde{Q}$ is $131 \times 0.1950 = 25.5$, and on the hypothesis of adequacy of the model, deviations greater than this would be expected in about 27% of cases. The check does not provide any evidence of inadequacy in the model.

Periodogram Check The cumulative periodogram (see Section 8.2.5) for the residuals is shown in Figure 9.6. The Kolmogorov–Smirnov 5 and 25% probability limits, which as we have seen in Section 8.2.5 supply a very rough guide to the significance of apparent deviations, fail in this instance to indicate any significant departure from the assumed model.

9.3 SOME ASPECTS OF MORE GENERAL SEASONAL ARIMA MODELS

9.3.1 Multiplicative and Nonmultiplicative Models

In previous sections we discussed methods of dealing with seasonal time series, and in particular, we examined an example of a multiplicative model. We have seen how this can provide a useful representation with remarkably few parameters. It now remains to study other seasonal models of this kind, and insofar as new considerations arise, the associated processes of identification, estimation, diagnostic checking, and forecasting.

Suppose, in general, that we have a seasonal effect associated with period s. Then the general class of multiplicative models may be typified in the manner

Table 9.7 Estimated Autocorrelations of the Residuals from Fitting the Model $\nabla\nabla_{12}z_t = (1 - 0.40B)(1 - 0.61B^{12})a_t$ to Logged Airline Data (Series G)

Lag k	Autocorrelation $r_k(\hat{a})$												Standard Error (Bound)
1–12	0.02	0.02	−0.13	−0.14	0.05	0.06	−0.07	−0.04	0.10	−0.08	0.02	−0.01	0.09
13–24	0.03	0.04	0.05	−0.16	0.03	0.00	−0.11	−0.10	−0.03	−0.03	0.22	0.03	0.09
25–36	−0.02	0.06	−0.04	−0.06	−0.05	−0.08	−0.05	0.12	−0.13	0.00	−0.06	−0.02	0.09
37–48	0.11	0.07	−0.02	−0.05	−0.10	−0.02	−0.04	0.00	−0.08	0.03	0.04	0.06	0.09

$$(n+2)\sum_{k=1}^{24} r_k^2(\hat{a})/(n-k) = 0.1950$$

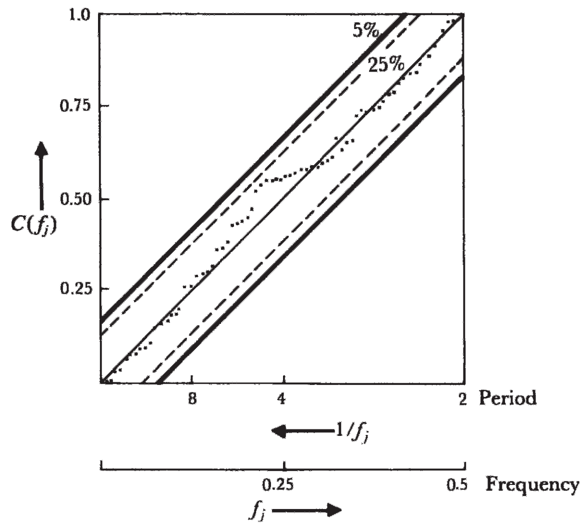

FIGURE 9.6 Cumulative periodogram check on residuals from the model $\nabla\nabla_{12}z_t = (1 - 0.4B)(1 - 0.6B^{12})a_t$ fitted to Series G.

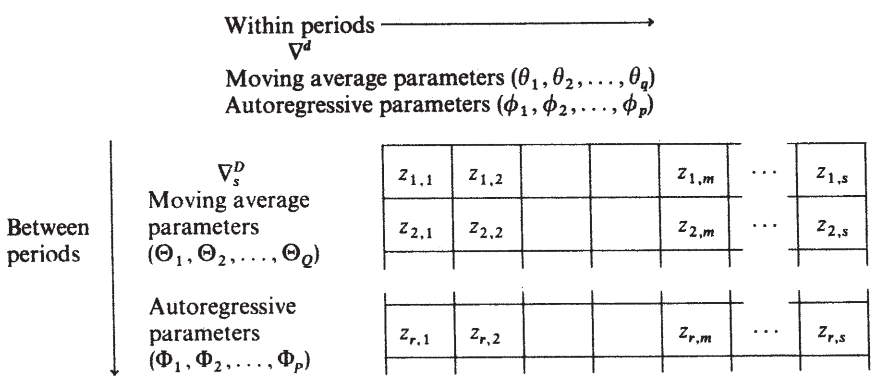

FIGURE 9.7 Two-way table for multiplicative seasonal model.

shown in Figure 9.7. In the multiplicative model it is supposed that the "between periods" development of the series is represented by some model

$$\Phi_P(B^s)\nabla_s^D z_{r,m} = \Theta_Q(B^s)\alpha_{r,m}$$

while "within periods" the α's are related by

$$\phi_p(B)\nabla^d \alpha_{r,m} = \theta_q(B)a_{r,m}$$

Obviously, we could change the order in which we considered the two types of models and in either case obtain the general multiplicative model

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_{r,m} = \theta_q(B)\Theta_Q(B^s)a_{r,m} \tag{9.3.1}$$

where $a_{r,m}$ is a white noise process with zero mean. In practice, the usefulness of models such as (9.3.1) depends on how far it is possible to parameterize actual time series parsimoniously in these terms. In fact, this has been possible for a variety of seasonal time series coming from widely different sources [25].

It is not possible to obtain a completely adequate fit with multiplicative models for all series. One modification that is sometimes useful allows the mixed moving average operator to be nonmultiplicative. By this is meant that we replace the operator $\theta_q(B)\Theta_Q(B^s)$ on the right-hand side of (9.3.1) by a more general moving average operator $\theta^*_{q^*}(B)$. Alternatively, or in addition, it may be necessary to replace the autoregressive operator $\phi_p(B)\Phi_P(B^s)$ on the left by a more general autoregressive operator $\phi^*_{p^*}(B)$. Some specimens of nonmultiplicative models are given in Appendix A9.1. These are numbered 4, 4a, 5, and 5a.

In those cases where a nonmultiplicative model is found necessary, experience suggests that the best-fitting multiplicative model can provide a good starting point from which to construct a better nonmultiplicative model. The situation is reminiscent of the problems encountered in analyzing two-way analysis of variance tables, where additivity of row and column constants may or may not be an adequate assumption, but may provide a good point of departure.

Our general strategy for relating multiplicative or nonmultiplicative models to data is that which we have already discussed and illustrated in some detail in Section 9.2. Using the autocorrelation function for guidance:

1. The series is differenced with respect to ∇ and/or ∇_s, so as to produce stationarity.
2. By inspection of the autocorrelation function of the suitably differenced series, a tentative model is selected.
3. From the values of appropriate autocorrelations of the differenced series, preliminary estimates of the parameters are obtained. These can be used as starting values in the search for the least squares or maximum likelihood estimates.
4. After fitting, the diagnostic checking process applied to the residuals either may lead to the acceptance of the tentative model or, alternatively, may suggest ways in which it can be improved, leading to refitting and repetition of the diagnostic checks.

As a few practical guidelines for model specification, we note that for seasonal series the order of seasonal differencing D needed would almost never be greater than one, and especially for monthly series with $s = 12$, the orders P and Q of the seasonal AR and MA operators $\Phi(B^s)$ and $\Theta(B^s)$ would rarely need to be greater than 1. This is particularly so when the series length of available data is not sufficient to warrant such a complicated form of model with $P > 1$ or $Q > 1$.

9.3.2 Identification

A useful aid in model identification is the list in Appendix A9.1, giving the autocovariance structure of $w_t = \nabla^d \nabla_s^D z_t$ for a number of simple seasonal models. This list makes no claim to be comprehensive. However, it is believed that it does include some of the frequently encountered models, and the reader should have no difficulty in discovering the characteristics of others that seem representationally useful. It should be emphasized that rather simple models (such as models 1 and 2 in Appendix A9.1) have provided adequate representations for many seasonal series.

Since the multiplicative seasonal ARMA models for the differences $w_t = \nabla\nabla_s z_t$ may be viewed merely as special forms of ARMA models with orders $p + sP$ and $q + sQ$, their autocovariances can be derived from the principles of Chapter 3, as was done in the previous section for the MA model $w_t = a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta\Theta a_{t-13}$. To illustrate the derivation of autocovariances for models such as those given in Appendix A9.1, consider the model

$$(1 - \phi B)w_t = (1 - \Theta B^s)a_t$$

which is a special form of ARMA model with AR order 1 and MA order s. First, since the ψ weights for this model for w_t satisfy $\psi_j - \phi\psi_{j-1} = 0$, $j = 1, \ldots, s-1$, we have $\psi_j = \phi^j$, $j = 1, \ldots, s-1$, as well as $\psi_s = \phi^s - \Theta$ and $\psi_j = \phi\psi_{j-1}$, $j > s$. From Section 3.4, we know that the autocovariances for w_t will satisfy

$$\begin{aligned} \gamma_0 &= \phi\gamma_1 + \sigma_a^2(1 - \Theta\psi_s) \\ \gamma_j &= \phi\gamma_{j-1} - \sigma_a^2\Theta\psi_{s-j} \qquad j = 1, \ldots, s \\ \gamma_j &= \phi\gamma_{j-1} \qquad j > s \end{aligned} \tag{9.3.2}$$

Solving the first two equations for γ_0 and γ_1, we obtain

$$\gamma_0 = \sigma_a^2 \frac{1 - \Theta(\phi^s - \Theta) - \phi^s\Theta}{1 - \phi^2} = \sigma_a^2 \frac{1 + \Theta^2 - 2\phi^s\Theta}{1 - \phi^2}$$

$$\gamma_1 = \sigma_a^2 \frac{\phi[1 - \Theta(\phi^s - \Theta)] - \phi^{s-1}\Theta}{1 - \phi^2} = \sigma_a^2 \frac{\phi(1 + \Theta^2 - \phi^s\Theta) - \phi^{s-1}\Theta}{1 - \phi^2}$$

with $\gamma_j = \phi\gamma_{j-1} - \sigma_a^2\Theta\phi^{s-j} = \phi^j\gamma_0 - \sigma_a^2\Theta\phi^{s-j}(1 - \phi^{2j})/(1 - \phi^2)$, $j = 1, \ldots, s$ and $\gamma_j = \phi\gamma_{j-1} = \phi^{j-s}\gamma_s$, $j > s$. Hence, in particular, for monthly data with $s = 12$ and $|\phi|$ not too close to one, the autocorrelation function ρ_j for this process will behave, for low lags, similarly to that of a regular AR(1) process, $\rho_j \simeq \phi^j$ for small j, while the value of ρ_{12} will be close to $-\Theta/(1 + \Theta^2)$.

A fact of considerable utility in deriving autocovariances of a multiplicative process is that for such a process, the autocovariance generating function (3.1.11) is the product of the generating functions of the components. Thus, in (9.3.1) if the component models for $\nabla^d z_t$ and $\nabla_s^D \alpha_t$,

$$\phi_p(B)\nabla^d z_t = \theta_q(B)\alpha_t \qquad \Phi_P(B^s)\nabla_s^D \alpha_t = \Theta_Q(B)a_t$$

have autocovariance generating function $\gamma(B)$ and $\Gamma(B^s)$, the autocovariance generating function for $w_t = \nabla^d \nabla_s^D z_t$ in (9.3.1) is

$$\gamma(B)\Gamma(B^s)$$

Another point to be remembered is that it may be useful to parameterize more general models in terms of their departures from related multiplicative forms in a manner now illustrated.

The three-parameter nonmultiplicative operator

$$1 - \theta_1 B - \theta_{12} B^{12} - \theta_{13} B^{13} \tag{9.3.3}$$

employed in models 4 and 5 may be written

$$(1 - \theta_1 B)(1 - \theta_{12} B^{12}) - \kappa B^{13}$$

where

$$\kappa = \theta_1 \theta_{12} - (-\theta_{13})$$

An estimate of κ that was large compared with its standard error would indicate the need for a nonmultiplicative model in which the value of θ_{13} is not tied to the values of θ_1 and θ_{12}. On the other hand, if κ is small, then on writing $\theta_1 = \theta$, $\theta_{12} = \Theta$, the model approximates the multiplicative $(0, 1, 1) \times (0, 1, 1)_{12}$ model.

9.3.3 Estimation

No new problems arise in the estimation of the parameters of general seasonal models. The unconditional sum of squares is computed quite generally by the methods set out fully in Section 7.1.5 and illustrated further in Section 9.2.4. As always, contour plotting can illuminate difficult situations. In well-behaved situations, iterative least squares with numerical determination of derivatives yield rapid convergence to the least squares estimates, together with approximate variances and covariances of the estimates. Recursive procedures can be derived in each case which allow direct calculation of derivatives, if desired.

Large-Sample Variances and Covariances of the Estimates The large-sample information matrix $\mathbf{I}(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta})$ is given by evaluating $E[\mathbf{X}'\mathbf{X}]$, where, as in

Section 7.2.6, $\mathbf{X}$ is the $n \times (p+q+P+Q)$ matrix of derivatives with reversed signs. Thus, for the general multiplicative model

$$a_t = \theta^{-1}(B)\Theta^{-1}(B^s)\phi(B)\Phi(B^s)w_t$$

where $w_t = \nabla^d \nabla_s^D z_t$, the required derivatives are

$$\frac{\partial a_t}{\partial \theta_i} = \theta^{-1}(B)B^i a_t \qquad \frac{\partial a_t}{\partial \Theta_i} = \Theta^{-1}(B^s)B^{si} a_t$$
$$\frac{\partial a_t}{\partial \phi_j} = -\phi^{-1}(B)B^j a_t \qquad \frac{\partial a_t}{\partial \Phi_j} = -\Phi^{-1}(B^s)B^{sj} a_t$$

Approximate variances and covariances of the estimates are obtained as before, by inverting the matrix $\mathbf{I}(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta})$.

9.3.4 Eventual Forecast Functions for Various Seasonal Models

We now consider the characteristics of the eventual forecast functions for a number of seasonal models. For a seasonal model with single periodicity s, the eventual forecast function at origin t for lead time l is the solution of the difference equation

$$\phi(B)\Phi(B^s)\nabla^d \nabla_s^D \hat{z}_t(l) = 0$$

Table 9.8 shows this solution for various choices of the difference equation; also shown is the number of initial values on which the behavior of the forecast function depends.

In Figure 9.8 the behavior of each forecast function is illustrated for $s = 4$. It will be convenient to regard the lead time $l = rs + m$ as referring to a forecast r years and m quarters ahead. In the diagram, an appropriate number of initial values (required to start the forecast off and indicated by bold dots) has been set arbitrarily and the course of the forecast function traced to the end of the fourth period. When the difference equation involves an autoregressive parameter, its value has been set equal to 0.5.

The constants $b_{0,m}$, b_1, and so on, appearing in the solutions in Table 9.8, should strictly be indicated by $b_{0,m}^{(t)}$, $b_1^{(t)}$, and so on, since each one depends on the origin t of the forecast, and these constants are adaptively modified each time the origin changes. The superscript t has been omitted temporarily to simplify notation.

The operator labeled (1) in Table 9.8 is stationary, with the model containing a fixed mean μ. It is autoregressive in the seasonal pattern, and the forecast function decays with each period, approaching closer and closer to the mean.

Operator (2) in Table 9.8 is nonstationary in the seasonal component. The forecasts for a particular quarter are linked from year to year by a polynomial of degree 0. Thus, the basic forecast of the seasonal component is exactly reproduced in forecasts of future years.

Table 9.8 Eventual Forecast Functions for Various Generalized Autoregressive Operators

Generalized Autoregressive Operator	Eventual Forecast Function $\hat{z}(r, m)^a$	Number of Initial Values on which Forecast Function Depends
(1) $1 - \Phi B^s$	$\mu + (b_{0,m} - \mu)\Phi^r$	s
(2) $1 - B^s$	$b_{0,m}$	s
(3) $(1 - B)(1 - \Phi B^s)$	$b_0 + (b_{0,m} - b_0)\Phi^r + b_1 \left\{ \dfrac{1 - \Phi^r}{1 - \Phi} \right\}$	$s + 1$
(4) $(1 - B)(1 - B^s)$	$b_{0,m} + b_1 r$	$s + 1$
(5) $(1 - \phi B)(1 - B^s)$	$b_{0,m} + b_1 \phi^{m-1} \left\{ \dfrac{1 - \phi^{sr}}{1 - \phi^s} \right\}$	$s + 1$
(6) $(1 - B)(1 - B^s)^2$	$b_{0,m} + b_{1,m} r + \frac{1}{2} b_2 r(r - 1)$	$2s + 1$
(7) $(1 - B)^2(1 - B^s)$	$b_{0,m} + [b_1 + (m - 1)b_2]r$ $+\frac{1}{2} b_2 sr(r - 1)$	$s + 2$

[a] Coefficients b are all adaptive and depend upon forecast origin t.

Operator (3) in Table 9.8 is nonstationary with respect to the basic time interval but stationary in the seasonal component. Operator (3) in Figure 9.8 shows the general level of the forecast approaching asymptotically the new level

$$b_0 + \frac{b_1}{1 - \Phi}$$

where, at the same time, the superimposed predictable component of the stationary seasonal effect dies out exponentially.

In Table 9.8, operator (4) is the limiting case of the operator (3) as Φ approaches unity. The operator is nonstationary with respect to both the basic time interval and the periodic component. The basic initial forecast pattern is reproduced, as is the incremental yearly increase. This is the type of forecast function given by the multiplicative $(0, 1, 1) \times (0, 1, 1)_{12}$ process fitted to the airline data.

Operator (5) is nonstationary in the seasonal pattern but stationary with respect to the basic time interval. The pattern approaches exponentially an asymptotic basic pattern

$$\hat{z}_t(\infty, m) = b_{0,m} + \frac{b_1 \phi^{m-1}}{1 - \phi^s}$$

Operator (6) is nonstationary in both the basic time interval and the seasonal component. An overall quadratic trend occurs over years, and a particular kind of modification occurs in the seasonal pattern. Individual quarters not only have their

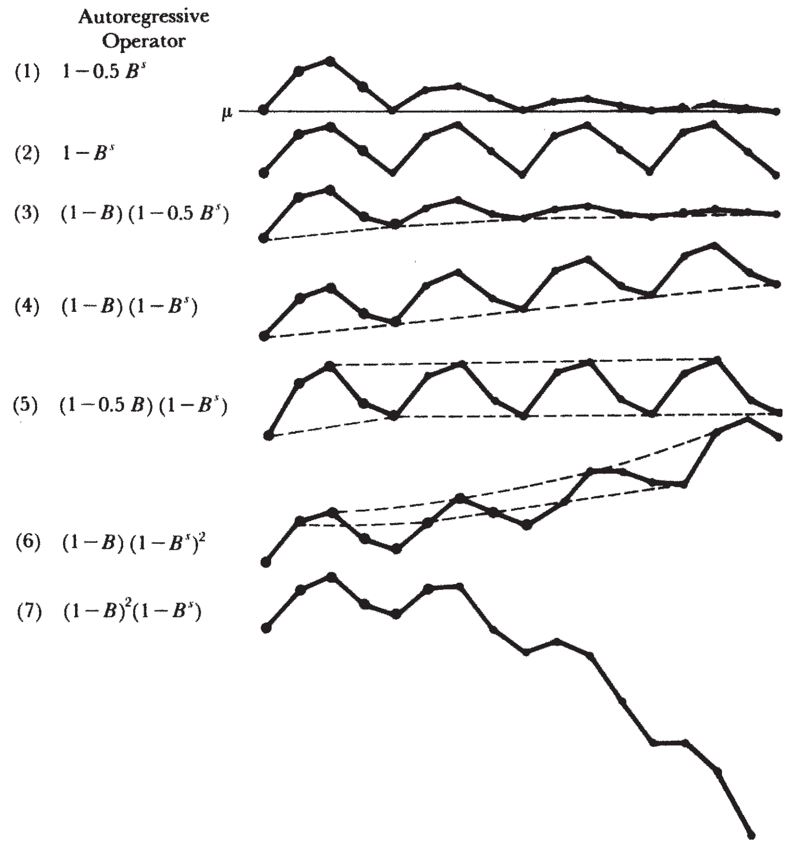

FIGURE 9.8 Behavior of the seasonal forecast function for various choices of the general seasonal autoregressive operator.

own level $b_{0,m}$ but also their own rate of change of level $b_{1,m}$. Therefore, when this kind of forecast function is appropriate, we can have a situation where, for example, as the lead time is increased, the difference in summer over spring sales can be forecast to increase from one year to the next, while at the same time the difference in autumn over summer sales can be forecast to decrease.

In Table 9.8, Operator (7) is again nonstationary in both the basic time interval and in the seasonal component, and there is again a quadratic tendency over years with the incremental changes in the forecasts from one quarter to the next changing linearly. However, in this case they are restricted to have a common *rate* of change.

9.3.5 Choice of Transformation

It is particularly true for seasonal models that the weighted averages of previous data values, which comprise the forecasts, may extend far back into the series.

Care is therefore needed in choosing a transformation in terms of which a parsimonious linear model will closely apply over a sufficient stretch of the series. Simple graphical analysis can often suggest such a transformation. Thus, an appropriate transformation may be suggested by determining in what metric the amplitude of the seasonal component is roughly independent of the level of the series. To illustrate how a data-based transformation may be chosen more exactly, denote the *untransformed* airline data by x, and let us assume that some power transformation [$z = x^\lambda$ for $\lambda \neq 0$, $z = \ln(x)$ for $\lambda = 0$] may be needed to make the model (9.2.1) appropriate. Then, as suggested in Section 4.1.3, the approach of Box and Cox [52] may be followed, and the maximum likelihood value obtained by fitting the model to $x^{(\lambda)} = (x^\lambda - 1)/\lambda \dot{x}^{\lambda - 1}$ for various values of λ, and choosing the value of λ that results in the smallest residual sum of squares S_λ. In this expression $\dot{x}$ is the geometric mean of the series x, and it is easily shown that $x^{(0)} = \dot{x}\ \ln(x)$.

For the airline data we find

λ	S_λ	λ	S_λ	λ	S_λ
−0.4	13,825.5	−0.1	11,627.2	0.2	11,784.3
−0.3	12,794.6	0.0	11,458.1	0.3	12,180.0
−0.2	12,046.0	0.1	11,554.3	0.4	12,633.2

The maximum likelihood value is thus close to $\lambda = 0$, confirming for this particular example the appropriateness of the logarithmic transformation.

9.4 STRUCTURAL COMPONENT MODELS AND DETERMINISTIC SEASONAL COMPONENTS

As mentioned at the beginning of Section 9.1.1, a traditional method to represent a seasonal time series has been to decompose the series into trend, seasonal, and noise components, as $z_t = T_t + S_t + N_t$, where the trend T_t and seasonal S_t are represented as deterministic functions of time using polynomial and sinusoidal functions, respectively. More recently, such form of models but with the trend and seasonal components following stochastic rather than deterministic models, referred to as *structural component* models, have become increasingly popular for modeling, forecasting, and seasonal adjustment of time series (e.g., Harvey and Todd [148], Gersch and Kitagawa [124], Kitagawa and Gersch [180], and Hillmer and Tiao [154]). We touch upon this topic in the following sections.

9.4.1 Structural Component Time Series Models

In general, a univariate structural component time series model is one in which an observed series z_t is formulated as the sum of unobservable component or “signal” time series. Although the components are unobservable and cannot be uniquely specified, they will usually have direct meaningful interpretation, such as representing the “seasonal” component behavior or the long-term “trend” component behavior of an observed economic time series or the physical “signal” that

Table A9.1 Autocovariances for Some Seasonal Models

Model	(Autocovariances of w_t)/σ_a^2	Special Characteristics
(1) $w_t = (1-\theta B)(1-\Theta B^s)a_t$ $w_t = a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1}$ $s \geq 3$	$\gamma_0 = (1+\theta^2)(1+\Theta^2)$ $\gamma_1 = -\theta(1+\Theta^2)$ $\gamma_{s-1} = \theta\Theta$ $\gamma_s = -\Theta(1+\theta^2)$ $\gamma_{s+1} = \gamma_{s-1}$ All other autocovariances are zero.	(a) $\gamma_{s-1} = \gamma_{s+1}$ (b) $\rho_{s-1} = \rho_{s+1} = \rho_1\rho_s$
(2) $(1-\Phi B^s)w_t = (1-\theta B)(1-\Theta B^s)a_t$ $w_t - \Phi w_{t-s} = a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1}$ $s \geq 3$	$\gamma_0 = (1+\theta^2)\left[1+\frac{(\Theta-\Phi)^2}{1-\Phi^2}\right]$ $\gamma_1 = -\theta\left[1+\frac{(\Theta-\Phi)^2}{1-\Phi^2}\right]$ $\gamma_{s-1} = \theta\left[\Theta-\Phi-\frac{\Phi(\Theta-\Phi)^2}{1-\Phi^2}\right]$ $\gamma_s = -(1+\theta^2)\left[\Theta-\Phi-\frac{\Phi(\Theta-\Phi)^2}{1-\Phi^2}\right]$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_j = \Phi\gamma_{j-s} j \geq s+2$ For $s \geq 4$, $\gamma_2, \gamma_3, \ldots, \gamma_{s-2}$ are all zero.	(a) $\gamma_{s-1} = \gamma_{s+1}$ (b) $\gamma_j = \Phi\gamma_{j-s} j \geq s+2$
(3) $w_t = (1-\theta_1 B - \theta_2 B^2)$ $\times (1-\Theta_1 B^s - \Theta_2 B^{2s})a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Theta_1 a_{t-s}$ $+ \theta_1\Theta_1 a_{t-s-1} + \theta_2\Theta_1 a_{t-s-2}$ $- \Theta_2 a_{t-2s} + \theta_1\Theta_2 a_{t-2s-1}$ $+ \theta_2\Theta_2 a_{t-2s-2}$ $s \geq 5$	$\gamma_0 = (1+\theta_1^2+\theta_2^2)(1+\Theta_1^2+\Theta_2^2)$ $\gamma_1 = -\theta_1(1-\theta_2)(1+\Theta_1^2+\Theta_2^2)$ $\gamma_2 = -\theta_2(1+\Theta_1^2+\Theta_2^2)$ $\gamma_{s-2} = \theta_2\Theta_1(1-\Theta_2)$ $\gamma_{s-1} = \theta_1\Theta_1(1-\theta_2)(1-\Theta_2)$ $\gamma_s = -\Theta_1(1-\Theta_2)(1+\theta_1^2+\theta_2^2)$	(a) $\gamma_{s-2} = \gamma_{s+2}$ (b) $\gamma_{s-1} = \gamma_{s+1}$ (c) $\gamma_{2s-2} = \gamma_{2s+2}$ (d) $\gamma_{2s-1} = \gamma_{2s+1}$

	$\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{s+2} = \gamma_{s-2}$ $\gamma_{2s-2} = \theta_2\Theta_2$ $\gamma_{2s-1} = \theta_1(1-\theta_2)\Theta_2$ $\gamma_{2s} = -\Theta_2(1+\theta_1^2+\theta_2^2)$ $\gamma_{2s+1} = \gamma_{2s-1}$ $\gamma_{2s+2} = \gamma_{2s-2}$ All other autocovariances are zero.	
(3a) *Special case of model 3* $w_t = (1-\theta_1 B - \theta_2 B^2)(1-\Theta B^s)a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Theta a_{t-s}$ $\quad + \theta_1\Theta a_{t-s-1} + \theta_2\Theta a_{t-s-2}$ $s \geq 5$	$\gamma_0 = (1+\theta_1^2+\theta_2^2)(1+\Theta^2)$ $\gamma_1 = -\theta_1(1-\theta_2)(1+\Theta^2)$ $\gamma_2 = -\theta_2(1+\Theta^2)$ $\gamma_{s-2} = \theta_2\Theta$ $\gamma_{s-1} = \theta_1(1-\theta_2)\Theta$ $\gamma_s = -\Theta(1+\theta_1^2+\theta_2^2)$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{s+2} = \gamma_{s-2}$ All other autocovariances are zero.	(a) $\gamma_{s-2} = \gamma_{s+2}$ (b) $\gamma_{s-1} = \gamma_{s+1}$
(3b) *Special case of model 3* $w_t = (1-\theta B)(1-\Theta_1 B^s - \Theta_2 B^{2s})a_t$ $w_t = a_t - \theta a_{t-1} - \Theta_1 a_{t-s} + \theta\Theta_1 a_{t-s-1}$ $\quad - \Theta_2 a_{t-2s} + \theta\Theta_2 a_{t-2s-1}$ $s \geq 3$	$\gamma_0 = (1+\theta^2)(1+\Theta_1^2+\Theta_2^2)$ $\gamma_1 = -\theta(1+\Theta_1^2+\Theta_2^2)$ $\gamma_{s-1} = \theta\Theta_1(1-\Theta_2)$ $\gamma_s = -\Theta_1(1-\Theta_2)(1+\theta^2)$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{2s-1} = \theta\Theta_2$ $\gamma_{2s} = -\Theta_2(1+\theta^2)$ $\gamma_{2s+1} = \gamma_{2s-1}$ All other autocovariances are zero.	(a) $\gamma_{s-1} = \gamma_{s+1}$ (b) $\gamma_{2s-1} = \gamma_{2s+1}$

(*continued*)

Table A9.1 (*continued*)

Model	(Autocovariances of w_t)/σ_a^2	Special Characteristics
(4) $w_t = (1 - \theta_1 B - \theta_s B^s - \theta_{s+1} B^{s+1}) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $- \theta_{s+1} s_{t-s-1}$ $s \geq 3$	$\gamma_0 = 1 + \theta_1^2 + \theta_s^2 + \theta_{s+1}^2$ $\gamma_1 = -\theta_1 + \theta_s \theta_{s+1}$ $\gamma_{s-1} = \theta_1 \theta_s$ $\gamma_s = -\theta_s + \theta_1 \theta_{s+1}$ $\gamma_{s+1} = -\theta_{s+1}$ All other autocovariances are zero.	(a) In general, $\gamma_{s-1} \neq \gamma_{s+1}$ $\gamma_1 \gamma_s \neq \gamma_{s+1}$
(4a) *Special case of model 4* $w_t = (1 - \theta_1 B - \theta_s B^s) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $s \geq 3$	$\gamma_0 = 1 + \theta_1^2 + \theta_s^2$ $\gamma_1 = -\theta_1$ $\gamma_{s-1} = \theta_1 \theta_s$ $\gamma_s = -\theta_s$ All other autocovariances are zero.	(a) Unlike model 4, $\gamma_{s+1} = 0$
(5) $(1 - \Phi B^s) w_t = (1 - \theta_1 B - \theta_s B^s$ $- \theta_{s+1} B^{s+1}) a_t$ $w_t - \Phi w_{t-s} = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $- \theta_{s+1} a_{t-s-1}$ $s \geq 3$	$\gamma_0 = 1 + \theta_1^2 + \dfrac{(\theta_s - \Phi)^2}{1 - \Phi^2} + \dfrac{(\theta_{s+1} + \theta_1 \Phi)^2}{1 - \Phi^2}$ $\gamma_1 = -\theta_1 + \dfrac{(\theta_s - \Phi)(\theta_{s+1} + \theta_1 \Phi)}{1 - \Phi^2}$ $\gamma_{s-1} = (\theta_s - \Phi) \left[\theta_1 + \Phi \dfrac{\theta_{s+1} + \theta_1 \Phi}{1 - \Phi^2} \right]$ $\gamma_s = -(\theta_s - \Phi) \left[1 - \Phi \dfrac{\theta_s - \Phi}{1 - \Phi^2} \right]$ $+ (\theta_{s+1} + \theta_1 \Phi) \left[\theta_1 + \Phi \dfrac{\theta_{s+1} + \theta_1 \Phi}{1 - \Phi^2} \right]$ $\gamma_{s+1} = -(\theta_{s+1} + \theta_1 \Phi) \left[1 - \Phi \dfrac{\theta_s - \Phi}{1 - \Phi^2} \right]$ $\gamma_j = \Phi \gamma_{j-s} j \geq s + 2$ For $s \geq 4$, $\gamma_2, \ldots, \gamma_{s-2}$ are all zero.	(a) $\gamma_{s-1} \neq \gamma_{s+1}$ (b) $\gamma_j = \Phi \gamma_{j-s} j \geq s + 2$

(5a) *Special case of model 5* $(1-\Phi B^s)w_t = (1-\theta_1 B - \theta_s B^s)a_t$ $w_t - \Phi w_{t-s} = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $s \geq 3$	$\gamma_0 = 1 + \frac{\theta_1^2 + (\theta_s - \Phi)^2}{1-\Phi^2}$ $\gamma_1 = -\theta_1 \left[1 - \Phi \frac{\theta_s - \Phi}{1-\Phi^2}\right]$ $\gamma_{s-1} = \frac{\theta_1(\theta_s - \Phi)}{1-\Phi^2}$ $\gamma_s = \frac{\Phi\theta_1^2 - (\theta_s - \Phi)(1-\Phi\theta_s)}{1-\Phi^2}$ $\gamma_j = \Phi\gamma_{j-s} j \geq s+1$ For $s \geq 4$, $\gamma_2, \ldots, \gamma_{s-2}$ are all zero.	(a) Unlike model 5, $\gamma_{s+1} = \Phi\gamma_1$