1

Introduction

1.1 BASIC DESCRIPTION

The study of time series is concerned with time correlation structures. It has diverse applications ranging from oceanography to finance. The celebrated CAPM model and the stochastic volatility model are examples of financial models that contain a time series component. When we think of a time series, we usually think of a collection of values $\{X_t : t = 1, \ldots, n\}$ in which the subscript t indicates the time at which the datum X_t is observed. Although intuitively clear, a number of nonstandard features of X_t can be elaborated.

UNEQUALLY SPACED DATA (MISSING VALUES). For example, if the series is about daily returns of a security, values are not available during nontrading days such as holidays.

CONTINUOUS-TIME SERIES. In many physical phenomena, the underlying quantity of interest is governed by a continuously evolving mechanism and the data observed should be modeled by a continuous time series $X(t)$. In finance, we can think of tick-by-tick data as a close approximation to the continuous evolution of the market.

AGGREGATION. The series observed may represent an accumulation of underlying quantities over a period of time. For example, daily returns can be thought of as the aggregation of tick-by-tick returns within the same day.

REPLICATED SERIES. The data may represent repeated measurements of the same quantity across different subjects. For example, we might monitor the total weekly spending of each of a number of customers of a supermarket chain over time.

MULTIPLE TIME SERIES. Instead of being a one-dimensional scalar, X_t can be a vector with each component representing an individual time series. For example, the returns of a portfolio that consist of p equities can be expressed as $\mathbf{X}_t = (X_{1t}, \ldots, X_{pt})'$, where each X_{it}, $i = 1, \ldots, p$, represents the returns of each equity in the portfolio. In this case, we will be interested not only in the serial correlation structures within each equity, but also the cross-correlation structures among different equities.

NONLINEARITY, NONSTATIONARITY, AND HETEROGENEITY. Many of the time series encountered in practice may behave nonlinearly. Sometimes transformation may help, but we often have to build elaborate models to account for such nonstandard features. For example, the asymmetric behavior of stock returns motivates the study of GARCH models.

Although these features are important, in this book we deal primarily with standard scalar time series. Only after a thorough understanding of the techniques and difficulties involved in analyzing a regularly spaced scalar time series will we be able to tackle some of the nonstandard features.

In classical statistics, we usually assume the X's to be independent. In a time series context, the X's are usually serially correlated, and one of the objectives in time series analysis is to make use of this serial correlation structure to help us build better models. The following example illustrates this point in a confidence interval context.

Example 1.1 *Let X_t be generated by the following model:*

$$X_t = \mu + a_t - \theta a_{t-1}, \quad a_t \sim \mathrm{N}(0,1) \text{ i.i.d.}$$

Clearly, $\mathrm{E}(X_t) = \mu$ *and* $\mathrm{var} X_t = 1 + \theta^2$. *Thus,*

$$\begin{aligned} \mathrm{cov}(X_t, X_{t-k}) &= \mathrm{E}(X_t - \mu)(X_{t-k} - \mu) \\ &= \mathrm{E}(a_t - \theta a_{t-1})(a_{t-k} - \theta a_{t-k-1}) \\ &= \begin{cases} -\theta, & |k| = 1, \\ 1 + \theta^2, & k = 0, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Let $\overline{X} = (\sum_{t=1}^{n} X_t)/n$. *By means of the formula*

$$\mathrm{var}\left(\frac{1}{n}\sum_{t=1}^{n} X_t\right) = \frac{1}{n^2}\sum_{t=1}^{n}\mathrm{var}(X_t) + \frac{2}{n^2}\sum_{t=1}^{n}\sum_{j=1}^{t-1}\mathrm{cov}(X_t, X_j),$$

Table 1.1 Lengths of Confidence Intervals for $n = 50$

θ	$L(\theta)$
-1	$L(-1) = (4 - \frac{2}{50})^{1/2} \cong 2$
-0.5	1.34
0	1
0.5	0.45
1	0.14

it is easily seen that

$$\begin{aligned}\sigma^2_{\overline{X}} &= \operatorname{var} \overline{X} \\ &= \frac{1}{n^2} n(1+\theta^2) - \frac{2}{n^2}(n-1)\theta \\ &= \frac{1}{n}\left(1+\theta^2 - 2\theta + \frac{2\theta}{n}\right) \\ &= \frac{1}{n}\left[(1-\theta)^2 + \frac{2\theta}{n}\right].\end{aligned}$$

Therefore, $\overline{X} \sim \mathrm{N}(\mu, \sigma^2_{\overline{X}})$. *Hence, an approximate* 95% *confidence interval* (*CI*) *for* μ *is*

$$\overline{X} \pm 2\sigma^2_{\overline{X}} = \overline{X} \pm \frac{2}{\sqrt{n}}\left[(1-\theta)^2 + \frac{2\theta}{n}\right]^{1/2}.$$

If $\theta = 0$, *this CI becomes*

$$\overline{X} \pm \frac{2}{\sqrt{n}},$$

coinciding with the independent identically distributed (*i.i.d.*) *case. The difference in the CIs between* $\theta = 0$ *and* $\theta \neq 0$ *can be expressed as*

$$L(\theta) = \left[(1-\theta)^2 + \frac{2\theta}{n}\right]^{1/2}.$$

Table 1.1 *gives numerical values of the differences for* $n = 50$. *For example, if* $\theta = 1$ *and if we were to use a CI of zero for* θ, *the wrongly constructed CI would be much longer than it is supposed to be. The time correlation structure given by the model helps to produce better inference in this situation.* □

Example 1.2 *As a second example, we consider the equity-style timing model discussed in Kao and Shumaker* (1999). *In this article the authors try to explain the spread between value and growth stocks using several fundamental quantities. Among them, the most interesting variable is the earnings–yield*

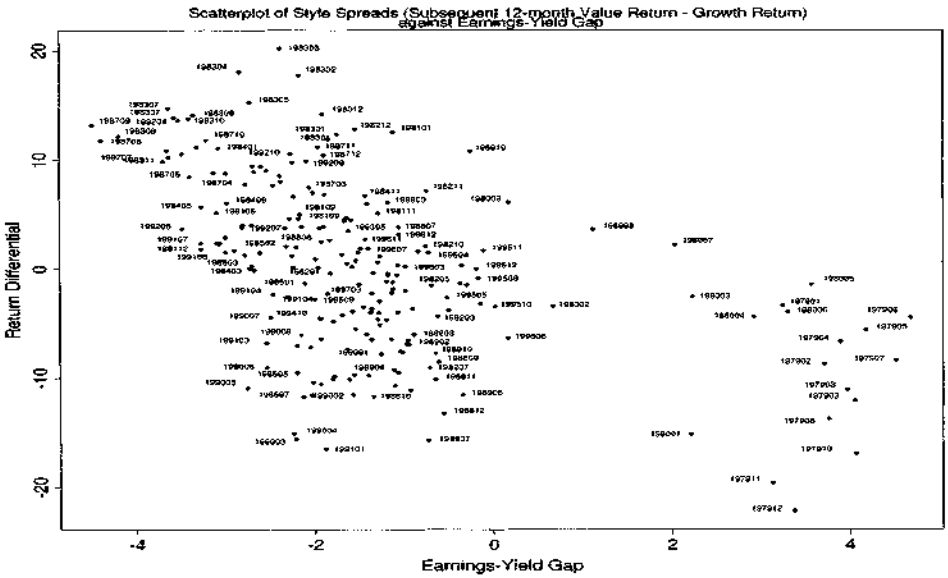

Fig. 1.1 Equity-style timing.

gap reported in Figure 4 *of their paper. This variable explains almost* 30% *of the variation of the spread between value and growth and suggests that the earnings–yield gap might be a highly informative regressor. Further description of this data set is given in their article. We repeat this particular analysis, but taking into account the time order of the observations. The data between January* 79 *to June* 97 *are stored in the file* `eygap.dat` *on the Web page for this book, which can be found at*

`http://www.sta.cuhk.edu.hk/data1/staff/nhchan/tsbook.html`

For the time being, we restrict our attention to reproducing Figure 4 *of Kao and Shumaker* (1999). *The plot and* SPLUS /R *commands are as follows:*

```
>eyield<-read.table("eygap.dat",header=T)
>plot(eyield[,2],eyield[,3],xlab="Earnings-Yield Gap",
+ ylab="Return Differential")
>title("Scatterplot of Style Spreads (Subsequent
+ 12-month Value Return - Growth Return)
+ against Earnings-Yield Gap, Jan 79- Jun 97",cex=0.6)
>identify(eyield[,2],eyield[,3],eyield[,1],cex=0.5)
```

As illustrated in Figure 1.1, *the scattergram can be separated into two clouds, those belonging to the first two years of data and those belonging to subsequent years. When time is taken into account, it seems that finding* *an $R^2 = 0.3$ depends crucially on the data cloud between 79 and 80 at the lower right-hand corner of Figure 1.1. Accordingly, the finding of such a high explanatory power from the earnings–yield gap seems to be spurious. This example demonstrates that important information may be missing when the time dimension is not taken properly into account.* □

1.2 SIMPLE DESCRIPTIVE TECHNIQUES

In general, a time series can be decomposed into a macroscopic component and a microscopic component. The macroscopic component can usually be described through a trend or seasonality, whereas the microscopic component may require more sophisticated methods to describe it. In this section we deal with the macroscopic component through some simple descriptive techniques and defer the study of the microscopic component to later chapters. Consider in general that the time series $\{X_t\}$ is decomposed into a time trend part T_t, a seasonal part S_t, and a microscopic part given by the noise N_t. Formally,

$$\begin{aligned} X_t &= T_t + S_t + N_t \\ &\cong \mu_t + N_t. \end{aligned} \tag{1.1}$$

1.2.1 Trends

Suppose that the seasonal part is absent and we have only a simple time trend structure, so that T_t can be expressed as a parametric function of t, $T_t = \alpha + \beta t$, for example. Then T_t can be identified through several simple devices.

Least squares method. We can use the least squares (LS) procedure to estimate T_t easily [i.e., find α and β such that $\sum (X_t - T_t)^2$ is minimized]. Although this method is convenient, there are several drawbacks.

1. We need to assume a fixed trend for the entire span of the data set, which may not be true in general. In reality, the form of the trend may also be changing over time and we may need an adaptive method to accommodate this change. An immediate example is the daily price of a given stock. For a fixed time span, the prices can be modeled pretty satisfactorily through a linear trend. But everyone knows that the fixed trend will give disastrous predictions in the long run.

2. For the LS method to be effective, we can only deal with a simple restricted form of T_t.

Filtering. In addition to using the LS method, we can filter or smooth the series to estimate the trend, that is, use a smoother or a moving average filter, such as

$$Y_t = \mathrm{Sm}(X_t) = \sum_{r=-q}^{s} a_r X_{t+r}.$$

We can represent the relationship between the output Y_t and the input X_t as

$$X_t \rightarrow \boxed{\text{filter}} \rightarrow \text{Sm}(X_t) = Y_t.$$

The weights $\{a_r\}$ of the filters are usually assumed to be symmetric and normalized (i.e., $a_r = a_{-r}$ and $\sum a_r = 1$). An obvious example is the simple moving average filter given by

$$Y_t = \frac{1}{2q+1} \sum_{r=-q}^{q} X_{t+r}.$$

The length of this filter is determined by the number q. When $q = 1$ we have a simple three-point moving average. The weights do not have to be the same at each point, however. An early example of unequal weights is given by the Spencer 15-point filter, introduced by an English actuary, Spencer, in 1904.

The idea is to use the 15-point filter to approximate the filter that passes through a cubic trend. Specifically, define the weights $\{a_r\}$ as

$$\begin{aligned} a_r &= a_{-r}, \\ a_r &= 0, \qquad |\, r \,| > 7, \\ (a_0,\ a_1, \ldots, a_7) &= \tfrac{1}{320}(74,\ 67,\ 46,\ 21,\ 3,\ -5,\ -6,\ -3). \end{aligned}$$

It can easily be shown that the Spencer 15-point filter does not distort a cubic trend; that is, for $T_t = at^3 + bt^2 + ct + d$,

$$\begin{aligned} \text{Sm}(X_t) &= \sum_{r=-7}^{7} a_r T_{t+r} + \sum_{r=-7}^{7} a_r N_{t+r} \\ &\cong \sum_{r=-7}^{7} a_r T_{t+r} \\ &= T_t. \end{aligned}$$

In general, it can be shown that a linear filter with weights $\{a_r\}$ passes a polynomial of degree k in t, $\sum_{i=0}^{k} c_i t^i$, without distortion if and only if the weights $\{a_r\}$ satisfy two conditions, as described next.

Proposition 1.1 *$T_t = \sum_r a_r T_{t+r}$, for all kth-degree polynomials $T_t = c_0 + c_1 t + \cdots + c_k t^k$ if and only if*

$$\begin{aligned} \sum_{r=-s}^{s} a_r &= 1, \\ \sum_{r=-s}^{s} r^j a_r &= 0, \quad \text{for } j = 1, \ldots, k. \end{aligned}$$

The reader is asked to provide a proof of this result in the exercises. Using this result, it is straightforward to verify that the Spencer 15-point filter passes a cubic polynomial without distortion. For the time being, let us illustrate the main idea on how a filter works by means of the simple case of a linear trend where $X_t = T_t + N_t$, $T_t = \alpha + \beta t$. Consider applying a $(2q+1)$-point moving average filter (smoother) to X_t :

$$\begin{aligned} Y_t &= \mathrm{Sm}(X_t) = \frac{1}{2q+1}\sum_{r=-q}^{q} X_{t+r} \\ &= \frac{1}{2q+1}\sum_{r=-q}^{q} [\alpha + \beta(t+r)] + N_{t+r} \\ &\cong \alpha + \beta t \end{aligned}$$

if $\frac{1}{2q+1}\sum_{r=-q}^{q} N_{t+r} \cong 0$. In other words, if we use Y_t to estimate the trend, it does a pretty good job. We use the notation $Y_t = \mathrm{Sm}(X_t) = \hat{T}_t$ and $\mathrm{Res}(X_t) = X_t - \hat{T}_t = X_t - \mathrm{Sm}(X_t) \cong N_t$. In this case we have what is known as a low-pass filter [i.e., a filter that passes through the low-frequency part (the smooth part) and filters out the high-frequency part, N_t]. In contrast, we can construct a high-pass filter that filters out the trend. One drawback of a low-pass filter is that we can only use the middle section of the data. If end-points are needed, we have to modify the filter accordingly. For example, consider the filter

$$\mathrm{Sm}(X_t) = \sum_{j=0}^{\infty} \alpha(1-\alpha)^j X_{t-j},$$

where $0 < \alpha < 1$. Known as the exponential smoothing technique, this plays a crucial role in many empirical studies. Experience suggests that α is chosen between 0.1 and 0.3. Finding the best filter for a specific trend was once an important topic in time series. Tables of weights were constructed for different kinds of lower-order trends. Further discussion of this point can be found in Kendall and Ord (1990).

DIFFERENCING. The preceding methods aim at estimating the trend by a smoother $\hat{T}_t$. In many practical applications, the trend may be known in advance, so it is of less importance to estimate it. Instead, we might be interested in removing its effect and concentrate on analyzing the microscopic component. In this case it will be more desirable to eliminate or annihilate the effect of a trend. We can do this by looking at the residuals $\mathrm{Res}(X_t) = X_t - \mathrm{Sm}(X_t)$. A more convenient method, however, will be to eliminate the trend from the series directly. The simplest method is differencing. Let B be the backshift operator such that $BX_t = X_{t-1}$. Define

$$\begin{aligned} \Delta X_t &= (1-B)X_t = X_t - X_{t-1}, \\ \Delta^j X_t &= (1-B)^j X_t, \; j = 1, 2, \ldots . \end{aligned}$$

If $X_t = T_t + N_t$, with $T_t = \sum_{j=0}^{p} a_j t^j$, then $\Delta^j X_t = j! a_j + \Delta^j N_t$ and T_t is eliminated. Therefore, differencing is a form of high-pass filter that filters out the low-frequency signal, the trend T_t, and passes through the high-frequency part, N_t. In principle, we can eliminate any polynomial trend by differencing the series enough times. But this method suffers one drawback in practice. Each time we difference the series, we lose one data point. Consequently, it is not advisable to difference the data too often.

LOCAL CURVE FITTING. If the trend turns out to be more complicated, local curve smoothing techniques beyond a simple moving average may be required to obtain good estimates. Some commonly used methods are spline curve fitting and nonparametric regression. Interested readers can find a lucid discussion about spline smoothing in Diggle (1990).

1.2.2 Seasonal Cycles

When the seasonal component S_t is present in equation (1.1), the methods of Section 1.2.1 have to be modified to accommodate this seasonality. Broadly speaking, the seasonal component can be either additive or multiplicative, according to the following formulations:

$$X_t = \begin{cases} T_t + S_t + N_t, & \text{additive case,} \\ T_t S_t N_t, & \text{multiplicative case.} \end{cases}$$

Again, depending on the goal, we can either estimate the seasonal part by some kind of seasonal smoother or eliminate it from the data by a seasonal differencing operation. Assume that the seasonal part has a period of d (i.e., $S_{t+d} = S_t,\ \sum_{j=1}^{d} S_j = 0$).

(A) Moving average method. We first estimate the trend part by a moving average filter running over a complete cycle so that the effect of the seasonality is averaged out. Depending on whether d is odd or even, we perform one of the following two steps:

1. If $d = 2q$, let $\hat{T}_t = \frac{1}{d}\left(\frac{1}{2}X_{t-q} + X_{t-q+1} + \cdots + X_{t+q-1} + \frac{1}{2}X_{t+q}\right)$ for $t = q+1, \ldots, n-q$.
2. If $d = 2q+1$, let $\hat{T}_t = \frac{1}{d}\sum_{r=-q}^{q} X_{t+r}$ for $t = q+1, \ldots, n-q$.

After estimating T_t, filter it out from the data and estimate the seasonal part from the residual $X_t - \hat{T}_t$. Several methods are available to attain this last step, the most common being the moving average method. Interested readers are referred to Brockwell and Davis (1991) for further discussions and illustrations of this method. We illustrate this method by means of an example in Section 1.4.

(B) Seasonal differencing. On the other hand, we can apply seasonal differencing to eliminate the seasonal effect. Consider the dth differencing of the data $X_t - X_{t-d}$. This differencing eliminates the effect of S_t in equation (1.1). Again, we have to be cautious about differencing the data seasonably since we will lose data points.

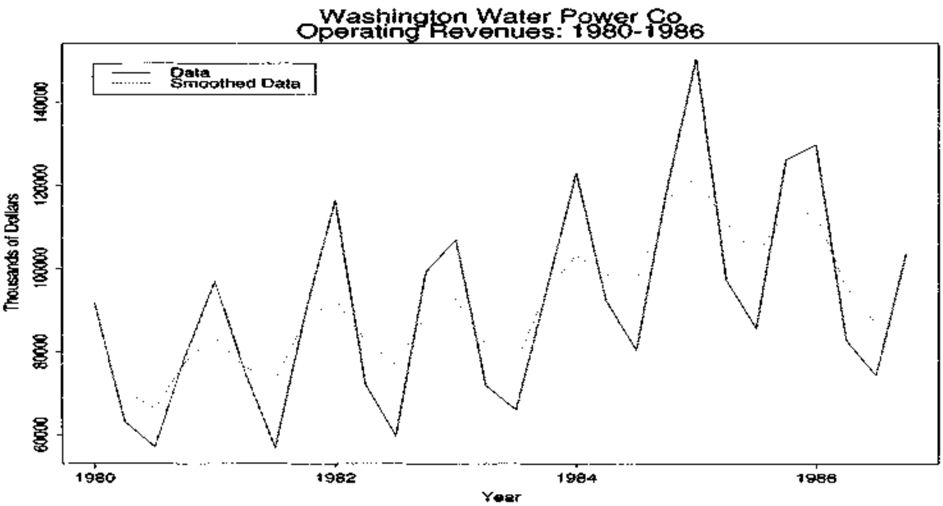

Fig. 1.2 Time series plots.

1.3 TRANSFORMATIONS

If the data exhibit an increase in variance over time, we may need to transform the data before analyzing them. The Box–Cox transformations can be applied here. Experience suggests, however, that log is the most commonly found transformation. Other types of transformations are more problematic, which can lead to serious difficulties in terms of interpretations and forecasting.

1.4 EXAMPLE

In this section we illustrate the idea of using descriptive techniques to analyze a time series. Figure 1.2 shows a time series plot of the quarterly operating revenues of Washington Water Power Company, 1980–1986, an electric and natural gas utility serving eastern Washington and northern Idaho. We start by plotting the data. Several conclusions can be drawn by inspecting the plot.

- As can be seen, there is a slight increasing trend. This appears to drop around 1985–1986.

- There is an annual (12-month) cycle that is pretty clear. Revenues are almost always lowest in the third quarter (July–September) and highest in the first quarter (January–March). Perhaps in this part of the country there is not much demand (and hence not much revenue) for electrical power in the summer (for air conditioning, say), but winters are cold and there is a lot of demand (and revenue) for natural gas and electric heat at that time.

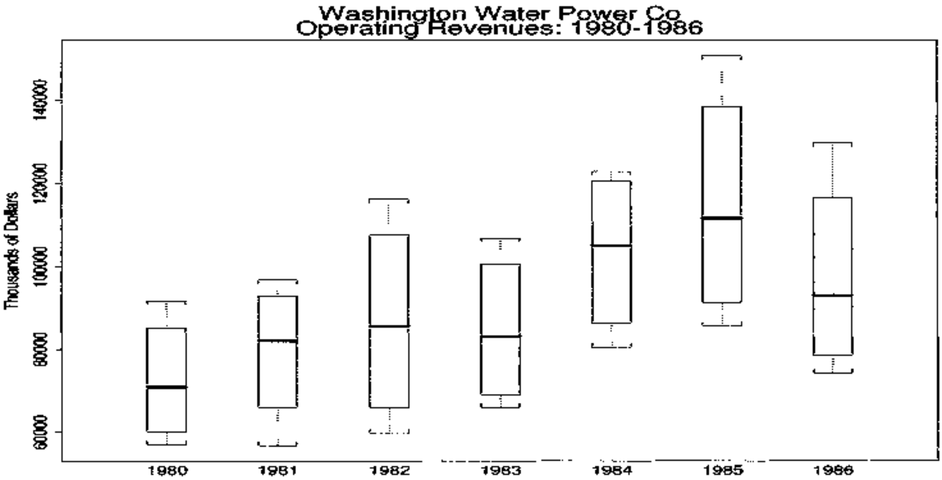

Fig. 1.3 Annual box plots.

- Figure 1.3 shows box plots for each year's operating revenues. The medians seem to rise from year to year and then fall back after the third year. The interquartile range (IQR) gets larger as the median grows and gets smaller as the median falls back; the range does the same. Most of the box plots are symmetric or very slightly positively skewed. There are no outliers.

- In Figure 1.3 we can draw a smooth curve connecting the medians of each year's quarterly operating revenues. We have already described the longer cycle about the medians; this pattern repeats once over the seven-year period graphed. This longer-term cycle is quite difficult to see in the original time series plot.

Assume that the data set has been stored in the file named `washpower.dat`. The SPLUS program that generates this analysis is listed as follows. In the case of R, replace the command "rts" in SPLUS by the command "ts". The rest of the commands in R are exactly the same. Readers are encouraged to work through these commands to get acquainted with the SPLUS /R program. Further explanations of these commands can be found in the books of Krause and Olson (1997) and Venables and Ripley (1999).

```
>wash<-rts(scan(''washpower.dat''),start=1980,freq=4)
>wash.ma<-filter(wash,c(1/3,1/3,1/3))
>leg.names<-c('Data','Smoothed Data')
>ts.plot(wash,wash.ma,lty=c(1,2),
+ main='Washington Water Power Co
Continue string: Operating Revenues: 1980-1986',
+ ylab='Thousands of Dollars',xlab='Year')
```

```
>legend(locator(1),leg.names,lty=c(1,2))
>wash.mat<-matrix(wash,nrow=4)
>boxplot(as.data.frame(wash.mat),names=as.character(seq(1980,
+ 1986)), boxcol=-1,medcol=1,main='Washington Water Power Co
Continue string: Operating Revenues: 1980-1986',
+ ylab='Thousands of Dollars')
```

To assess the seasonality, we perform the following steps in the moving average method.

1. Estimate the trend through one complete cycle of the series with $n = 28, d = 4$, and $q = 2$ to form $X_t - \hat{T}_t : t = 3, \ldots, 26$. The $\hat{T}_t$ is denoted by `washsea.ma` in the program.

2. Compute the averages of the deviations $\{X_t - \hat{T}_t\}$ over the entire span of the data. Then estimate the seasonal part $\hat{S}_i : i = 1, \ldots, 4$ by computing the demeaned values of these averages. Finally, for $i = 1, \ldots, 4$ let $\hat{S}_{i+4j} = \hat{S}_i : j = 1, \ldots, 6$. The estimated seasonal component $\hat{S}_i$ is denoted by `wash.sea` in the program, and the deseasonalized part of the data $X_t - \hat{S}_t$ is denoted by `wash.nosea`.

3. The third step involves reestimating the trend from the deseasonalized data `wash.nosea`. This is accomplished by applying a filter or any convenient method to reestimate the trend by $\hat{T}_t$, which is denoted by `wash.ma2` in the program.

4. Finally, check the residual $X_t - \hat{T}_t - \hat{S}_t$, which is denoted by `wash.res` in the program, to detect further structures. The SPLUS /R code follows.

```
> washsea.ma<-filter(wash,c(1/8,rep(1/4,3),1/8))
> wash.sea<-c(0,0,0,0)
> for(i in 1:2){
+     for(j in 1:6) {
+         wash.sea[i]<-wash.sea[i]+
+         (wash[i+4*j][[1]]-washsea.ma[i+4*j][[1]])
+     }
+ }
> for(i in 3:4){
+    for (j in 1:6){
+        wash.sea[i]<-wash.sea[i]+
+        (wash[i+4*(j-1)][[1]]-washsea.ma[i+4*(j-1)][[1]])
+    }
+ }
> wash.sea<-(wash.sea-mean(wash.sea))/6
> wash.sea1<-rep(wash.sea,7)
> wash.nosea<-wash-wash.sea
> wash.ma2<-filter(wash.nosea,c(1/8,rep(1/4,3),1/8))
```

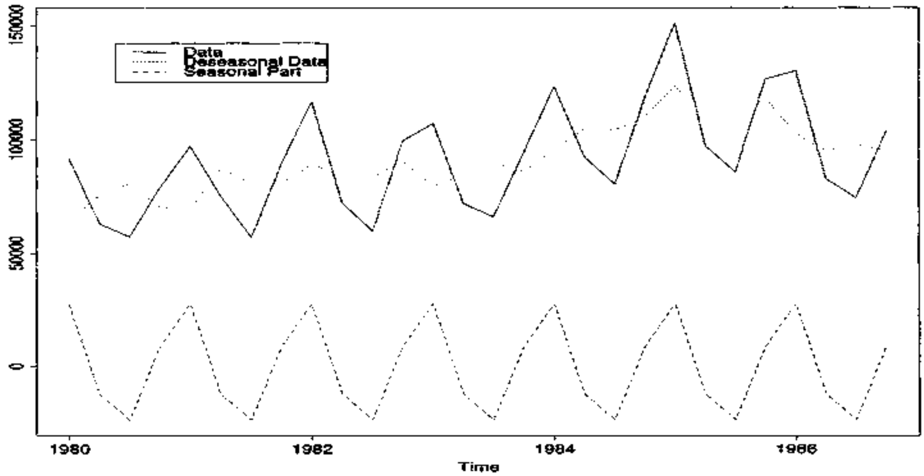

Fig. 1.4 Moving average method of seasonal decomposition.

```
> wash.res<-wash-wash.ma2-wash.sea
> write(wash.sea1, file='out.dat')
> wash.seatime<-rts(scan('out.dat'),start=1980,freq=4)
% This step converts a non-time series object into a time
% series object.
> ts.plot(wash,wash.nosea,wash.seatime)
```

Figure 1.4 gives the time series plot, which contains the data, the deseasonalized data, and the seasonal part. If needed, we can also plot the residual `wash.res` to detect further structures. But it is pretty clear that most of the structures in this example have been identified.

Note that SPLUS /R also has its own seasonal decomposition function `stl`. Details of this can be found with the `help` command. To execute it, use

```
> wash.stl<-stl(wash,'periodic')
> dwash<-diff(wash,4)
> ts.plot(wash,wash.stl$sea,wash.stl$rem,dwash)
```

Figure 1.5 gives the plot of the data, the deseasonal part, and the seasonal part. Comparing Figures 1.4 and 1.5 indicates that these two methods accomplish the same task of seasonal adjustments. As a final illustration we can difference the data with four lags to eliminate the seasonal effect. The plot of this differenced series is also drawn in Figure 1.5.

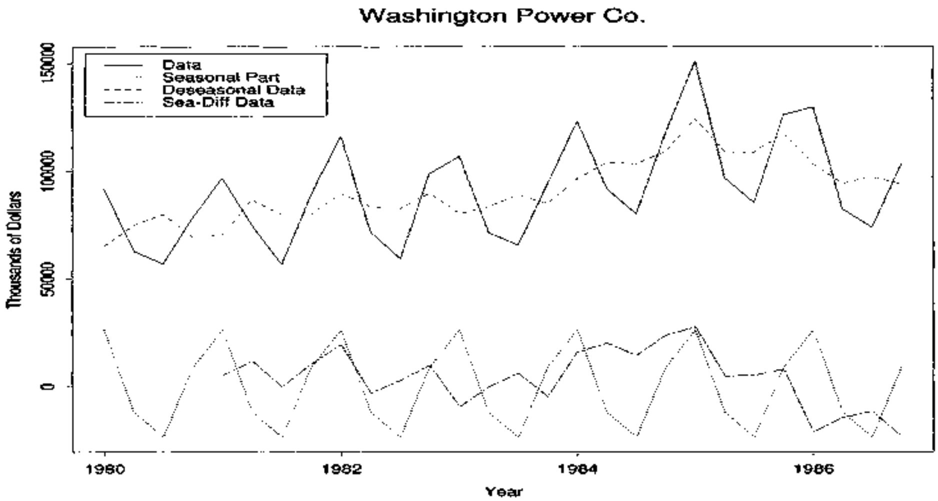

Fig. 1.5 SPLUS `stl` seasonal decomposition.

1.5 CONCLUSIONS

In this chapter we studied several descriptive methods to identify the macroscopic component (trend and seasonality) of a time series. Most of the time, these components can be identified and interpreted easily and there is no reason to fit unnecessarily complicated models to them. From now on we will assume that this preliminary data analysis step has been completed and we focus on analyzing the residual part N_t for microscopic structures. To accomplish this goal, we need to build more sophisticated models.

1.6 EXERCISES

1. (a) Show that a linear filter $\{a_j\}$ passes an arbitrary polynomial of degree k without distortion, that is,

$$m_t = \sum_j a_j m_{t-j},$$

for all kth-degree polynomials $m_t = c_0 + c_1 t + \cdots + c_k t^k$ if and only if

$$\sum_j a_j = 1, \quad \text{and} \quad \sum_j j^r a_j = 0 \ \text{ for } r = 1, \ldots, k.$$

(b) Show that the Spencer 15-point moving average filter does not distort a cubic trend.

2. If $m_t = \sum_{k=0}^{p} c_k t^k, t = 0, \pm 1, \ldots$, show that Δm_t is a polynomial of degree $(p-1)$ in t and hence $\Delta^{p+1} m_t = 0$.

3. In SPLUS, get hold of the yearly airline passenger data set by assigning it to an object. You can use the command

```
x<-rts(scan('airline.dat'),freq=12,start=1949)
```

The data are now stored in the object x, which forms the time series $\{X_t\}$. This data set consists of monthly totals (in thousands) of international airline passengers from January 1949 to December 1960 [details can be found in Brockwell and Davis (1991)]. It is stored under the file **airline.dat** on the Web page for this book.

(a) Do a time series plot of this data set. Are there any obvious trends?

(b) Is it necessary to transform the data? If a transformation is needed, what would you suggest?

(c) Do a yearly running median for this data set. Sketch the box plots for each year to detect any other trends.

(d) Find a trend estimate by using a moving average filter. Plot this trend.

(e) Estimate the seasonal component S_k, if any.

(f) Consider the deseasonalized data $d_t = X_t - \hat{S}_t, t = 1, \ldots, n$. Reestimate a trend from $\{d_t\}$ by applying a moving average filter to $\{d_t\}$; call it $\hat{m}_t$, say.

(g) Plot the residuals $r_t = X_t - \hat{m}_t - \hat{S}_t$. Does it look like a white noise sequence? If not, can you make any suggestions?