TIME SERIES ANALYSIS

Chapter 6: Forecasting

Forecasting the future values of an observed time series is an important problem in many areas including economics, production planning and sales forecasting and stock control.

1 Simple exponential smoothing

Exponential smoothing is the name given to the a general class of forecasting procedures that rely on simple updating equations to calculate forecasts. The most basic form, introduced here, is called **simple exponential smoothing**, but this should only be used for non-seasonal time series.

Given a non-seasonal time series, say $x_1, x_2, \cdots, x_N$, with no systematic trend, it is natural to forecast x_{N+1} by means of a weighted sum of the past observations

$$\hat{x}_N(1) = c_0 x_N + c_1 x_{N-1} + c_2 x_{N-2} + \cdots \tag{1.1}$$

where $\{c_i\}$ are weights. It seems sensible to give more weight to recent observations and less weight to observations further in the past. An intuitively appealing set of weights are geometric weights, which decrease by a constant ratio for every unit increase in the lag. Therefore in order that the weights sum to one we take

$$c_i = \alpha(1-\alpha)^i, \quad i = 0, 1, \cdots,$$

where α is a constant such that $0 < \alpha < 1$. Then equation (1.1) becomes

$$\begin{aligned}
\hat{x}_N(1) &= \alpha x_N + \alpha(1-\alpha)x_{N-1} + \alpha(1-\alpha)^2 x_{N-2} + \cdots \\
&= \alpha x_N + (1-\alpha)\Big[\alpha x_{N-1} + \alpha(1-\alpha)x_{N-2} + \cdots\Big] \\
&= \alpha x_N + (1-\alpha)\hat{x}_{N-1}(1).
\end{aligned}$$

If we set $\hat{x}_1(1) = x_1$, then the above equation can be used recursively to compute forecasts. Weights are exponentially decaying (hence the name). Choose α by minimizing squared one-step prediction error.

2 Holt's Linear Method

In Holt's Linear Method, we begin with describing the trend component, which decomposes into a level term and a growth term. For simplicity, we denote ℓ and g by the level term and the growth term, respectively. Also, we let T_h denote the forecast trend over the next h time periods. Using different combinations between the level term and the growth term, there are five future trend types. For example, Simple Exponential Smoothing method can be rewritten as

$$\begin{aligned} \ell_t &= \alpha x_t + (1-\alpha)\ell_{t-1} \\ T_h &= \ell_t \\ \hat{x}_{t+1} &= T_h. \end{aligned} \tag{2.1}$$

From (2.1), we see that there is only a level term in T_h, and it refers to one simplest trend type (None).

Holt extended Simple exponential smoothing to Linear Exponential Smoothing by using an additive trend type (Additive), i.e., $T_h = \ell + gh$. The forecast follows three equations:

$$\begin{aligned} \ell_t &= \alpha x_t + (1-\alpha)(\ell_{t-1} + g_{t-1}) \\ g_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta)g_{t-1} \\ \hat{x}_{t+h} &= \ell_t + g_t h, \end{aligned} \tag{2.2}$$

where α and β are two smoothing constants between 0 and 1. Compared with Simple Exponential Smoothing method, Holt's Linear Method considers a linear trend for the data.

3 Holt-Winters' Trend and Seasonality Method

Apart from considering a trend of data, seasonality is another important factor in time series analysis. There are two main types of time series with respect to seasonality:

(a) purely additive model

$$x = T + S + E, \tag{3.1}$$

(b) purely multiplicative model

$$x = T \times S \times E, \tag{3.2}$$

where T, S and E represent trend, seasonal and noise components, respectively.

Sometimes to make our data closer to stationary, we first remove the seasonal component. A seasonally adjusted series is one where the seasonal component has been removed, leaving only the trend and error components. In the additive model, the seasonally adjusted series is $x - S$. In the multiplicative model, the seasonally adjusted series is x/S.

Now, we introduce Holt-Winters' Trend and Seasonality (H-W) method with respect to these two models. The governing equations for the H-W method with respect to the additive model are given as follows:

$$\begin{aligned} \ell_t &= \alpha(x_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + g_{t-1}), \\ g_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta)g_{t-1}, \\ s_t &= \gamma(x_t - \ell_{t-1} - g_{t-1}) + (1+\gamma)s_{t-m}, \\ \hat{x}_{t+h} &= \ell_t + g_t h + s_{t+h-m(k+1)}, \end{aligned} \tag{3.3}$$

where k is the integer part of $(h-1)/m$, α, β, γ are three smoothing constants between 0 and 1, s_t is the seasonal component, m is the seasonal period and $h_m^+ = \lfloor (h-1) \mod m \rfloor + 1$. Similarly, based on (3.2), we also provide the H-W method with respect to the multiplicative model:

$$\begin{aligned} \ell_t &= \alpha(\tfrac{x_t}{s_{t-m}}) + (1-\alpha)(\ell_{t-1} + g_{t-1}), \\ g_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta)g_{t-1}, \\ s_t &= \gamma(\tfrac{x_t}{\ell_{t-1}+g_{t-1}}) + (1+\gamma)s_{t-m}, \\ \hat{x}_{t+h} &= (\ell_t + g_t h) \times s_{t+h-m(k+1)}. \end{aligned} \tag{3.4}$$

4 Real data analysis

We first consider the oil production in Saudi Arabia from 1996 to 2007. From Figure 1, we see that there is no obvious seasonal trend. Hence, for this dataset, we apply Simple Exponential Smooth method and Holt's Linear method to do the forecast. For Simple Exponential Smooth method, the smooth parameter α plays a significant role. To choose a best α, we we loop through α values from 0.01 to 0.99 and identify the level that minimizes our test RMSE. While for Holt's Linear method, the parameters can be selected by "holt" in R automatically with the minimum RMSE. Figure 2 displays the performances of these two methods.

Similarly, we also consider the data of Air passengers in Australia. From Figure 3, we also find that there is no obvious seasonal trend. Still, we apply Simple Exponential Smooth method and Holt's Linear method to do the forecast. Figure 4 shows the the performances of these two methods.

Now, we consider the dataset of international tourist visitor nights in Australia. From Figure 5, we see that there is an obvious seasonal trend, and hence we apply Holt-Winters' Trend and Seasonality Method. Moreover, the overall trend of the data is increasing with a slow rate. In such a case, one may use Holt-Winters' Trend and Seasonality Method with respect to the additive model (HWADD). Figure 6 demonstrates the forecasts for the period 2011Q1-2012Q4. For comparison, we also add the Holt-Winters' Trend and Seasonality Method

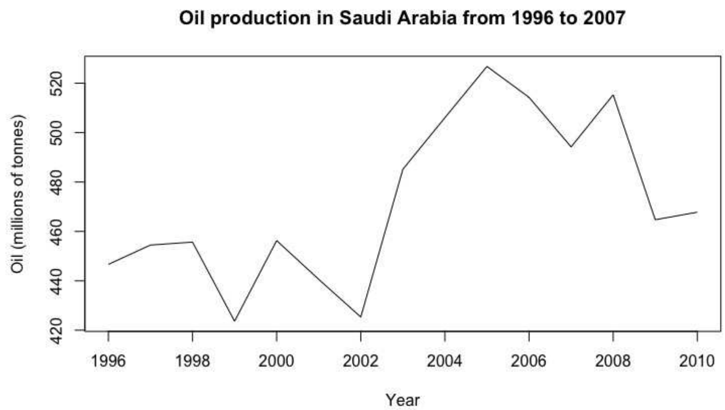


Figure 1: Fig 1

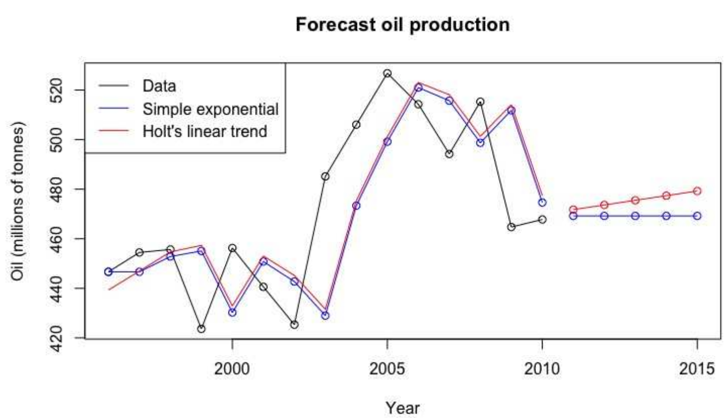


Figure 2: Fig 2

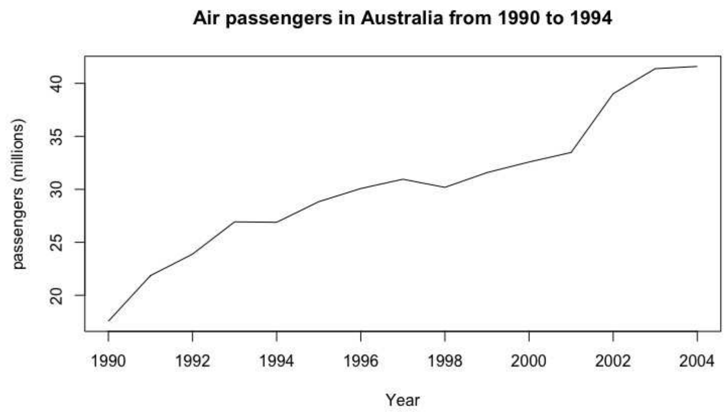


Figure 3: Fig 3

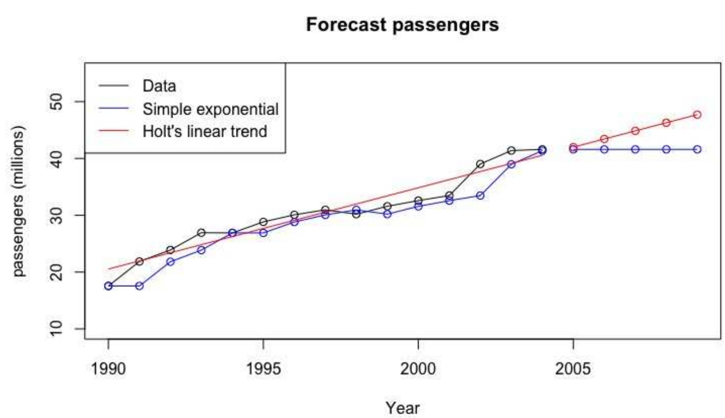


Figure 4: Fig 4

with respect to the multiplicative model (HWMUL). From Figure 4, we see that both methods performs similarly.

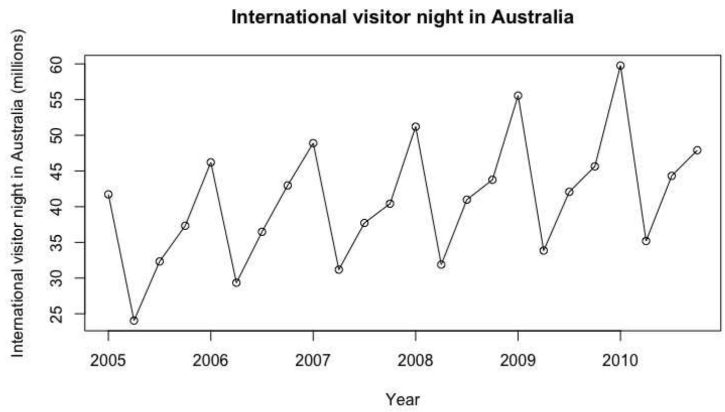


Figure 5: Fig 5

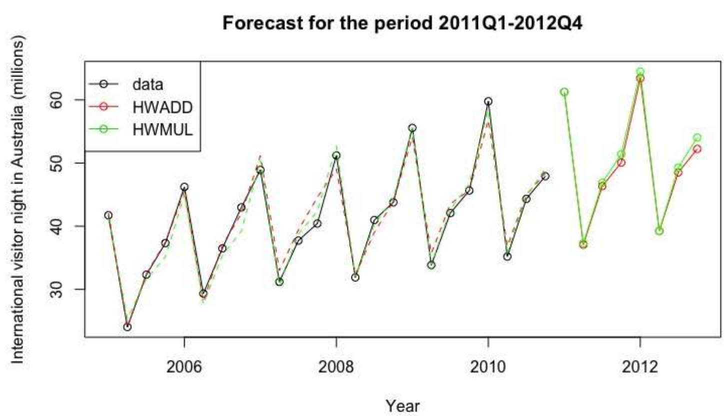


Figure 6: Fig 6

```
library(fpp)
oildata <- window(oil, start=1996)
plot(oildata, ylab="Oil (millions of tonnes)",xlab="Year",
main="Oil production in Saudi Arabia from 1996 to 2007")

###simple exponential method
###choose the best alpha
# identify optimal alpha parameter
alpha <- seq(.01, .99, by = .01)
```

```
RMSE <- NA
for(i in seq_along(alpha)) {
  fit <- ses(oildata, alpha = alpha[i], h = 10)
  RMSE[i] <- accuracy(fit)[2]
}
ind <- which( RMSE == min(RMSE))
alp=alpha[ind]
###use alp to forecast
fc1=ses(oildata, alpha=alp, initial="simple", h=5)
plot(fc1,PI=F, ylab="Oil (millions of tonnes)",
     xlab="Year", main="Forecast oil production", fcol="white", type="o")
lines(fitted(fc1), col="blue", type="o")
lines(fc1$mean, col="blue", type="o")


###holt's linear
##Within holt you can manually set the Îś and Îš.
##However, if you leave those parameters as NULL,
##the holt function will actually identify the optimal model parameters
##by minimizing AIC and BIC values.
fc2=holt(oildata,h=5)
fc2$model

###AIC      AICc      BIC
###148.4423 155.1090 151.9826
lines(fitted(fc2), col="red")
lines(fc2$mean, col="red", type="o")

legend("topleft", lty=1, col=c("black","blue","red"),
       c("Data","Simple exponential","Holt's linear trend"))

aust <- window(austourists,start=2005)
fit1 <- hw(aust,seasonal="additive")
fit2 <- hw(aust,seasonal="multiplicative")
plot(aust,ylab="International visitor night in Australia (millions)", xlab="Ye
     type="o",main="International visitor night in Australia")
plot(fit2,ylab="International visitor night in Australia (millions)",
     PI=FALSE, type="o", fcol="white", xlab="Year",
     main="Forecast for the period 2011Q1-2012Q4")
lines(fitted(fit1), col="red", lty=2)
lines(fitted(fit2), col="green", lty=2)
lines(fit1$mean, type="o", col="red")
lines(fit2$mean, type="o", col="green")
legend("topleft",lty=1, pch=1, col=1:3,
```

```
c("data","HWADD","HWMUL"))
```

5 Minimum Mean Square Error Prediction

Suppose Y is a random variable with mean $\mu = EY$ and variance σ^2. If our object is to predict Y using only a constant c, what is the best choice for c? Clearly, we must first define best. A common (and convenient) criterion is to choose c to minimize the mean square error of prediction, that is, to minimize

$$g(c) = E(Y - c)^2.$$

Theorem 1 The minimum of $E(Y - c)^2$ is obtained when $c = EY$.

Proof If we expand $g(c)$, we have

$$g(c) = E(Y)^2 - 2cEY + c^2.$$

Since $g(c)$ is quadratic in c and opens upward, solving $g'(c) = 0$ will produce the required minimum. Note that

$$g'(c) = -2EY + 2c = 0$$

so that the optimal c is

$$c = EY.$$

And

$$\min_c g(c) = E(Y - EY)^2.$$

Now consider the situation where a second random variable X is available and we wish to use the observed value of X to help predict Y. Again, our criterion will be to minimize the mean square error of prediction. We need to choose the function $h(X)$, say, that minimizes

$$E(Y - h(X))^2.$$

Rewrite

$$E(Y - h(X))^2 = E\Big[E\Big((Y - h(X))^2|X\Big)\Big].$$

The inner conditional expectation can be written as

$$E\Big((Y - h(X))^2|X = x\Big) = E\Big((Y - h(x))^2|X = x\Big).$$

For each value of x, $h(x)$ is a constant and hence we can apply Theorem 1 to the conditional distribution of Y given $X = x$. Thus the best choice of $h(x)$ is

$$h(x) = E(Y|X = x).$$

It follows that $h(X) = E(Y|X)$ is the best predictor of Y of all functions of X.

Theorem 2 The minimum of $E(Y - g(X))^2$ is obtained when $g(X) = E(Y|X)$.

Based on the available history of the series up to time t, namely $Y_1, Y_2, \cdots, Y_{t-1}, Y_t$, we would like to forecast the value of Y_{t+l} that will occur l time units into the future. In view of Theorem 2 the minimum mean square error forecast is given by

$$E(Y_{t+l}|Y_1, Y_2, \cdots, Y_t). \tag{5.1}$$

When evaluating the conditional expectation we use the fact that the conditional expectation of Z_{N+h} given data up to time N, is zero for all $h > 0$.

6 Prediction using AR model

Suppose we fit an AR(p) model to data $X_1, X_2, \cdots, X_n$ and the estimate of the model is

$$\hat{X}_t = \hat{\delta} + \hat{\phi}_1 X_{t-1} + \cdots + \hat{\phi}_p X_{t-p}$$

where $\hat{\delta} = (1 - \hat{\phi}_1 - \cdots - \hat{\phi}_p)\hat{\mu}$.]

Point prediction To predict X_{n+1}: use (5.1) to obtain

$$E(X_{n+1}|X_1, \cdots, X_n) = E\Big(\delta + \phi_1 X_n + \cdots + \phi_p X_{n+1-p} + Z_{n+1}|X_1, \cdots, X_n\Big)$$

$$= \delta + \phi_1 X_n + \cdots + \phi_p X_{n+1-p} + E\Big(Z_{n+1}|X_1, \cdots, X_n\Big)$$

$$= \delta + \phi_1 X_n + \cdots + \phi_p X_{n+1-p}.$$

It follows that

$$\hat{X}_{n+1} = \hat{\delta} + \hat{\phi}_1 X_n + \cdots + \hat{\phi}_p X_{n-p+1}.$$

Likewise, to predict X_{n+2}: use

$$\hat{X}_{n+2} = \hat{\delta} + \hat{\phi}_1 \hat{X}_{n+1} + \hat{\phi}_n X_n + \cdots + \hat{\phi}_p X_{n-p+2}.$$

Continue the procedure to make any τ step ahead prediction $\hat{X}_{n+\tau}$.

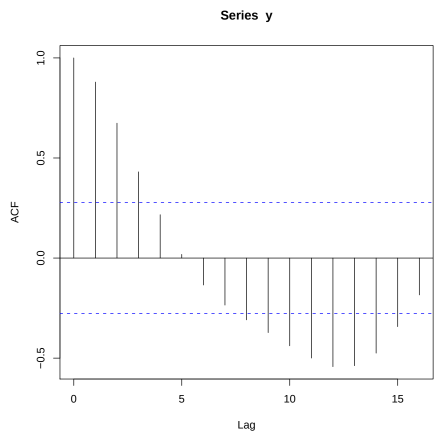
Prediction intervals with 95% confidence

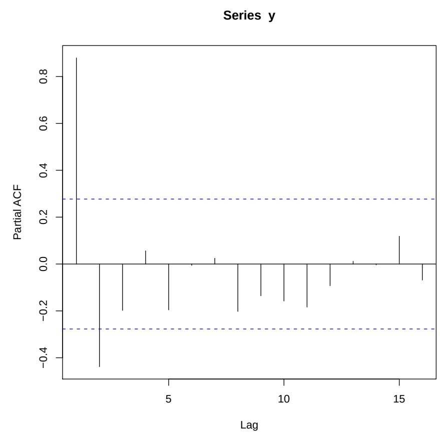
$$[\hat{y}_{n+\tau} - 1.96 s_{n+\tau}(n), \hat{y}_{n+\tau} + 1.96 s_{n+\tau}(n)]$$

Here $s_{n+\tau}(n)$ is the *standard error of the forecast error*. The calculation of $s_{n+\tau}(n)$ is beyond the scope of this module. However, $\mathbb{S}AS(orR)$ can help us to calculate it.

Example 1 y_t: 3.91, 3.86, 3.81, 3.02, 2.62, 1.89, -1.13, -3.82, -5.08, -4.42, -1.99, 0.70, 1.86, 2.98, 1.78, 3.01, 2.13, 3.23, 3.17, 4.64, 5.20, 6.76, 5.79, 5.08, 1.88, -0.72, -2.00, -3.03, -2.35, -3.34, -3.21, -3.57, -4.28, -3.54, -3.16, -1.41, 0.48, 1.61, 2.42, 2.11, 2.45, 1.39, 2.04, 1.71, 3.26, 3.20, 1.43, 1.68, 4.17, 4.75

In $\mathbb{R}$, using acf(y) and pacf(y) can produce the following plots.





Because the SPACF has a clear cut-off after 2, we can choose AR(2) model for y_t.

fit = arima(y, order=c(2,0,0))

fit

Call: arima(x = x, order = c(2, 0, 0))

Coefficients:

	ar1	ar2	intercept
	1.3734	-0.5233	1.3621
s.e.	0.1171	0.1187	1.0352

sigma^2 estimated as 1.272: log likelihood = -78.13, aic = 164.26

The estimates are $\hat{\mu} = 1.3621, \hat{\phi}_1 = 1.3734, \hat{\phi}_2 = -0.5233$. Thus

$$\hat{\delta} = 1.3621 * (1 - 1.3734 - (-0.5233)) = 0.2042$$

The estimated model is

$$\hat{y}_t = 0.2042 + \underset{(0.1171)}{1.3734 y_{t-1}} - \underset{(0.1187)}{0.5233 y_{t-2}}$$

Fitted values

$$\hat{y}_3 = 0.2042 + 1.3734 y_2 - 0.5233 y_1 = 3.46$$
$$\hat{y}_4 = 0.2042 + 1.3734 y_3 - 0.5233 y_2 = 3.42$$
$$\ldots$$
$$\hat{y}_{50} = 0.2042 + 1.3734 y_{49} - 0.5233 y_{48} = 5.05.$$

Residuals

$$
\begin{aligned}
e_3 &= y_3 - \hat{y}_3 = 0.35 \\
e_4 &= y_4 - \hat{y}_4 = -0.40 \\
&... \\
e_{50} &= y_{50} - \hat{y}_{50} = -0.30
\end{aligned}
$$

Predict

$$
\begin{aligned}
\hat{y}_{51} &= 0.2042 + 1.3734 y_{50} - 0.5233 y_{49} = 4.55 \\
\hat{y}_{52} &= 0.2042 + 1.3734 \hat{y}_{51} - 0.5233 y_{50} = 3.96 \\
\hat{y}_{53} &= 0.2042 + 1.3734 \hat{y}_{52} - 0.5233 \hat{y}_{51} = 3.27 \\
\hat{y}_{54} &= 0.2042 + 1.3734 \hat{y}_{53} - 0.5233 \hat{y}_{52} = 2.62 \\
\hat{y}_{55} &= 0.2042 + 1.3734 \hat{y}_{54} - 0.5233 \hat{y}_{53} = 2.09
\end{aligned}
$$

$\mathbb{R}$:

```
predict(fit, n.ahead=5)
predict(fit, n.ahead=5)
$pred
Time Series:
Start = 51
End = 55
Frequency = 1
[1] 4.545503 3.961132 3.265595 2.616178 2.088275
$se
Time Series:
Start = 51
End = 55
Frequency = 1
[1] 1.128049 1.916399 2.456816 2.779828 2.948068
```

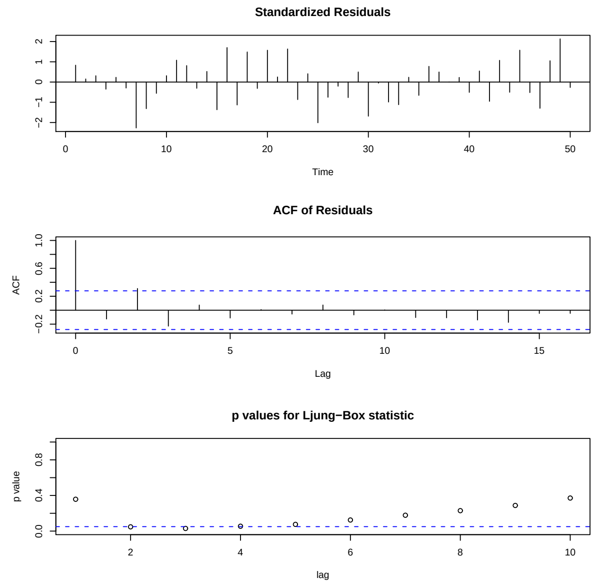
Predict intervals with 95% confidence

$$
\begin{aligned}
y_{51} &: [4.55 - 1.96 * 1.13, 4.55 + 1.96 * 1.13] = [2.3352, 6.7648] \\
y_{52} &: [3.96 - 1.96 * 1.92, 3.96 + 1.96 * 1.92] = [0.1968, 7.7232] \\
y_{53} &: [3.27 - 1.96 * 2.46, 3.27 + 1.96 * 2.46] = [-1.5516, 8.0916] \\
y_{54} &: [2.62 - 1.96 * 2.78, 2.62 + 1.96 * 2.78] = [-2.8288, 8.0688] \\
y_{55} &: [2.09 - 1.96 * 2.94, 2.09 + 1.96 * 2.94] = [-3.6724, 7.8524]
\end{aligned}
$$

Further topics: (diagnostic) checking of the model (we will discuss this later)
Check whether there is autocorrelation in the residuals.

$\mathbb{R}$:

```
tsdiag(fit)
```



7 Seasonal ARIMA models

In practice, it may not be reasonable to assume that the seasonality component repeats itself precisely in the same way cycle after cycle. Seasonal ARIMA models to be introduced allow for randomness in the seasonal pattern from one cycle to the next.

Suppose that we have r years of monthly data which we tabulate as follows:

Month				
Year	1	2	$\cdots$	12
1	X_1	X_2	$\cdots$	X_{12}
2	X_{13}	X_{14}	$\cdots$	X_{24}
3	X_{25}	X_{26}	$\cdots$	X_{36}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r	$X_{1+12(r-1)}$	$X_{2+12(r-1)}$	$\cdots$	$X_{12+12(r-1)}$

Each column may be viewed as a time series. Suppose that each one of these time series is generated by the same ARMA(P,Q) model. Then at the j–th month and year t, we have

$$\begin{aligned} X_{j+12t} &= \Phi_1 X_{j+12(t-1)} + \cdots + \Phi_P X_{j+12(t-P)} + U_{j+12t} \\ &+ \Theta_1 U_{j+12(t-1)} + \cdots + \Theta_Q U_{j+12(t-Q)}, \end{aligned} \tag{7.1}$$

for $j = 1, 2, \cdots, 12$ and $t = 0, 1, \cdots, r-1$, where $\{U_s\} \sim WN(0, \sigma_U^2)$.

Since the same ARMA(P,Q) model is assumed to apply to each month, model (7.1) can be written for all t as

$$\begin{aligned} X_t &= \Phi_1 X_{t-12} + \cdots + \Phi_P X_{t-12P} + U_t \\ &+ \Theta_1 U_{t-12} + \cdots + \Theta_Q U_{t-12Q}, \end{aligned} \tag{7.2}$$

which is equivalent to

$$\Phi(B^{12}) X_t = \Theta(B^{12}) U_t. \tag{7.3}$$

Since it is likely that the 12 series corresponding to the different months are correlated. To incorporate dependence between these series we assume that $\{U_t\}$ follows an ARMA(p,q) model of the form

$$\phi(B) U_t = \theta(B) Z_t, \tag{7.4}$$

where $Z_t \sim WN(0, \sigma^2)$.

Combining models (7.3) and (7.4) implies

$$\phi(B)\Phi(B^{12}) X_t = \theta(B)\Theta(B^{12}) Z_t, \tag{7.5}$$

which motives us to introduce a definition for general seasonal ARIMA models.

Definition 1 : Consider a general seasonal ARIMA (SARIMA) model of the form

$$\phi_p(B)\Phi_P(B^s) Y_t = \theta_q(B)\Theta_Q(B^s) Z_t, \tag{7.6}$$

where B denotes the backward shift operator, ϕ_p, Φ_P, θ_q and Θ_Q are polynomials of order p, P, q and Q, respectively, $\{Z_t\} \sim WN(0, \sigma^2)$, and

$$Y_t = \nabla^d \nabla_s^D X_t = (1-B)^d (1-B^s)^D X_t$$

denotes the differenced series. This model is called a SARIMA model of order $(p,d,q) \times (P,D,Q)_s$ for $\{X_t\}$.

Remark 1 : Note that ∇^d is used to remove the trend while ∇_s^D is to remove seasonality. ■

Note also that

$$\{Y_t\} \sim \text{ARMA}(p + sP, q + sQ)$$

and

$$W_t = \nabla_s^D X_t \sim \text{ARIMA}(p,d,q).$$

When there is no seasonality ($s = D = 0$), $X_t \sim \text{ARIMA}(p,d,q)$.

EXAMPLE 2 : SARIMA models look rather complicated at first sight, so let us consider some simple cases.

Note that when $d = D = 1$ and $s = 12$, $\{Y_t\}$ becomes

$$\begin{aligned} Y_t &= \bigtriangledown \bigtriangledown_{12} X_t = \bigtriangledown_{12} X_t - \bigtriangledown_{12} X_{t-1} \\ &= (X_t - X_{t-12}) - (X_{t-1} - X_{t-13}). \end{aligned}$$

Also consider a SARIMA model of order $(1, 0, 0) \times (0, 1, 1)_{12}$. This means that we have one non–seasonal AR term, one seasonal MA term and one seasonal difference. We then may simplify (7.6) as

$$(1 - \alpha B)Y_t = (1 + \Theta B^{12})Z_t, \tag{7.7}$$

where $Y_t = \bigtriangledown_{12} X_t$. This implies that model (7.7) is equivalent to

$$X_t - X_{t-12} - \alpha(X_{t-1} - X_{t-13}) = Z_t + \Theta Z_{t-12}. \tag{7.8}$$

■

Implementation:

General guidelines for identifying model SARIMA(p,d,q,P,D,Q,s) are as follows.

(a) Identify d and D so as to make the differenced observations

$$Y_t = \bigtriangledown^d \bigtriangledown_s^D X_t = (1 - B)^d (1 - B^s)^D X_t$$

stationary in appearance.

(b) Examine the SACF and SPACF of Y_t at lags which are multiples of s in order to identify the orders of P and Q in the model. For example, P and Q should be chosen so that $r(ks), k = 1, 2, \cdots$ is compatible with the ACF of ARMA(P,Q).

(c) Examine the SACF and SPACF of Y_t at lags $1, \cdots, s - 1$. The p and q should be chosen so that they are compatible with ARMA(p,q).

(d) Ultimately, the AIC and diagnostic checking are used to identify the best SARIMA model among competing alternatives.

Typical R codes are the following.

```
> fit = arima(y, order = c(p,d,q), seasonal = list(order=c(P,D,Q), period=s))
```

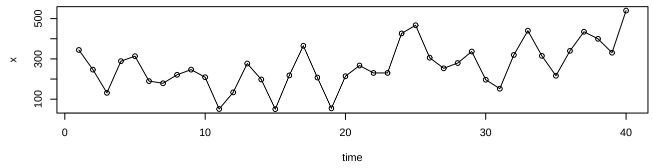
e.g.

```
> tsdiag(fit)
```

e.g.

```
> predict(fit, n.ahead=10)
```

EXAMPLE 3 Quarterly Propane gas bills for the Farmer's Bureau Co-op X_t: 344.39, 246.63, 131.53, 288.87, 313.45, 189.76, 179.1, 221.1, 246.84, 209, 51.21, 133.89, 277.01, 197.98, 50.68, 218.08, 365.1, 207.51, 54.63, 214.09, 267, 230.28, 230.32, 426.41, 467.06, 306.03, 253.23, 279.46, 336.56, 196.67, 152.15, 319.67, 440, 315.04, 216.42, 339.78, 434.66, 399.66, 330.8, 539.78. ■



In this data, the seasonal period is s = 4.

Figure 7, the acf of X_t dies down very slowly. We need to take differences.

$$u_t = X_t - X_{t-1}$$

and

$$v_t = X_t - X_{t-4}.$$

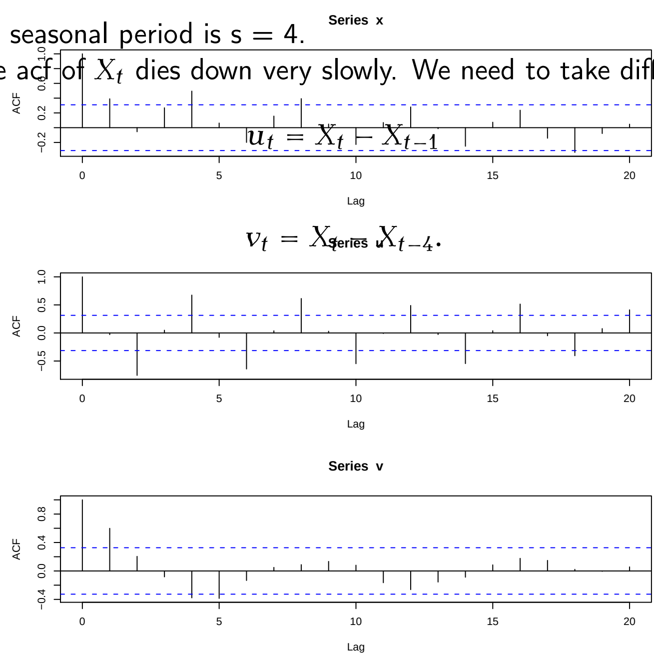


Figure 7:

The acf of u_t dies down slowly and thus it is not stationary.

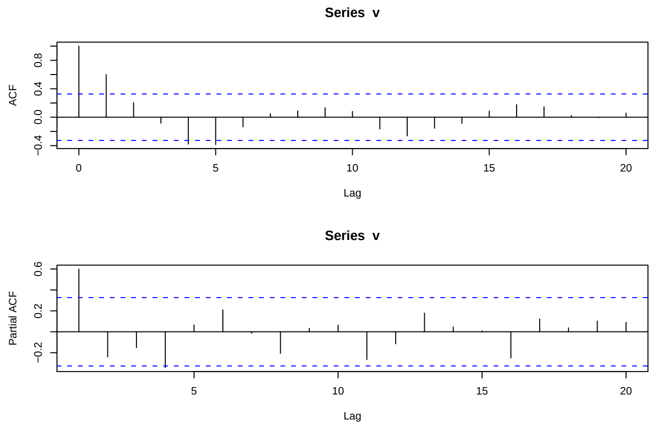


Figure 8:

The acf of v_t dies down quickly. Based on figure 8, we can try models

$$\text{model A:} \quad v_t = (1 + \theta_1 B)(1 + \theta_4 B^4) Z_t$$

or

$$\text{model B: :} \quad (1 + \phi_1 B) v_t = (1 + \theta_4 B^4) Z_t.$$

i.e.

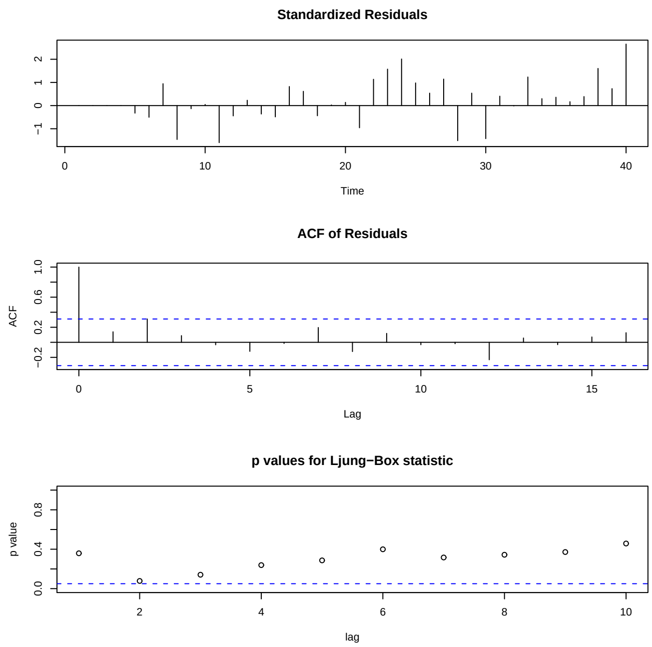
$$\text{model A: SARIMA(0,0,1, 0,1,1, 4) :}$$
$$(1 - B^4) X_t = (1 + \theta_1 B)(1 + \theta_4 B^4) Z_t$$

or

$$\text{model B: SARIMA(1,0,0, 0,1,1, 4) :}$$
$$(1 + \phi_1 B)(1 - B^4) X_t = (1 + \theta_4 B^4) Z_t.$$

Model A:
Try model SARIMA(0,0,1, 0,1,1, 4).

```
> fita1 = arima(x, order=c(0,0,1), seasonal=list(order=c(0,1,1),period=4))
> tsdiag(fita1)
```



Try model SARIMA(0,0,2, 0,1,1, 4).

> fita2 = arima(x, order=c(0,0,2), seasonal=list(order=c(0,1,1),period=4))

> tsdiag(fita2)

Write down the fitted model.

> fita2

Call: arima(x = x, order = c(0, 0, 2), seasonal = list(order = c(0, 1, 1), period = 4))

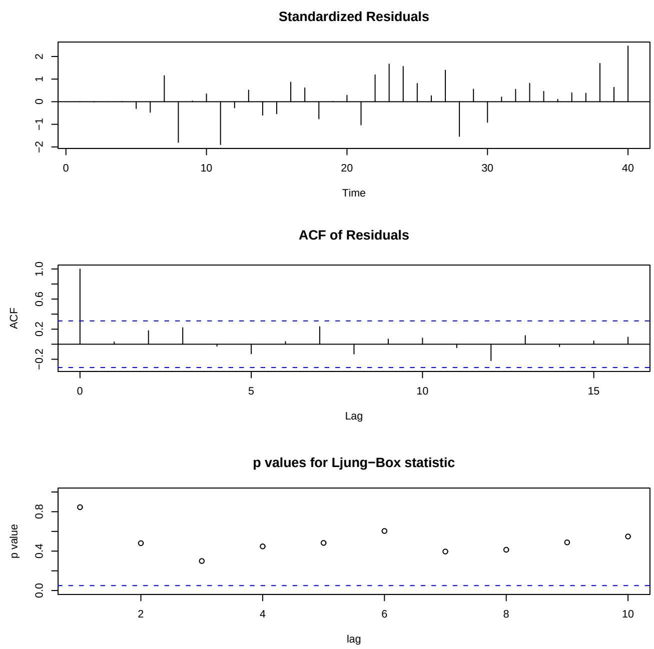
Coefficients:

	ma1	ma2	sma1
	0.8810	0.2739	-0.6132
s.e.	0.1756	0.1662	0.1874

σ^2 estimated as 4130: log likelihood = -202.24, aic = 412.47

The fitted model is

$$(1 - B^4)X_t = (1 + 0.881B + 0.2739B^2)(1 - 0.6132B^4)Z_t.$$


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung−Box statistic
p value
lag