Statistics 435/535/711 30 March 2022
P. Shaman

Differencing Time Series for ARIMA Fitting

In part D of the 16 March class notes (page 8), the use of differencing to achieve a stationary ARMA model is discussed. If a given time series y_t is not stationary and if the d th-order difference,

$$w_t = (1 - \mathrm{B})^d y_t ,$$

is a stationary ARMA(p,q) process, we say y_t is an ARIMA(p,d,q) process.

A time series which includes a trend component is not stationary. Many series which include a trend can be approximated by an ARMA(p,q) model with small p and q after a suitable amount of differencing. For such series, as noted previously, one chooses the amount of differencing d by inspecting the plot of the data vs. time, the sample acf, and the sample pacf. When differencing is needed, the sample acf typically shows a smooth pattern which does not tail off or cut off rapidly as the lag increases. Also, a value near 1 for the lag 1 sample partial autocorrelation is a signal that differencing is required. After a proper amount of differencing, the plot of the data vs. time will show absence of trend, and the sample acf and sample pacf will show nonsmooth patterns which tail off or cut off rapidly. When differencing is effectively used with data series, most commonly d = 1; occasionally one encounters the need for d = 2, and use of $d > 2$ almost never occurs. This use of differencing will work effectively for many economic and financial time series, but not for all. However, such differencing and subsequent use of an ARMA(p,q) model are sometimes not effective for time series arising in physical science and medical applications.

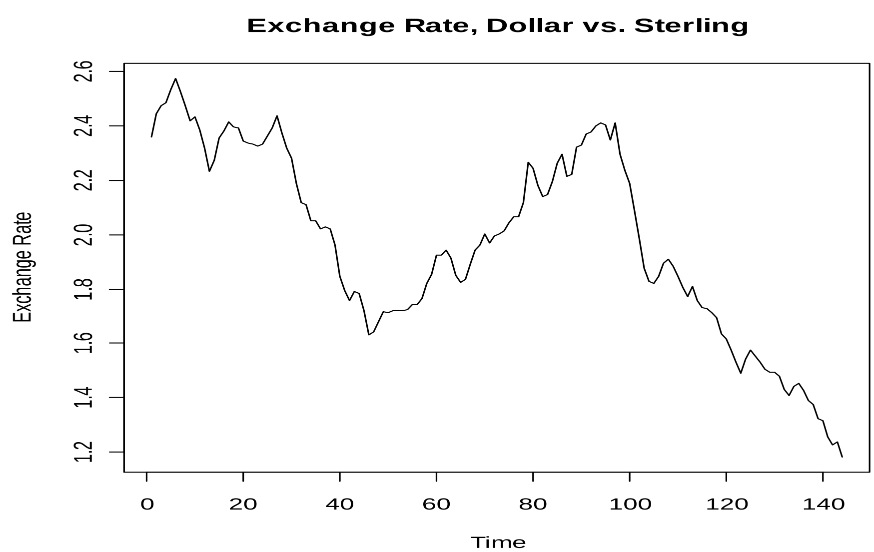
It should be noted that while differencing may eliminate a trend from a time series, it does change the structure of the other components of the series. Thus, a seasonal component, when present, and the remaining structure are altered by differencing. Specifically, differencing magnifies high frequency fluctuations and attenuates low frequency fluctuations. This can be observed by comparison of the spectrum after differencing to the spectrum after trend removal but without differencing.

In use of a statistical program which supports ARIMA analysis, one can either input the original data series to the program and fit an ARIMA(p,d,q) model, or one can input the dth difference of the data series to the program and fit an ARMA(p,q) model. If one wants the program to supply forecasts for the original data series, then the original data series should be input to the program. Of course, if one inputs the differenced series to the program, the program will supply forecasts for the differenced series.

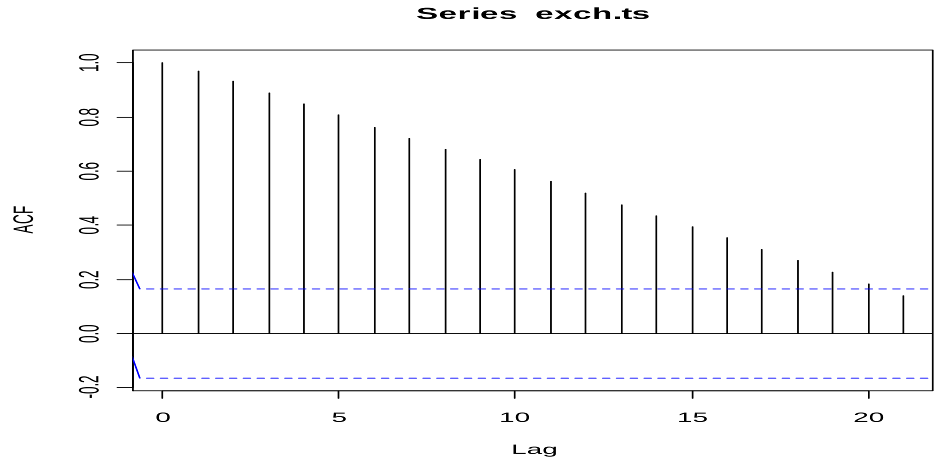
In some previous examples (U.S. quarterly GNP, CRSP equal-weighted index, FTA All Share index) we have fit an ARMA(p,q) model to the log return series. This is, of course, the same as fitting an ARIMA(p,1,q) model to the log series, and the log transformation in the examples is used to adjust for the increase in the volatility of the series as the level rises. For many time series, differencing is employed without first logging the data, which is appropriate when the volatility is stable over time. When one analyzes the difference of the logged series, one is essentially considering the percentage changes from one time to the next, and when one analyzes the difference of the series, one is considering the changes from one time to the next.

The plot directly below shows the U.S. dollar-U.K. sterling exchange rate, in dollars per pound, for 144 consecutive months. The data are in Exchange.txt.

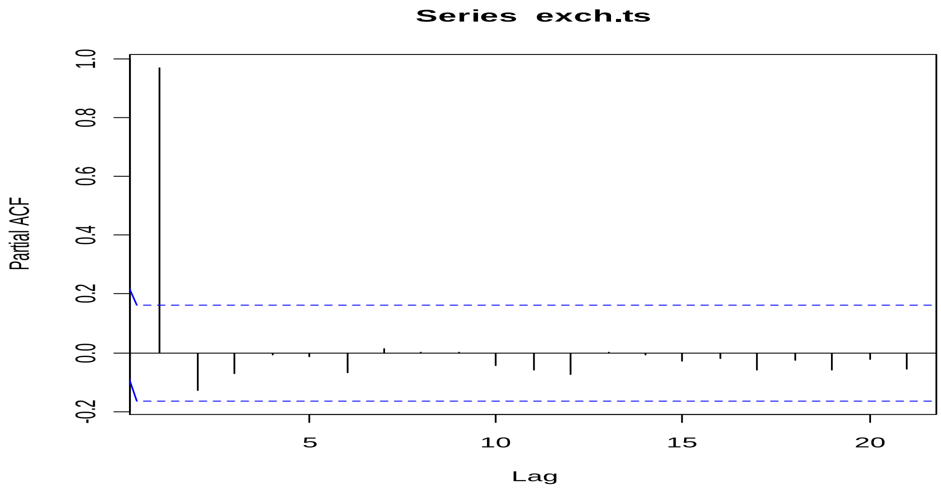
```
> exch<-read.csv("F:/Stat71122Spring/Exchange.txt")
> attach(exch)
> head(exch)
  exchange
1    2.359
2    2.446
3    2.474
4    2.485
5    2.533
6    2.574

> exch.ts<-ts(exchange)
> plot(exch.ts,xlab="Time",ylab="Exchange Rate",main="Exchange Rate,
Dollar vs. Sterling")
```



```
> acf(exch.ts)
```



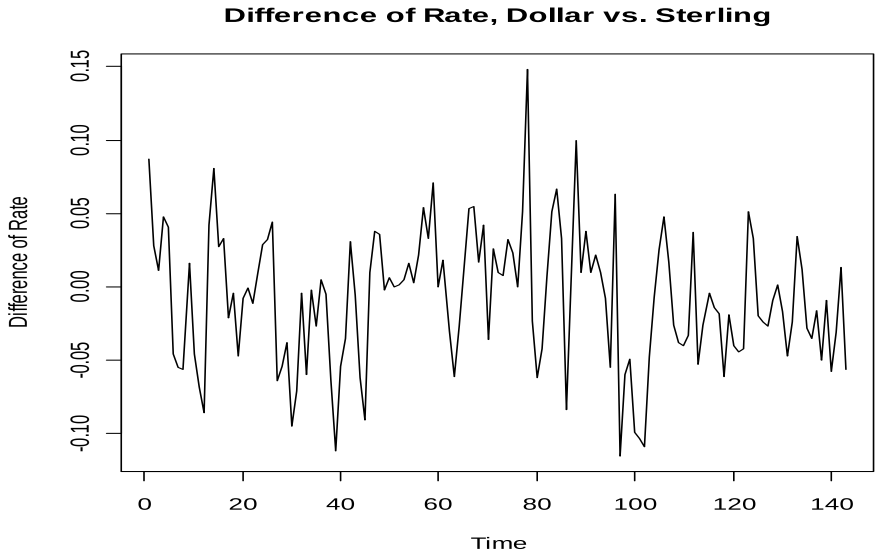

```
> pacf(exch.ts)
```



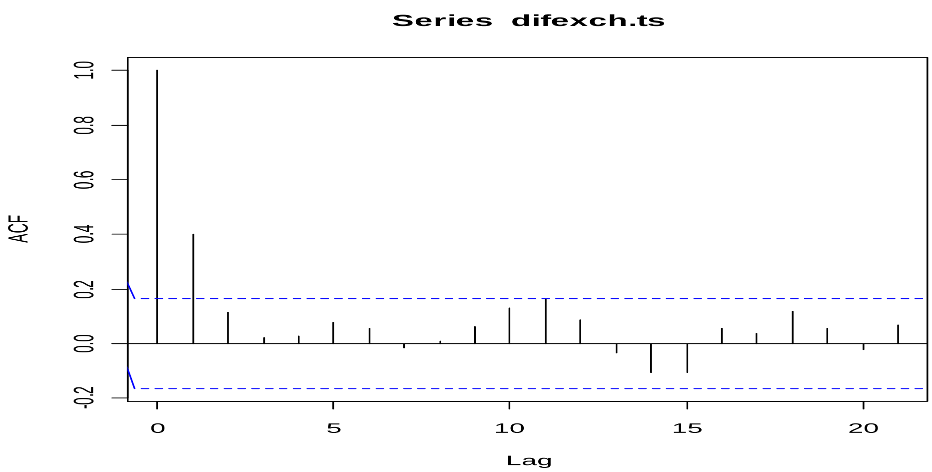
The smooth pattern of the sample acf and the large value of the lag 1 sample pacf value show differencing is needed. Of course, the plot of the data also indicates we should difference, because it shows presence of a trend.

Here is the corresponding output after first-order differencing.

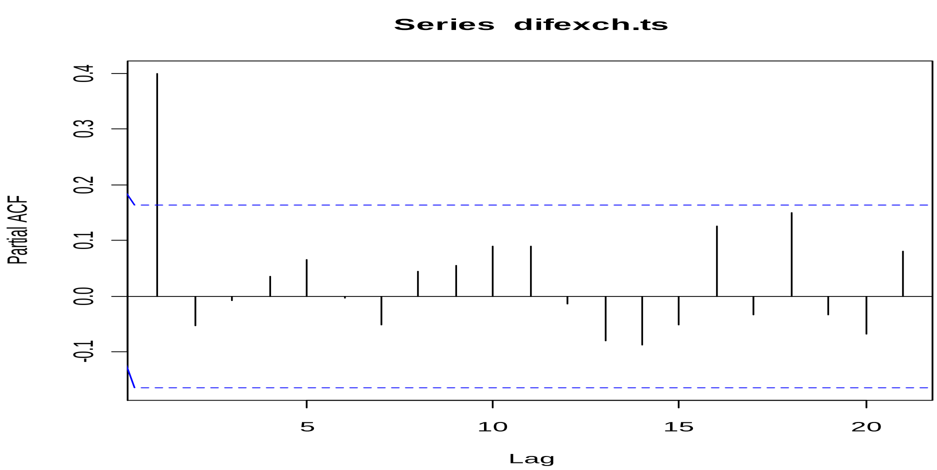
```
> difexch.ts<-ts(diff(exchange))
> plot(difexch.ts,xlab="Time",ylab="Difference of
Rate",main="Difference of Rate, Dollar vs. Sterling")
```




```
> acf(difexch.ts)
```




```
> pacf(difexch.ts)
```

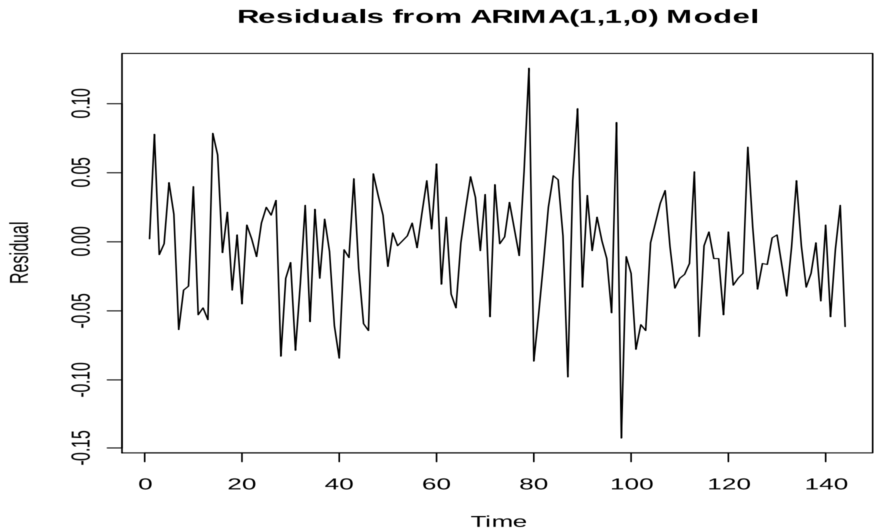


The differencing has produced a stationary series, for which AR(1) and MA(1) fits are reasonable selections. That is, the original series is fit well by ARIMA(1,1,0) and ARIMA(0,1,1) models.

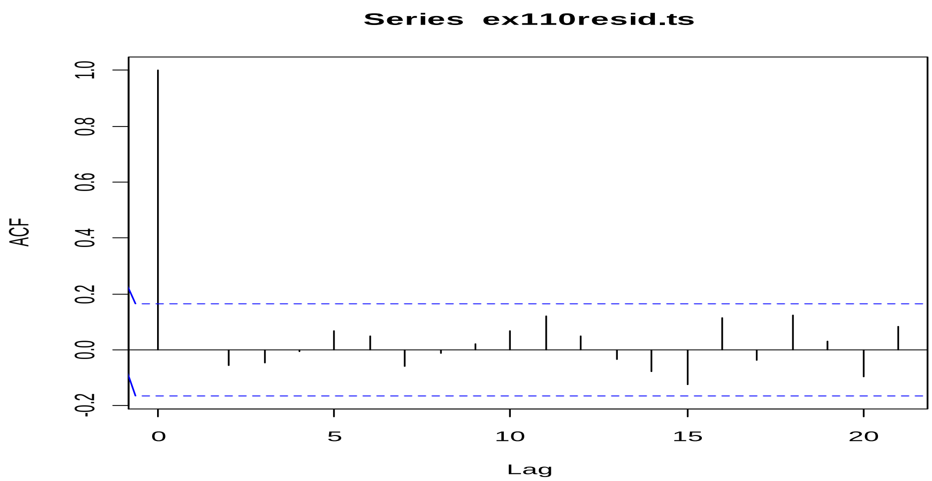
```
> exarima110<-arima(exch.ts,order=c(1,1,0))
> exarima110

Call:
arima(x = exch.ts, order = c(1, 1, 0))

Coefficients:
         ar1
      0.4317
s.e.  0.0765

sigma^2 estimated as 0.001756:  log likelihood = 250.65,  aic = -497.3

> coeftest(exarima110)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1 0.431679   0.076527  5.6409 1.692e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> exarima011<-arima(exch.ts,order=c(0,1,1))
> exarima011

Call:
arima(x = exch.ts, order = c(0, 1, 1))

Coefficients:
         ma1
      0.3982
s.e.  0.0693

sigma^2 estimated as 0.001791:  log likelihood = 249.25,  aic = -494.51

> coeftest(exarima011)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ma1 0.398225   0.069336  5.7434 9.277e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

AIC favors the ARIMA(1,1,0) model fit. Here are its residual diagnostics:

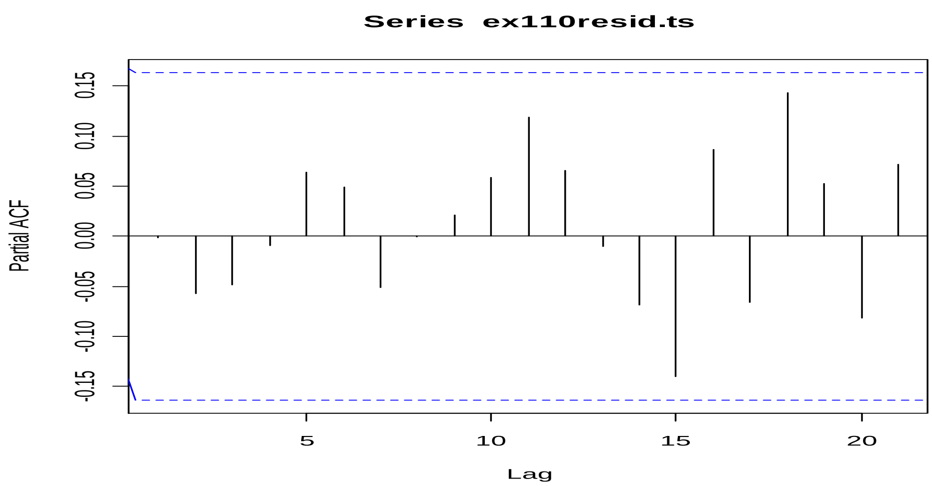
```
> ex110resid.ts<-ts(resid(exarima110))
> plot(ex110resid.ts,xlab="Time",ylab="Residual",main="Residuals from
ARIMA(1,1,0) Model")
```




```
> acf(ex110resid.ts)
```

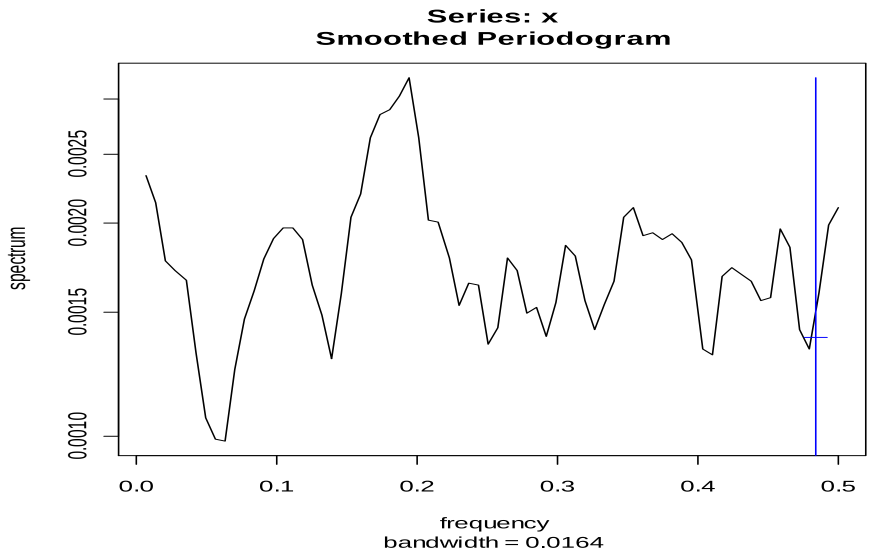



```
> pacf(ex110resid.ts)
```



According to these diagnostics, there is reduction to white noise. The Kolmogorov–Smirnov statistic and residual spectral density plot follow.

```
> spectrum(ex110resid.ts,span=8)
```




```
> library("hwwntest")
> bartlettB.test(ex110resid.ts)

        Bartlett B Test for white noise

data:
= 0.58744, p-value = 0.8805
```

The residual spectral density plot contains a peak near frequency 0.2. However, the standard error indicator is a long segment. We do not reject the white noise hypothesis.

Let's write the form of the fitted ARIMA(1, 1,0) model and interpret it. The model is

$$(1-0.4317B)(1-B)y_t=\varepsilon_t,$$

or

$$\left(1-1.4317\mathrm{B}+0.4317\mathrm{B}^2\right)y_t=\varepsilon_t,$$

or

$$y_t=1.4317y_{t-1}-0.4317y_{t-2}+\varepsilon_t.$$

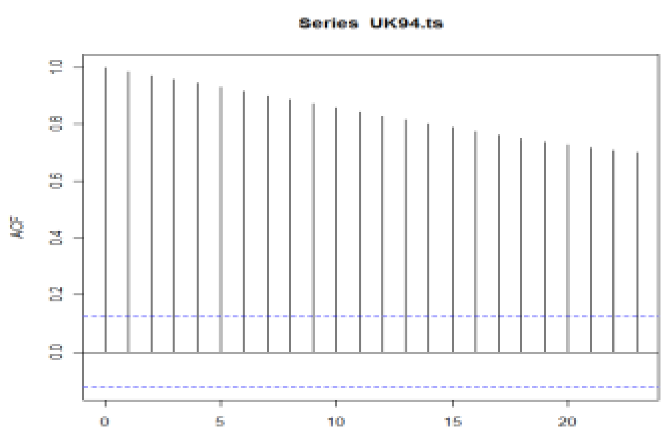
That is, the exchange rate this month is modeled as 1.43 times the rate in the previous month, minus 0.43 times the rate two months previously, plus a new random shock in the present month. Thus, one uses a linear combination of the rates in the last two months to forecast the rate in the next month.

Let us now turn to another data set of exchange rates, the daily time series of the U.S. dollar-U.K.sterling exchange rate, in dollars per pound, for 1994-2000. The data are in forex9401.txt. Note there are no readings for Saturdays, Sundays, and holidays.

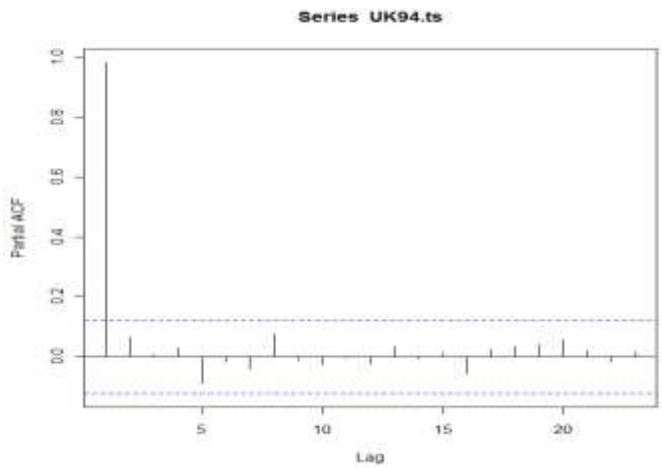
We begin with the daily series of the exchange rate for the year 1994.

```
> forex<-read.csv("F:/Stat71122Spring/forex9401.txt")
> attach(forex)
> head(forex)
        Date Canadiandollar   Mark UKpound    Yen Frenchfranc
1 01/03/1994         1.3155 1.7420  1.4760 112.50       5.928
2 01/04/1994         1.3173 1.7380  1.4840 112.75       5.905
3 01/05/1994         1.3180 1.7410  1.4870 113.10       5.907
4 01/06/1994         1.3215 1.7425  1.4855 112.75       5.920
5 01/07/1994         1.3235 1.7317  1.4900 111.97       5.891
6 01/10/1994         1.3181 1.7343  1.4923 112.38       5.911

> UKpound94<-UKpound[1:251]
> UK94.ts<-ts(UKpound94)
```


```
> acf(UK94.ts)
```

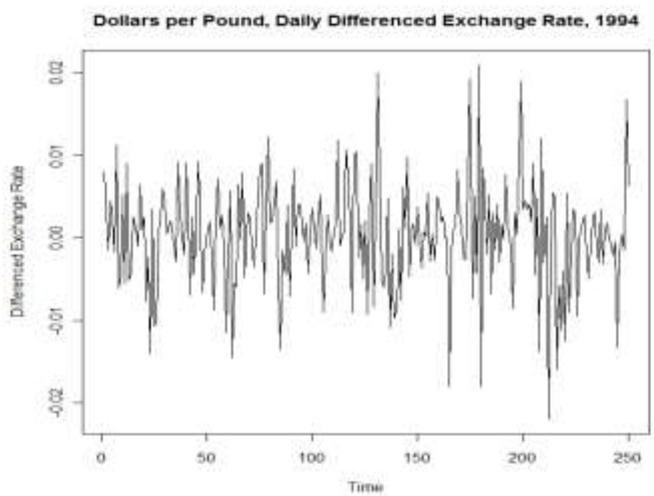


```
> pacf(UK94.ts)
```

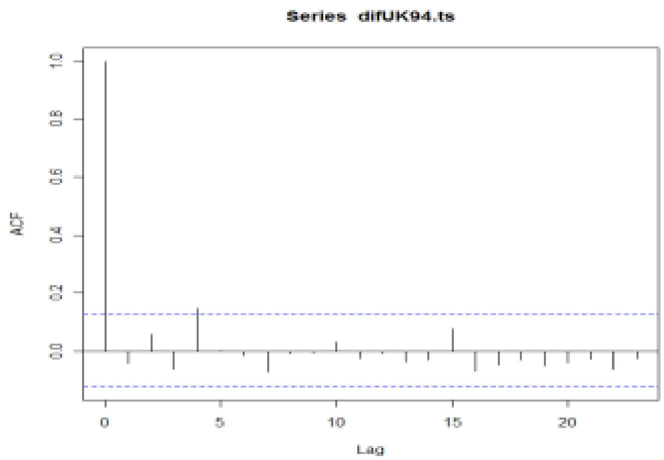


The plots of the data and the sample acf and sample pacf clearly show the need for differencing. (There is trending of the data, the sample acf is very slow to attenuate, and the pacf has a very large lag 1 value.)

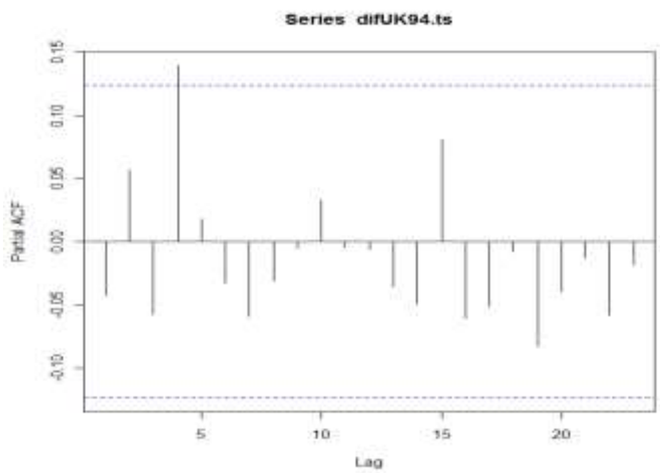
```
> difUK94.ts<-ts(diff(UKpound94))
> plot(difUK94.ts,xlab="Time",ylab="Differenced Exchange
Rate",main="Dollars per Pound, Daily Differenced Exchange Rate, 1994")
```



```
> acf(difUK94.ts)
```



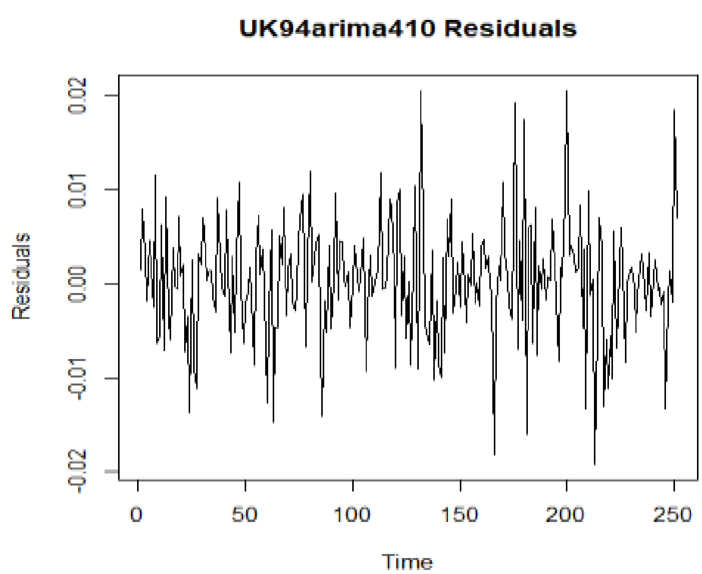

```
> pacf(difUK94.ts)
```



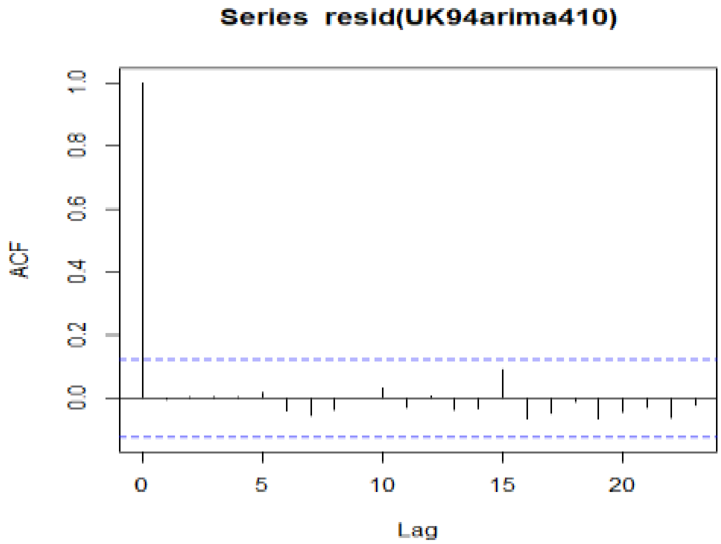
ARIMA(4,1,0) and ARIMA(0,1,4) models both fit the data well (*fit the models to the original data, not to the differenced data*). The ARIMA(4,1,0) fit follows.

```
> UK94arima410<-arima(UK94.ts,order=c(4,1,0))
> UK94arima410

Call:
arima(x = UK94.ts, order = c(4, 1, 0))

Coefficients:
          ar1     ar2      ar3     ar4
      -0.0248  0.0491  -0.0500  0.1446
s.e.   0.0628  0.0635   0.0634  0.0634

sigma^2 estimated as 4.036e-05:  log likelihood = 909.93,  aic = -
1809.86

> library("lmtest")

> coeftest(UK94arima410)

z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)
ar1 -0.024839   0.062814 -0.3954  0.69252
ar2  0.049140   0.063450  0.7745  0.43866
ar3 -0.050039   0.063415 -0.7891  0.43007
ar4  0.144633   0.063370  2.2823  0.02247 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

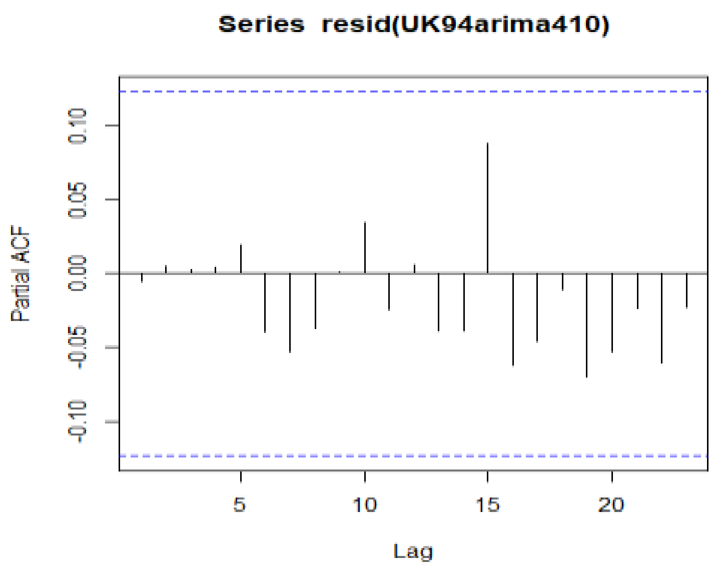
Residual analysis for this ARIMA(4,1,0) model follows.



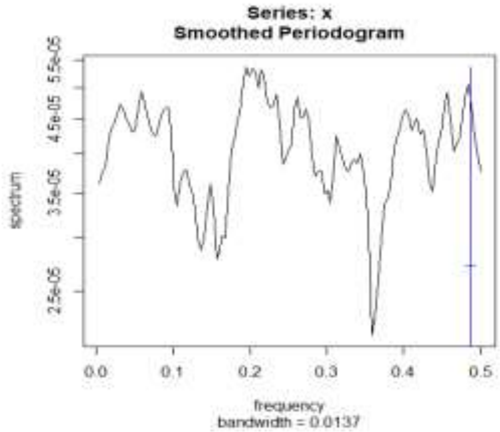
```
> acf(resid(UK94arima410))
```




```
> pacf(resid(UK94arima410))
```



```
> spectrum(resid(UK94arima410),span=12)
```



The fitted ARIMA(4,1,0) model has achieved adequate reduction to white noise. The model is

$$(1+0.024839B-0.049140B^2+0.050039B^3-0.144633B^4)(1-B)y_t=\varepsilon_t\,.$$

```
> zeros<-1/polyroot(c(1,-coef(UK94arima410)))
> zeros
[1]  0.0269931-0.5981613i -0.6757795+0.0000000i  0.0269931+0.5981613i
[4]  0.5969545-0.0000000i

> #amplitude
> Mod(zeros)[3]
[1] 0.59877
> #phase
> 2*pi/Arg(zeros)[3]
[1] 4.118231
```

The associated autoregressive polynomial is therefore

$$\begin{aligned} &z^4+0.024839z^3-0.049140z^2+0.050039z-0.144633 \\ &\quad =(z+0.6758)(z-0.5970)(z-0.5988\exp(i\frac{2\pi}{4.1182}))(z-0.5988\exp(-i\frac{2\pi}{4.1182})). \end{aligned}$$

The model estimates a pseudocycle of length about 4 days.

Let's examine other time stretches. For the 1996 data an ARIMA(4,1,0) model adequately achieves reduction to white noise. The fitted model is

$$(1-0.014593B-0.165742B^2-0.037715B^3+0.147212B^4)(1-B)y_t=\varepsilon_t\,.$$

```
> zeros<-1/polyroot(c(1,-coef(UK96arima410)))
> zeros
[1]  0.4868351-0.3611355i -0.4795388-0.4131567i -0.4795388+0.4131567i
[4]  0.4868351+0.3611355i
```

```
> #amplitude
> Mod(zeros)[3:4]
[1] 0.6329739 0.6061578
> #phase
> 2*pi/Arg(zeros)[3:4]
[1] 2.585227 9.844641
```

The associated autoregressive polynomial is therefore

$$\begin{aligned}
&z^4 - 0.014593z^3 - 0.165742z^2 - 0.037715z + 0.147212 \\
&\quad = (z - 0.6330\exp(i\frac{2\pi}{2.5852}))(z - 0.6330\exp(-i\frac{2\pi}{2.5852})) \\
&\qquad \times (z - 0.6062\exp(i\frac{2\pi}{9.8446}))(z - 0.6062\exp(-i\frac{2\pi}{9.8446})).
\end{aligned}$$

There are two estimated pseudocycles, with lengths about 2.6 days and about 9.8 days.

Next consider 1998. The fit is an ARIMA(2,1,0) model. It is

$$(1 - 0.167568B + 0.203251B^2)(1 - B)y_t = \varepsilon_t .$$

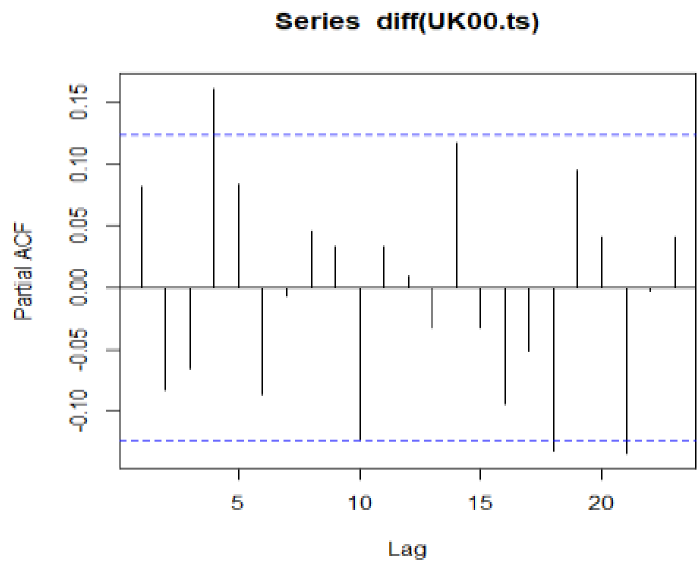
```
> zeros<-1/polyroot(c(1,-coef(UK98arima210)))
> zeros
[1] 0.0837841-0.4429799i 0.0837841+0.4429799i

> #amplitude
> Mod(zeros)[2]
[1] 0.4508336
> #phase
> 2*pi/Arg(zeros)[2]
[1] 4.54031
```

The associated autoregressive polynomial is

$$z^2 - 0.167568z + 0.203251 = (z - 0.4508\exp(i\frac{2\pi}{4.5403}))(z - 0.4508\exp(-i\frac{2\pi}{4.5403})).$$

There is an estimated pseudocyle with length about 4.5 days, but it is very weak. The model does achieve adequate reduction to white noise.

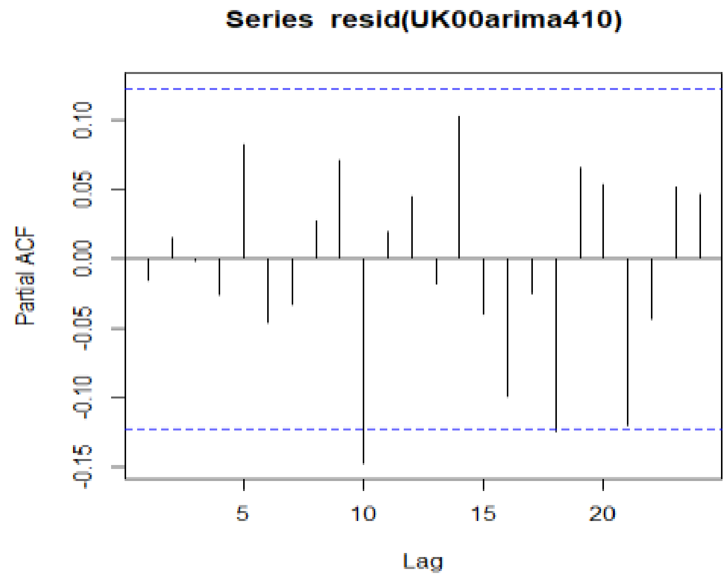
Finally, let's also analyze the exchange rate for the year 2000. The plot of the pacf of the differenced series follows.



The lag 4 partial autocorrelation is significant, suggesting we start by fitting an ARIMA(4,1,0) model. Partial correlations at lags 18 and 21 are also significant, and perhaps the lag 10 result merits attention.

```
> UK00arima410<-arima(UK00.ts,order=c(4,1,0))
> UK00arima410

Call:
arima(x = UK00.ts, order = c(4, 1, 0))

Coefficients:
         ar1      ar2      ar3     ar4
      0.0965  -0.0597  -0.0746  0.1663
s.e.  0.0623   0.0623   0.0626  0.0624

sigma^2 estimated as 6.259e-05:  log likelihood = 858.48,  aic = -
1706.96
> coeftest(UK00arima410)

z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)
ar1  0.096488   0.062343  1.5477 0.121697
ar2 -0.059732   0.062289 -0.9590 0.337582
ar3 -0.074646   0.062572 -1.1929 0.232890
ar4  0.166312   0.062447  2.6633 0.007739 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> pacf(resid(UK00arima410))
```



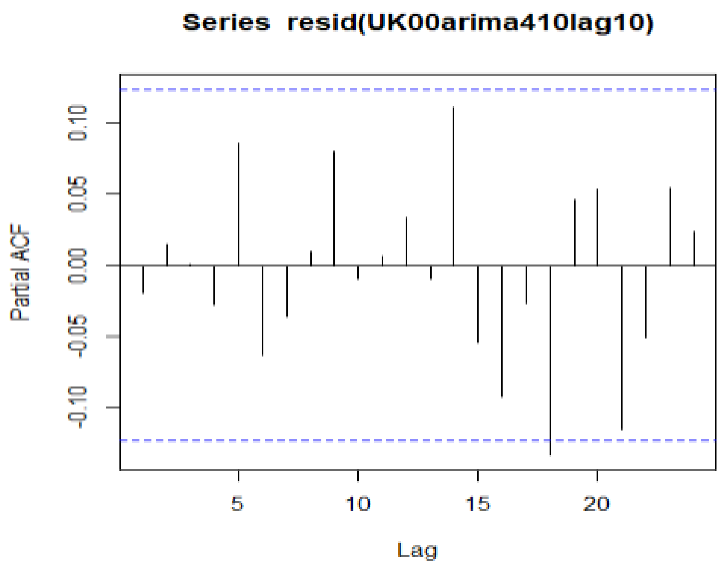
The residual partial autocorrelation at lag 10 is significant. We can modify the fitted model by using the seasonal ARIMA estimation option to add a lag 10 autoregressive coefficient. The code to do this follows.

```
> UK00arima410lag10<-
arima(UK00.ts,order=c(4,1,0),seasonal=list(order=c(1,0,0),period=10))
> UK00arima410lag10

Call:
arima(x = UK00.ts, order = c(4, 1, 0), seasonal = list(order = c(1, 0,
0), period = 10))

Coefficients:
         ar1      ar2      ar3     ar4     sar1
      0.1105  -0.0502  -0.0831  0.1784  -0.1388
s.e.  0.0626   0.0624   0.0627  0.0626   0.0638

sigma^2 estimated as 6.138e-05:  log likelihood = 860.81,  aic = -
1709.63
```

Note that the AIC value for this model is −1709.63, compared to −1706.96 for the ARIMA(4,1,0) model.

```
> coeftest(UK00arima410lag10)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1   0.110536   0.062640  1.7646 0.077624 .
ar2  -0.050183   0.062437 -0.8037 0.421543
ar3  -0.083066   0.062692 -1.3250 0.185176
ar4   0.178442   0.062605  2.8503 0.004368 **
sar1 -0.138824   0.063848 -2.1743 0.029684 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> coef(UK00arima410lag10)
        ar1         ar2         ar3         ar4        sar1
 0.11053622 -0.05018331 -0.08306609  0.17844238 -0.13882427
```

Let's examine the residual pacf plot for this model.



The added factor has produced improvement. There is a modestly significant lag 18 residual partial autocorrelation, but I won't try to make a further adjustment for it.

This new model is denoted by ARIMA(4,1,0)(1,0,0)$_{10}$. It is written as

$$(1-0.110536B+0.050183B^2+0.083066B^3-0.178442B^4)(1+0.138824B^{10})(1-B)y_t=\varepsilon_t\,.$$

Note the multiplicative structure. Let's determine the zeros of the AR(4) polynomial part of the model.

```
> zeros<-1/polyroot(c(1,-coef(UK00arima410lag10)[1:4]))
> zeros
[1]  0.0785574-0.6678182i -0.6519333+0.0000000i  0.0785574+0.6678182i
[4]  0.6053548+0.0000000i
```

There are two real zeros and one complex conjugate pair.

```
> #amplitude
> Mod(zeros)[3]
[1] 0.6724228
> #phase
> 2*pi/Arg(zeros)[3]
[1] 4.322198
```

The AR(4) polynomial is

$$z^4 - 0.110536z^3 + 0.050183z^2 + 0.083066z - 0.178442$$
$$= (z + 0.6519)(z - 0.6054)(z - 0.6724\exp(i\frac{2\pi}{4.3222}))(z - 0.6724\exp(-i\frac{2\pi}{4.3222})).$$

Thus, the model estimates the presence of a pseudocycle of length about 4.3 days.

In summary, there are some similarities between the fits for the different years, primarily the presence of a pseudocycle with period length close to 4 days. Overall the evidence points to signals which change somewhat over time, and which are sometimes rather weak.

Summary and additional remarks

1. If a time series has trend structure, it is necessary to remove the trend before fitting an ARMA model. Trend removal is accomplished by differencing the time series. To determine whether differencing is needed, plot the time series to judge visually if trending is present. If it is, difference the data series and plot and inspect the differenced series to verify that the trend has been removed.

2. If the sample autocorrelations of the time series die down smoothly and slowly as the lag increases, differencing is needed. A large lag one partial autocorrelation (a value close to 1) is another indication that differencing is required.

3. An ARIMA(p, d, q) model has autoregressive structure of order p, moving average structure of order q, and differencing of order d to remove the trend.

4. If one wants to fit an ARIMA(p, 1, q), there are two choices to accomplish this fit. One can fit an ARMA(p, q) model *to the differenced data*, or one can choose an ARIMA(p, 1, q) model for the data without first differencing the data. These two lead to the same result. Usually one chooses the latter approach, so that forecasts from the model fit will be calculated for the data, rather than for the differenced data.

5. Time series of exchange rates are analyzed in these notes. The series contain trends, and differencing is required.