Ch 11: Non-financial time series models

Background reading

Gujarati and Porter – Chapter 22
Venables and Ripley – Chapter 14

Venables, W. N. and Ripley, B. D. (2003) *Modern applied statistics with S-Plus*, 4th edition. Springer

Outline

1. Overview
2. ARIMA modelling
3. S-ARIMA modelling
4. Example 1: ARIMA modellling
5. Example 2: Box-Jenkins method
6. Example 3: SARIMA modelling

1.1 Purpose of time series models

- Very generally, a time series refers to a set of observations that have been collected over time
- Time series models essentially solve two basic problems
  1. Accounting for correlations in related observations that have been collected close together in time.
  2. Characterising seasonal and cyclical behaviour that may be often be found in data sets in, for example, business economics and in the natural sciences.
- As discussed in Venables and Ripley (2003) there is a lot of mileage in simple graphical plots to characterise general time series problems e.g.
  - Are seasonal patterns present?
  - Is the series generally increasing or decreasing over time?

1.2 Specific purposes of time series models

1. Accounting for correlations in related observations that have been collected close together in time.
   - AR, MA, ARMA, ARIMA models
   - Specialised financial time series models such as ARCH/GARCH
2. Characterising seasonal and cyclical behaviour that may be often be found in data sets in, for example, business economics and in the natural sciences.
   - SARIMA modelling

1.3 Business examples

• Two obvious ways in which nonstationary time series can occur in financial and economic time series

- Seasonality e.g. different seasons and time of the year affect customer behaviour and consumption
- Most financial and economic time series show (exponential?) growth trends over time

• **It is prudent to take steps to try and account for these stylised empirical facts – albeit imperfectly!**

1.4 Purpose of the lecture

- In this lecture we consider general ARIMA modelling of economic time series
- There is an entire subject in its own right that discusses financial time series models derived from prices on financial markets like stock and currency markets
- Financial time series models are thus an entirely different set of models in their own right but share foundational motivations
  - e.g. ARCH models are AR models for price volatility
  - e.g. GARCH models are ARMA models for price volatility

2.1 Overview of ARIMA models

• ARIMA models provide a systematic solution to the following problem:
How do you systematically account for correlations caused by observations being collected close together in time e.g. correlations in say Monday's and Tuesday's sales figures?
• The construction of ARIMA models has a modular structure
• In practice model selection can be performed rigorously using a graphical approach known as the Box-Jenkins method
• in practical terms much of the R commands and model interpretation shares much in common with standard regression approaches

2.2 Hierarchy of Time Series models

1. **Basic models.** Autoregressive (AP(p)) and Autoregressive Moving Average (MA(q))
2. **Hybrid I: ARMA Models.** ARMA models=AR(p) combined with MA(q)
3. **Hybrid II: ARIMA(p, q, q) Models.** Difference d times to get an approximately stationary series and then fit an ARMA model to the resultant series.

- ARMA($p, 0$)=AR(p)
- ARMA($0, q$)=MA(q)
- ARIMA($p, 0, q$)=ARMA(p, q)

2.3 Differencing

• A discussion of time series models and differencing involves some maths and a polynomial in the differencing shift operator $(1 - B)^d$
• This terminology used can make the subject sound more difficult than it really is
• However, the terminology is probably justified as it leads to as systematic approach as possible which will be important for problems of a practical size
• The main thing to remember is that as the name suggests the effect of the backward shift operator B is to shift the series backwards!

2.4 Differencing and the backward shift operator

- **The main thing to remember is that as the name suggests the effect of the backward shift operator B is to shift the series backwards!**
- $d = 0$ gives the original series

$$(1 - B)^0 Z_t = Z_t$$

- $d = 1$ gives the first differences:

$$(I - B)^1 Z_t = (I - B) Z_t = Z_t - Z_{t-1}$$

- $d = 2$ gives

$$(1 - B)^2 Z_t = (I - 2B + B^2) Z_t = Z_t - 2Z_{t-1} + Z_{t-2}$$

- **Usually low values of d are sufficient. Higher values of d begin to lose interpretability.**

2.5 Differencing in R

- In R the basic function to perform differencing is `diff`

$$\text{diff(series, number of lags)}$$

- The default is diff(series) will return the first differences
- This makes senses as first differences are the most commonly encountered scenario associated with practical modelling work

2.6 R Example I

```
z<-seq(1, 10)
z2<-z∧2
z2
1 4 9 16 25 36 49 64 81 100
diff(z2)
3 5 7 9 11 13 15 17 19
```

- Sequence above constructed as

$$4-1, 9-4, 16-9, 25-16, 36-25, 49-36, 64-49, 81-61, 100-81$$

- To give

$$3, 5, 7, 9, 13, 15, 17, 19$$

- As an alternative consider

```
diff(z2, 2)
8 12 16 20 24 28 32 36
```

- Sequence above constructed by subtracting from observations two previously:

$$9-1, 16-4, 25-9, 36-16, 49-25, 64-36, 81-49, 100-64$$

- To give

$$8, 12, 16, 20, 24, 28, 32, 36$$

2.8 ARMA models

- ARMA(p, q) model defined by the equation

$$\phi_\alpha(B)Z_t = \phi_\beta(B)\epsilon_t$$

- ϵ_t=uncorelated White Noise sequence with $E[\epsilon_t] = 0$, Var(ϵ_t) $= \sigma^2$
- $\phi_\alpha(B)$=order p autoregressive (AR) polynomial in B

$$\phi(B) = 1 - \alpha_1 B^1 - ... - \alpha_p B^p$$

- $\phi_\beta(B)$=order q Moving Average (MA) poynomial in B

$$\theta(B) = 1 - \beta_1 B^1 - ... - \beta_p B^p$$

2.9 ARIMA models

- Let $\nabla = (1 - B)$
- Then $W_t = \nabla^d Z_t = (1 - B)^d Z_t$ is an ARMA process satisfying the earlier definition with Z_t replaced by $\nabla^d Z_t$
- The original series Z_t is then said to follow an ARIMA model
- The underlying maths actually being a whole lot simpler than it might first seem. In words
  - Take the dth difference
  - Once you do this should end up with an ARMA model
  - If $d = 0$ you already have an ARMA model without having to do any additional differencing

2.10 ARIMA models

- The ARIMA(p, d, q) model is thus defined by the equation

$$\phi_\alpha(B)\nabla^d Z_t = \phi_\beta(B)\epsilon_t$$

- ϵ_t=uncorelated White Noise sequence with $E[\epsilon_t] = 0$, Var(ϵ_t) $= \sigma^2$

- $\phi_\alpha(B)$=order p autoregressive (AR) polynomial in B

$$\phi(B) = 1 - \alpha_1 B^1 - ... - \alpha_p B^p$$

- $\phi_\beta(B)$=order q Moving Average (MA) poynomial in B

$$\phi_\beta(B) = 1 - \beta_1 B^1 - ... - \beta_p B^p$$

2.11 ARIMA=Autoregressive Integrated Moving Average

• The basic idea is that by taking differences an ARIMA model reduces to a regular ARMA model
• As a result of this the name that is often used for this is an Autoregressive Integrated Moving Average model
• By analogy with regular calculus this is assumed to be an integrated model since you need to take differences ("differentiate") to get back to an ARMA model
• Maybe integrated ARMA model would be a better term but unfortunately ARIMA is the term that is used

2.12 Example: Differencing is not a panacea ...

- Differencing a stationary time series will produce another stationary time series
- However it is important not to over-difference a series

  - Want as faithful statistical description of data as possible
  - Inappropriate differencing can lead to numerical problems with computational software
  - Standard techniques usually work best for data that is only moderately correlated
  - Numerical examples exist whereby inappropriate differencing can be shown to lead increase correlations
  - This can have particular relevance for theoretical financial modelling as conventional models typically assume that series of asset market returns should be uncorrelated e.g. random walk model

2.13 Spurious correlations added to a white noise sequence

- ϵ_t=uncorelated White Noise sequence with $E[\epsilon_t] = 0$, $\text{Var}(\epsilon_t) = \sigma^2$
- Consider the first differences W_t of ϵ_t: $W_t = \epsilon_t - \epsilon_{t-1}$

$$\begin{aligned}
\text{Var}(W_t) &= \text{Var}(\epsilon_t - \epsilon_{t-1}) \\
&= \text{Var}(\epsilon_t) + \text{Var}(\epsilon_{t-1}) - 2\text{Cov}(\epsilon_t, \epsilon_{t-1}) \\
&= \sigma^2 + \sigma^2 - 0 = 2\sigma^2
\end{aligned}$$

$$\begin{aligned}
\gamma_1 = \text{Cov}(W_t, W_{t-1}) &= \text{Cov}(\epsilon_t - \epsilon_{t-1}, \epsilon_{t-1} - \epsilon_{t-2}) \\
&= -\text{Cov}(\epsilon_{t-1}, \epsilon_{t-1}) + 0 \\
&= -\text{Var}(\epsilon_{t-1}) = -\sigma^2
\end{aligned}$$

$$\text{Correlation}(W_t, W_{t-1}) = \frac{\text{Cov}(W_t, W_{t-1})}{\sqrt{\text{Var}(W_t)\text{Var}(W_{t-1})}} = \frac{-\sigma^2}{\sqrt{2\sigma^2 . 2\sigma^2}} = -\frac{1}{2}$$

- In this case we can calculate Covariance(W_t, W_{t-2}) as

$$\begin{aligned}\mathrm{Cov}(W_t, W_{t-2}) &= \mathrm{Cov}(\epsilon_t - \epsilon_{t-1}, \epsilon_{t-2} - \epsilon_{t-3}) \\ &= 0 \text{ since none of the subscripts match}\end{aligned}$$

2.15 Example: model construction for ARIMA(0, 1, 1)

- ARIMA(0, 1, 1)
  - $p = 0$ so no autoregressive bit
  - $d = 1$ so need to take a first difference

$$Z_t - Z_{t-1}$$

- $q = 1$ so there is one Moving Average term in addition to the observation error:

$$\epsilon_t - \beta_1 \epsilon_{t-1}$$

- The presence of random statistical error means an ϵ_t term is always present
- The full model becomes

$$Z_t - Z_{t-1} = \epsilon_t - \beta_1 \epsilon_{t-1}$$

2.16 Example: forecasting for ARIMA(0, 1, 1)

- ARIMA models generalise ARMA models so that the same general forecasting and modelling principles apply
- Consider the following model for a time series of 197 observations

$$Z_t - Z_{t-1} = \epsilon_t - \beta_1 \epsilon_{t-1}$$

- Suppose that we have estimates $\hat{\beta}_1 = 0.699$, $\hat{\epsilon}_{197} = -0.15$ and the last observation is $z_{197} = 17.4$
- The next three forecasts for this model are

$$Z_{198} - Z_{197} = \epsilon_{198} - \hat{\theta} a_{197}$$

$$E[Z_{198}] = Z_{197} + 0 - (0.699)(-0.15) = 17.505$$

$$E[Z_{199}] = E[Z_{198}] + E[\epsilon_{199}] - \hat{\theta} E[\epsilon_{198}] = E[Z_{198}]$$

$$E[Z_{200}] == E[Z_{199}] + E[\epsilon_{200}] - \hat{\theta} E[\epsilon_{199}] = E[Z_{199}]$$

3.1 SARIMA and modelling seasonal effects

- A second problem that time series models focus upon is the modelling of seasonal effects
- This can be achieved using SARIMA models
  - Combine an ARIMA-type modelling approach discussed earlier
  - Together with an explicit link to the underlying human calendar to quantify the impact upon financial and social systems
- The original motivation behind SARIMA models was modelling passenger demand for airlines (Venables and Ripley, 2003)
  - However, SARIMA models are thought to have a particularly broad range of application.
  - See e.g. Fry et al. (2021) for an application to business-cycle effects in corporate bank accounts

3.2 Seasonal time series

- Seasonal effects are very important in forecasting
- Examples might include monthly ice cream sales (dependent on the weather), intraday stock price data (dependent on the time of day and the opening of US markets)
- If you have monthly data, as there are 12 months in a year it would make sense to model

$$B^{12} Z_t = Z_{t-12}; \ \nabla_{12} = Z_t - Z_{t-12}$$

rather than looking at

$$BZ_t = Z_{t-1}; \ \nabla = Z_t - Z_{t-1}$$

- **When you consider models of this form...**
  - Need a clear link to the human calendar
  - **And** a realistic physical mechanism e.g. seasonal effects of temperature will obviously affect ice cream sales, heating gas consumption etc.

3.3 SARIMA model construction

- Start with a general ARIMA model defined in terms of the seasonal effects:

$$\Phi_{SAR}(B^S)(1 - B^S)^D Z_t = \Phi_{SMA}(B^S)\epsilon_t, \tag{1}$$

where the S in equation (1) refers to a seasonal term term and D reflects the order of differencing (usually $D = 0$ or $D = 1$)
- Next apply a regular ARIMA model to the deseasonalised series
- Apply the differencing operator $(I - B)^d$ to the LHS of equation (1)

$$(I - B)^d \Phi_{SAR}(B^S)(1 - B^S)^D Z_t = \Phi_{SMA}(B^S)\epsilon_t, \tag{2}$$

Since polynomials in the backward shift operator B commute re-write equation (2) as

$$\Phi_{SAR}(B^S)(I - B)^d (1 - B^S)^D Z_t = \Phi_{SMA}(B^S)\epsilon_t.$$

3.4 Constructing the SARIMA model

- Starting with the deseasonalised and differenced series

$$\Phi_{SAR}(B^S)(I-B)^d(1-B^S)^D Z_t = \Phi_{SMA}(B^S)\epsilon_t.$$

- Now apply an AR term to the LHS and a MA term to the RHS

$$\Phi_{AR}(B)\Phi_{SAR}(B^S)(I-B)^d(1-B^S)^D Z_t = \Phi_{MA}(B)Phi_{SMA}(B^S)\epsilon_t. \quad (3)$$

- Equation (3) thus defines a SARIMA $(p,d,q)\times(P,D,Q)_S$ model
- S denotes the length of the seasonal difference and the capital letters refer to the components of the seasonal term

3.5 Practical model construction

- Venables and Ripley (2003) give the example of an SARIMA $(0,1,1) \times (0,1,1)_{12}$ model fitted to monthly series of aircraft passenger numbers
- This has the interpretation of two moving average models one associated with the time of the year and one that remains once seasonal effects are adjusted for
- Using equation (3) the model can be constructed as follows
- In the absence of any autoregressive components the LHS of equation (3) reduces to

$$
\begin{aligned}
(I-B)(I_B^{12})Z_t &= Z_t - BZ_t - B^{12}Z_t + B^{13}Z_t \\
&= Z_t - Z_{t-1} - Z_{t-12} + Z_{t-13}. \qquad (4)
\end{aligned}
$$

- The RHS of equation (3) reduces to

$$
\begin{aligned}
(I-\beta_1 B)(I-\beta_{12}B^{12})\epsilon_t &= (I-\beta_1 B-\beta_{12}B^{12}+\beta_1\beta_{12}B^{13}) \\
&= \epsilon_t-\beta_1\epsilon_{t-1}-\beta_{12}\epsilon_{t-12}+\beta_1\beta_{12}\epsilon_{t-13}
\end{aligned} \tag{5}
$$

Finally, combining equations (4-5) gives

$$
Z_t = Z_{t-1}+Z_{t-12}-Z_{t-13}+\epsilon_t-\beta_1\epsilon_{t-1}-\beta_{12}\epsilon_{t-12}+\beta_1\beta_{12}\epsilon_{t-13}
$$

4.1 Example 3: ARIMA modelling

- Data in the file `Airlines.txt` gives data on monthly airline passenger figures
- Simple summary plots give clear evidence of a monthly seasonal effect
- Can get a lot of joy out of simple descriptive plots
- Suggestion is that passenger numbers are generally increasing over time subject to significant seasonal variations

```
Airlines<-read.table(‘‘E:Airlines.txt")
passengers<-Airlines[,1]
ts.plot(passengers)
```

4.2 Summary plot using `ts.plot`

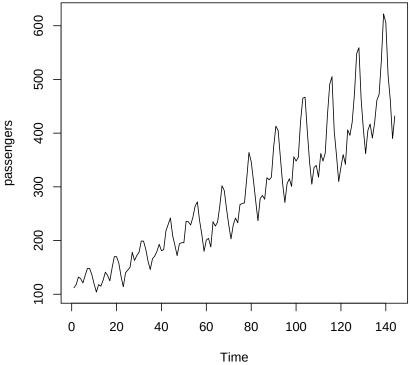

Figure: Descriptive time series plot for airline data.

4.3 A simplified ARIMA approach

- In this section we will fit simple ARIMA models to this data

  - In reality this reduces to fitting ARMA models to the first-differenced series

  - First-differencing is an artifical way of generating an approximately stationary time series without a clear underlying trend

  - We will see later in the lecture that in this case a more specialised SARIMA model is ultimately needed to capture underlying seasonal effects in this data

4.4 ARMA modelling

- We apply ARMA models applied to the first differences constructed in R using
`firstdifferences<-diff(passengers)`
- We consider three simple models
  1. AR(1) model
  2. MA(1) model
  3. ARMA(1, 1) model
- Lots of similar models should ultimately give similar interpretations
- It is often easiest to just assume (as here) that AR and MA effects may be present and just fit the lowest possible first-order model
- The Box-Jenkins approach shown in Section 5 gives a principled way of selecting the order of the models

4.5 AR(1)/ARIMA(1, 0, 0) model

The model fitted in this case is

$$Z_t = \mu + \alpha_1(Z_{t-1} - \mu)$$

```
 arima(firstdifferences, order = c(1, 0, 0))
Coefficients:
ar1 intercept
0.3037 2.3700
s.e.  0.0797 3.8369
```

- Conduct a t-test using

```
length(firstdifferences)
143
```

- Residual d.f. $=$143-2 estimated parameters

```
1-pt(0.3037/0.0797, 141)
0.0001032107
```

- So significant evidence of an autoregressive effect

4.6 MA(1)/ARIMA(0, 0, 1) model

The model fitted in this case is

$$Z_t = \mu + \beta_1 \epsilon_{t-1} + \epsilon_t.$$

```
 arima(firstdifferences, order = c(0, 0, 1))
Coefficients:
ma1 intercept
0.4012 2.4213
s.e.  0.0893 3.6858
```

- Conduct a t-test using

```
length(firstdifferences)
143
```

- Residual d.f. $=$143-2 estimated parameters

```
1-pt(0.4012/0.0893, 141)
7.264574e-06
```

- So significant evidence of a moving average effect

4.7 ARMA(1, 1)/ARIMA(1, 0, 1) model

The model fitted in this case is

$$Z_t = \mu + \alpha_1(Z_{t-1} - \mu)\beta_1\epsilon_{t-1} + \epsilon_t.$$

```
 arima(firstdifferences, order = c(1, 0, 1))
Coefficients:
ar1 ma1 intercept
-0.4767 0.8645 2.4509
s.e.  0.1153 0.0714 3.2660
```

- Conduct a t-test using

```
length(firstdifferences)
143
```

- Residual d.f. =143-3 estimated parameters

```
c(2*(1-pt(0.4767/0.1153, 140)),
2*(1-pt(0.8645/0.0714, 140)))
6.101428e-05 0.000000e+00
```

- So significant evidence of an autoregressive effect and a moving average effect

5.1 Example 2: Box-Jenkins methodology

- Model selection for time series is difficult
- Various models will perform approximately as well as each other
- Basic idea is that you can choose a model based on the ACF and PACF though this will be a bit of an artform and may work imperfectly with real datasets

Model	Typical ACF	Typical PACF
AR(p)	Exponential decay or damped sine wave pattern	Significant spikes through lag p
MA(q)	Significant spikes through lag q	Exponential decay or damped sine wave pattern
ARMA(p, q)	Exponential decay	Exponential decay

Table: Graphical ACF and PACF model selection procedure

5.2 A note on the ACF and PACF

• The autocorrelation function measure correlation of Z_t with itself τ time periods later:

$$\rho_t = \text{Corr}(X_t, X_{t+\tau}). \tag{6}$$

• The partial autocorrelation function essentially repeats equation (6) but makes a further adjustment for correlations in intervening lags

• In R the relevant functions are `acf` and `pacf`

• For our example in R

```
acf(firstdifferences)
pacf(firstdifferences)
```

5.3 ACF plot

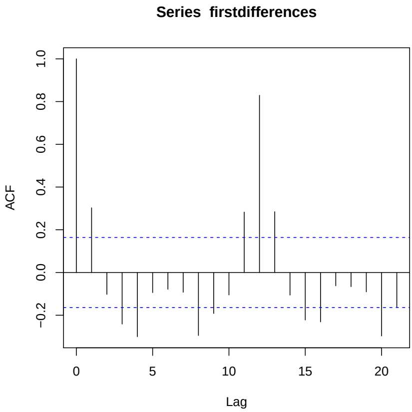

Figure: ACF plot for the first-differenced series.

5.4 PACF plot

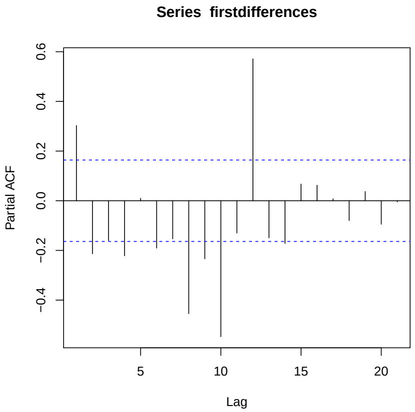

Figure: PACF plot for the first-differenced series.

5.5 Interpretation of the Box-Jenkins technique

- Interpret these ACF and PACF plots as follows
  - ACF plot shows damped sine wave pattern with a spike at Lag 1
  - PACF plot shows damped sine wave pattern with a spike at Lag 1
  - Suggestion would then be to fit at ARMA(1, 1) to the first-differenced series
  - This graphical interpretation would then tally with the numerical results in Slide 4.7

- Plotting the first differences of the data suggests that some residual seasonality may be present
- Suggestion in this case is that more specialised SARIMA models may ultimately be needed
- In R

```
firstdifferences<-diff(passengers)
ts.plot(firstdifferences)
```

6.2 Summary plot using ts.plot

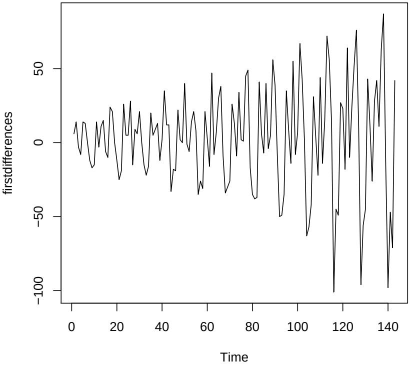

Figure: Residual seasonality remaining in first-differenced series.

6.3 Model fitting

- Venables and Ripley (2003) describe an SARIMA $(0,1,1) \times (0,1,1)_{12}$ as being a classical model for airline passenger data
- If you look at the figure on the previous slide suggestion is need a differencing procedure to get rid of the time trend
- Simplest to include a single Moving Average term both for the seasonal component and for the de-seasonalised series
- SARIMA models can be fitted using the function `ARIMA`.
- Need to specify separate ARIMA components for the seasonal component and for the de-seasonalised series
- Need to specify the period – here 12 months in a year
- Note the `summary` command **does not work** with the `arima` command in R

6.4 SARIMA in R

```
arima(passengers, order=c(0, 1, 1),
seasonal=list(order=c(0, 1, 1), period=12))
Coefficients:
ma1 sma1
-0.3087 -0.1074
s.e.  0.0890 0.0828
length(passengers)
144
```

- Result tells there are 140 residual degrees of freedom: 144 observations -2 differencing paramters - 2 estimated parameters

```
c(2*(1-pt(3087/0.0890, 140)), 2*(1-pt(0.1074/0.0828,
140)))
0.0000000 0.1967298
```

- Results thus give significant evidence of a Moving Average effect ($p = 0.000$) but no formal evidence of a monthly effect in this case ($p = 0.197$)