Modelling seasonality

Beáta Stehlíková
Time series analysis
Modelling seasonality

- We have seen models with seasonal character - for example AR(2) model with complex roots
- These models are, however, not sufficient to model all seasonal data
- Also, we might want to include seasonality (quarterly data, monthly data) into specification
- There are models which are specifically for modelling seasonal data - SARIMA models (seasonal ARIMA models)
Example - data

- Number of airline passengers by Box and Jenkins - founders of ARIMA modelling
- Monthly data, January 1949 - December 1960
- We work with logarithms, they stabilize variance
- In R: `data(AirPassengers); x <- log(AirPassengers)`
Example - differences

- We take differences - in R: `diff(x)` - they also have seasonality:
Example - differences

- ACF and PACF for these differences - in R
  \texttt{acf2(diff(x))}: 

![ACF and PACF plots for diff(x)](image)
Example - differences

- We can take seasonal differences $x_t - x_{t-12}$ - in R

```
diff(x, 12):
```

Modelling seasonality – p.6/14
Example - differences

- Both classical and seasonal differences - classical because of the trend, seasonal because of the differences - in R `diff(diff(x,12))`:
Example - differences

- What to do with that:

![ACF and PACF plots for the series: diff(diff(x, 12))](image)
Example - seasonal AR and MA terms

- Based on ACF we may try using terms up to $ma(12)$
- Box and Jenkins:
  - not all $ma(1), ma(2), ..., ma(12)$
  - neither only $ma(1)$ nor $ma(12)$
  - but multiply polynomials of order 1 and 12:
    $$ (1 - \beta L)(1 - \theta L^{12})u_t $$
    - we get 13 ma terms but we need only 2 coefficients
- In the same way, we can combine them:
  - seasonal ma terms of higher order:
    $$ 1 - \theta_1 L^{12} - \theta_2 L^{24} $$
  - seasonal ar term with an ordinary one:
    $$ (1 - \alpha L)(1 - \theta_1 L^{12})x_t $$

**SARIMA models - terminology**

- Recall **ARIMA** \((p, d, q)\) models:
  - \(p\) - number of AR terms
  - \(d\) - how many times we take a difference
  - \(q\) - number of MA terms

- **SARIMA** \((p, d, q) \times (P, D, Q)_s\) has also:
  - \(P\) - number of seasonal AR terms
  - \(D\) - how many seasonal differences
  - \(Q\) - number of seasonal AR terms
  - \(s\) - period of the data

- We need to check that the data we use - after differencing - does not have a unit root
Example - model in R

• For our data: SARIMA \((0, 1, 1) \times (0, 1, 1)_s\), where \(s = 12\)

• Time series \(\text{diff(diff(x,12))}\) does not have neither trend nor a unit root

• In R: \(\text{sarima(x,0,1,1,0,1,1,12)}\)

• We get:

```r
> sarima(x,0,1,1,0,1,1,12,details="FALSE")
$fit
Series: xdata
ARIMA(0,1,1)(0,1,1)[12]

Coefficients:
      ma1    sma1
ma1  -0.4018  -0.5569
s.e.  0.0896   0.0731

sigma^2 estimated as 0.001348:  log likelihood=244.7
AIC=-483.4  AICc=-483.21  BIC=-474.77
```

Modelling seasonality  – p.11/14
Example - model in R

- Residuals are ok:

![ACF of Residuals](image1)

![Normal Q-Q Plot of Std Residuals](image2)

![p values for Ljung-Box statistic](image3)
Example - predictions in R

- We have a model **SARIMA** $(0, 1, 1) \times (0, 1, 1)_s$
- Prediction for the following 2 years: 
  \[ \text{sarima.for}(x, 24, 0, 1, 1, 0, 1, 1, 12) \]
Exercises

Find a suitable SARIMA model (data on course webpage):

- `spain.txt` - number of tourists in Spain, monthly data from January 1970 to March 1989
- `souvenirs.txt` - sales in a souvenir shop on a beach in Australia, monthly data from January 1987 to December 1993