MW82: Time Series Analysis, Tutorial IV

Dr. Giulia Sabbadini

DICE

2023/24

Recap: Tutorial III

We did:

- Properties of MA(q)
- ACF, PACF
- Empirical use of ACF, PACF for lag, model selection

Today we will go through:

- nonseasonal ARIMA(p,d,q) models
- "Box-Jenkins approach"
- model selection (AIC, BIC)
- model diagnostics

Recap: ACF and PACF

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off to 0	Cuts off after lag q	Tails off to 0
PACF	Cuts off after lag p	Tails off to 0	Tails off to 0

ARMA/ARIMA models

- Generalization of AR and MA models, with 'built-in' differencing (the **I** part)
- Recall Wold's theorem: any stationary time series can be 'well' represented by an ARMA model.
- They include:
  - autoregressive (AR) terms
  - moving average (MA) terms
  - differencing (**I** stands for Integrated)
- "Integrated"=reverse of differencing

ARIMA

An ARIMA(p,d,q) model has *p* AR-terms, *d* first-differences applied, and *q* MA-terms.

Examples and Special Cases:

- AR(2) model is an ARIMA(2,0,0)
- MA(2) model is an ARIMA(0,0,2)
- White Noise is ARIMA(0,0,0)
- Random Walk is ARIMA(0,1,0)
- $y_t = \phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ is ARIMA(1,0,1) if $|\phi| < 1$

Box Jenkins Approach

(0) Preliminary analysis: Is the time series stationary?

1. Model identification (plot, ACF, PACF) : What class of models probably produced the (transformed) series?
2. Model estimation (OLS, ML): What are the model parameters (e.g. $\hat{\phi}$)?
3. Model checking: Are the residuals from the estimated model white noise?

Preliminary Analysis

Time series plot, (P)ACF plots and intuition:

- Trends, seasonality, non-constant variance, outliers, unit roots, level changes, ...

Trends:

- Estimate and remove (simple linear trend) **or** first-difference (careful with over-differencing)

Non-constant variance:

- upwards trend accompanied by increasing variance
- Box-Cox transformations

Seasonality:

- can be modeled as part of an seasonal ARIMA model (more on this in Tutorial 6)

Data Transformation

Box-Cox transformations:

- log: if variance is increasing quadratically with the mean
- sqrt: if variance is increasing linearly with the mean
- More in general: power transformations

$$y_t = \begin{cases} \left(x_t^{\lambda} - 1\right)/\lambda & \lambda \neq 0 \\ \log x_t & \lambda = 0 \end{cases}$$

- log and sqrt are special cases ($\lambda = 0$ and $\lambda = 0.5$)

No transformation

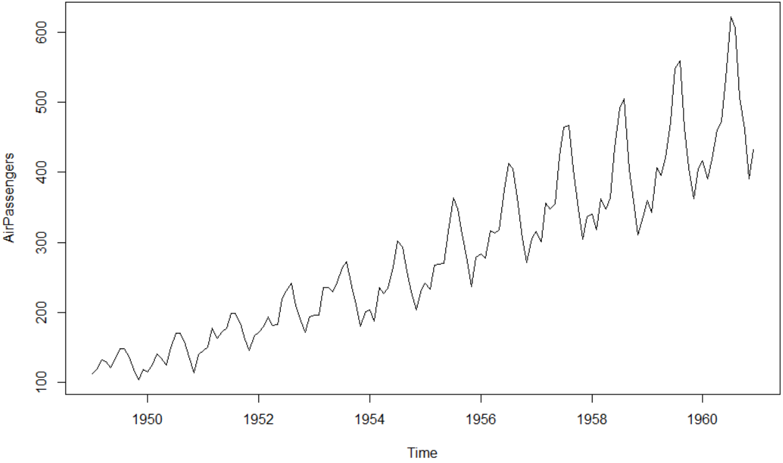

Figure: Monthly airline passengers from 1949 to 1960

Log(y_t)

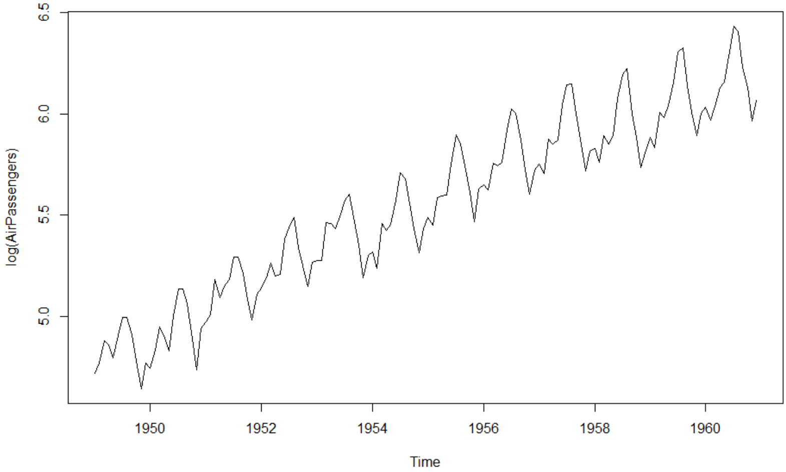

Figure: Monthly airline passengers from 1949 to 1960 (logged)

Log(y_t) plus First Difference

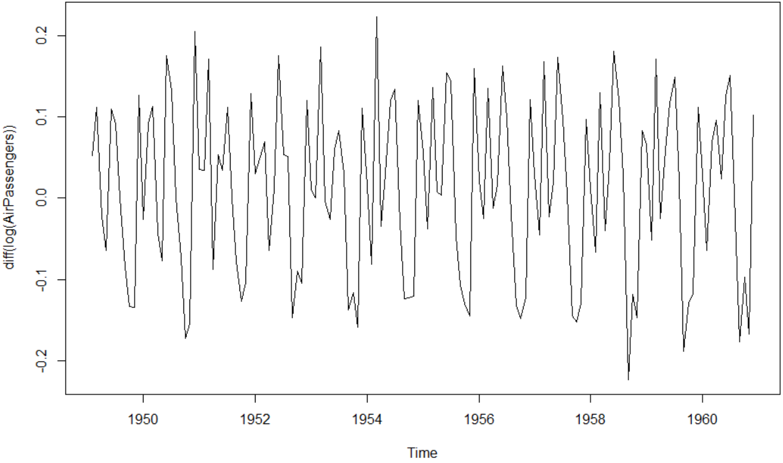

Figure: Monthly airline passengers from 1949 to 1960 (logged and Δ)

Model identification

- Which lag order to use?
  - Consider ACF and PACF plots together
  - Can be tricky sometimes (theoretical $\neq$ sample)

- If competing models, there are several criteria:
  - Mean squared error of the residuals (error variance):

$$s^2 = \frac{\sum_{t=1}^{T} e_t^2}{T - p}$$

  - (adjusted) R-squared statistic:

$$R^2 = 1 - \frac{\sum_{t=1}^{T} e_t^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

  - **Information Criterion (AIC, AICc, BIC)**

Information Criteria for ARMA(p,q) models

Akaike Information Criterion

$$AIC(p,q) = \log(\hat{\sigma}^2) + \frac{2(p+q)}{T}$$

where $\hat{\sigma}^2$ is our estimate of the residual variance and T is the sample size/length.

- $\hat{\sigma}^2$ measures the goodness of fit.
- $2(p+q)$ penalizes additional parameters (to stop overfitting).
- Clearly prefers more parsimonious models (as few parameters as possible).
- Models with *lower* AIC values are preferred.

Bayesian information criterion (BIC) and Akaike's Information Corrected Criterion (AICc)

$$BIC(p,q) = \log(\hat{\sigma}^2) + \frac{log(T)(p+q)}{T}$$

$$AICc(p,q) = \log(\hat{\sigma}^2) + \frac{2T(p+q)}{T-p-q-2}$$

- $\hat{\sigma}^2$ is our estimate of the residual variance and T is the sample size/length.
- Information Criterion can be rewritten ito likelihood $\mathcal{L}$:

$$AIC = -2log\hat{\mathcal{L}} + 2(p+q+1)$$

Model estimation

- AR(p) models can be estimated by OLS (Tutorial 2)
- MA(q) models can be estimated by ML (Lecture 6)
- ARMA models are usually estimated by ML

In R (with time series y):

- `arima(y,order = c(p,d,q))`
- `forecast::Arima(y,order = c(p,d,q))` includes BIC and AICc
- P-values are not provided. Calculate t-statistic by hand or use `lmtest::coeftest` .

Model checking

- Check your estimates and significance levels
- Check the ACF plot of your residuals (we want white noise!), should be no significant time dependence 'left'.
- Ljung-Box/Box-Pierce tests for residual autocorrelation over number of lags.

Ljung-Box/Box-Pierce test

- Tests the time-independence of our residuals, which is what we want (white noise). *
- A more formal approach than ACF plot of residuals
- To evaluate a set of autocorrelations jointly to determine if they indicate that the time series is white noise
- Box–Pierce test:

$$Q = T \sum_{k=1}^{K} \hat{\rho_k}^2$$

- Ljung-Box test is the same, with some weighting:

$$Q_{\mathrm{LB}} = T(T+2) \sum_{k=1}^{K} \left(\frac{1}{T-k} \right) \rho_k^2$$

- H_0: The residuals are independently distributed.
- In R: `Box.test` . Set parameter `fitdf` $= p + q$.

Exercise I

- The dataset *copper_prices.csv* includes monthly World copper prices from 1960 to 2006. Load the data into R.
- Discuss the stationarity of the time series using plots. Do you observe a trend, increasing variance or seasonality?
- If needed, apply an appropriate Box-Cox transformation.
- Estimate an ARIMA model on this time series. Choose the appropriate order of differencing and lags using ACF/PACF plots.
- Check your model (significance, residuals, ...). Run a Ljung-Box test for independence of residuals.
- The function `auto.arima` from the `forecast` package automates the selection of p, d, q (estimates up to 5 lags each, then uses AIC to select one). Use the function on the time series; is the same model recommended?