Module 9.3: Time Series Analysis
Fall Term 2022

Week 3:

ARMA Models

Lucerne University of Applied Sciences and Arts

Outline in Weeks

1. Introduction; Descriptive Modelling
2. Returns; Autocorrelation; Stationarity
3. ARMA Models
4. Unit Roots; Regressions between Time Series
5. Volatility Modelling
6. Value at Risk
7. Cointegration
8. Panel Data

Outline

White Noise and Random Walk

- Last week, we encountered two important stochastic processes: the *white noise* process

$$Y_t = U_t,$$

and the *random walk* process

$$Y_t = Y_{t-1} + U_t.$$

- Observe that we can unify the notation by writing

$$Y_t = \phi_1 Y_{t-1} + U_t,$$

where $\phi_1 = 0$ produces white noise, and $\phi_1 = 1$ results in the random walk.

The AR(1) Process

- In fact, there is no reason to restrict the parameter ϕ_1 to just these two values
- Allowing arbitrary values results in the *autoregressive model of order 1*, or *AR(1)* model.
- Usually, an intercept is also addded. The full model is then

$$Y_t = \alpha + \phi_1 Y_{t-1} + U_t, \quad \text{with } Y_t \text{ white noise.}$$

- We will see soon that the model is stationary if and only if $-1 < \phi_1 < 1$.
- The AR(1) process is a member of a very powerful class of models, the *autoregressive-moving average* (*ARMA*) models, with its special cases *autoregressive* (*AR*) and *moving average* (*MA*) models.
- Goal: find the right model (for forecasting etc.) by matching the correlogram.

Mean of Stationary AR(1)

- A first intuition for the *stationarity condition* is obtained if we try to find the (constant) mean and variance of Y_t. The mean of Y_t is to be solved from

$$\mathbb{E}[Y_t] = \alpha + \phi_1 \mathbb{E}[Y_{t-1}] + \mathbb{E}[U_t] = \alpha + \phi_1 \mathbb{E}[Y_t],$$

which implies

$$\mathbb{E}[Y_t] = \frac{\alpha}{1 - \phi_1},$$

and this requires $\phi_1 \neq 1$.

Variance of Stationary AR(1)

- Next, because $\{U_t\}$ is white noise, U_t is uncorrelated with Y_{t-1}, so

$$\mathrm{var}(Y_t) = \phi_1^2 \mathrm{var}(Y_{t-1}) + \mathrm{var}(U_t) + 2\phi_1 \mathrm{cov}(Y_{t-1}, U_t) = \phi_1^2 \mathrm{var}(Y_t) + \sigma^2,$$

  so that, if and only if $|\phi_1| < 1$,

$$\mathrm{var}(Y_t) = \frac{\sigma^2}{1 - \phi_1^2}.$$

- Note that $\mathrm{var}(Y_t) > \mathrm{var}(Y_{t-1})$ if $|\phi_1| \geq 1$, i.e., the variance grows without bounds in that case.

ACF / PACF of Stationary AR(1)

- It can be shown (see optional exercise) that if $|\phi_1| < 1$, then

$$\tau_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k,$$

i.e., if the process is stationary, then the ACF decays exponentially (or *geometrically*).

- The PACF satisfies

$$\tau_{11} = \phi_1 \quad \text{and} \quad \tau_{kk} = 0, k > 1;$$

i.e., it drops to zero after the first lag.

Summary: Moments of AR(1)

- In summary, we have obtained the following properties for the stationary AR(1) process:

Properties of AR(1)

- $\mathbb{E}(Y_t) = \dfrac{\alpha}{1-\phi_1}$;
- $\text{var}(Y_t) = \dfrac{\sigma^2}{1-\phi_1^2}$;
- $\tau_k = \phi_1^k, \quad k = 1, 2, \ldots$;
- $\tau_{kk} = \begin{cases} \phi_1, & k = 1, \\ 0, & k > 1. \end{cases}$

- Hence, if we find a geometrically declining acf, but a pacf which suddenly cuts off when $k > 1$, then we seem to have an AR(1) process.
- Try playing with this in the sheet “AR(1)“ in `simulation.xlsx`.

The Random Walk

The AR(1) process is non-stationary if:

- $\phi_1 = 1$:
  - now $Y_t = \alpha + Y_{t-1} + U_t$: a *random walk with drift*. It satisfies

$$\begin{aligned} \mathbb{E}[Y_t] &= Y_0 + \alpha t, \\ \mathrm{var}(Y_t) &= \sigma^2 t, \\ \mathrm{corr}(Y_t, Y_{t-k}) &= \sqrt{(t-k)/t}. \end{aligned}$$

  - Like for a random walk without drift, the correlogram typically stays close to 1, and decreases slowly, approximately linearly, while the first difference ΔY_t is stationary (see exercises).
- $\phi_1 > 1$: this is a so-called *explosive* process. The mean and variance increase very fast (exponentially). We usually do not consider this because it is unrealistic (at least for long periods of time).

AR(p) Models

- To fit more complicated ACF patterns, the AR(1) model can be extended to the *AR(p)* model if necessary:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + U_t.$$

- A *necessary* (but not sufficient) *condition for stationarity* is

$$\phi_1 + \phi_2 + \ldots + \phi_p < 1.$$

- If the model is stationary, then the mean is $\mathbb{E}[Y_t] = \alpha/(1 - \sum_{i=1}^{p} \phi_i)$.
- The ACF should gradually approach zero, but not necessarily with a clear pattern.
- The PACF satisfies

$$\tau_{kk} = 0, \quad k > p.$$

- Because all autoregressive models are basically regressions (with lagged variables as regressors), they can simply be estimated by ordinary least-squares as usual.

Example

Is the AR(2) process

$$Y_t = .5Y_{t-1} + .5Y_{t-2} + U_t$$

stationary?

Outline

The MA(1) Model

- The first-order *moving average model* — MA(1) — is given by

$$Y_t = \alpha + U_t + \theta_1 U_{t-1},$$

  where $\{U_t\}$ is again a white noise process.

- This process is stationary for all values of θ_1. It can be shown that

$$\begin{aligned} \mathbb{E}[Y_t] &= \alpha, \\ \mathrm{var}(Y_t) &= \sigma^2(1+\theta_1^2), \\ \tau_1 = \mathrm{corr}(Y_t, Y_{t-1}) &= \theta_1/(1+\theta_1^2). \end{aligned}$$

ACF and PACF of the MA(1) Model

- The most important property is that, for all $k > 1$,

$$\tau_k = \mathrm{corr}(Y_t, Y_{t-k}) = 0,$$

  i.e., the ACF has a *cut-off point*; shocks that happened longer than 1 period ago have no effect on Y_t.
- Also, the PACF of an MA(1) process is $\tau_{kk} = -(-\theta_1)^k \rightarrow$ *geometric decay*.
- Note that the patterns of the ACF and PACF are *reversed* compared to the AR(1) process!
- We will use this to *identify* the correct model from the SACF/SPACF.

MA(q) process

- Again we may generalize to the MA(q) model

$$Y_t = \alpha + U_t + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q}.$$

- Now $\tau_k = 0$ for $k > q$, and τ_{kk} still decays exponentially.
- This will again be used to identify the correct model from the correlogram.

Some more properties of MA(q) Models

- *Estimation*: since the past values of U_t are unobserved, we cannot estimate the coefficient θ_1 by OLS. Eviews has a built-in method, based on *non-linear least-squares*.
- *Stationarity*: All MA(q) processes are stationary, regardless of their parameter values.

ARMA(1, 1) Process

- The most general class of processes is the mixed autoregressive-moving average (*ARMA*) class of models.
- The simplest one is the ARMA(1,1):

$$Y_t = \phi_1 Y_{t-1} + U_t + \theta_1 U_{t-1}.$$

- The model is stationary if $|\phi_1| < 1$.
- The ACF turns out to be

$$\begin{aligned} \tau_1 &= \frac{(1+\phi_1\theta_1)(\phi_1+\theta_1)}{1+\theta_1^2+2\phi_1\theta_1}, \\ \tau_k &= \phi_1^{k-1}\tau_1, \quad k > 1 \rightarrow \textit{geometric decay}. \end{aligned}$$

- The PACF is also gradually declining (no cut-off point).

ARMA(p, q) Process

- The general ARMA(p, q) model is

$$Y_t = \alpha + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} U_t + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q}.$$

- Like for the ARMA(1, 1), both its ACF and PACF decay exponentially (provided it is stationary), without a clear pattern.
- *Estimation*: Like pure MA models, ARMA models cannot be estimated by OLS because the past values of U_t are unobserved. Eviews has built-in methods however.
- *Stationarity*: The stationarity or otherwise of an ARMA process depends only on the AR component, i.e., a necessary condition is that

$$\phi_1 + \phi_2 + \ldots + \phi_p < 1.$$

A Theorem

- The following theorem explains why ARMA processes are so important:

Theorem

Any stationary process can be represented as an ARMA process.

- So as long as we can find the right model, we can predict *any* stationary process[1]!
- The properties of the ACF and PACF, summarized on the next slide, will help us do this.

[1] And even integrated processes, by taking the first difference and modelling that.

Summary of Properties

- We have found the following properties of the different processes:
  - AR(p): geometrically decaying ACF, PACF is zero after p lags;
  - MA(q): ACF is zero after q lags, geometrically decaying PACF;
  - ARMA(p, q): geometrically decaying ACF and PACF.
- So we can infer the model type and order from the SACF/SPACF for pure AR and MA models.
- A full ARMA model would be required if both SACF and PACF decline geometrically, but we won't be able to infer the orders then.
- The usual procedure is to try an ARMA(1,1) in that case, and test whether the model needs to be extended.
- The above is called the *Box-Jenkins* approach.

Outline

Box-Jenkins

- In their 1970 textbook, Box and Jenkins proposed an approach to empirical ARMA modelling that soon became and still is the predominant approach to univariate time-series analysis and forecasting. Their procedure consists of three steps:
  1. **Identification**. This refers to the problem of selecting an initial ARMA model, i.e., the choice of the orders p and q. This is based on inspection of the graph and the correlogram of Y_t. Also, so-called *information criteria* may be used for model selection.
  2. **Estimation**. The unknown autoregressive and moving average parameters, as well as the variance σ^2 of the disturbances, need to be estimated.
  3. **Diagnostic checking**. A correctly specified model should not display any autocorrelation in the residuals. Therefore, the main model check is a test for residual autocorrelation. Also other misspecification tests (heteroskedasticity, normality, structural change) may be used.
- If we find some problem with the model in Step 3, then we return to Step 1 and go through the cycle again, until the tests indicate no further problem.

Identification: Stationarity

- Before even choosing p and q, it must be ensured that the data are *stationary*. An integrated time series displays very large autocorrelations, which converge to zero only slowly. In contrast, stationary time series have an autocorrelation function that decays exponentially, or is not significantly different from zero after a few lags.
- Additional information is available from the graph of the time series. Stationary time series should display *mean-reversion*, i.e., they should fluctuate around a constant mean (or a linear trend, in case of trend-stationarity). If a series does not display this property, and behaves more like a random walk, then it may not be stationary.
- A formal procedure to test this (and how to proceed in case of non-stationarity) will be introduced later.

Identification: p and q

- The second step is choosing p and q.
- Recall that the correlograms of AR, MA and ARMA processes are characterized thusly:
  - AR(p): geometrically decaying ACF, PACF is zero after p lags;
  - MA(q): ACF is zero after q lags, geometrically decaying PACF;
  - ARMA(p, q): geometrically decaying ACF and PACF.
- If neither ACF nor PACF have a clear cut-off point, start with an ARMA(1,1), estimate it, and test whether the model needs to be extended.
- General goal: *parsimony*. Find the smallest possible model that describes the data well.

Estimation

- Eviews has built-in routines for estimating ARMA models.
- The specification of, e.g., an ARMA(3, 2) model for a time series $\{y_t\}$ is (see exercises)

Y c ar(1) ar(2) ar(3) ma(1) ma(2)

- Note: for pure AR models, we may also use (e.g., for AR(2)):

Y c Y(-1) Y(-2)

- The difference lies in the interpretation of the constant term: Somewhat confusingly, when using the "`ar`" specification, then the estimated coefficient `c` corresponds to $\mathbb{E}(Y_t) = \alpha/(1 - \sum_{i=1}^{p} \phi_i)$ rather than α.
- Yet in practice the "`ar`" specification is preferred, because only in that case does EViews compute the residual Q-tests appropriately.

Diagnostic Testing

- After estimating our favorite ARMA model, we obtain the *residuals* $\hat{u}_t$. These residuals should look like *white noise*.
- If there is significant autocorrelation left in $\hat{u}_t$, we should *extend* the model. E.g.,, if the residuals of an AR(1) model look like an MA(1) process, then we might try an ARMA(1,1) model instead.
- The most often used test for residual autocorrelation in ARMA models is the Ljung-Box Q-statistic, based on the residuals $\hat{u}_t$ instead of the original time series y_t.
- It can be shown that if $\hat{u}_t$ is a residual from an ARMA(p, q) model, then the Q-statistic with m correlations has an approximate χ^2_{m-p-q} distribution under the null hypothesis (for m large enough). This explains why in EViews, the first $(p+q)$ Q-statistics have no p-value.
- The tests cannot reliably be used to find out in which direction the model should be *extended*. This means that we may have to try different alternative specifications before we find a satisfactory choice.

Model Selection Criteria

- If more than one specification passes the diagnostic tests (e.g., both an AR(2) and an ARMA(1,1)), then the decision is often based on the *Akaike information criterion* (AIC) and the *Schwarz criterion* (SC; a.k.a. the *Bayesian* information criterion, BIC):

$$
\begin{aligned}
AIC &= 1 + \log 2\pi + \log\left(\frac{1}{n}\sum_{t=1}^{n}\hat{u}_t^2\right) + \frac{2k}{T}, \\
SC &= 1 + \log 2\pi + \log\left(\frac{1}{n}\sum_{t=1}^{n}\hat{u}_t^2\right) + \frac{k\log T}{T}.
\end{aligned}
$$

- $k =$ number of parameters ($p + q + 2$ including intercept and σ_u^2).
- Constant $1 + \log 2\pi$ is irrelevant in comparisons and sometimes deleted.
- We choose the model with the smallest AIC or SC. Typically AIC leads to higher p and q than SC.
- The idea is similar to the adjusted R^2: the second term is a penalty for including too many parameters. Trade-off between *goodness of fit* and *parsimony*. Smaller models are often better for forecasting.

Outline

Forecasting Terminology

- The main purpose of ARIMA models is forecasting.
- Denote the forecast of y_{t+s} based on information available at time t as $f_{t,s}$.
- We distinguish *in-sample* and *out-of-sample* forecasts.
- In-sample means that the same data are forecasted that were used for estimating the model.
- Out-of sample forecasting means that the sample is divided into two subsamples, say $1, \ldots, T_1$ (the *estimation period*) and $T_1 + 1, \ldots, T$ (the *holdout period*). The former is used for selecting a model and estimating its parameters. Based on this information at time T_1, we forecast $y_{T_1+1}, \ldots, y_T$.

One-Step Ahead Forecasts

- Stationary implies that for, e.g., stationary ARMA$(1,1)$ model,

$$Y_{t+1} = \alpha + \phi_1 Y_t + U_{t+1} + \theta_1 U_t.$$

- Strategy: replace parameters with estimates, errors U_t with residuals $\hat{u}_t$, and unobserved future U_t with zero. Hence

$$f_{t,1} = \hat{\alpha} + \widehat{\phi_1} y_t + 0 + \theta_1 \hat{u}_t.$$

- Analogous procedure for general ARMA(p, q) models.

Multi-Step Forecasts

- When forecasting multiple steps ahead, not only the future U_t are unknown, but also the future y_t.
- Solution: recursive procedure. Forecast y_{t+1} first, then use this to forecast y_{t+2}, etc. All future errors are replaced with zero.

Remarks

- EViews distinguishes between *static* and *dynamic* forecasts. This is essentially the same as the difference between one-step and multi-step forecasts: "static" uses information up to time $t-1$ to forecast y_t, for any $t > T_1$; this is only possible for in-sample forecasting. "Dynamic"uses only the information up to T (or T_1) to forecast $T+s$ (or T_1+s).

Outline

Learning Goals

Students are able to

- understand the essential characteristics of AR, MA, and ARMA processes;
- identify the correct model type from the correlogram;
- estimate ARMA models with EViews;
- apply diagnostic tests to evaluate a fitted model;
- choose between competing models using information criteria, and
- use the final model for forecasting.

Homework

- Exercise 3
- Questions 3 and 6–12 from Chapter 6 of Brooks (2019)