Time Series

Peter Lukianchenko

14 January 2023

Today

- Time series: definition
- Structural/non-structural modeling
- Stationarity, weak and strong
- Autocovariance, autocorrelation, partial autocorrelation
- Lag operator
- ARMA models (beginning)
- Tsay "Analysis of Financial Time Series." (1.2, 2.1-2.6)

  Hamilton "Time Series Analysis" (2.1, 3.1-3.5)

  Stock and Watson "Introduction to Econometrics" (14.1, 14.2)

  Diebold "Forecasting" (online version:

  http://www.ssc.upenn.edu/ fdiebold/Teaching221/Forecasting.pdf

  (6.5, 7.1, 7.2)

Time Series

Cross-sectional data:

- The sample is **i.i.d.** (or at least independent)
- Useful for answering questions about **causal effects** of one variable on another

Time Series:

- The sample is **not i.id.**, observe variable(s) over time
- Useful for answering questions about **dynamic causal effects**
- Useful for forecasting **future** values of a variable

Time Series

Cross-sectional data:

- The sample is **i.i.d.** (or at least independent)
- Useful for answering questions about **causal effects** of one variable on another

Time Series:

- The sample is **not i.id.**, observe variable(s) over time
- Useful for answering questions about **dynamic causal effects**
- **Useful for forecasting future values of a variable**

Time Series

Useful for forecasting future values of a variable

We will study quantitative (non-structural) models of time series to use them later for forecasting:

Structural

- Has some **economic theory** behind it
- Parameters **have meaning and causal interpretation**
- Will briefly touch in VAR and ADL topic

Non-structural

- Models based on **fitting data**
- Coefficients **do not have causal interpretation**
- Will be the main topic of the course

Time Series: definition

- **Informaly:** a set of realizations of a random variable ordered according to time
- **Formally:**

Definition 1

Collection of random variables defined on the sample space $\{Y_t, t \in \boldsymbol{T}\}$ is called a *stochastic process*

We will consider $\boldsymbol{T} = \{\dots, -1, 0, 1, 2, \dots\} = \mathbf{Z}$

Definition 2

A *time series* is a realization of a stochastic process: $\{y_t, t \in \boldsymbol{Z}\}$

Definition 3

A *time series sample* is $\{y_t, t = 1, \dots, T\}$ for some $T < \infty$.

But 'time series' can be used as a synonym of 'stochastic process'

Important concepts

- **Goal:** forecast values of a random variable using the time series sample
- So, we need the future to be like the past
- Reflected in the concept of ***stationarity***

Stationarity

Definition 4

A process $\{Y_t, t \in Z\}$ is *strictly stationary* if, for any k, s and any $t_1, \dots, t_k$, the *distributions* of $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k})$ and $(Y_{t_1+s}, Y_{t_2+s}, \dots, Y_{t_k+s})$ are *the same*.

In other words, the following distributions are the same:

- of Y_1 and Y_{100}
- of (Y_1, Y_2) and (Y_5, Y_6)
- of (Y_3, Y_{10}, Y_{22}) and (Y_{13}, Y_{20}, Y_{32})
- and so on ...

Stationarity

Strict stationarity is a complicated concept

$\Delta y_t = y_t - y_{t-1}$

Very often people consider *weak stationarity*

$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$

Definition 5

A process $\{Y_t, t \in Z\}$ is *weakly, or covariance-, stationary* if, for any $t_1, t_2, s \in \boldsymbol{Z}$

- $E[Y_{t_1}] = E[Y_{t_2}]$, $Var(y_{t_i}) = Var(y_{t_j})$
- $Cov\left(Y_{t_1}, Y_{t_1+s}\right) = Cov\left(Y_{t_2}, Y_{t_2+s}\right)$

So, only the following has to be the same:

- mean of all Y_t
- variance of all Y_t
- covariances between all of the possible pairs of Y_t that are fixed number of periods away from each other

Stationarity

Question

If $\{Y_t, t \in \boldsymbol{Z}\}$ is *weakly stationary*, is it also *strictly stationary*?

Stationarity

Question

If $\{Y_t, t \in \boldsymbol{Z}\}$ is *strictly* *stationary*, is it also *weakly* *stationary*?

Stationarity: extra remarks

Not all *weakly stationary* process *strictly stationary.*

But if $\{Y_t\}$ is gaussian, then it is *weakly stationary*, it is also *strictly stationary*.

Autocovariance and autocorrelation function

- Want to forecast future by exploring the relation between r.v. corresponding to consecutive periods of time
- Autocovariance is a way to quantify this relation

Definition 6

- Autocovariance of order k is $\gamma(k) = Cov(Y_t, Y_{t+k})$
- Autocorrelation of order k is $\rho(k) = corr\ (Y_t, Y_{t+k}) = \frac{\gamma(k)}{Var(Y_t)} = \frac{\gamma(k)}{\gamma(0)}$
- $\gamma(\cdot)$ is called *autocovariance function* (ACF)
- $\rho(\cdot)$ is called *autocorrelation function* (also ACF)

Estimated ACF

$$\overline{Y_T} = \frac{1}{T}\sum_{t=1}^{T} Y_t$$

$$\widehat{\gamma(0)} = \frac{1}{T}\sum_{t=1}^{T}(Y_t - \overline{Y_T})^2$$

$$\widehat{\gamma(k)} = \frac{1}{T}\sum_{t=k+1}^{T}(Y_t - \overline{Y_T})(Y_{t-k} - \overline{Y_T})$$

Sample autocorrelation function:

$$\hat{\rho}(k) = \frac{\widehat{\gamma(k)}}{\widehat{\gamma(0)}}$$

Correlogram: a graph of sample ACF

ACF / PACF

Partial Autocorrelation Function (PACF)

- Autocorrelation measures how dependent the data is
- If Y_1 and Y_2 are related, and Y_2 and Y_3 are related, then Y_1 and Y_3 have to be related at least indirectly
- *Partial Autocorrelation Function (PACF)* measures direct relation between

different Y_t.

Definition 6

Partial Autocorrelation Function (PACF) at lag k is

$$\alpha(k) = corr(\, Y_1 - P(1, Y_2, \ldots, Y_k)Y_1, \quad Y_{k+1} - P(1, Y_2, \ldots, Y_k)Y_{k+1}),$$

where $P(1, Y_2, \ldots, Y_k)Y_j$ is the linear projection of Y_j on a constant, $Y_2, \ldots,$ and Y_k.