Time Series

Peter Lukianchenko

10 January 2022

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 $\boldsymbol{T} = \{\ldots, -1, 0, 1, 2, \ldots\} = \mathbf{Z}$

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 $\{y_t,\ t = 1, \ldots, T\}$ 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 $t_1,\ \ldots,\ t_k$, the *distributions* of $(Y_{t_1}, Y_{t_2},\ \ldots, Y_{t_k})$ and $(Y_{t_1+\ s}, Y_{t_2+s},\ \ldots, Y_{t_k+s})$ 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

Very often people consider *weak stationarity*

Definition 5

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

1) $E\left[Y_{t_1}\right] = E\left[Y_{t_2}\right]$,

2) $Cov\left(Y_{t_1},\ Y_{t_1+s}\right) = Cov\left(Y_{t_2},\ Y_{t_2+s}\right)$ $\neq f(t_1)$ $\forall s$

So, only the following has to be the same:

- mean of all Y_t $E[Y_t]$
- variance of all Y_t $Var(Y_t)$
- covariances between all of the possible pairs of Y_t that are fixed number of periods away from each other $Cov(Y_t, Y_{t+s})$ $\forall s$

MUST BE indep. on t

Stationarity

Question

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

Stationarity

Question

If $\{Y_t,\ t \in \mathbb{Z}\}$ 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

$\gamma(0) = Var(y_t)$
$\gamma(1) = Cov(y_t, y_{t+1})$

Definition 6

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

Estimated ACF

$$\bar{Y}_T = \frac{1}{T} \sum_{t=1}^{T} Y_t$$

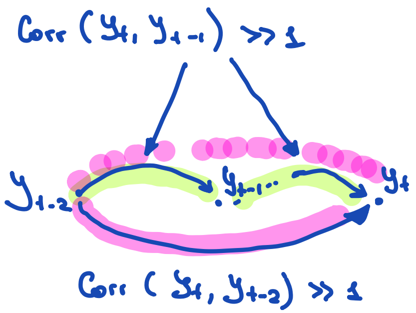
$$\hat{\gamma}(0) = \frac{1}{T} \sum_{t=1}^{T} (Y_t - \bar{Y}_T)^2$$

$$\hat{\gamma}(k) = \frac{1}{T} \sum_{t=k+1}^{T} (Y_t - \bar{Y}_T)(Y_{t-k} - \bar{Y}_T)$$

Sample autocorrelation function:

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

Correlogram: a graph of sample ACF



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

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

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

Partial Autocorrelation Function (PACF)

Write the linear projection of $Y_{t+k+1} - \mu$ on $Y_{t+k} - \mu, \ \ldots, \ Y_{t+1} - \mu$:

$$\hat{Y}_{t+k+1} - \mu = \alpha(1)\left(Y_{t+k} - \mu\right) + \ \ldots + \alpha(k)Y_{t+1} - \mu)$$

PACF of order 1 to k can be found as

$$\begin{pmatrix} \alpha(1) \\ \alpha(2) \\ \cdots \\ \alpha(k) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \ldots & \gamma(k) \\ \gamma(1) & \gamma(0) & \ldots & \gamma(k-2) \\ \cdots & \cdots & \cdots & \cdots \\ \gamma(k-1) & \gamma(k-2) & \ldots & \gamma(0) \end{pmatrix}^{-1} \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \cdots \\ \gamma(k) \end{pmatrix}$$

Partial Autocorrelation Function (PACF)

Write the linear projection of $Y_{t+k+1} - \mu$ on $Y_{t+k} - \mu, \ \dots, \ Y_{t+1} - \mu$:

$$\hat{Y}_{t+k+1} - \mu = \alpha(1)\left(Y_{t+k} - \mu\right) + \ \dots + \alpha(k)(Y_{t+1} - \mu)$$

PACF of order 1 to k can be found as

$$\begin{pmatrix} \alpha(1) \\ \alpha(2) \\ \dots \\ \alpha(k) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(k) \\ \gamma(1) & \gamma(0) & \dots & \gamma(k-2) \\ \dots & \dots & \dots & \dots \\ \gamma(k-1) & \gamma(k-2) & \dots & \gamma(0) \end{pmatrix}^{-1} \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \dots \\ \gamma(k) \end{pmatrix}$$

Sample PACF: estimate by OLS

$$Y_{t+k+1} - \mu = \alpha(1)\left(Y_{t+k} - \mu\right) + \ \dots + \alpha(k)(Y_{t+1} - \mu) + \varepsilon_{t+k+1}$$

Lag operator (a.k.a back-shift operator)

Definition 7

The lag operator L is a linear operator such that for all t

$LY_t = Y_{t-1}$ $\quad L(y_t) = y_{t-1}$

Properties:

$LLy_t = y_{t-2}$

- $L^2Y_t = L(LY_t) = LY_{t-1} = Y_{t-2}$
- $L^jL^iY_t = L^{j+i}Y_t = Y_{t-i-j}$
- $Lc = c$
- $(L^j + L^i)Y_t = Y_{t-i} + Y_{t-j}$
- L to a negative power is a *lead operator:* $L^{-i}Y_t = Y_t$
- For $|a| < 1$,

$y_t - y_{t-1} + y_{t-2} - y_{t-3} + y_{t-4} = y_t(1 - L + L^2 - L^3 + L^4)$

? if $1 - aL = 0$

$$y_t\left[\underbrace{1 + aL + a^2L^2 + \ldots + a^kL^k}_{y_t + ay_{t-1} + a^2y_{t-2} + \ldots + a^k y_{t-k}} + \ldots\right] = (1 - aL)^{-1}Y_t = \frac{Y_t}{1 - aL}$$

ARMA models

Starting with basics

- $\{\varepsilon_t,\ t \in Z\}: \varepsilon_t$ **are iid**
  - is it stationary?
- MDS (Martingale Difference Sequence): $\{\varepsilon_t,\ t \in Z\}$:

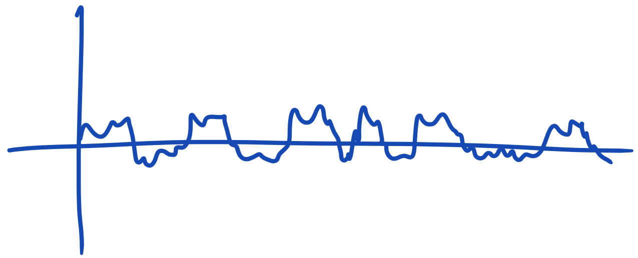
$$E[\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] = 0$$

- A process like that is closer to economics: many dynamic optimization problems result in a condition of this type.
- White noise: (A more statistical description of innovations) $\{\varepsilon_t,\ t \in Z\}$:

s.t.

$Cov(\varepsilon_t, \varepsilon_s) = 0,\ for\ any\ t \neq s$

$E[\varepsilon_t] = 0,\ \cancel{Var(\varepsilon_t)\sigma^2 < \infty}$ $Var(\varepsilon_t) = \sigma^2 < \infty$



Moving-Average Models

- Start with $\{\varepsilon_t\}$, a white noise.

- **MA(1) – Moving average of order 1**

$$Y_t = \varepsilon_t + \varphi\varepsilon_{t-1}$$

$y_t = \varepsilon_t + \varphi \cdot \varepsilon_{t-1}$

- **MA(q) – Moving average of order q**

$$Y_t = \varepsilon_t + \varphi_1\varepsilon_{t-1} + \varphi_2\varepsilon_{t-2} \ldots + \varphi_q\varepsilon_{t-q}$$