Time Series Analysis

Prof. Lea Petrella

MEMOTEF Department
Sapienza, University of Rome, Italy

Class 2

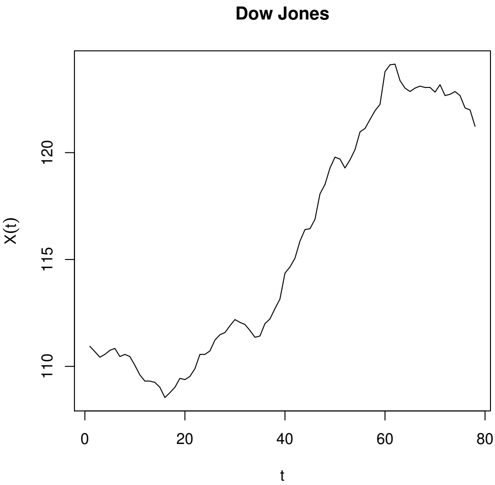

Figure: Dow Jones: average growth.

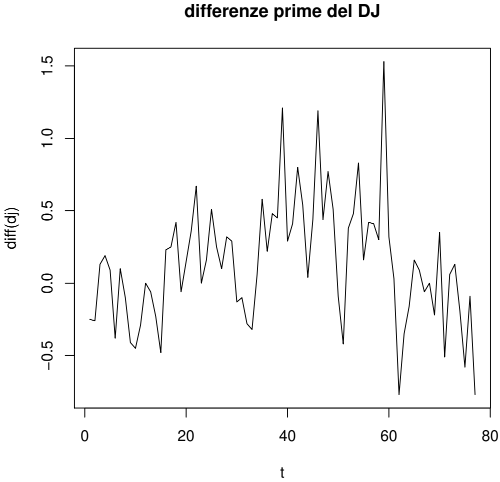

Figure: First differences of Dow Jones: stationary.

- Dow Jones is a stock index that allows to evaluate the overall performance of stock markets.
- It is approximately computed as an average of the 30 most capitalized stocks (although does not account for the different weights of those stocks.

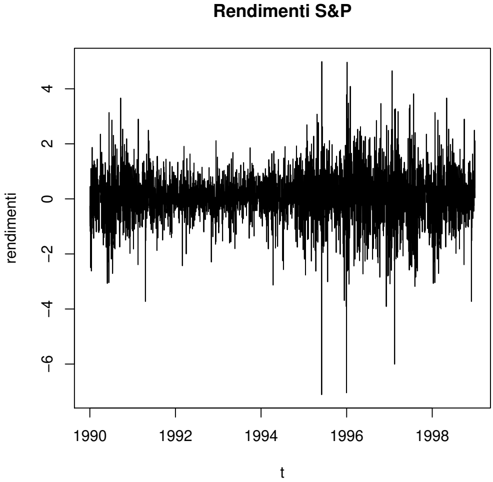

Figure: S&P 500 returns from 1990 to 1999. High volatility.

- Standard and Poor's is a financial company that discloses some indexes about the US financial market.
- The S&P 500 is computed as a weighted arithmetic mean of 500 stocks of US high capitalized companies.
- Returns are computed: $\log \frac{p_t}{p_{t-1}} = \log p_t - \log p_{t-1}$.

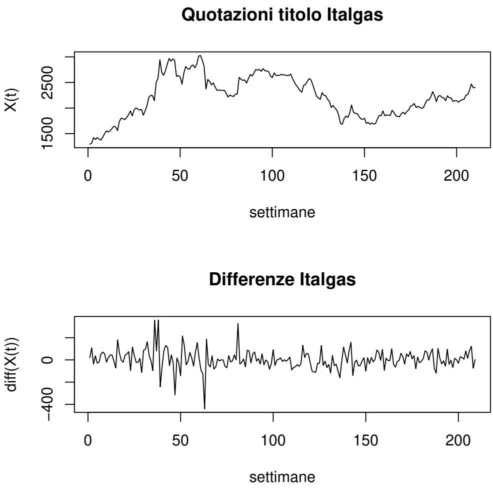

Figure: Close prices for ITALGAS between July 1985 and June 1989. Series and difference series.

Stationary property

- Stationarity represents a link among past, present and future.
- Intuitively, a process $\{X_t, t = 0, 1 \dots\}$ is said to be stationary if it possesses statistical properties similar to the shifted process $\{X_{t+h}, t = 0, 1 \dots\}$, where h is a positive or negative integer.
- Def. A process is said to be strong stationary if the joint distribution $(X_{t_1}, \dots, X_{t_k})$ is the same of the joint distribution $\left(X_{t_{1+h}}, \dots, X_{t_{k+h}}\right)$ for any $h, k > 0$.
- This property represents a kind of homogeneity on the probabilistic structure of the process with respect to time.
- Example. We have a strong stationary process if $(X_{1985}, X_{1986}, X_{1987})$ has the same distribution as $(X_{1990}, X_{1991}, X_{1992})$, that is $(X_{1985+5}, X_{1986+5}, X_{1987+5})$.

- However, such property is very difficult to be verified !
- We prefer to require a lighter concept of homogeneity based on the second moment.
- A process X_t is said to be weakly stationary if:

1. $\mathbb{E}(X_t) = \mu$, the expected value is constant $\forall t$,
2. $\mathbb{V}ar(X_t) : \gamma(0) = \sigma^2 < \infty$, the variance is constant $\forall t$,
3. $Cov(X_t, X_{t+h}) = \gamma(h)$, the variance at lag h only depends on the distance h, not on t, $\forall t$.

- An immediate implication is the fact that the correlation structure of the variables does not change in time with respect to the same lag h.
- Strong stationary $\Longrightarrow$ Weak stationary. Viceversa not true.
- The first two moments do not identify a process except for the Gaussian case. Therefore, for a Gaussian process strong and weak stationary property coincide.

Autocorrelation function

- The autocovariance function of X_t is defined as

$$\gamma(h) = Cov(X_{t+h}, X_t), \qquad h \in \mathbb{R}.$$

- The autocorrelation function (ACF) of X_t is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = Corr(X_{t+h}, X_t), \qquad h \in \mathbb{R}.$$

- The ACF measures the correlation (linear dependence) between values of the process at different lags h, and it indicates the amplitude and length of the memory of the process.
- $\gamma(h) = \gamma(-h)$ and $\rho(h) = \rho(-h)$, i.e., the autocovariance function and ACF are even functions (symmetric respect to zero) as the distance in time between X_t e X_{t-h} is the same as X_t e X_{t+h}. For this reason the ACF is usually plotted for positive lags.
- Since $\rho(h)$ is a correlation, it follows $-1 < \rho(h) < 1$.
- Intuition suggests that for a stationary process $\rho(h) \to 0$ (fast or slow) as h increases, otherwise the process would explode.

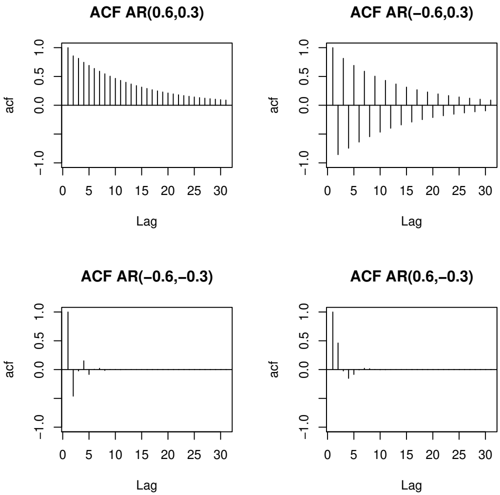

Figure: Examples of ACF.

Partial autocorrelation function

- Besides the ACF a quantity of interest is the correlation between X_t e X_{t+h} netting for the effect of intermediate variables $X_{t+1}, X_{t+2}, \ldots X_{t+h-1}$. This is the partial autocorrelation function (PACF).
- The PACF measures the autocorrelation between X_t e X_{t+h} after removing their linear dependence with the other intermediate variables (recall the partial correlation coefficient in multiple regression).

- It takes to compute

$$\phi_{hh} = Corr(X_t, X_{t+h}|X_{t+1}, X_{t+2}, \ldots X_{t+h-1}).$$

- PACF can be derived as follows: consider a regression model where the dependent variable X_{t+h} is regressed against $X_{t+h-1}, X_{t+h-2}, \ldots X_t$, i.e.,

$$X_{t+h} = \phi_{h1}X_{t+h-1} + \phi_{h2}X_{t+h-2} + \ldots + \phi_{hh}X_t + e_{t+h},$$

  where ϕ_{hj} represents the parameter of the regression of X_{t+h} with respect to the variable X_{t+h-j}, and e_{t+h} is the shock uncorrelated with X_{t+h-j} for $j \geq 1$.
- Note that we are considering a zero mean process.

- Multiplying both sides by X_{t+h-j} and taking the expected values, we get:

$$\gamma(j) = \phi_{h1}\gamma(j-1) + \phi_{h2}\gamma(j-2) + \ldots + \phi_{hh}\gamma(j-h)$$

thus,

$$\rho(j) = \phi_{h1}\rho(j-1) + \phi_{h2}\rho(j-2) + \ldots + \phi_{hh}\rho(j-h).$$

- For $j = 1, 2, \ldots, h$ we obtain the following system of equations, known as Yule-Walker equations:

$$\begin{aligned} \rho(1) &= \phi_{h1}\rho(0) + \phi_{h2}\rho(1) + \ldots + \phi_{hh}\rho(h-1) \\ \rho(2) &= \phi_{h1}\rho(1) + \phi_{h2}\rho(0) + \ldots + \phi_{hh}\rho(h-2) \end{aligned} \tag{1}$$

$$\rho(h) = \phi_{h1}\rho(h-1) + \phi_{h2}\rho(h-2) + \ldots + \phi_{hh}\rho(0)$$

- It can be shown that, after some computations for $h = 1, 2, \ldots$ we get:

$$\phi_{11} = \rho(1)$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}}$$

$$\ldots$$

$$\phi_{hh} = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(h-2) & \rho(1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(h-3) & \rho(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \rho(h-3) & \dots & \rho(1) & \rho(h) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(h-2) & \rho(1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(h-3) & \rho(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \rho(h-3) & \dots & \rho(1) & 1 \end{vmatrix}}$$

- ACF and PACF are often unknown. We will show how to estimate them. Clearly, to estimate $\hat{\phi}_{hh}$ it takes to estimate $\hat{\rho}(h)$.

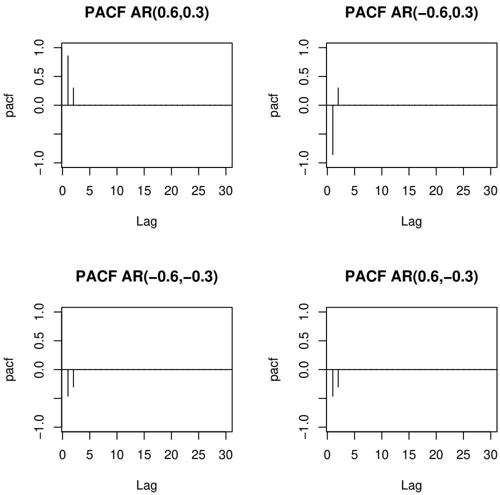

Figure: Examples of PACF.

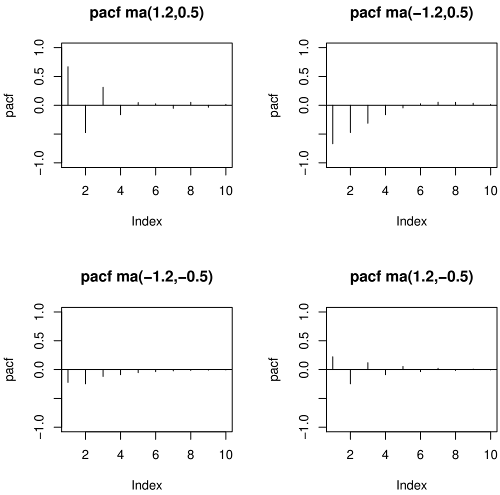

Figure: Examples of PACF.

- Given a stationary process it is possible to compute a unique ACF and PACF.
- We can ask the following: given an ACF is the process that possess that ACF unique ?
- The answer is negative, in general because the ACF does not fully characterize a process. It can be shown that there may exists more proceses with same ACF
- The answer is positive, instead, if besides the stationarity we require the invertibility conditions.

- Invertibility relates to the possibility of expressing the process X_t as a function of past random variables.
- More formally, a process $\{X_t\}$ is said to be invertibile if it exists a linear function $h(\cdot)$ and a white noise $\{\epsilon_t\}$, such that for each t:

$$X_t = h(X_{t-1}, X_{t-2}, \cdots) + \epsilon_t.$$

- We will discuss the role of ACF and PACF in finding the model or the underlying process for the observed time series.
- Depending on the specified model we will have different ACF and PACF.

- Example 1. White Noise. The easiest stationary process is the White Noise: $\{X_t\}$ is a sequence of uncorrelated random variables, $Cov(X_t, X_{t+h}) = 0 \; \forall h \neq 0$, zero mean a variance equal to σ^2.
- Therefore,

$$\rho(h) = 1 \quad per \quad h = 0$$

$$\rho(h) = 0 \quad per \quad h \neq 0.$$

- Note that the PACF and ACF coincide since the components are serially uncorrelated.
- We write $X_t \sim WN(0, \sigma^2)$.
- *WN* is a benchmark to assess if the bserved series shows autocorrelation. That is, if a series is autocorelated we can compare it to the ACF of a White Noise.

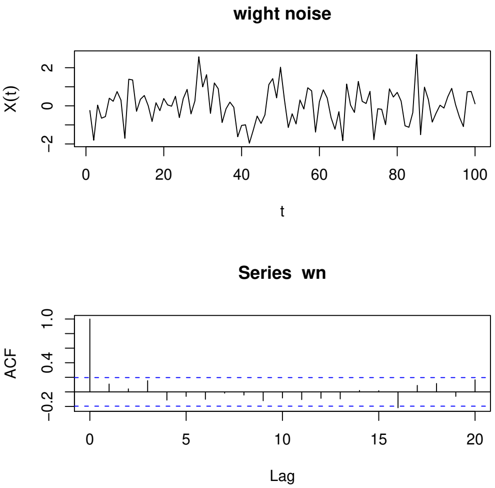

Figure: Representation of a WN series and its ACF.

- Note that Gaussian assumption is not required for a process to be *WN*.
- When adding the Gaussian hypothesis we end up with a sequence of random variables

$$X_t \overset{i.i.d}{\sim} N(0, \sigma^2)$$