Econometría (II / Práctica)

Magíster en Economía

Tema 7: Breve Introducción a Series de Tiempo (TS)

Prof. Luis Chancí

www.luischanci.com

Apuntes Econometría Prof. Luis Chancí

$E(Y_i|X_i = x) = \int z\, f(z|X_i = x)dz$

Introduction

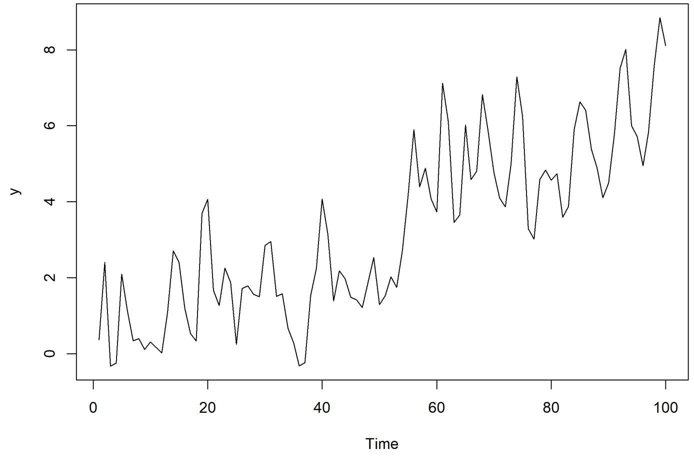
Previously: we were interested in $\mathbb{E}(y_i|x) = f(x_i)$ using cross-sectional data.

In this section: we will (briefly) introduce methods for time series analysis.

That is, a set of observations $y_1, \ldots, y_t, \ldots, y_T$ with t as the time index. We will focus on discrete time series, with a natural temporal ordering, $1 < \ldots < t < \ldots < T$), where y_{t-1} is realized when y_t is determined.

Most data in macroeconomics and finance come in this form.

A researcher's interest typically lies in modeling, forecasting, and studying the effects of shocks (and whether these effects will dissipate).



Time Series as a Stochastic Process

A time series is a stochastic process (the sequence of random variables $\{Y_t\}$), where observations close in time will be dependent. Their study requires a different distributional theory than the one we used in cross-sectional.

- While a deterministic process will always produce the same output from a given starting condition, a stochastic process has some indeterminacy that relates to the future evolution of the process.
- **Stochastic Process:** the probability law governing $\{Y_t\}$.
- **Realization:** One draw from the process (would be $\{y_t\}$).

For instance, 'if we could re-run history', one result would be $\sum_r y_r/M$ and another $\sum_t y_t/T$,

Stochastic process	$Y_1,$	$\ldots$	$Y_t,$	$\ldots$	Y_T
Realization 1:	$y_1^{(1)},$	$\ldots$	$y_t^{(1)},$	$\ldots$	$y_T^{(1)}$
	$\vdots$		$\vdots$		$\vdots$
Realization r:	$y_1^{(r)},$	$\ldots$	$y_t^{(r)},$	$\ldots$	$y_T^{(r)}$
	$\vdots$		$\vdots$		$\vdots$
Realization R:	$y_1^{(R)},$	$\ldots$	$y_t^{(R)},$	$\ldots$	$y_T^{(R)}$

Introduction (cont.)

We will review:

1. Univariate Time Series (single, scalar, observations recorded sequentially over equal time increments). For instance, $Y_t = 0.7Y_{t-1} + u_t$.

2. Non-Stationary Time Series (unit root). For instance, $Y_t = Y_{t-1} + u_t$.
   (+ a note on models for the Variance of a Time Series ARCH/GARCH, for instance, $\sigma_t^2 = 0.2Y_{t-1}^2$).

3. Vector Autoregressive models (VAR).
   For instance, $Y_t = 0.7Y_{t-1} + 0.2X_{t-1} + u_t$ and $X_t = 0.3X_{t-1} + 0.1Y_{t-1} + \nu_t$.

Concepts

Before we start, some concepts:

- Conditional first moment: $\mathbb{E}(Y_t|Y_{t-1}) \equiv f(y_{t-1})$.
- **Autocovariances:** $\gamma_{t,k} = cov(Y_t, Y_{t+k})$; $\gamma_0 = Var(Y_t) = \mathbb{E}(Y_t - \mathbb{E}(Y_t))^2$.
- **Autocorrelations:** $\rho_{t,k} = cor(Y_t, Y_{t+k})$.
- **Strict Stationarity (strong):** The process is strictly stationary if the probability distribution of $(Y_t, Y_{t+1}, \ldots, Y_{t+k})$ is identical to the probability distribution of $(Y_\tau, Y_{\tau+1}, \ldots, Y_{\tau+k})\ \forall\ t, \tau, k$ (joint distributions are time invariant).
- **Covariance Stationarity (weak):** The process is covariance stationary if $\mathbb{E}(Y_t) = \mu_t = \mu = \text{cons}$ and $\gamma_{t,k} = \gamma_k\ \forall\ t, k$ (mean and autocovariances are time invariant).

The central point will be whether the TS of interest are stationary or not (e.g., whether the series will return to its mean after a shock). This determines the technique to use.

Concepts (cont.)

- **White noise.** serially uncorrelated random variables with zero mean and finite variance. For example, the Gaussian white noise process, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, which implies $\mathbb{E}(\varepsilon_t) = \mathbb{E}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2} \dots) = 0$, $\mathbb{E}(\varepsilon_t \varepsilon_{t-j}) = Cov(\varepsilon_t \varepsilon_{t-j}) = 0$, and $\mathbb{E}(\varepsilon_t^2) = Var(\varepsilon_t) = \sigma^2$ (cons.). A related idea is the concept of **innovation** (whether the information set is involved).

- **Martingale:** Y_t follows a martingale process if $\mathbb{E}(Y_{t+1} | \mathcal{I}_t) = Y_t$ where $\mathcal{I}_t$ is the t information set.

- **Martingale Difference Process:** Y_t follows a martingale difference process if $\mathbb{E}(Y_{t+1} | \mathcal{I}_t) = 0$. $\{Y_t\}$ is called a martingale difference sequence ('MDS'). A related concept is **Brownian Motion** (a continuous version of an MDS).

- The Lag Operator L lags the elements of a sequence by one period: $Ly_t = y_{t-1}$; $L^2 y_t = y_{t-2}$.

1. Univariate Time Series

The ARMA Process

Autoregressive Process (AR). The present value of a time series is a linear function of previous observations,

$$Y_t = \sum_j^p \phi_j Y_{t-j} + u_t$$

or $a(L)Y_t = u_t$ where $a(L) = (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)$ and u_t is sometimes called an innovation.
For instance, the AR(1) is $Y_t = \phi Y_{t-1} + u_t$.

Moving Average process (MA). The (weighted) sum of the current and previous errors,

$$Y_t = \sum_j^q \theta_j u_{t-j} + u_t$$

or $Y_t = b(L)u_t$ where $b(L) = (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_p L^p)$. For instance, the MA(1) is $Y_t = \theta_1 u_{t-1} + u_t$.

The **ARMA(p,q)** is $a(L)Y_t = b(L)u_t$. For instance, ARMA(1,1): $(1 - \phi L)Y_t = (1 + \theta L)u_t$. This model was popularized by Box and Jenkins, who also developed a methodology that I will mention later.

AR to MA

Let's explore an interesting link between the AR and MA models, which will be useful later in our discussion of stationarity. We'll start with the AR(1) model and then progress to the AR(2) to establish a more general case.

Notice that the AR(1), after repeated substitutions, can take the following form

$$\begin{aligned} Y_t &= \phi Y_{t-1} + u_t \\ &= \phi(\phi Y_{t-2} + u_{t-1}) + u_t \\ &= \phi^r Y_{t-r} + \phi^{r-1} u_{t-r+1} + \ldots + \phi u_{t-1} + u_t \end{aligned}$$

therefore, if $|\phi| < 1$,

- $\lim_{r\to\infty} \phi^r Y_{t-r} = 0$
- and, hence,

$$\left. Y_t = \sum_{j=0}^{\infty} \phi^j \, u_{t-j} \quad \right\} \text{MA}(\infty)$$

AR to MA (cont.)

Alternatively, using the lag operator, $(1-\phi L)Y_t = u_t$, the question would be like

$$Y_t \stackrel{?}{=} (1-\phi L)^{-1} u_t$$

Thus, for $|\phi| < 1$, and honoring Brook Taylor,

$$\begin{aligned} Y_t &= (1-\phi L)^{-1} u_t \\ &= (1 + \phi L + \phi^2 L^2 + \dots) u_t \\ &= \sum_{j=0}^{\infty} \phi^j \, u_{t-j} \qquad , \quad \mathrm{MA}(\infty) \end{aligned}$$

Brook Taylor (1685-1731)

AR to MA (cont.)

For the AR(2), $(1-\phi_1 L-\phi_2 L^2)Y_t=u_t$, the term $(1-\phi_1 L-\phi_2 L^2)$ looks like a second order polynomial; that is, because L is an operator, replacing L by ψ, we have $(1-\phi_1\psi-\phi_2\psi^2)$. Factoring second degree polynomials,

$$
\begin{gathered}
(1-\phi_1\psi-\phi_2\psi^2)\equiv(1-\lambda_1 L)(1-\lambda_2 L)\\
\lambda_1*\lambda_2=-\phi_2\\
\lambda_1+\lambda_2=\phi_1
\end{gathered}
$$

thus,

$$Y_t=(1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1}u_t$$

and, therefore, for $|\lambda_i|<1$,

$$\left. Y_t=\left(\sum_{j=0}^{\infty}\lambda_1^j L^j\right)\left(\sum_{j=0}^{\infty}\lambda_2^j L^j\right)u_{t-j}\quad\right\}\ \mathrm{MA}(\infty)$$

In other words, instead of imposing conditions on ϕ, the requirements are placed on λ. Generally speaking, **having eigenvalues with a modulus less than one**, which is equivalent, as we will state later, to all **the roots of the characteristic polynomial having a modulus greater than one.**

AR to MA (cont.)

Example 1:

$$Y_t = 0.6Y_{t-1} + 0.2Y_{t-2} + u_t$$

we have

$$(1 - 0.6L - 0.2L^2)Y_t = u_t$$

Thus, to find the eigenvalues,

$$(\lambda^2 - 0.6\lambda - 0.2) = 0$$

which implies

$$\lambda_i = \frac{-(-0.6) \pm \sqrt{(-0.6)^2 - 4(-0.2)}}{2}$$

hence, $\lambda_1 = 0.84$ and $\lambda_2 = -0.24$.

Example 2:

$$Y_t = 0.5Y_{t-1} - 0.8Y_{t-2} + u_t$$

we have

$$(1 - 0.5L + 0.8L^2)Y_t = u_t$$

Thus,

$$(\lambda^2 - 0.5\lambda + 0.8) = 0$$

which implies

$$\lambda_i = 0.25 \pm 0.86i$$

hence,

$$R = \sqrt{0.25^2 + 0.86^2} = 0.9 < 1$$

Moments and Stationarity - MA

Let's begin with the MA(1)

$$Y_t = \theta u_{t-1} + u_t \quad , \quad u_t \sim (0, \sigma^2)$$

one can show that:

- $\mathbb{E}(Y_t) = \theta\mathbb{E}(u_{t-1}) + \mathbb{E}(u_t) = 0$
- $\gamma_0 = V(Y_t) = V(\theta u_{t-1}) + V(u_t) + 2Cov(\theta u_{t-1}, u_t) = (\theta^2 + 1)\sigma_2$, which is a constant (does not depend on time).
- $\gamma_1 = \mathbb{E}(Y_t Y_{t-1}) = \theta\sigma^2$, which is also a constant (does not depend on time).
- $\gamma_s = 0$ for $s > 1$
- $\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta}{(1+\theta^2)}$
- $\rho_s = 0$ for $s > 1$

Hence, **the MA(1) process is said to be covariance (weakly) stationary.**

Moments and Stationarity - MA (cont.)

Now, for the MA(q) process:

$$Y_t = \sum_{j=0}^{q} \theta_j L^j u_{t-j} \quad , \quad u_t \sim (0, \sigma^2)$$

one can show that

- $\mathbb{E}(Y_t) = 0$,
- $\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2)\sigma_2$,
- and

$$\gamma_s = \begin{cases} \sigma^2(\theta_j + \theta_{j+1}\theta_1 + \ldots + \theta_{q-j}\theta_q) & if \quad s = 1, \ldots, q \\ 0 & if \quad s > q \end{cases}$$

In other words, **the MA(q) process is (weakly) stationary:** Y_t is a combination of stationary terms, where the mean and variance are constant, and the autocovariances depend on s but not on t.

Moments and Stationarity - MA (cont.)

Lastly, let's review the MA(∞),

$$Y_t = \sum_{j=0}^{\infty} \theta_j L^j u_{t-j} \quad , \quad u_t \sim (0, \sigma^2)$$

In this case,

- $\mathbb{E}(Y_t) = 0$,
- $\gamma_0 = \sigma_2(1 + \theta_1^2 + \theta_2^2 + \dots) \equiv \sigma_2 \sum_j^{\infty} \theta_j^2$,
- and $\gamma_s = \sigma^2 \sum_j^{\infty} \theta_j \theta_{j+s}$

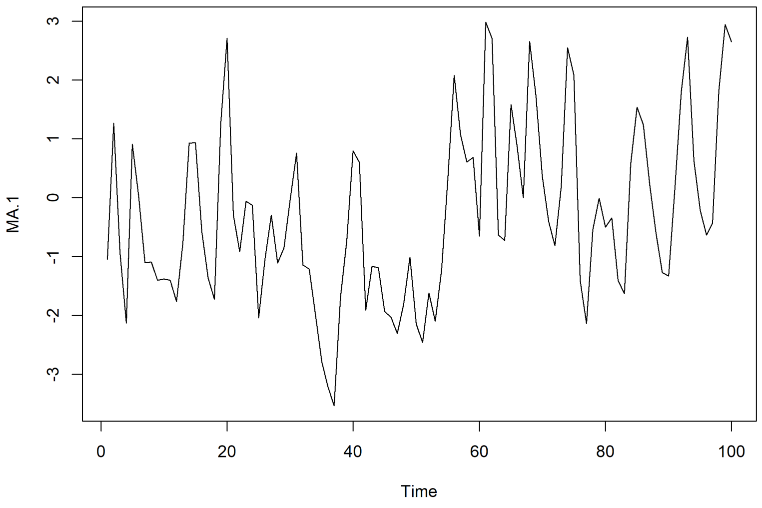
Thus, **the process is covariance (weakly) stationary** under the following assumption: **Square Summability**, $\sum_j \theta_j^2 < \infty$.

An alternative (stronger) requirement would be: **Absolute Summability**, $\sum_j |\theta_j| < \infty$.

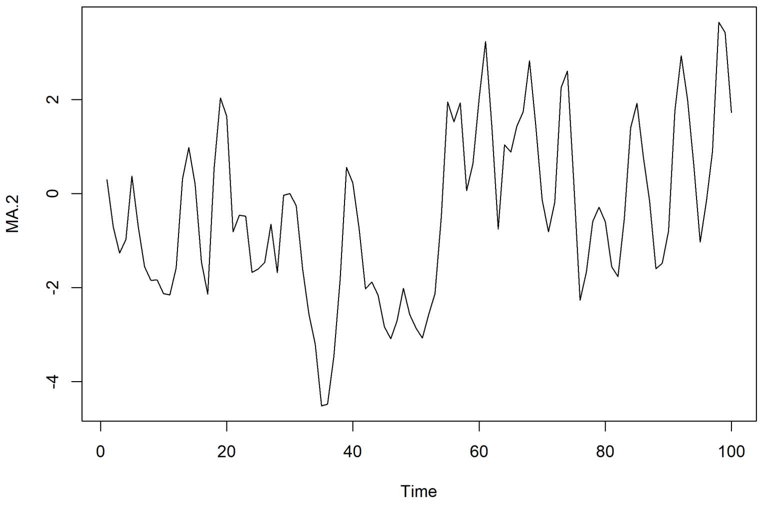
Simulating MA models

Simulations of MA(q) processes, with $u_t \sim (0, 0.8^2)$.

MA(1): $Y_t = 1.2u_{t-1} + u_t$



MA(2): $Y_t = 1.2u_{t-1} + 0.9u_{t-2} + u_t$



Moments and Stationarity - The AR Model

We just reviewed conditions for a MA process to be covariance stationary. Thus, the focus **for the AR model** will be on whether one can **write the infinite moving average representation**.

Let's start with the **AR(1) model**, $Y_t = \phi Y_{t-1} + u_t$.

- As showed, the process can be expressed as $Y_t = \sum_{j=0}^{\infty} \phi^j u_{t-j}$ and, therefore, the **stationarity condition for the AR(1) is** $|\phi| < 1$, so that the MA sum converges, $1 + \phi + \phi^2 + \ldots \to \frac{1}{1-\phi}$.
- Similarly, for $(1 - \phi L)Y_t = u_t$ the characteristic equation is $(1 - \phi\psi) = 0$. Thus, its one characteristic root is $\psi = 1/\phi$. Therefore, the series is stationary as long as $|\phi| < 1$ which is the same condition as $|\psi| > 1$.

Thus,

$$Y_t = (1 - \phi L)^{-1} u_t$$

Moments and Stationarity - The AR Model (cont.)

Thus, for the **AR(1) model** $Y_t = \phi Y_{t-1} + u_t$ with $|\phi| < 1$, we can find that the first moments are as follows:

- the unconditional expectation is $\mathbb{E}(Y_t) = 0$,
- the unconditional variance is $\gamma_0 = V(u_t + \phi u_{t-1} + \phi^2 u_{t-2} + \dots) = (1 + \phi^2 + \phi^4 + \dots)\sigma^2 \rightarrow \frac{\sigma^2}{1-\phi^2}$,
- the autocovariances are
  - $\gamma_1 = cov(Y_t, Y_{t-1}) = (\phi\sigma^2 + \phi^3\sigma^2 + \phi^5\sigma^2 + \dots) = \phi\gamma_0$
  - $\gamma_j = \phi^j \gamma_0$

Moments and Stationarity - The AR Model (cont.)

For the AR(p) model,

$$(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p) Y_t = u_t$$

recall that $a(L) = (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)$ is a polynomial in L.

Define the characteristic equation as $\phi(\psi) = (1 - \phi_1 \psi - \phi_2 \psi^2 - \ldots - \phi_p \psi^p) = 0$. The p solutions, $\lambda_1, \ldots, \lambda_p$, can be used to factorize the polynomial,

$$\phi(\psi) = (1 - \lambda_1 \psi)(1 - \lambda_2 \psi)\ldots$$

The relationship between eigenvalues or **inverse roots** and the **roots** is $\psi_j = \lambda_j^{-1}$. Therefore, $\phi(z)$ is invertible if each factor is invertible; that is, if $|\psi_j| > 1$ (outside the unit circle) or $|\lambda_j| > 1$ (inside the unit circle). Notice that this condition considers that some roots may be complex, $\lambda_j = r_j \pm c_j \sqrt{-1}$.

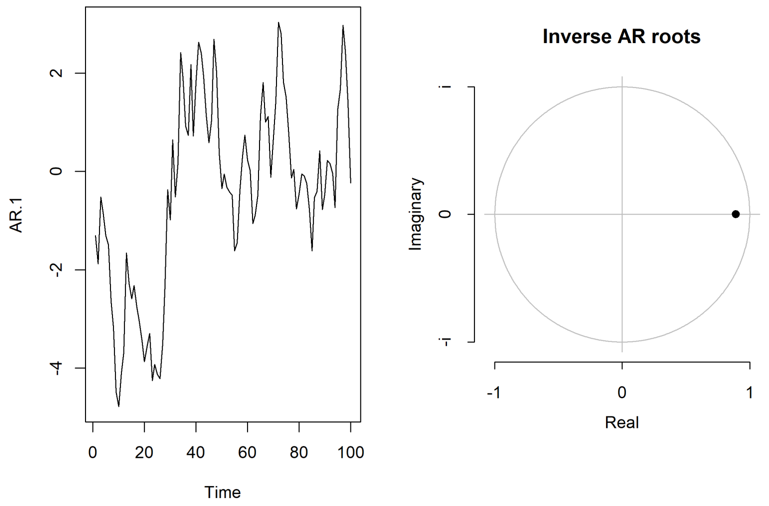
Hence, the condition can be stated as that **all the roots of the characteristic polynomial** (ψ_j) **having a modulus greater than one (outside the unit circle).** In such case,

$$Y_t = (1 - \lambda_1 \psi)^{-1} (1 - \lambda_2 \psi)^{-1} \ldots (1 - \lambda_p \psi)^{-1} u_t$$

Simulating AR models

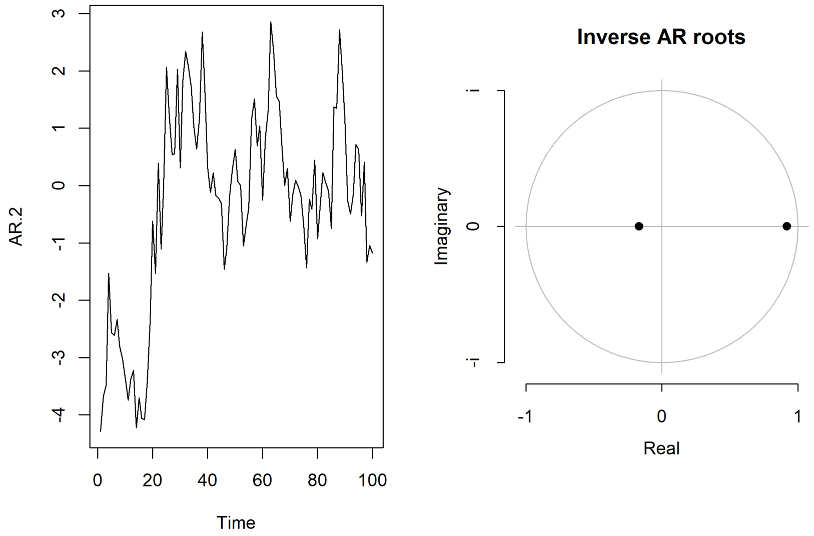
Simulations of AR(p) processes, with $u_t \sim (0, 0.8^2)$.

AR(1): $Y_t = 0.8Y_{t-1} + u_t$



Inverse root: $\lambda = 0.8$.

AR(2): $Y_t = 0.6Y_{t-1} + 0.2Y_{t-2} + u_t$



Inverse roots: $\lambda_1 = -0.24$ and $\lambda_2 = 0.84$.

Autocorrelation function (ACF)

Researchers (empirically) study the time dependence by the correlation

$$Corr(Y_t, Y_{t-h}) = \rho_{Y_t,Y_{t-h}} = \frac{Cov(Y_t, Y_{t-h})}{\sqrt{V(Y_t) \cdot V(Y_{t-h})}} \quad ; \quad h = \ldots, -2, -1, 0, 1, 2, \ldots$$

As a function of h, $\rho(h)$, it is called a **correlogram**.

Under stationarity, $V(Y_t) = V(Y_{t-h})$, the function is known as **the autocorrelation function (ACF)**, thus $\rho_h = \gamma_h/\gamma_0$.

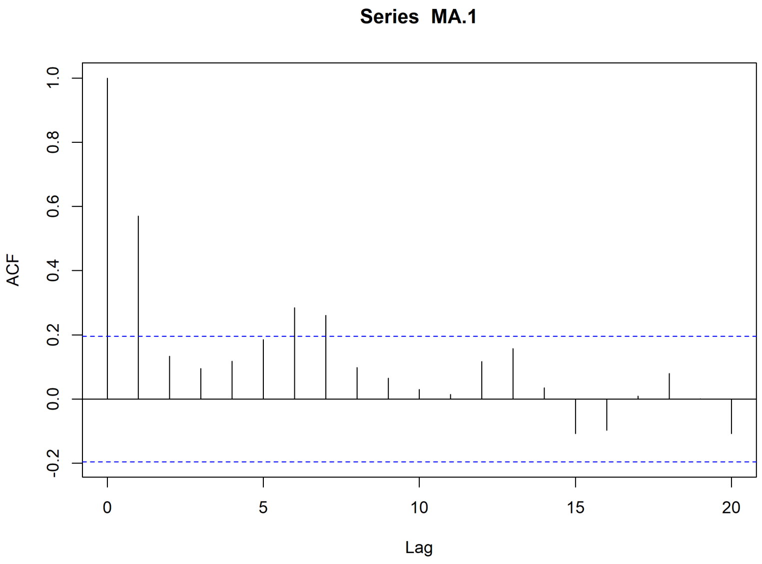
The population moments are replaced with sample moments,

$$\hat{Cov}(y_t, y_{t-h}) = \frac{1}{T-h} \sum_{t=h+1}^{T} (y_t - \bar{y})(y_{t-h} - \bar{y}) \quad ; \quad h = \ldots, -2, -1, 0, 1, 2, \ldots$$

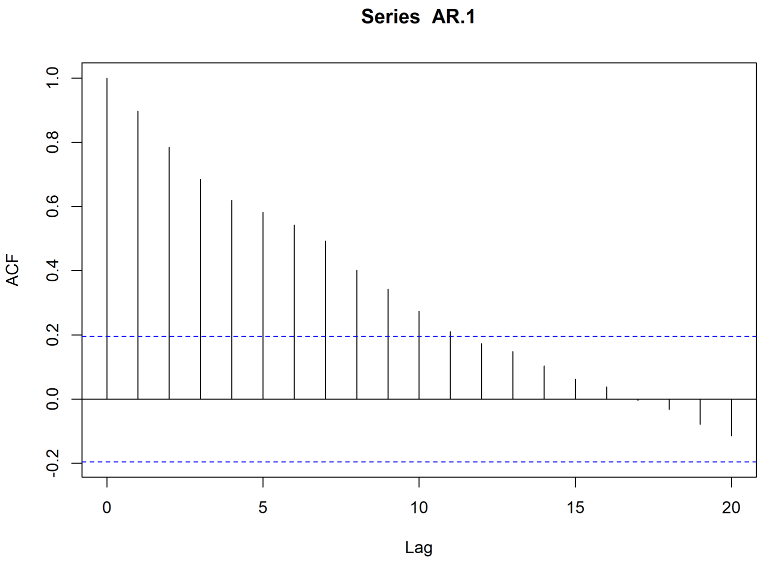
Approaching Time Dependence (cont.)

The autocorrelation function (ACF) for the simulated MA and AR processes, with $u_t \sim (0, 0.8^2)$.

MA(1): $Y_t = 1.2u_{t-1} + u_t$



AR(1): $Y_t = 0.8Y_{t-1} + u_t$



Understanding Partial Autocorrelation Function (PACF)

The Partial Autocorrelation Function (PACF) is an empirical tool for identifying the number of lags in Autoregressive (AR) components of ARMA processes in time series. It measures the correlation between observations in a time series separated by a certain number of lags (k), while controlling for the correlations at all shorter lags.

Lag-by-Lag Analysis: PACF examines the correlation of a time series with its lagged values for various lags, one at a time, removing the effects of previous lags.

Regression Approach: Each lag in PACF corresponds to a regression of the time series on its past values up to that lag:

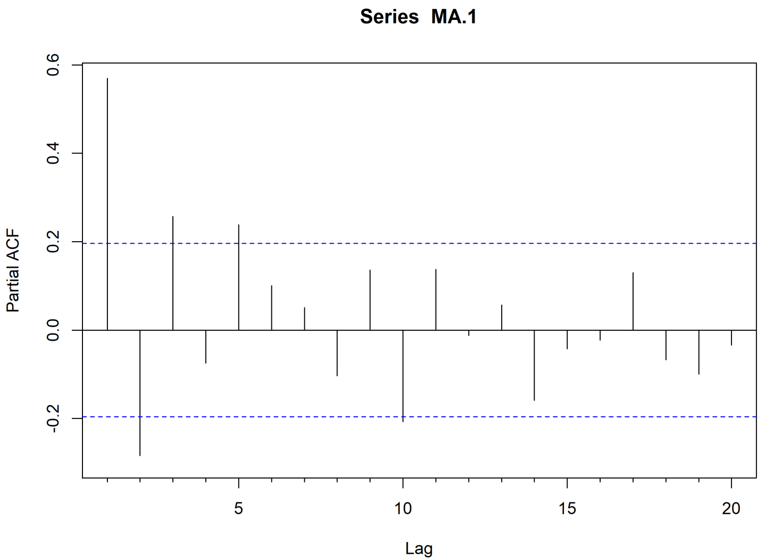
(Lag) k	Regression Equation	PACF
1	$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$	β_1
2	$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + u_t$	β_2
$\vdots$	$\vdots$	$\vdots$

The coefficient (β_k) of the highest lagged term in each regression represents the partial correlation at that lag. A significant spike in the PACF plot at lag k followed by non-significant values suggests an AR(k) model.

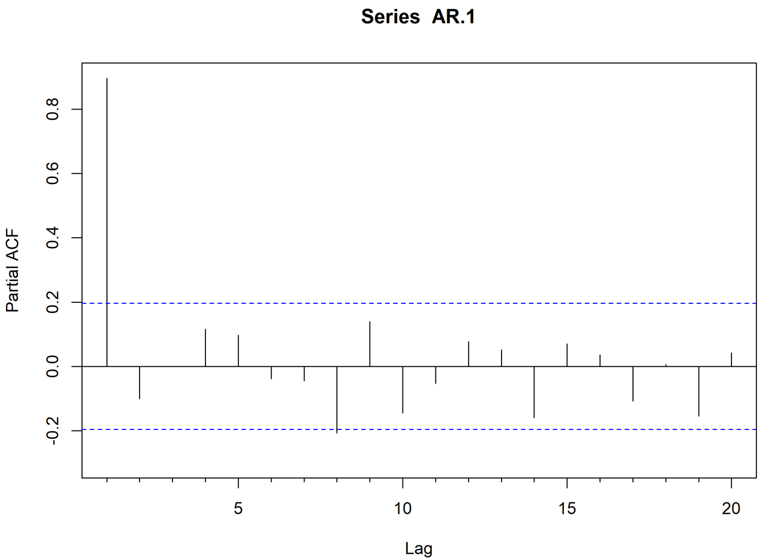
Approaching Time Dependence (cont.)

The Partial Autocorrelation Function (PACF) for the simulated MA and AR processes, with $u_t \sim (0, 0.8^2)$.

MA(1): $Y_t = 1.2u_{t-1} + u_t$



AR(1): $Y_t = 0.8Y_{t-1} + u_t$



The Box-Jenkins Approach

It is a systematic (empirical) methodology for analyzing and forecasting time series data. Primarily used for forecasting time series.

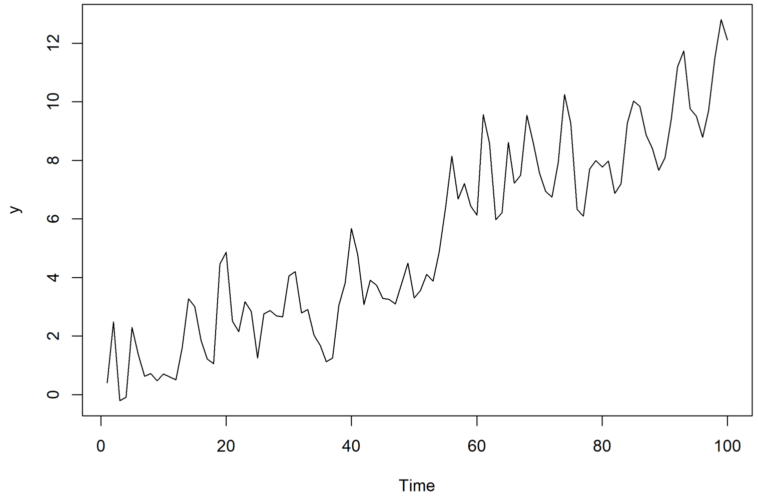
Steps:

- **Model Identification.** Analyze the time series plot. Also, the ACF and PACF plots. Identify the appropriate ARIMA (Autoregressive Integrated Moving Average) model (determine the order of differencing, the number of AR terms, and the number of MA terms). In the followig subsection we review the meaning of 'Integrated'.
- **Estimation.** Estimate the parameters of the identified ARIMA(p, d, q) model.
- **Model Checking.** Validate the fitted model (i.e., checking for autocorrelation in the residuals or the Ljung-Box test to explore no significant autocorrelation in residuals).
- **Model Refinement.** If the model does not fit well, return to step 1 for re-identification.

2. Non-Stationary Time Series

Non-Stationary Time Series

The main assumption on the time series data so far has been stationarity. However, many macro-economic variables are trending.



Stationarity can be violated in different ways:

- **Deterministic trends** - or trend stationarity.
- Unit roots - or **stochastic trends**
- Level shifts - breaks
- Variance changes.

Non-Stationary Time Series

Deterministic Trend

Some examples are: (i) $Y_t = \psi + \beta \cdot t + \varepsilon_t$; or (ii) $X_t = \phi X_{t-1} + u_t$, where $|\phi| < 1$ and $Y_t = X_t + \psi + \beta \cdot t$

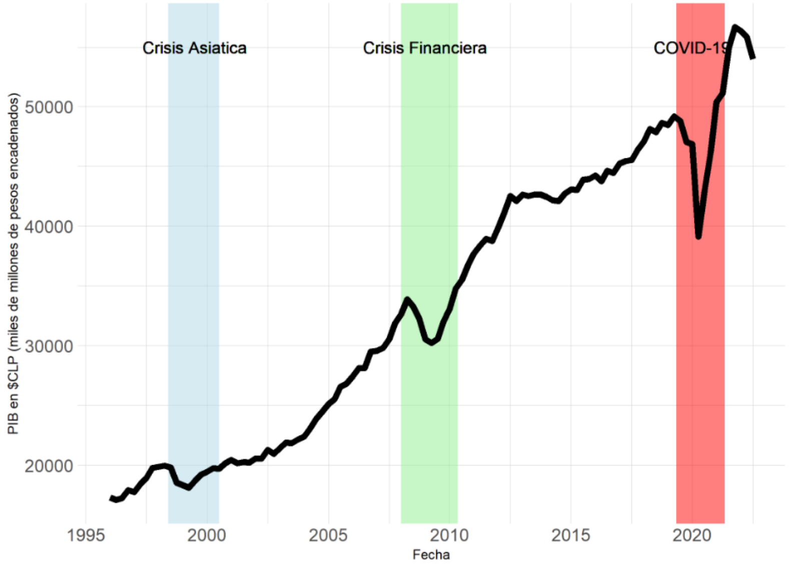


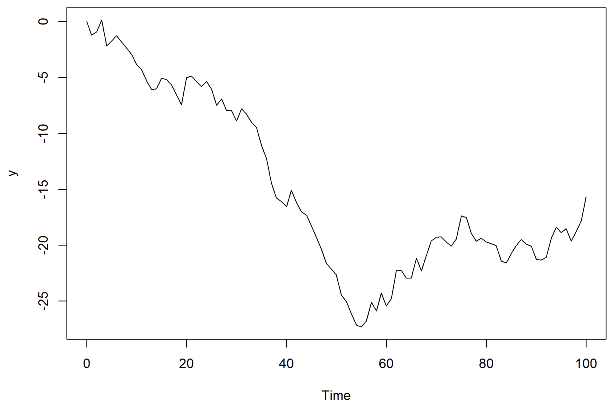
Figure: Quarterly GDP - Chile. 'Volumen a precios del año anterior encadenado, series empalmadas, desestacionalizado, referencia 2018.' Source: Construcción del autor usando R y la API del BCentral.

Stochastic Trend: Random Walk

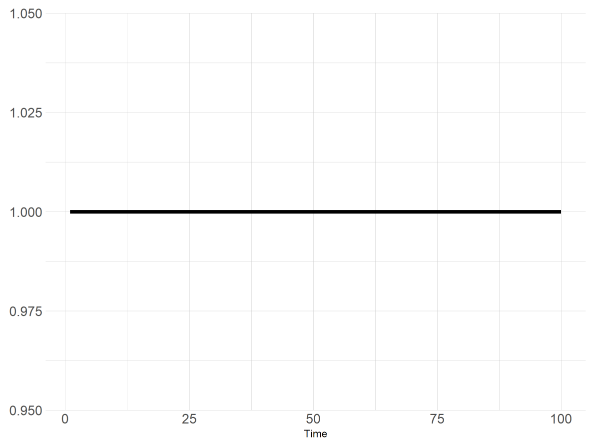
A radom walk process is: $y_t = y_{t-1} + \varepsilon_t$

This structure would imply that the effect of a shock is permanent: $y_t = \sum_{\tau=1}^{t} \varepsilon_\tau$

```
# R code to simulate the Random Walk
set.seed(1234)
t  <- 1:100
y <- arima.sim(list(order = c(0, 1, 0)), n = length(t))
plot(y)
```



The effect of a shock ($y_0 = 0$, $\varepsilon_1 = 1$, and $\varepsilon_2 = \ldots = \varepsilon_T = 0$)

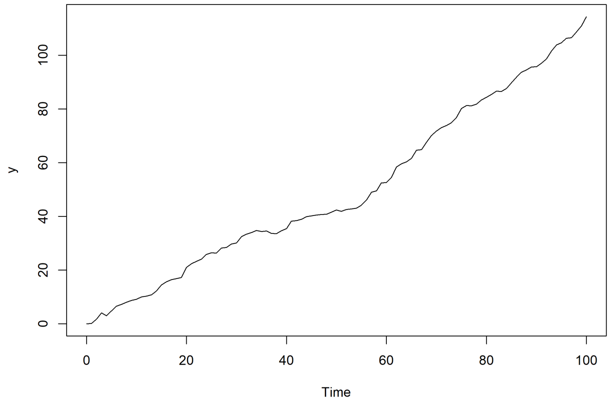


Stochastic Trend: Random Walk with Drift

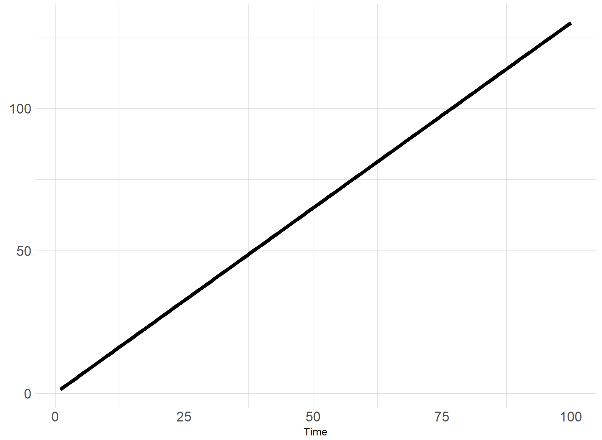
It is a random walk plus a constant term: $y_t = \mu + y_{t-1} + \varepsilon_t$

This structure would imply that shocks have permanent effects and are influenced by the drift: $y_t = \mu \cdot t + \sum_{\tau=1}^{t} \varepsilon_\tau$

```
# R code to simulate the Random Walk with drift
set.seed(1234)
y <- arima.sim(model= list(order = c(0, 1, 0)), n=100, mean=1.3 )
plot(y)
```



The effect of a shock ($y_0 = 0$, $\mu = 1.3$, $\varepsilon_1 = 1$, and $\varepsilon_2 = \ldots = \varepsilon_T = 0$)



Deterministic vs. Stochastic trend and De-trending

Stochastic Trend. There are important implications for $\phi = 1$ rather than $\phi < 1$ in $Y_t = \phi Y_{t-1} + \varepsilon_t$:

- The effect of the initial value stays in the process. $\mathbb{E}(Y_t|Y_0) = Y_0$.
- Shocks have permanent effects. Accumulate to a random walk component $\sum \varepsilon_t$ called a stochastic trend.
- The variance increases $V(\sum \varepsilon_t|Y_0) = t \cdot \sigma^2$.
- The covariance is $\mathbb{E}((Y_t - Y_0)(Y_{t-s} - Y_0)|Y_0) = (t-s)\sigma^2$ and the autocorrelation is $\sqrt{(t-s)/t}$ (which dies out very slowly with s)

De-trending. When dealing with non-stationary time series, a common approach is to transform the series to achieve stationarity. This transformation is often referred to as 'de-trending'.

First Order Integration (I(1)). If the first difference, $\Delta Y_t = (Y_t - Y_{t-1}) = \varepsilon_t$, is stationary, the series is called integrated of first order, denoted as I(1); hence, it is named **Integrated Series** with $d = 1$.

Non-Stationary Time Series

A caution in treating trends

1. Using the transformation $(Y_t - \beta \cdot t)$ for
   - **Deterministic Trend:** After the transformation $(Y_t - \beta \cdot t) = \phi + \varepsilon_t$, the series becomes stationary.
   - **Stochastic Trend:** After the transformation $(Y_t - \beta \cdot t) = Y_0 + \sum_{j=0}^{t} \varepsilon_j$, the series remains non-stationary.

2. Using First Difference $(Y_t - Y_{t-1})$ for
   - **Deterministic Trend:** The first difference is $(Y_t - Y_{t-1}) = (\beta \cdot t - \varepsilon_t) - (\beta \cdot (t-1) - \varepsilon_{t-1}) = \beta + (\varepsilon_t - \varepsilon_{t-1})$. This is akin to a Moving Average process with a kind of 'unit root.'
   - **Stochastic Trend:** The first difference is $(Y_t - Y_{t-1}) = \varepsilon_t$. The first difference is stationary.

It is crucial to conduct statistical tests to verify the presence of unit roots and correctly identify the nature of the trend (deterministic or stochastic) in the time series.

Unit Root Testing

The key approach in unit root testing is to test for a unit root in an autoregressive model.

To illustrate, let's review the **Dickey-Fuller Test** - for the AR(1) $Y_t = \phi Y_{t-1} + \varepsilon_t$:

- **Hypothesis Testing**. The null Hypothesis (H_0) is 'tests against stationarity', specifically $H_0 : \phi = 1$; thus, the alternative hypothesis (H_1) implies stationarity.
- **Equivalent Formulation**. Reformulate it as $\Delta Y_t = \pi Y_{t-1} + \varepsilon_t$, where $\pi = \phi - 1$. Thus, the null hypothesis becomes $H_0 : \pi = 0$, and the alternative is $H_1 : -2 < \pi < 0$.
- **Dickey-Fuller Test Statistic**: The test statistic is calculated as the t-ratio: $\hat{\pi}/\mathrm{se}(\hat{\pi})$. The asymptotic distribution for this test statistic follows the Dickey-Fuller distribution, not the standard normal distribution $N(0, 1)$.

Unit Root Testing: ADF

Augmented Dickey–Fuller Test in AR(p) Model. The Augmented Dickey–Fuller (ADF) test extends the Dickey-Fuller test to higher-order autoregressive processes, $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$.

- The inclusion of lagged difference terms accounts for serial correlation and makes the test robust for higher-order AR processes.
- **ADF Test Formulation**: Reformulate the model to include lagged difference terms
  $\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \delta_2 \Delta Y_{t-2} + \cdots + \delta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t$.
- **Null and Alternative Hypotheses**: (H_0) is that 'the time series has a unit root (non-stationary)', i.e., $\gamma = 0$. The alternative Hypothesis is that the time series is stationary.
- **ADF Test Statistic**: The test statistic is $\text{t-statistic} = \hat{\gamma}/\text{se}(\hat{\gamma})$. And the critical values for this test are specific to the ADF distribution. Rejecting H_0 suggests that the series is stationary.

Implementation Notes: The selection of p is crucial and can be determined based on information criteria like AIC or BIC.

Unit Root Testing: ADF (cont.)

There are, however, some weaknesses of the ADF test:

- Assumption of i.i.d. residuals. But many time series exhibit, for instance, time-varying volatility or conditional heteroskedasticity, violating this assumption.
- Power Issues: The ADF test can suffer from low statistical power, especially when the series is close to being non-stationary but not exactly so (the test may fail to reject the null hypothesis even when the series is actually stationary).
- Small Samples: The test may have size distortion. The probability of rejecting H_0 when it is true (type I error) can be higher than the nominal level.

Unit Root Testing: Alternative Unit Root Tests

- **Phillips-Perron Test** is a variation of the ADF test. Focuses on correcting for any autocorrelation and heteroskedasticity in the error terms utilizing non-parametric statistical methods.

- **Zivot-Andrews Test** accounts for the possibility that the time series may appear to have a unit root but is actually stationary around a changing mean (structural break). The null hypothesis of a unit root with a one-time structural break in level or trend.

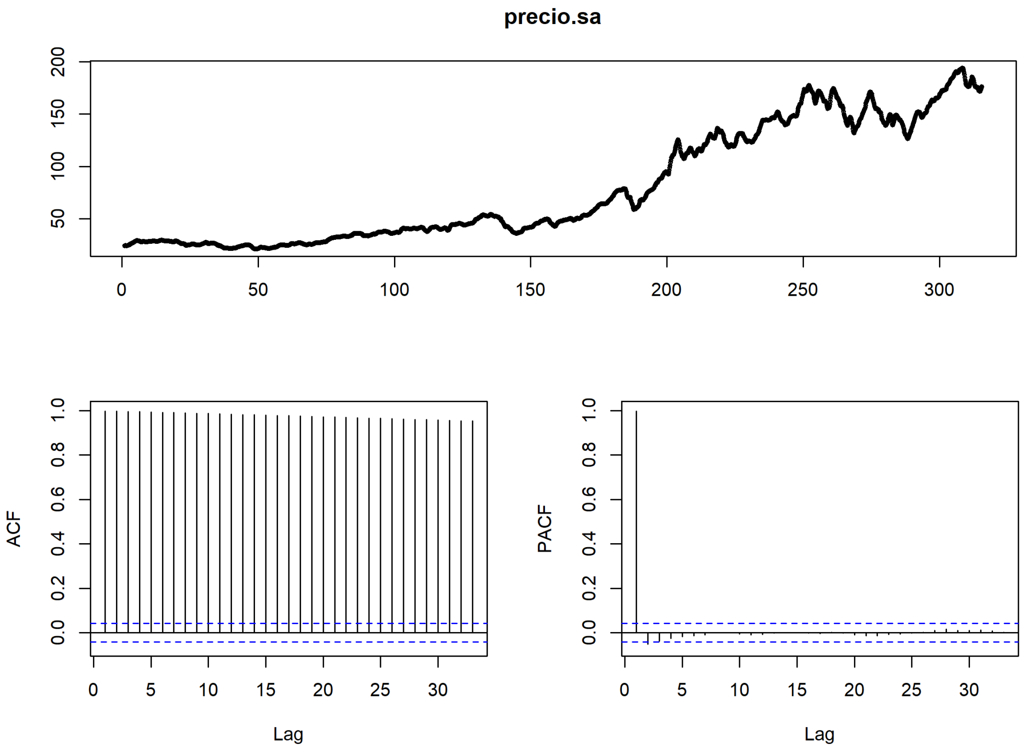
- **Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test** aims to determine if a time series is stationary around a deterministic trend. The null hypothesis is that the series is stationary (trend stationary or level stationary).

Example: Time Series in R

Apple Inc. stock prices (note: this is only an illustration of commands and it is not aimed at discussing whether is a good empirical model).

```r
# Código de programación en R:
library(quantmod) # acceder a datos
library(forecast) # methods and tools for displaying and analysing univariate ti
library(tseries)  # Time Series Analysis and Computational Finance

apple <- getSymbols( "AAPL",
                      from=as.Date("2015-01-01"),
                      to  =as.Date("2023-10-12"),
                      auto.assign=F)
df      <- data.frame(date= index(apple), apple, row.names = NULL)
precio.sa<- seasadj(stl( ts( na.omit(ma(df$AAPL.Adjusted, order=7)),
                             frequency=7), s.window="periodic"))
tsdisplay(precio.sa)
```



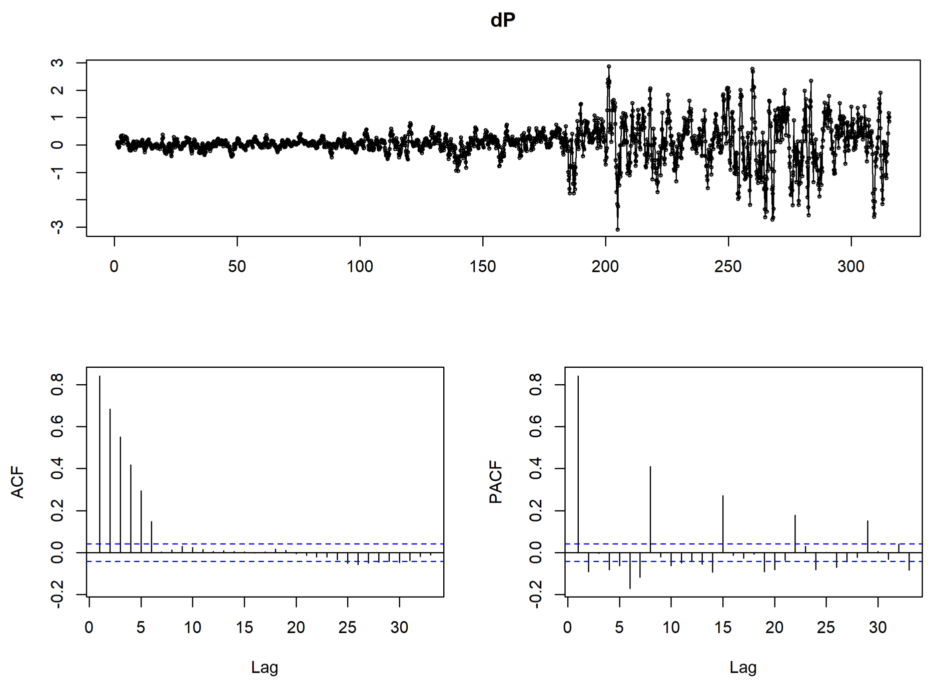
Example: Time Series in R (cont.)

```
# Unit Root:
adf.test(precio.sa)
##
##     Augmented Dickey-Fuller Test
##
## data:  precio.sa
## Dickey-Fuller = -2.2511, Lag order = 13, p-value = 0.472
## alternative hypothesis: stationary

# First Diff
dP <- diff(precio.sa, differences = 1)

# Estimar el mejor modelo ARIMA
auto.arima(dP, seasonal=FALSE) # Then, ARIMA(1,1,1)
## Series: dP
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1     ma1    mean
##       0.8116  0.1084  0.0691
## s.e.  0.0147  0.0253  0.0428
##
## sigma^2 = 0.1168:  log likelihood = -759.87
## AIC=1527.74   AICc=1527.76   BIC=1550.53
```

```
#ACF PACF
tsdisplay(dP)
```




```
# Unit Root
adf.test(dP)$p.value
## [1] 0.01
```

A short note on ARCH Models

Autoregressive Conditional Heteroscedasticity (ARCH)

Up to now, we modeled Y_t using ARIMA structures, but it is also possible to model the residuals or the variance.

ARCH models are used to model and forecast time-varying volatility in time series data.

Widely used for modeling and forecasting the volatility of financial assets like stocks, currencies, and derivatives.

The ARCH(q) model is defined for a time series Y_t with the following structure for the variance:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2$$

where, σ_t^2 is the conditional variance and $\varepsilon_t = Y_t - \mu_t$ is the residual at time t.

The parameters of an ARCH model are typically estimated using Maximum Likelihood Estimation (MLE). Choosing the correct order q is often based on statistical tests like the Lagrange Multiplier (LM) test.

Extensions of ARCH Models

Some extensions are aimed at offering more flexibility and accuracy in modeling complex volatility patterns observed in real-world financial time series data.

- **Generalized ARCH (GARCH)**: An extension of ARCH that includes lagged conditional variances in the model.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

  where p and q are the orders of the GARCH model.

- **Exponential GARCH (EGARCH)**: Models the log of the variance, capturing the asymmetric effects of positive and negative shocks on volatility.

- **Threshold GARCH (TGARCH)**: Allows different responses of volatility to positive and negative shocks, useful in financial markets where volatility tends to increase more with negative shocks.

4. Vector Autoregressive models (VAR)

Introduction to VAR

Stock and Watson (2021, in *Journal of Economic Perspectives*) describe the role of macroeconomists as encompassing a list of tasks, for which Vector Autoregressive models (VARs) serve as a useful statistical tool. These tasks include:

- Describe and summarize macroeconomic time series
- Make forecasts
- Recover the structure of the macroeconomy from the data
- Advise macroeconomic policy-makers

Vector Autoregression (VAR) is used to capture the linear interconnections among multiple time series. It generalizes the univariate autoregressive model (the one we just reviewed) to allow for more than one evolving variable.

State-Space Notation in Vector Autoregression (VAR)

To illustrate, consider the following system of OLS regressions for two variables modeled using two lags:

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \theta_1 m_{t-1} + \theta_2 m_{t-2} + \epsilon_{1t} \\
m_t &= \gamma_1 m_{t-1} + \gamma_2 m_{t-2} + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \epsilon_{2t}
\end{aligned}
$$

Where, y_t and m_t are the time series variables (for instance, GDP and monetary policy); $\{\phi, \theta, \gamma, \alpha\}$ are the coefficients; ϵ_{1t}, ϵ_{2t} are the error terms.

Vector Autoregression (VAR) models can be represented using state-space notation, which is useful for dealing with multivariate time series data in a compact and structured form. (Note: A **State-Space model** comprises two main equations the state equation and the observation equation. The model is conducive to advanced forecasting methods and can be used in conjunction with Kalman filtering for dynamic updates and analysis.)

For the example: firstly, notice that the system can be expressed as a kind of AR(2):

$$
[\begin{matrix} y_t \\ m_t \end{matrix}] = [\begin{matrix} \phi_1 & \theta_1 \\ \gamma_1 & \alpha_1 \end{matrix}][\begin{matrix} y_{t-1} \\ m_{t-1} \end{matrix}] + [\begin{matrix} \phi_2 & \theta_2 \\ \gamma_2 & \alpha_2 \end{matrix}][\begin{matrix} y_{t-2} \\ m_{t-2} \end{matrix}] + [\begin{matrix} \epsilon_{1t} \\ \epsilon_{1t} \end{matrix}]
$$

$$
\mathrm{Y}_t = \Phi_1 \mathrm{Y}_{t-1} + \Phi_2 \mathrm{Y}_{t-1} + \Phi \epsilon_t
$$

State-Space Notation in Vector Autoregression (VAR)

Secondly, the 'kind of AR(2)' is expressed as follows:

$$[\begin{matrix} Y_t \\ Y_{t-1} \end{matrix}] = [\begin{matrix} \Phi_1 & \Phi_2 \\ I_2 & 0 \end{matrix}][\begin{matrix} Y_{t-1} \\ Y_{t-2} \end{matrix}] + [\begin{matrix} \epsilon_t \\ 0 \end{matrix}]$$

$$X_t = \Phi X_{t-1} + \xi_t$$

where, Φ is a $2p \times 2p$ (or, generally speaking, $np \times np$) matrix of coefficients to be estimated.

Stationarity. Notice that the last equation looks like an AR(1). Thus, it is possible to state conditions, similar to those we reviewed for AR(1), to ensure that the system is stationary. In particular, the eigenvalues of Φ are inside the unit circle (i.e., $|I_2\lambda^2 - \Phi_1\lambda - \Phi_2| = 0$).

Estimation of VARs.

- VAR models require a large number of observations due to the number of parameters being estimated.
- After defining the appropriate lag length and variables, estimation of the parameters (coefficients) is based on an OLS approach.
- Choosing the correct lag length is crucial. Information criteria like AIC, BIC can be used for this purpose.

Impulse Response Functions in Vector Autoregression

Impulse Response Functions (IRFs) play an esential role in VAR analysis by illustrating the effect of one-time shocks on the system. The key question IRFs address is how an exogenous shock, such as a monetary policy change, impacts other variables in the VAR model (e.g., GDP).

From the Moving Average (MA) representation of a VAR model, the response of future values to a current shock is given by:

$$\frac{\partial \mathrm{X}_{t+s}}{\partial \xi_t} = \Phi^s$$

To continue, we need some assumptions for the shock variables ξ_t,

- **Expectation**: $\mathbb{E}\{\xi_t\} = 0$ (Mean of shocks is zero).
- **Covariance Matrix**: $\mathbb{E}\{\xi_t \xi_t'\} = \Omega$ (Covariance matrix of shocks).
- **Independence Over Time**: $\mathbb{E}\{\xi_t \xi_{t-j}'\} = 0 \text{ for all } j > 0$ (Shocks are uncorrelated over time).

Impulse Response Functions in Vector Autoregression

Non-Diagonal Ω: If Ω is not diagonal, a shock in one equation impacts other equations, creating interdependencies.

Alternative Representation The VAR process can also be represented as:

$$\mathrm{X}_t = \Phi \mathrm{X}_{t-1} + \mathrm{C} \cdot \eta_t$$

where η_t are standardized shocks with:

- $\mathbb{E}\{\eta_t\} = 0$
- $\mathbb{E}\{\eta_t \eta_t'\} = I$
- $\mathbb{E}\{\eta_t \eta_{t-j}'\} = 0$ for all $j > 0$

Identifying Shocks in VAR. To 'identify' the relationship between shocks, a transformation of the shock variables is necessary.

Approaches for Shock Identification

- **Orthogonalization**: Utilizing a transformation $\Omega = B\Sigma B'$, with B being a triangular matrix and Σ a diagonal matrix.
- **Standardization Using Cholesky Decomposition**: Here, $\Omega = CC'$, where C is a lower triangular matrix. The order of equations in the VAR becomes crucial.
- **Generalization by Pesaran and Shin**: This approach offers a more generalized method for constructing IRFs as is less sensitive to the ordering of variables in the VAR, reducing the bias that can occur due to arbitrary ordering.

Software Considerations: Different software packages may compute IRFs differently. Some default to using C as a lower triangular matrix, while others may not. **Important**: Be cautious about the approach used, as results can vary significantly based on the method of shock identification.

VAR in Practice using R

Consider a bivariate VAR(2) model with GDP growth (y_t) and inflation (π_t):

$$y_t = \alpha_y + \Phi_{yy1} y_{t-1} + \Phi_{y\pi 1} \pi_{t-1} + \Phi_{yy2} y_{t-2} + \Phi_{y\pi 2} \pi_{t-2} + \epsilon_{yt}$$
$$\pi_t = \alpha_\pi + \Phi_{\pi y1} y_{t-1} + \Phi_{\pi\pi 1} \pi_{t-1} + \Phi_{\pi y2} y_{t-2} + \Phi_{\pi\pi 2} \pi_{t-2} + \epsilon_{\pi t}$$

Simulating the data:

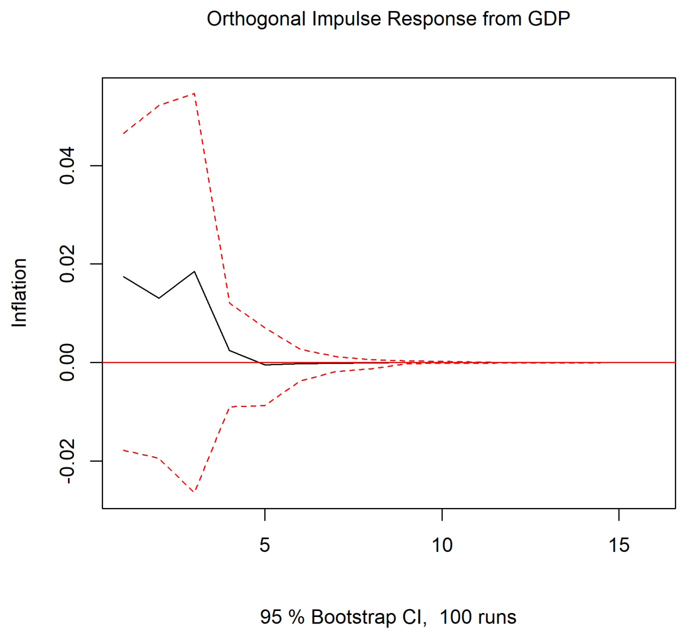
```r
# Simular los datos de PIB e Inflación

GDP       <- rnorm(100, mean = 2, sd = 0.5)  # Simulated GDP
Inflation <- rnorm(100, mean = 1, sd = 0.2)  # Simulated Inflation
data_var  <- ts(cbind(GDP, Inflation), start = c(2010, 1), frequency = 4) # Quarterly, from 2010
```

Example: Estimation VAR(2) in R

```r
# Estimación de un VAR con dos rezagos
var_gdp_pi <- VAR(data_var, p=2) # p=2 lag order
var_gdp_pi
##
## VAR Estimation Results:
## =======================
##
## Estimated coefficients for equation GDP:
## ========================================
## Call:
## GDP = GDP.l1 + Inflation.l1 + GDP.l2 + Inflation.l2 + const
##
##       GDP.l1 Inflation.l1       GDP.l2 Inflation.l2        const
##   0.09179460   0.10613572  -0.02195858  -0.23427881   2.01168221
##
##
## Estimated coefficients for equation Inflation:
## ==============================================
## Call:
## Inflation = GDP.l1 + Inflation.l1 + GDP.l2 + Inflation.l2 + const
##
##       GDP.l1 Inflation.l1       GDP.l2 Inflation.l2        const
##   0.02259444   0.06618442   0.03181936  -0.01385804   0.86477643
```

```r
# Función de Impulso Respuesta (IRF)
IRF1 <- irf(var_gdp_pi,
            impulse = "GDP",
            response = "Inflation",
            n.ahead = 15)
plot(IRF1)
```



Cierre

¿Preguntas?

O vía E-mail: lchanci1@binghamton.edu