MACROECONOMÍA I: CÁTEDRA 1

SERIES DE TIEMPO.PARA MACROECONOMISTAS

Eduardo Engel
ENECO 630. Macroeconomía I
Magíster en Economía, FEN, U. de Chile.

Abril 8, 2021.

Time Series

Rational Expectations

Time Series

Rational Expectations

Macro: Methodology

Take dynamics seriously:

- Expectations
- Adjustment costs
- Commitment problems

Take equilibrium concept seriously: often hard in dynamic settings

Take aggregation seriously

Why Time Series?

- Propagation mechanism: Frisch and Slutzky (1930s)
- Input to the model: shocks (technology, demand, taxes, ...):

$$v_{1t}, v_{2t}, ..., v_{nt}$$

  where v_{kt}: value of k-th shock in period t. These are the **exogenous** variables
- Output of the model: macro variables (output, inflation, employment, ...):

$$x_{1t}, x_{2t}, ..., x_{mt}$$

  where x_{kt}: value of k-th macro variable in period t. These are the **endogenous** variables.

Macro Models in a Nutshell

Shocks

v_{1t}
v_{2t}
v_{3t}
$\ldots$
$\ldots$
v_{nt}

Macro Model

- Optimizing Households
- Optimizing Firms
- Equilibrium Concept

Time Series

x_{1t}
x_{2t}
x_{3t}
$\ldots$
$\ldots$
x_{mt}

Examples:
TFP
-Monetary Policy
-Fiscal Policy
-Demand Shock

Examples:
-Consumption
-Investment
-Output

Comments

- Model inputs ≡ **exogenous** variables ≡ shocks. n.
- Model output ≡ **endogenous** variables. m.
- Usually have $n \leq m$, otherwise model may be hard to falsify.
- Model inputs **cause** model outputs, i.e., exogenous variables cause endogenous variables.
- Sometimes we assume you observe the shocks, other times you want to infer values of the shocks from the model.
- Typically inputs are the "new" component of the shock at time t, thus also referred to as **innovations**.

Modeling choices

Which variables exogenous? Which endogenous?

In reality, all variables are endogenous.

The answer should depend on the question at hand.

Productivity shocks:

- exogenous in most macro models
- a good starting point when modeling long run growth
- not a great choice if you want to understand economic fluctuations: unrealistic to believe the financial crisis of 2008 or the current coronavirus crisis was caused by a productivity shock

Basics

- Most macro data come in the form of a time-series:

$$x_1, x_2, x_3, ...$$

- We will use x_t to denote the **observed** series and also to denote the corresponding random variable (that could have taken many values different from the observed x_t)
- When viewed as random variables, time-series models are about the joint distribution of big vectors of variables:

$$[x_1, x_2, ..., x_T]$$

- What is special is that the subindices of the x's correspond to time
- Until further notice, we consider the **univariate** case

Moments of a Time-Series Process

- Given a time-series process $x_1, x_2, x_3, \ldots$, the first moments are given by $\mathrm{E}[x_t]$, the second moments by $\mathrm{E}[x_t x_s]$ for all combinations of t and s, the third moments by $\mathrm{E}[x_t x_s x_u]$ and so on.
- A time-series process x_t is Gaussian, if all linear combinations of the x_t have a normal distribution (of course, the mean and variance of this distribution can depend on the linear combination)
- Since a normal random variable is determined by its mean and variance, higher (than the second) moments of a Gaussian process are determined by its first and second moments

1. Stationarity

Stationarity is important because we need some sort of regularity in the observed time-series to be able to estimate the underlying process (and hence, the underlying macroeconomic model). If the series we are interested in (or some simple transformation thereof) is not stationary, it is difficult to take macro models to the data.

Strong Stationarity

The process x_t in discrete time (t takes only integer values) is **strongly stationary** if for all positive integers k, $t_1, \dots, t_k$ and h the joint distributions of

$$(x_{t_1}, x_{t_2}, \dots, x_{t_k}) \text{ and } (x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h})$$

are the same.

It follows that:

- the distribution of x_t does not depend on t
- the distributions of (x_t, x_s) and (x_{t+h}, x_{s+h}) are the same

Strong Stationarity

- That the distribution of x_t does not depend on t means that the probabilities

$$\Pr\{x_t \le u\}$$

  do not depend on t (the observed values for x_t will vary with t).
- It follows that $\mathrm{E}[x_t]$ does not depend on t
- Similarly, that the distribution of (x_{t+h}, x_{s+h}) does not depend on h means that the probabilities

$$\Pr\{x_{t+h} \le u_1, x_{s+h} \le u_2\}$$

  do not depend on h and only depend on $t - s$
- It follows that $\mathrm{Cov}(x_t, x_{t+h}) \equiv \mathrm{E}[x_t x_{t+h}] - \mathrm{E}[x_t]\mathrm{E}[x_{t+h}]$ depends only on h.
- Thus the definition of strong stationarity is equivalent to the distribution of the process being invariant under time-shifts

Weak Stationarity

The process x_t in discrete time is **weakly stationary** or **covariance stationary** if $\mathrm{E}|x_t| < \infty$, $\mathrm{E}x_t^2 < \infty$ and:

- $\mathrm{E}x_t$ does not depend on t
- $\mathrm{E}x_t x_{t+h}$ does not depend on t for $h = 0, 1, 2, \ldots$

Note that:

- Strong stationarity and finite second moments $\Longrightarrow$ weak stat.
- Gaussian process and weak stat. $\Longrightarrow$ strong stationarity

2. ARMA Models

1. White noise
2. Basic ARMA models
3. Lag operators and lag polynomials

White Noise

- Building block for linear time-series models
- The ε_t follow a white noise process if the ε_t are i.i.d. $N(0, \sigma_\varepsilon^2)$.
- The normality assumption will be relaxed soon.

Properties of a White Noise

1. Unpredictable (1):

$$\mathrm{E}(\varepsilon_t) = \mathrm{E}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = \mathrm{E}(\varepsilon_t | \text{all infromation at } t-1) = 0.$$

2. Unpredictable (2):

$$\mathrm{E}(\varepsilon_t \varepsilon_{t-j}) = \mathrm{Cov}(\varepsilon_t, \varepsilon_{t-j}) = 0.$$

3. Conditional homoscedasticity:

$$\mathrm{Var}(\varepsilon_t) = \mathrm{Var}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = \mathrm{Var}(\varepsilon_t | \text{all infromation at } t-1) = \sigma_\varepsilon^2.$$

- Normality is not essential
- Property 1 is essential: **martingale difference** property

Basic ARMA Models

Use a white noise ε_t to build more interesting models:

$$\begin{aligned}
\text{AR(1)}: \quad x_t &= a x_{t-1} + \varepsilon_t, \\
\text{MA(1)}: \quad x_t &= \varepsilon_t + b\varepsilon_{t-1}, \\
\text{ARMA(1,1)}: \quad x_t &= a_1 x_{t-1} + \varepsilon_t + b\varepsilon_{t-1}, \\
\text{AR(p)}: \quad x_t &= \textstyle\sum_{k=1}^{p} a_k x_{t-k} + \varepsilon_t, \\
\text{MA(q)}: \quad x_t &= \varepsilon_t + \textstyle\sum_{k=1}^{q} b_k \varepsilon_{t-k}, \\
\text{ARMA(p,q)}: \quad x_t &= \textstyle\sum_{k=1}^{p} a_k x_{t-k} + \varepsilon_t + \sum_{k=1}^{q} b_k \varepsilon_{t-k}.
\end{aligned}$$

Basic ARMA Models

- All of them correspond to stochastic difference equations, where x_t is a linear combination of past values of the x's and current and past values of the shocks (often referred to as the **innovation process**).
- ε should be thought of as the exogenous variable, x the as the endogenous variable.
- All these models have zero mean, they are used to represent the **deviation** from the mean value of x: $\overline{x}$ or, more generally, the deviation from some deterministic trend.

Lag operators and lag polynomials

- Lag operator moves the time-index back one unit:

$$Lx_t \equiv x_{t-1}.$$

- Hence:

$$L^2 x_t \equiv L(Lx_t) = Lx_{t-1} = x_{t-2}.$$

- In general, for $j = 0, 1, 2, 3, \ldots$:

$$\begin{aligned} L^j x_t &= x_{t-j}, \\ L^{-j} x_t &= x_{t+j}. \end{aligned}$$

- Strictly speaking, the L operator transforms time-series into time-series

Lag operators and lag polynomials

- With this notation we can rewrite ARMA models
- For example, the AR(1) process we saw earlier:

$$x_t = a x_{t-1} + \varepsilon_t \Longrightarrow x_t = aLx_t + \varepsilon_t \Longrightarrow (1 - aL)x_t = \varepsilon_t,$$

and defining the lag-polynomial $a(L) \equiv 1 - aL$ we have:

$$a(L)x_t = \varepsilon_t.$$

- For the general ARMA(p,q) process we define:

$$\begin{aligned} a(L) &= 1 - a_1 L - a_2 L^2 - \cdots - a_p L^p, \\ b(L) &= 1 + b_1 L + b_2 L^2 + \ldots + b_q L^q. \end{aligned}$$

Then:

$$a(L)x_t = b(L)\varepsilon_t.$$

Representing an AR(1) as an MA(∞)

- Assume x_t follows an AR(1):

$$x_t = a x_{t-1} + \varepsilon_t.$$

- Applying this expression recursively:

$$\begin{aligned}
x_t &= a(a x_{t-2} + \varepsilon_{t-1}) + \varepsilon_t, \\
&= a^2 x_{t-2} + a\varepsilon_{t-1} + \varepsilon_t, \\
&\vdots \\
&= a^k x_{t-k} + a^{k-1}\varepsilon_{t-k+1} + \cdots + a\varepsilon_{t-1} + \varepsilon_t.
\end{aligned}$$

- If $|a| < 1$, so that $\lim_{k\to\infty} a^k x_{t-k} = 0$, it follows that:

$$x_t = \sum_{k\geq 0} a^k \varepsilon_{t-k}.$$

- Thus an AR(1) can be expressed as an MA(∞).

Representing an AR(1) as an MA(∞)

- Next we repeat the derivation using lag-polynomials.
- To do this we use the following identity for lag-polynomials:

$$\frac{1}{1-aL} = \sum_{k\geq 0} a^k L^k,$$

  as long as $|a| < 1$.
- This identity can be proved formally and is analogous to the well known geometric series expression for real (and complex) numbers z that satisfy $|z| < 1$:

$$\frac{1}{1-z} = \sum_{k\geq 0} z^k.$$

Representing an AR(1) as an MA(∞)

- From
$$(1 - aL)x_t = \varepsilon_t,$$
we have:
$$\begin{aligned} x_t &= \frac{1}{1-aL}\varepsilon_t \\ &= \sum_{k\geq 0} a^k L^k \varepsilon_t \\ &= \sum_{k\geq 0} a^k \varepsilon_{t-k}. \end{aligned}$$
- The condition that $|a| < 1$ is important to apply the above trick

Representing an AR(1) as an MA(∞)

- We define the roots of the AR-polynomial

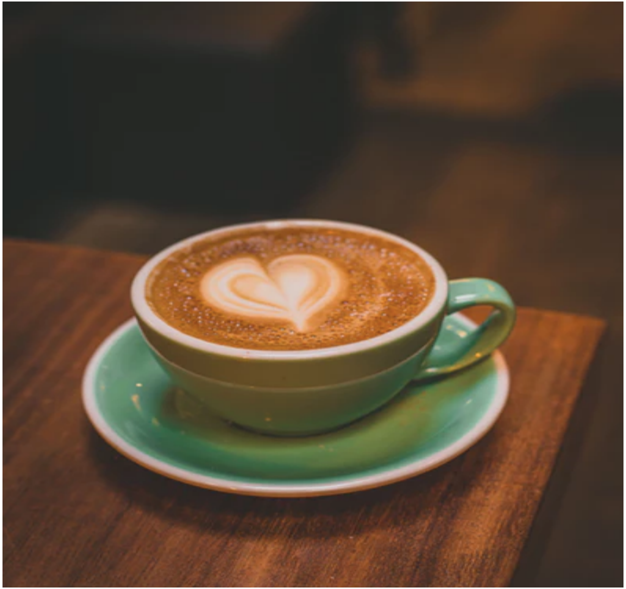
$$a(L) = 1 - aL$$

  as the values of z such that

$$a(z) = 1 - az = 0.$$

- It follows that the root of $a(z)$ is $1/a$
- Hence, the root of the AR-polynomial must have absolute value larger than one to be able to go from an AR(1) to an MA(∞)

Coffee Break



Representing an AR(p) as an MA(∞)

- The same idea can be used to represent the AR(p) process:

$$x_t = a_1 x_{t-1} + \cdots a_p x_{t-p} + \varepsilon_t$$

  as an MA(∞).

- This requires that all the roots of $a(z) = 1 - a_1 z - \cdots a_p z^p$ have absolute value larger than one
- In this case we say that the AR-polynomial is **invertible**
- Similarly, we can represent an MA(q) process

$$x_t = b(L)\varepsilon_t,$$

  with $b(L) = 1 + b_1 L + \cdots + b_q L^q$ as an AR(∞) process

- For such a representation to exist, the roots of the MA-polynomial $b(z) = 1 + b_1 z + \cdots b_q z^q$ must have absolute value larger than one

Stationary ARMA Processes

- An MA(∞), $x_t = \sum_{k\geq 0} b_k \varepsilon_{t-k}$, has finite second moments if $\sum_{k\geq 0} b_k^2 < \infty$. This condition ensures (strong and weak) stationarity.
- An AR(p):

$$a(L)x_t = \varepsilon_t$$

  is (strongly and weakly) stationary if and only if the roots of $a(z)$ are outside the unit circle. Roots inside the unit circle lead to an explosive process, roots on the unit circle to an integrated process (more on the latter shortly).
- An ARMA(p,q) is stationary if the AR-polynomial has all roots outside the unit circle.

Beyond Gaussian ARMA Processes

- The process ε_t is a (generalized) white noise if it is uncorrelated, with zero mean and finite variance σ_ε^2. Most time-series textbooks have this process in mind when they refer to white noise.
- Note that "uncorrelated" is weaker than "independent" for a non-Gaussian process. Even though "independent" captures better the economic idea of a **shock** or **innovation**, uncorrelated turns out to be the right concept when stating Wold's Representation (see below).

Beyond Gaussian ARMA Processes

- A (generalized) ARMA process differs from the ARMA processes we have seen so far only in that the innovation process is a generalized white noise. The conditions for stationarity are the same as in the Gaussian case.
- If the ARMA process x has a non-Gaussian white noise, then x itself is non-Gaussian.
- There is no ex-ante reason why macroeconomic time series should be Gaussian. For example, some macro series are driven by large and infrequent shocks, suggesting fatter tails than those of a Gaussian process.

3. Integrated processes

- Many macroeconomic time-series are non stationary processes with stationary growth rates: e.g., GDP, consumption, price level, stock of capital.
- x_t is an **integrated process of order 1** if x_t is not stationary and $\Delta x_t \equiv x_t - x_{t-1} \equiv (1-L)x_t$ is stationary. We then say that x_t is $I(1)$.
- x_t is an **integrated process of order 2** if x_t and Δx_t are not stationary and $\Delta^2 x_t \equiv \Delta(\Delta x_t)$ is stationary. We then say that x_t is $I(2)$.
- There exist statistical tests to see whether a series comes from an I(1) or a stationary process ('unit root tests'). There also exist tests to determine whether, given two (or more) non stationary series, there exists a (non-trivial) linear combination of these series that is stationary ('cointegration tests'). We won't cover these topics here.
- A stationary process is sometimes referred to as an I(0) process

Integrated Process

- When x_t is not stationary and Δx_t is ARMA(p,q), we say that x_t is ARIMA(p,1,q).
- If x_t and Δx_t are non stationary and $\Delta^2 x_t$ is ARMA(p,q), we say that x_t is ARIMA(p,2,q).
- Similarly, if x_t and Δx_t are non stationary and $\Delta^2 x_t$ is stationary, we say that x_t is I(2).

Random Walk

- x_t follows a random walk if there exists a white noise process ε_t and a constant g (drift or growth rate) s.t.:

$$x_t = g + x_{t-1} + \varepsilon_t. \tag{1}$$

- From (1):

$$\Delta x_t = g + \varepsilon_t.$$

- Hence, x_t follows a random walk $\Longleftrightarrow$ Δx_t follows a white noise (plus a constant) $\Longleftrightarrow$ x_t follows an ARIMA(0,1,0).
- The random walk is the most important (and simplest) I(1) process.

4. FORECASTING

- A natural candidate to forecast x_{t+k} based upon information available at time t, $\mathscr{I}_t$, is the conditional expectation:

$$\mathrm{E}[x_{t+k}|\mathscr{I}_t].$$

  which we denote by $\widehat{x}_{t+k|t}$

- Since we have a one dimensional setting: $\mathscr{I}_t = \{x_t, x_{t-1}, x_{t-2}, ...\}$.
- Precision of forecast measured by **mean-squared error**:

$$\mathsf{MSE}_k \equiv \mathrm{E}[(x_{t+k} - \widehat{x}_{t+k|t})^2|\mathscr{I}_t].$$

- Can show that $\mathrm{E}[x_{t+k}|\mathscr{I}_t]$ minimizes MSE_k over all possible functions $g(x_t, x_{t-1}, ...)$, i.e., it solves:

$$\min_g \mathsf{MSE}_k(g) \equiv \mathrm{E}[(x_{t+k} - g(x_t, x_{t-1}, ...))^2|\mathscr{I}_t]. \tag{2}$$

- When x is Gaussian, it suffices to consider **linear** functions g, i.e., the solution to the problem above is a linear function of $x_t, x_{t-1}, ...$
- If we only consider linear functions in (2) we obtain the **linear projection** of x_t on $\mathscr{I}_t$, denoted by $LP(x_t|\mathscr{I}_t)$.

Forecasting an AR(1)

In the derivations that follow, we use the fact that the properties for the expectations operator also hold for conditional expectations. We also use the fact that the εs are innovations, and therefore the best forecast of future values is their mean, that is, zero.

We have:

$$
\begin{aligned}
\mathrm{E}[x_{t+1}|\mathscr{I}_t] &= \mathrm{E}[ax_t + \varepsilon_{t+1}|\mathscr{I}_t] = ax_t, \\
\mathrm{E}[x_{t+2}|\mathscr{I}_t] &= \mathrm{E}[ax_{t+1} + \varepsilon_{t+2}|\mathscr{I}_t] = a\mathrm{E}[x_{t+1}|\mathscr{I}_t] = a^2 x_t, \\
\vdots \quad & \vdots \quad \vdots \\
\mathrm{E}[x_{t+k}|\mathscr{I}_t] &= \mathrm{E}[ax_{t+k-1} + \varepsilon_{t+k}|\mathscr{I}_t] = a\mathrm{E}[x_{t+k-1}|\mathscr{I}_t] = \cdots = a^k x_t.
\end{aligned}
$$

The corresponding forecast errors are:

$$\begin{aligned}
x_{t+1} - \mathrm{E}[x_{t+1}|\mathscr{I}_t] &= \varepsilon_{t+1}, \\
x_{t+2} - \mathrm{E}[x_{t+2}|\mathscr{I}_t] &= \varepsilon_{t+2} + a\varepsilon_{t+1}, \\
\vdots \quad &\vdots \quad \vdots \\
x_{t+k} - \mathrm{E}[x_{t+k}|\mathscr{I}_t] &= \varepsilon_{t+k} + a\varepsilon_{t+k-1} + \cdots + a^{k-1}\varepsilon_{t+1}.
\end{aligned}$$

And hence:

$$\begin{aligned}
\mathsf{MSE}_1 &= \sigma_\varepsilon^2, \\
\mathsf{MSE}_2 &= (1+a^2)\sigma_\varepsilon^2, \\
\vdots \quad & \vdots \quad \vdots \\
\mathsf{MSE}_k &= (1+a^2+\cdots+a^{2k-2})\sigma_\varepsilon^2,
\end{aligned}$$

and we have that:

$$\begin{aligned}
\lim_{k\to\infty} \mathrm{E}[x_{t+k}|\mathscr{I}_t] &= 0, \\
\lim_{k\to\infty} \mathsf{MSE}_k &= \frac{\sigma_\varepsilon^2}{1-a^2} = \sigma_x^2.
\end{aligned}$$

Where we used that

$$x_t = a x_{t-1} + \varepsilon_t \Longrightarrow \mathrm{Var}(x_t) = a^2 \mathrm{Var}(x_{t-1}) + \mathrm{Var}(\varepsilon_t) \Longrightarrow \sigma_x^2 = a^2\sigma_x^2 + \sigma_\varepsilon^2 \Longrightarrow \sigma_x^2 = \sigma_\varepsilon^2/(1-a^2).$$

Thus, long run forecasts of x are "trivial" (equal to the mean).

Forecasting an MA(1)

- The process:

$$x_t = \varepsilon_t + b\varepsilon_{t-1}.$$

- As usual, we have $\mathscr{I}_t = \{x_t, x_{t-1}, \ldots\}$. We assume $|b| < 1$ which implies that x_t can be expressed as a function of ε_t, ε_{t-1}, ε_{t-2}, ..., so that $\mathscr{I}_t = \{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$.
- Straightforward calculations then show that:

$$\begin{aligned} \widehat{x}_{t+1|t} &= b\varepsilon_t, \\ \widehat{x}_{t+k|t} &= 0, \qquad k = 2, 3, \cdots \end{aligned}$$

and therefore

$$\begin{aligned} \mathsf{MSE}_1 &= \sigma_\varepsilon^2, \\ \mathsf{MSE}_k &= (1 + b^2)\sigma_\varepsilon^2 = \sigma_x^2, \qquad k = 2, 3, \ldots \end{aligned}$$

Forecasting a Random Walk

Given a random walk

$$x_t = g + x_{t-1} + \varepsilon_t$$

we denote $\mathscr{I}_t = \{x_t, x_{t-1}, x_{t-2}, ...\}$.

We then have:

$$\widehat{x}_{t+k|t} \equiv \mathrm{E}[x_{t+k}|\mathscr{I}_t] = \mathrm{E}[x_t + \Delta x_{t+1} + \cdots + \Delta x_{t+k}|\mathscr{I}_t] = x_t + kg.$$

Therefore:

$$\mathrm{Var}[\widehat{x}_{t+k|t}|\mathscr{I}_t] = \mathrm{Var}[x_t + \Delta x_{t+1} \cdots + \Delta x_{t+k}|\mathscr{I}_t] = k\sigma_\varepsilon^2.$$

Thus, if we observe $x_1, \ldots, x_T$ for a random walk (with known drift), and we forecast x_{T+k}, we have:

- MSE_k grows linearly with k.
- If x_T increases by one unit, then $\widehat{x}_{T+k|t}$ also increases by one unit.

Both properties mentioned above hold approximately for any I(1) process: the variance of the forecast error grows without bound with k. By contrast, for a stationary process, the variance of the forecast error is bounded by the variance of the process (which is finite).

5. Impulse Response Function

An important objective of macroeconomics is to describe how macro variables respond to shocks of different sorts:

1. How do output and inflation respond to an oil shock?
2. How do consumption and inflation respond to an increase in the short term interest rate?
3. How do macro aggregates respond to the various shocks associated with the ongoing pandemic crisis?
4. How does investment respond to an investment tax credit
5. What is the effect of a fiscal stimulus on output, productivity, inflation?
6. How does Chile's output respond to a shock in the copper price.

Often we wish to distinguish between the response to the entire shock and to the **unexpected** component in the shock. Of course, this distinction is irrelevant when the shock is a white noise.

The concept of **impulse response function** (IRF) is very useful to answer the above questions. It is a central concept when taking macroeconomic models to the data and used as a convenient summary of macroeconomic models.

Impulse Response Function for an ARMA Process

Assume x is an ARMA process with innovation ε:

$$a(L)x_t = b(L)\varepsilon_t.$$

The **Impulse Response Function** (IRF) of x with respect to ε is defined as:

$$\mathrm{IRF}_k \equiv \frac{\partial x_{t+k}}{\partial \varepsilon_t}, \quad k = 0, 1, 2, \ldots$$

IRF of an MA Process

- Consider:

$$x_t = \sum_{j \geq 0} b_j \varepsilon_{t-j}.$$

- Then:

$$\text{IRF}_k = \frac{\partial x_{t+k}}{\partial \varepsilon_t} = b_k$$

- Hence, for an MA-process, the IRF is given by the coefficients of the MA representation.

IRF of an AR(1)

From

$$x_t = a x_{t-1} + \varepsilon_t$$

it follows that

$$\mathrm{IRF}_0 = \frac{\partial x_t}{\partial \varepsilon_t} = \frac{\partial (a x_{t-1} + \varepsilon_t)}{\partial \varepsilon_t} = 1.$$

Furthermore, for $k \geq 1$:

$$\mathrm{IRF}_k = \frac{\partial x_{t+k}}{\partial \varepsilon_t} = \frac{\partial (a x_{t+k-1} + \varepsilon_{t+k})}{\partial \varepsilon_t} = a \frac{\partial x_{t+k-1}}{\partial \varepsilon_t} = a \mathrm{IRF}_{k-1}$$

which combined with the expression above for IRF_0 leads to:

$$\mathrm{IRF}_k = a^k, \qquad k = 0, 1, 2, \cdots$$

An alternative derivation of the IRF of an AR(1) follows from its MA representation and the above expression:

$$x_{t+k} = \sum_{j\geq 0} a^j \varepsilon_{t+k-j}$$

We then have:

$$\mathrm{IRF}_k = \frac{\partial x_{t+k}}{\partial \varepsilon_t} = a^k, \quad k = 0, 1, 2, \ldots$$

IRF of an AR(p)

Consider an AR(p) process:

$$x_t = a_1 x_{t-1} + \cdots + a_p x_{t-p} + \varepsilon_t.$$

A similar derivation to the one we made for an AR(1) leads to:

$$\mathrm{IRF}_k = a_1 \mathrm{IRF}_{k-1} + \cdots + a_p \mathrm{IRF}_{k-p}.$$

Thus, from basic difference equations, assuming the roots of the p roots of the AR polynomial $a(z) = 1 - \sum_{k=1}^{p} a_k z^k$ are G_1, G_2, ..., G_p, all different, we have:

$$\mathrm{IRF}_k = \sum_{j=1}^{p} c_j G_j^{-k}$$

where the constants c_j are determined from noting that initial conditions are given by $\mathrm{IRF}_0 = 1$ and $\mathrm{IRF}_k = 0$ for $k < 0$.

Impulse Response for an AR(2)

Next we derive the IRF for a family of AR(2) processes.

The result we derive is important, since we establish necessary and sufficient conditions (within the family of AR(2) processes considered) for the IRF to be hump shaped.

We will apply this result often.

IRF for an AR(2): Result

Consider the following AR(2) process:

$$(1 - a_1 L)(1 - a_2 L)x_t = \epsilon_t,$$

where a_1 and a_2 are real numbers such that $0 < a_2 < a_1 < 1$.

We then have that the IRF of this process is hump shaped if and only if $a_1 + a_2 > 1$.

Comment: The result applies to any stationary AR(2) process where the AR polynomial has **real** (and different) roots.

Proof

We have that

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \epsilon_t$$

with $\phi_1 = a_1 + a_2$ and $\phi_2 = -a_1 a_2$

From what we saw for an AR(p) process:

$$\begin{aligned}
\mathrm{IRF}_k &= \phi_1 \mathrm{IRF}_{k-1} + \phi_2 \mathrm{IRF}_{k-2}, \;\; k \geq 2; \\
\mathrm{IRF}_0 &= 1; \\
\mathrm{IRF}_1 &= \phi_1 \mathrm{IRF}_0 = \phi_1 = a_1 + a_2.
\end{aligned}$$

The roots of the AR polynomial are $1/a_1$ and $1/a_2$ and therefore

$$\mathrm{IRF}_k = C_1 a_1^k + C_2 a_2^k \tag{3}$$

To find C_1 and C_2 we note that:

$$\text{IRF}_0 = 1 \Longrightarrow C_1 + C_2 \;\; = 1$$

and

$$\text{IRF}_1 = a_1 + a_2 \Longrightarrow C_1 a_1 + C_2 a_2 \;\; = \phi_1.$$

Solving we obtain

$$C_1 = \frac{a_1}{a_1 - a_2}, \qquad C_2 = -\frac{a_2}{a_1 - a_2}.$$

Using the above expressions it is easy to show that

$$\text{IRF}_k > \text{IRF}_{k+1} \iff \left(\frac{a_1}{a_2}\right)^{k+1} > \frac{1 - a_2}{1 - a_1}. \tag{4}$$

Since $a_1/a_2 > 1$, we have that $(a_1/a_2)^{k+1}$ increases without bound with k. It follows that there exists an integer K s.t. for all $k \geq K$ we have IRF_k decreasing.

All we need to show now to conclude that IRF_k is hump shaped is to prove that for small values of k we have IRF_k increasing. This will be the case if and only if $\text{IRF}_0 < \text{IRF}_1$. And, from the expressions for IRF_0 and IRF_1 we derived on the previous slide, we have

$$\text{IRF}_0 < \text{IRF}_1 \Longleftrightarrow 1 < a_1 + a_2.$$

And this concludes the proof. $\square$

IRF of a Random Walk

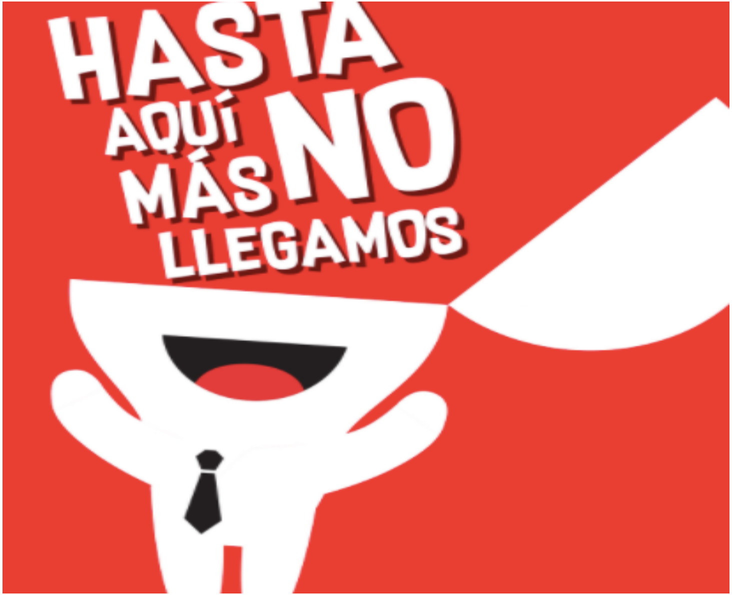
Consider the randon walk

$$x_t = g + x_{t-1} + \varepsilon_t. \tag{5}$$

We then have, for $k \geq 0$:

$$\begin{aligned}
\mathrm{IRF}_k &= \frac{\partial x_{t+k}}{\partial \varepsilon_t} \\
&= \frac{\partial [g + x_{t+k-1} + \varepsilon_{t+k}]}{\partial \varepsilon_t} \\
&\vdots \quad \frac{\partial [kg + x_t + \varepsilon_{t+k} + \varepsilon_{t+k-1} + \cdots + \varepsilon_t]}{\partial \varepsilon_t} \\
&= 1.
\end{aligned}$$

A marginal shock to ε_t of increases the current and future values of x one-for-one.


HASTA
AQUÍ
MÁS
NO
LLEGAMOS

Time Series

Rational Expectations

Time Series

Rational Expectations

Rational Expectations

Usually rational expectations means that economic agents know the true model.

Yet rational expectations is much more than that.

And, in solving the Calvo model in the preceding section, we have used the broader meaning for rational expectations.

We illustrate the main ideas with a simple example, close to the examples used by Muth (Econometrica, 1961)

And then explain where and how we used rational expectations in its more nuanced sense in the preceding section.

Supply of an agricultural good in period t, q_t^S, depends on the price expected by producers when they make production decisions, p_t^e, as captured by the following supply function:

$$q_t^S = c p_t^e + v_t \tag{6}$$

where v_t is i.i.d. and independent of previous prices and quantities.

Demand is given by:

$$q_t^D = 1 - p_t. \tag{7}$$

To determine equilibrium prices and quantities, we need to make assumptions about how producers form expectations about future prices.

Naive expectations

Assume $p_t^e = p_{t-1}$. That is, people expect the current price to prevail in the next period.

Solving for $q_t^S = q_t^d$ leads to

$$p_t = 1 - c p_{t-1} - v_t.$$

and p_t follows an AR(1).

Two problems (the first reason is more important):

- People assume $p_t^e = p_{t-1}$, yet when period t price occurs, they should realize that their expectation was wrong, which they do not. That is, expectations are not rational in the sense that producers do not learn for their forecast mistakes.
- The process p_t will be explosive if $|c| > 1$.

Rational expectations

We assume that when forming expectations about period t prices, producers use all information available in the best possible way, that is:

$$p_t^e = \mathrm{E}_{t-1} p_t \equiv \mathrm{E}[p_t | \mathscr{I}_{t-1}],$$

with $\mathscr{I}_{t-1} = \{p_{t-1}, q_{t-1}, p_{t-2}, q_{t-2}, ...\}$.

We also assume producers know the model and the corresponding parameters (c in our case).

Equating supply and demand:

$$c\mathrm{E}_{t-1} p_t + v_t = 1 - p_t.$$

Taking E_{t-1} on both sides yields

$$\mathrm{E}_{t-1} p_t = \frac{1}{1+c}.$$

Substituting this expression in (7) and (6):

$$\begin{aligned} p_t &= \frac{1}{1+c} - v_t, \\ q_t &= \frac{c}{1+c} + v_t. \end{aligned}$$

That is:

- producers assume $p_t = \frac{1}{1+c} - v_t$,
- therefore at time $t-1$ they set $p_t^e = \mathrm{E}_{t-1} p_t = 1/(1+c)$,
- the process of prices, p_t, they observe will follow the process they assumed,
- therefore they will not regret having made the expectations they made

Rational Expectations

Next we relax the i.i.d. assumption for v_t and assume an AR(1) process instead:

$$v_t = \phi v_{t-1} + e_t,$$

with $|\phi| < 1$ and e_t is i.i.d.

Assuming rational expectations and equating supply and demand yields

$$c\mathrm{E}_{t-1} p_t + v_t = 1 - p_t.$$

Taking E_{t-1} on both sides:

$$(1+c)\mathrm{E}_{t-1} p_t + \phi v_{t-1} = 1$$

and we obtain

$$\mathrm{E}_{t-1} p_t = \frac{1}{1+c} - \frac{\phi}{1+c} v_{t-1}.$$

Substituting the expression above in (6):

$$p_t = \frac{1}{1+c} - v_t + \frac{\phi c}{1+c} v_{t-1}.$$

Applying $1-\phi L$ on both sides:

$$(1-\phi L)p_t = \frac{1-\phi}{1+c} - e_t + \frac{\phi c}{1+c} e_{t-1}.$$

Letting $\widetilde{e}_t \equiv -e_t$, we trivially have that $\widetilde{e}_t$ will also be i.i.d. and

$$(1-\phi L)p_t = \frac{1-\phi}{1+c} + \widetilde{e}_t - \frac{\phi c}{1+c} \widetilde{e}_{t-1}.$$

We conclude that p_t follows an ARMA(1,1) with AR coefficient ϕ and MA coefficient $-\phi c/(1+c)$.

Rational Expectations

Using the ARMA(1,1) representation for p_t and (7) we conclude that q_t also follows an ARMA(1,1):

$$(1-\phi L)q_t = \frac{c+\phi}{1+c} + e_t - \frac{\phi c}{1+c} e_{t-1}.$$

As in the i.i.d. case, producers assume prices and quantities follow the ARMA(1,1) processes derived above, they form their expectations about future prices accordingly, and once prices and quantities are realized they see confirmation for the processes they assumed.

That is, producers have **model consistent expectations**, which over time became known as rational expectations.

Rational expectations and commodity markets

Summary

Agents assume
a process for p_t

$\nearrow$ $\searrow$

Actual process
for p_t determined

Against set
$p_t^e = \mathrm{E}_{t-1} p_t$

$\nwarrow$ $\swarrow$

Market clears:
p_t and q_t determined

Macroeconomía I: Cátedra 1

Series de tiempo.para macroeconomistas

Eduardo Engel

ENECO 630. Macroeconomía I

Magíster en Economía, FEN, U. de Chile.

Abril 8, 2021.