Modelización univariante

Miguel Jerez

Universidad Complutense de Madrid

Febrero 2015

Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Introducción (I): Objetivo

El análisis univariante de series temporales es distinto de la Econometría tradicional:

- Se considera una sola variable endógena (por eso es "…univariante")
- No consideran variables exógenas
- Se trata de construir un modelo sencillo que aproveche la inercia de los datos pasados para predecir los valores futuros de la serie

Ejemplos. Algunos modelos univariantes sencillos son:

- Autorregresivo de primer y segundo orden:

$$z_t = c + \phi z_{t-1} + a_t$$

$$z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

- Medias móviles de primer orden:

$$z_t = \mu_z + a_t - \theta a_{t-1}$$

- Paseo aleatorio con deriva:

$$y_t = c + y_{t-1} + a_t$$

Introducción (II): Enlace con el análisis de regresión

Aparte de su utilidad intrínseca, el análisis univariante de series temporales puede combinarse con las ideas básicas del análisis de regresión

Los modelos resultantes combinan dos elementos:

- El “modelo de relación” que formaliza la relación causal entre las variables explicativa y la variable endógena
- El “modelo de error”, que predice la parte de la variable endógena que no explican las exógenas

El modelo del error puede interpretarse como una representación de posibles variables omitidas

Ejemplo. Un modelo de regresión estándar

$$y_t = \mathbf{x}_t^T \boldsymbol{\beta} + \varepsilon_t$$

…puede combinarse con un modelo univariante para el error como, por ejemplo:

$$\varepsilon_t = \phi \varepsilon_{t-1} + a_t$$

$$\varepsilon_t = a_t - \theta a_{t-1}$$

Introducción (III): Definiciones y supuestos

El punto de partida del análisis de series temporales son las definiciones de proceso estocástico y serie temporal:

- Proceso estocástico, es un conjunto de variables aleatorias asociadas a distintos instantes de tiempo
- Serie temporal, es un conjunto de observaciones o medidas realizadas secuencialmente en intervalos predeterminados y de igual, o aproximadamente igual, duración

Por tanto, se trata analizar estadísticamente una serie (o vector de series) para ajustarle un modelo especificado a partir de propiedades de los datos

Para ello se realizan tres supuestos acerca del proceso estocástico generador:

- …es lineal, de manera que el valor de la variable hoy depende linealmente de sus valores pasados y de los valores presentes y pasados de otras variables
- …es (débilmente) estacionario, esto es, sus momentos de primer y segundo orden no dependen del tiempo
- …sigue un modelo normal de distribución de probabilidad (proceso "gaussiano")

Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Procesos elementales (I): Notación

Los momentos poblacionales de primer y segundo orden del proceso z_t ; $t = 1, 2, \ldots, n$ se denotan de la siguiente manera:

- Media: $\mathrm{E}(z_t) = \mu_z(t) = \mu_z$
- Varianza (autocovarianza de orden cero): $\mathrm{E}\left[(z_t - \mu_z)^2\right] = \sigma_z^2(t) = \sigma_z^2 = \gamma_0$
- Autocovarianza de orden *k*: $\mathrm{E}\left[(z_t - \mu_z)(z_{t-k} - \mu_z)\right] = \gamma_k(t) = \gamma_k \;\; ; \; k \geq 1$

…en donde en la expresión final estamos teniendo en cuenta el supuesto de estacionariedad débil

A partir de esta notación, la autocorrelación de orden *k* se define como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \;\; ; \; k \geq 1$$

…y puede estimarse a partir de datos muestrales mediante las expresiones:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_{\cdot 0}} \qquad \hat{\gamma}_k = \frac{1}{n} \sum_{t=k+1}^{n} \tilde{z}_t \tilde{z}_{t-k} \;\; ; \; \tilde{z}_t = z_t - \bar{z} \;\; , \; k = 1,2,\ldots$$

como veremos, estas autocorrelaciones muestrales son los instrumentos fundamentales que usaremos para “identificar” (especificar a partir de evidencia muestral) un modelo univariante

Procesos elementales (II): Ruido blanco

Un proceso de ruido blanco representa una variable que: (a) oscila en torno a una media constante, (b) con una volatilidad constante, y (c) cuyo pasado no contiene información útil para predecir valores futuros.

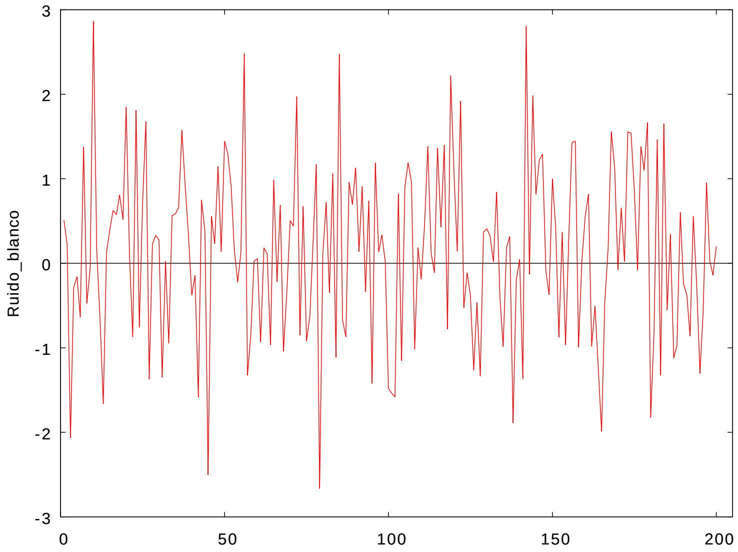
Podemos representar esta variable como $z_t = \mu_z + a_t$ con:

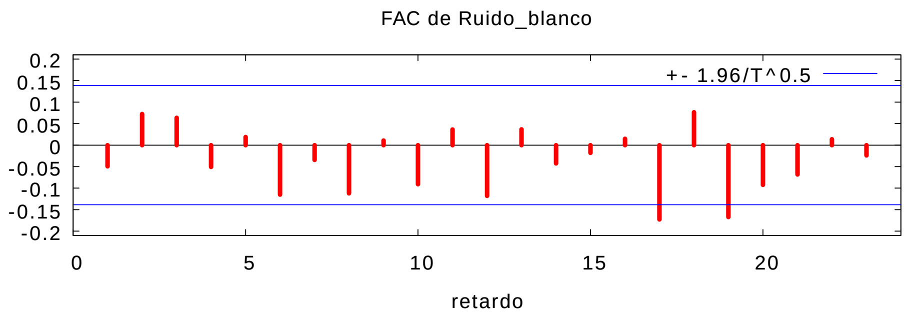
$$E(z_t) = \mu_z \; ; E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = \sigma_a^2 \; ; \rho_k = \frac{\gamma_k}{\gamma_0} = 0 \; ; \; k \geq 1$$

La figura muestra el perfil de 200 observaciones simuladas del proceso de ruido blanco:

$$z_t = a_t \; ; \; a_t \sim \text{iid N}(0,1)$$

…y una representación gráfica de su *función de autocorrelación muestral*





Procesos elementales (III): MA(1)

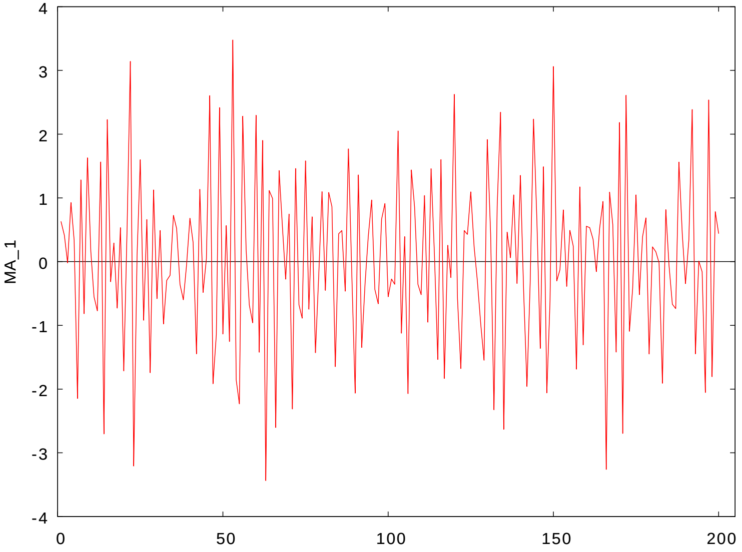
Un proceso de medias móviles de primer orden, MA(1), se define como:

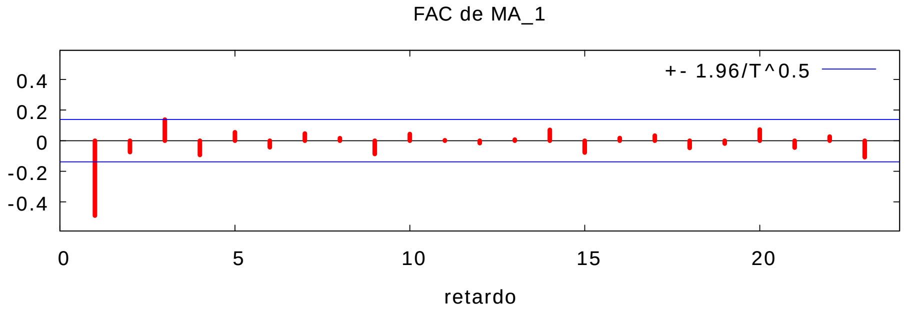
$$z_t = \mu_z + a_t - \theta a_{t-1} \; ; \; |\theta| < 1$$

…y representa una variable cuyo valor actual está correlado con su valor anterior y es independiente de todos los valores previos, ya que:

$$E(z_t) = \mu_z \; ; E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = (1+\theta^2)\sigma_a^2 \; ; \rho_k = \begin{cases} \dfrac{-\theta}{1+\theta^2} & si \; k = 1 \\ 0 & si \; k > 1 \end{cases}$$

Las figuras muestran los gráficos fundamentales de: $z_t = a_t - .8 a_{t-1} \; ; \; a_t \sim \text{iid N}(0,1)$





Procesos elementales (IV): AR(1)

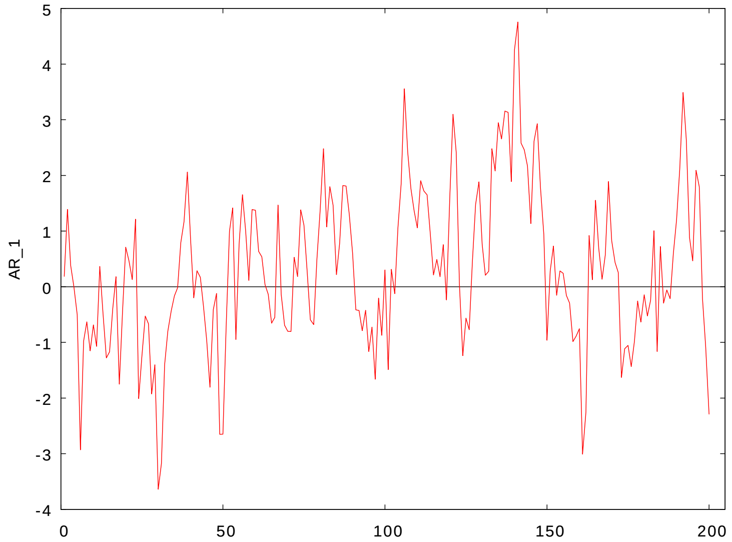
Un proceso autorregresivo de primer orden, AR(1), representa una variable cuyo valor actual está relacionado con su valor anterior mediante un modelo de regresión. Esto es:

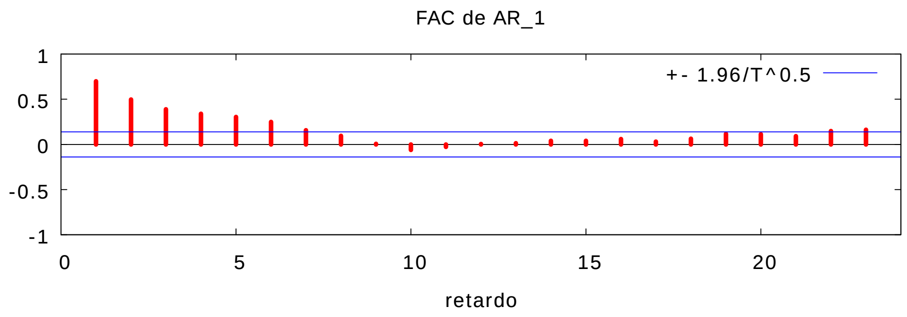
$$z_t = c + \phi z_{t-1} + a_t \;\; ; \;\; |\phi| < 1$$

con:

$$E(z_t) = \mu_z = \frac{c}{1-\phi} \; ; E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = \frac{\sigma_a^2}{1-\phi^2} \; ; \rho_k = \phi^k \;\; ; \; k \geq 0$$

Las figuras muestran los gráficos fundamentales de $z_t = .7 z_{t-1} + a_t \; ; \; a_t \sim \text{iid N}(0,1)$





Procesos elementales (V): Paseo aleatorio

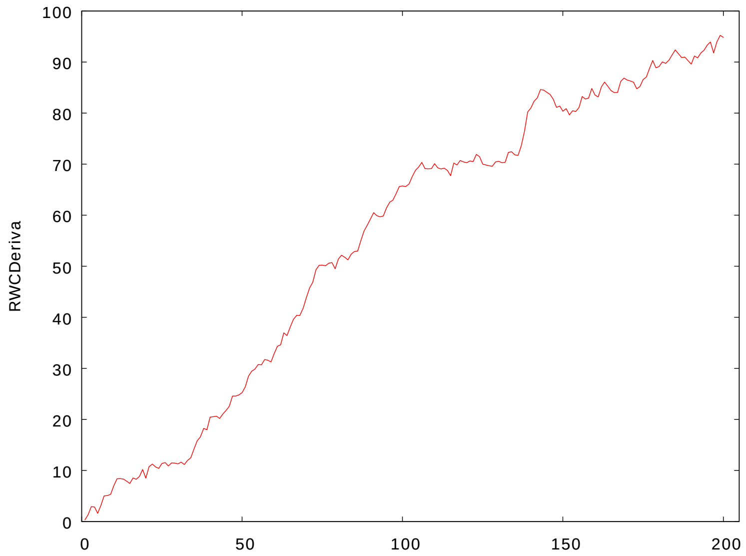
Un paseo aleatorio representa una variable cuyos cambios son ruido blanco y, por tanto, imprevisibles. Esto es:

$$y_t = c + y_{t-1} + a_t$$

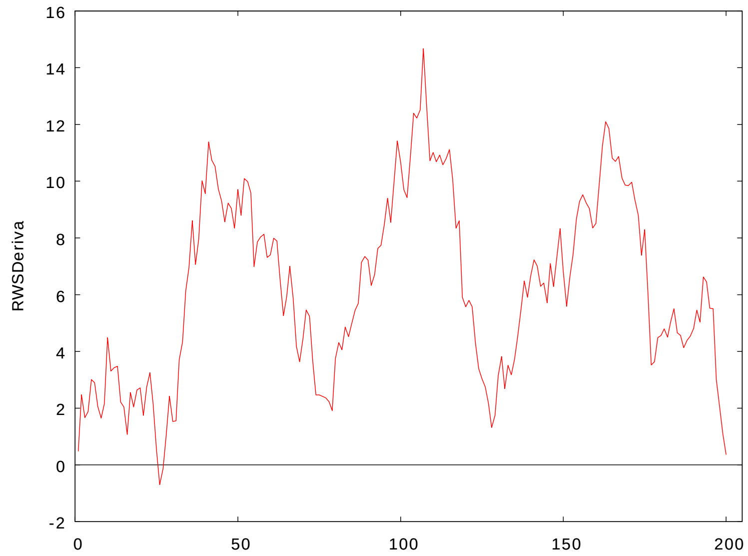
La figura muestra el perfil de dos series generadas por paseos aleatorios con y sin deriva. Como puede observarse:

- Este proceso no es estacionario (bastaría una diferencia para lograr estacionariedad)
- La presencia de un término constante (o "deriva") aporta direccionalidad

$y_t = .5 + y_{t-1} + a_t$; $a_t \sim$ iid N(0,1)



$y_t = y_{t-1} + a_t$; $a_t \sim$ iid N(0,1)



Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Instrumentos (I): Función de autocorrelación simple

La k-ésima autocorrelación muestral simple ($\hat{\rho}_k$) se define como:

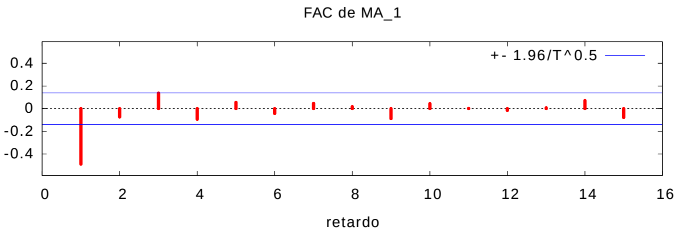
$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \qquad \hat{\gamma}_k = \frac{1}{n}\sum_{t=k+1}^{n} \tilde{z}_t \tilde{z}_{t-k} \;\; ; \; \tilde{z}_t = z_t - \bar{z} \;\; , \; k = 1,2,\ldots$$

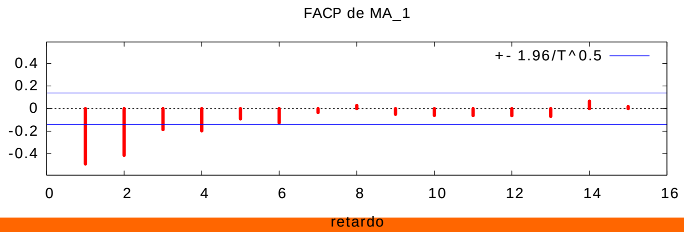
Para valorar la significatividad de las autocorrelaciones pueden usarse los siguientes resultados:

- Para muestras grandes, $s.e.(\hat{\rho}_k) \simeq 1/\sqrt{n}$
- La nula $H_0 : \rho_1 = \rho_2 = \ldots = \rho_K = 0$ puede contrastarse con el estadístico de Ljung-Box:

$$Q(K) = n(n+2)\sum_{k=1}^{K} \frac{1}{n-k}\hat{\rho}_k^2 \underset{H_0}{\sim} \chi^2_{K-p}$$

en donde K es el número de autocorrelaciones consideradas y p es el número de parámetros estimados si la serie está formada por residuos





La figura muestra la función de autocorrelación (ACF) de una muestra del proceso:

$$z_t = a_t - .8a_{t-1} \;\; ; \;\; a_t \sim \text{iid N}(0,1)$$

…que se denomina "media móvil de primer orden" o MA(1). Como veremos, la ACF proporciona pistas claras sobre la estructura del proceso generados de datos

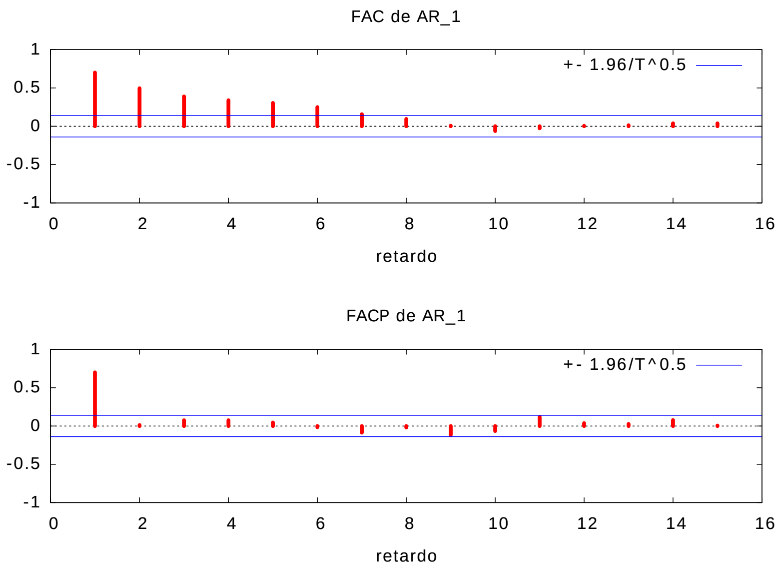
Instrumentos (II): Función de autocorrelación parcial

Otra herramienta del análisis de serie son las “autocorrelaciones parciales”

El coeficiente *k*-ésimo de autocorrelación parcial ($\hat{\phi}_{kk}$) se define como el *k*-ésimo coeficiente MCO de una autorregresión de orden *k*:

$$\tilde{z}_t = \hat{\phi}_{k1}\tilde{z}_{t-1} + \hat{\phi}_{k2}\tilde{z}_{t-2} + \ldots + \hat{\phi}_{kk}\tilde{z}_{t-k} + \hat{\varepsilon}_{kt} \;\; ; \; k = 1,2,\ldots$$

Este estadístico puede interpretarse como una medida de la dependencia lineal entre $\tilde{z}_t$ y $\tilde{z}_{t-k}$, tras descontar el efecto de los retardos intermedios



La figura muestra las autocorrelaciones simples (ACF) y parciales (PACF) del proceso AR(1) (autorregresivo de primer orden):

$$z_t = .7 z_{t-1} + a_t \; ; \; a_t \sim \text{iid N}(0,1)$$

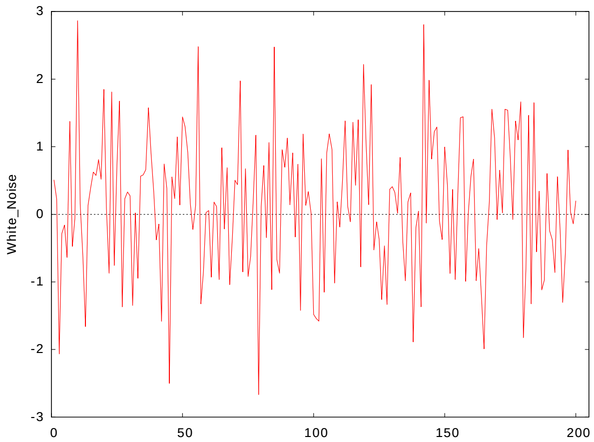
La PACF tiene un sólo valor significativo, lo que proporciona una pista clara sobre la estructura del proceso

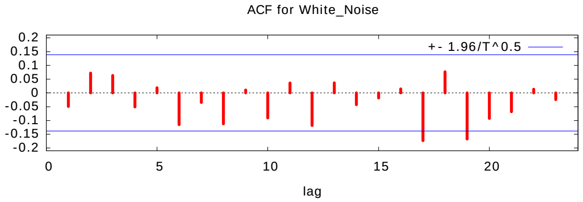
Índice

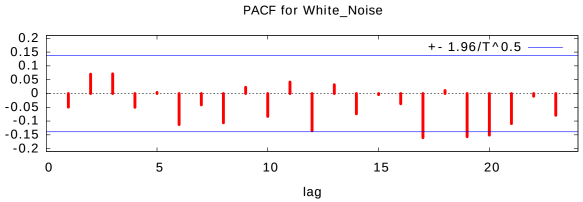
- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Identificación y diagnosis (I)

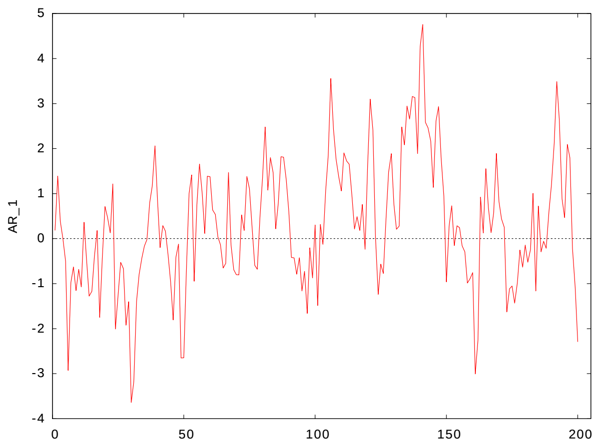
Ruido blanco: $z_t = a_t \ ; \ a_t \sim \text{iid N}(0,1)$

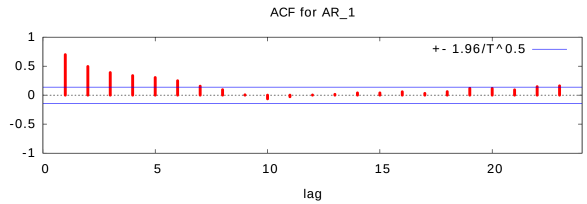


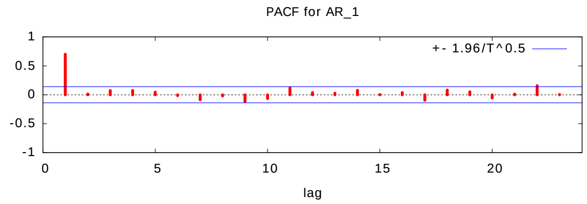




AR(1): $z_t = .7 z_{t-1} + a_t \ ; \ a_t \sim \text{iid N}(0,1)$





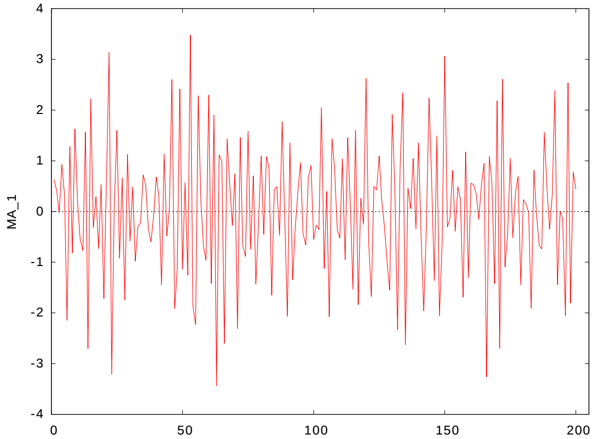


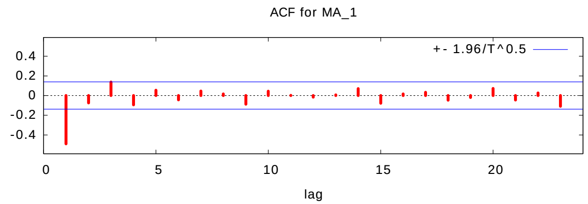
Combinando las herramientas gráficas y estadísticas es posible reconocer los patrones de autocorrelación que caracterizan a los diferentes procesos lineales

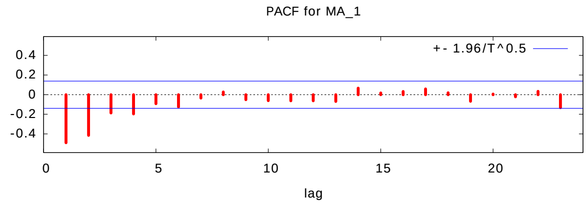
En el análisis de series temporales, este proceso de especificación empírica es conocido como "identificación"

Identificación y diagnosis (II)

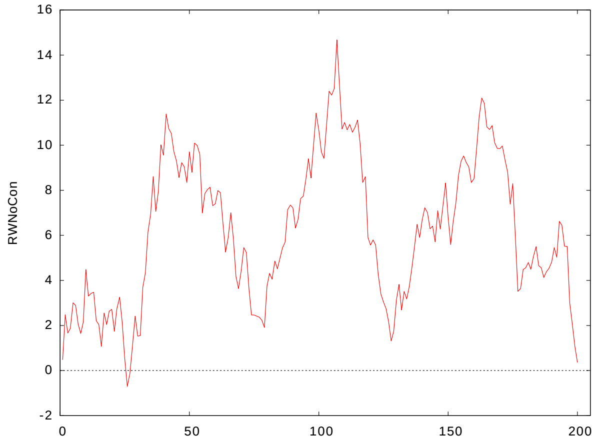
MA(1): $z_t = a_t - .8a_{t-1}$; $a_t \sim \text{iid N}(0,1)$

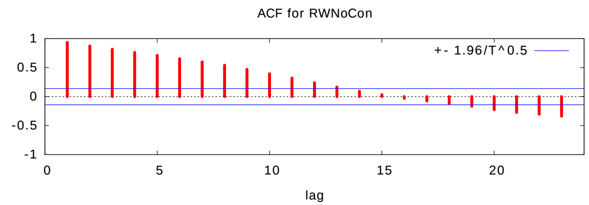


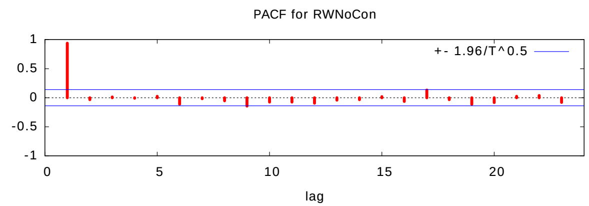




Paseo aleatorio: $y_t = y_{t-1} + a_t$; $a_t \sim \text{iid N}(0,1)$







El proceso de identificación puede estructurarse como una secuencia de preguntas:

- ¿Es estacionaria la serie?
- ¿Tiene una media significativa?
- ¿Es persistente la ACF?
- ¿Es persistente la PACF?

Identificación y diagnosis (III)

- La identificación se basa en estadísticos, como la media muestral o las autocorrelaciones, que sólo son válidos para series estacionarias
- Tras inducir estacionariedad mediante una transformación adecuada, se puede especificar un proceso tentativo decidiendo cuál de las dos funciones - ACF o PACF - es finita
- La siguiente tabla resume las combinaciones posibles:

		ACF	
		Finita	**Persistente**
PACF	**Finita**	**Ruido blanco** No hay ACFs ni PACFs significativas	**AR** El número de PACFs significativas indica el orden
	Persistente	**MA** El número de ACFs significativas indica el orden	**ARMA** En series económicas la parametrización de mayor orden suele ser ARMA(2,1)

Identificación y diagnosis (IV): Instrumentos de identificación

Parámetro	**Instrumento de identificación**	**Observaciones**
Transformación log	• Gráfico media-desviación típica • Gráfico de la serie temporal	Se trata de conseguir que la variabilidad de los datos sea independiente de su nivel. En series económicas las series suelen necesitar la transformación logarítmica
d, orden de diferenciación	• Gráfico de la serie temporal • ACF (decrecimiento lento y lineal)	Se trata de conseguir que los datos fluctúen en torno a una media aproximadamente estable
Término constante	• Media muestral de la serie diferenciada • Desviación típica de la media	Si la media de la serie transformada es significativa, el modelo debe incluir un término constante
p, orden del término AR	• PACF de orden p • ACF infinita	La PACF tiene p valores no nulos Un proceso AR finito y estacionario equivale a un MA(∞)
q, orden del término MA	• ACF de orden q • PACF infinita	La ACF tiene q valores no nulos Un proceso MA finito e invertible equivale a un AR(∞)

Identificación y diagnosis (V): Instrumentos de identificación

Parámetro	**Instrumento de identificación**	**Observaciones**
d, orden de diferenciación	• Raíces de los polinomios AR y MA	• Una raíz próxima a uno en la parte AR indica que conviene añadir una diferencia • Una raíz próxima a uno en la parte MA indica que conviene quitar una diferencia
	• Gráfico de la serie de residuos	Si muestra rachas largas de residuos positivos o negativos, puede ser necesaria una diferencia adicional
Término constante	• Media muestral de los residuos • Desviación típica de la media	Si la media de los residuos es significativa, debe añadirse un término constante
p y *q*	• Contrastes de significación de los parámetros estimados	Permiten eliminar parámetros irrelevantes
	• ACF y PACF residuales	Detectan pautas de autocorrelación no modelizadas
	• Test *Q* de Ljung-Box	Contrasta la hipótesis conjunta de que todos los coeficientes de autocorrelación son nulos
	• Correlaciones elevadas entre parámetros estimados	Puede ser un síntoma de sobreparametrización
	• Sobreajuste	Consiste en añadir parámetros AR y/o MA, para comprobar si resultan significativos y mejoran la calidad estadística del modelo

Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Notación de retardos (I): Operadores retardo y diferencia

A menudo resulta práctico representar los procesos estocásticos utilizando el operador retardo que se define de la siguiente manera:

$$B \;/\; i)\; Bz_t = z_{t-1}$$

$$ii)\; Bk = k\;;\; k\text{:constante}$$

$$iii)\; B^{-1} z_t = z_{t+1}$$

$$iv)\; B^{0} z_t = z_t$$

$$v)\; z_t = Bz_{t+1} = B^2 z_{t+2} = B^3 z_{t+3} = \ldots = B^l z_{t+l}\;\; (l > 0)$$

…en donde la letra "*B*" se refiere a la palabra inglesa "**B**ackward." En la literatura, a veces se denota este operador con la letra "*L*" (**L**ag)

El operador diferencia se define a partir del operador retardo como:

$$\nabla \;/\; \nabla y_t = (1 - B) y_t = y_t - y_{t-1}$$

…y en series estacionales de período *S*, a menudo se utiliza una variante de este operador que se conoce como diferencia estacional:

$$\nabla_S \;/\; \nabla_S y_t = (1 - B^S) y_t = y_t - y_{t-s}$$

Notación de retardos (II): Modelos ARMA(*p*,*q*) y ARIMA(*p*,*d*,*q*)

Los procesos anteriores pueden escribirse con órdenes generales:

AR(*p*): $z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t$

MA(*q*): $z_t = \mu_z + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}$

ARMA(*p*,*q*): $z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}$

... y expresarse en términos del operador retardo de la siguiente forma:

AR(*p*): $\phi_p(B) z_t = c + a_t$

MA(*q*): $z_t = \mu_z + \theta_q(B) a_t$

ARMA(*p*,*q*): $\phi_p(B) z_t = c + \theta_q(B) a_t$

... en donde los polinomios característicos de las componentes AR y MA del modelo son: $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ y $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$

Asimismo, si extendemos la representación para admitir raíces unitarias, que representan las diferencias que la serie necesita para ser estacionaria en media, obtenemos el proceso ARIMA (*p*,*d*,*q*):

$$\phi_p(B) \nabla^{\mathrm{d}} y_t = c + \theta_q(B) a_t$$

Notación de retardos (III): Raíces unitarias

El estudio de las raíces de los polinomios característicos AR y MA sirve para caracterizar la dinámica del proceso y puede ser útil para simplificarlo

Un caso especialmente importante es el de las raíces unitarias

- Cuando el polinomio AR tiene alguna raíz igual a uno, se dice que tiene "raíces unitarias". Si el polinomio corresponde a un modelo estimado, esto es síntoma de subdiferenciación
- Si la raíz unitaria está en el polinomio MA y este ha sido estimado, esto es síntoma de: (a) sobrediferenciación o (b) presencia de componentes deterministas en el proceso

La tabla muestra varios ejemplos de este tipo de simplificación

El proceso…	**…equivale a:**
$(1-1.5B+.5B^2)y_t=a_t$	$(1-.5B)\nabla y_t=a_t$
$(1-.5B+.7B^2)\nabla^2 y_t=(1-B)a_t$	$(1-.5B+.7B^2)\nabla y_t=a_t$
$y_t=\beta t+a_t$	$\nabla y_t=\beta+(1-B)a_t$

Notación de retardos (IV): Estacionariedad e invertibilidad

A partir del estudio de las raíces de los polinomios AR y MA, se definen dos conceptos:

Estacionariedad: Un proceso estocástico es estacionario si todas las raíces de la ecuación característica

$$1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p=0$$

…están fuera del círculo de radio unidad del plano complejo

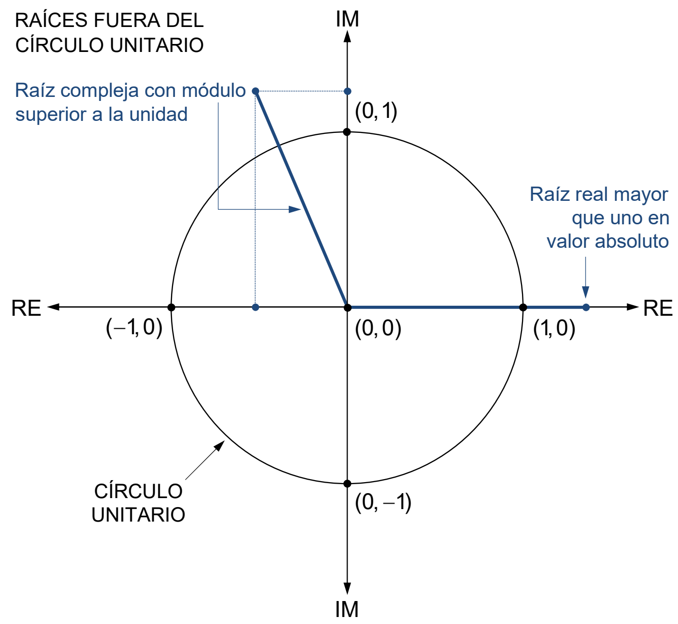
Si una raíz AR está dentro del círculo, se dice que el proceso es explosivo

Invertibilidad: Se dice que un proceso estocástico es invertible si todas las raíces de la ecuación característica

$$1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q=0$$

…están fuera del círculo de radio unidad del plano complejo

La figura muestra el círculo de radio unidad y distintos tipos de raíces



Ejemplos:

Las raíces de $1+.2B+.5B^2=0$ son -.20 ± 1.40*i* y su módulo (1.41) indica que se trata de raíces estacionarias (o invertibles)

Las raíces de $1+.5B+B^2=0$ son -.25 ± .968*i* y su módulo (1.00) indica que se trata de raíces en el círculo unitario

El polinomio $1-2B=0$ tiene una raíz explosiva (*B*=.5)

Notación de retardos (V): Modelos ARIMA en forma fraccional

Cuando un proceso estocástico no tiene raíces AR ni MA dentro del círculo de radio unidad, puede escribirse de como AR(∞) o MA(∞). Concretamente:

1) Un AR(1) puede escribirse como: $(1-\phi B)z_t = c + a_t$ y, si no es explosivo, como un proceso media móvil infinito:

$$z_t = \frac{c}{1-\phi} + \frac{1}{1-\phi B} a_t = \mu_z + (1+\phi B + \phi^2 B^2 + \ldots) a_t$$

2) Un MA(1) con raíces en o fuera del círculo de radio unidad puede escribirse en forma estándar: $z_t = \mu_z + (1-\theta B) a_t$ o, alternativamente, como:

$$\frac{1}{1-\theta B} z_t = \frac{\mu_z}{1-\theta} + a_t \ ; \ (1+\theta B + \theta^2 B^2 + \ldots) z_t = c + a_t$$

3) Si los polinomios AR y MA tienen sus raíces en o fuera del círculo de radio unidad, el proceso ARIMA(p,q):

$$\phi_p(B)\nabla^{\mathrm{d}} y_t = c + \theta_q(B) a_t$$

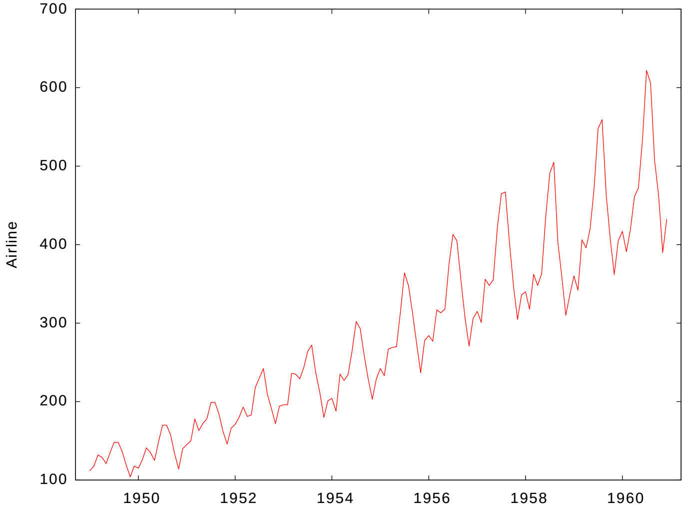
…puede escribirse de las siguientes formas:

$$\nabla^{\mathrm{d}} y_t = \mu_z + \frac{\theta_q(B)}{\phi_p(B)} a_t \quad \frac{\phi_p(B)}{\theta_q(B)}(\nabla^{\mathrm{d}} y_t - \mu_z) = a_t \quad \frac{\phi_p(B)}{\theta_q(B)} \nabla^{\mathrm{d}} y_t = \bar{c} + a_t$$

Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Series estacionales (I): Modelos ARIMA(p,d,q)x(P,D,Q)$_S$



Las series económicas a menudo muestran un estacionalidad, esto es, una pauta que se repite con una periodicidad fija, a menudo anual

El período estacional (S) se define como el número de observaciones necesarias para recorrer un ciclo estacional completo. Por ejemplo, S=12 para datos mensuales, S=4 para datos trimestrales…

Para captar este comportamiento, se define el modelo ARIMA(p,d,q)x(P,D,Q)$_S$:

$$\phi_p(B)\Phi_P(B^S)\nabla_S^D\nabla^d y_t = c + \theta_q(B)\Theta_Q(B^S)a_t$$

... que incluye tres nuevos factores diseñados para relacionar cada dato con el de S períodos atrás:

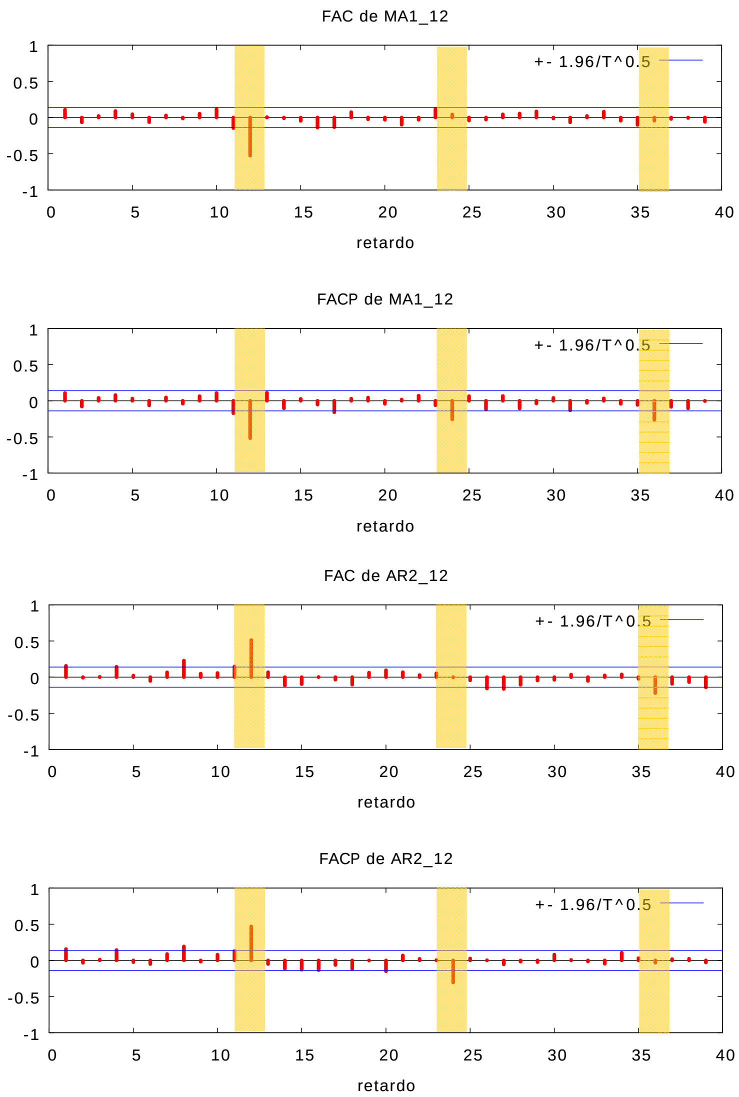
$$\Phi_P(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2\cdot S} - \ldots - \Phi_P B^{P\cdot S} \quad \text{[polinomio AR}(P)_S\text{]}$$

$$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2\cdot S} - \ldots - \Theta_Q B^{Q\cdot S} \quad \text{[polinomio MA}(Q)_S\text{]}$$

$$\nabla_S \equiv 1 - B^S \quad \text{[operador diferencia estacional]}$$

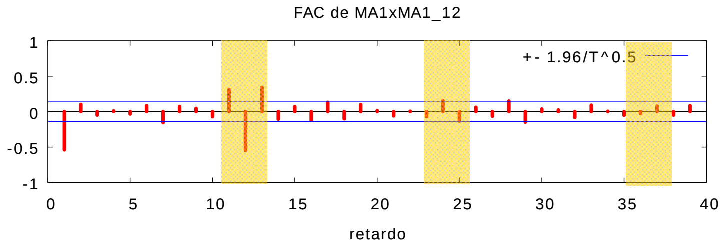
En los siguientes *slides* se muestra cómo especificar estos factores

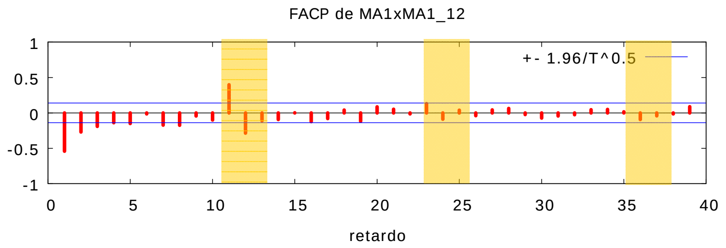
Series estacionales (II): Correlogramas de una serie estacional

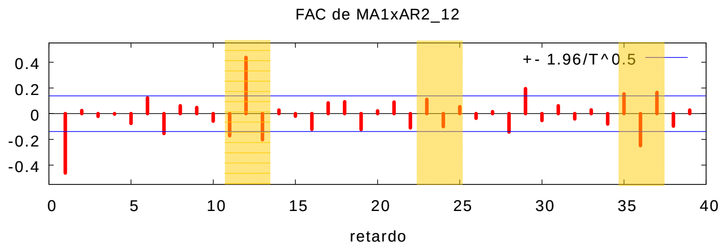


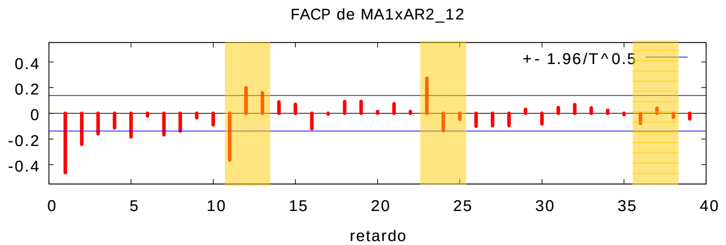
- La primera figura muestra los correlogramas de una serie simulada a partir de un proceso media móvil estacional (S=12) sin estructura regular
- La segunda figura muestra los resultados análogos para un proceso AR(2) estacional (S=12), nuevamente sin estructura regular
- Como puede observarse, las pautas de autocorrelación son las que caracterizan a las series generadas por procesos MA(1) y AR(2), pero con las autocorrelaciones situadas en los retardos múltiplos del período estacional
- Las correlaciones correspondientes a los "retardos regulares" (todos menos el 12, 24 y 36) son no significativas en general

Series estacionales (III): Satélites









- Las figuras muestran los correlogramas de un proceso media móvil estacional y un proceso AR(2) estacional, con S=12 en ambos casos
- En esta ocasión hemos especificado el proceso generador con una estructura MA(1) regular
- Como puede observarse, en el entorno de los retardos estacionales surgen una serie de coeficientes significativos ("satélites") que proceden de la interacción entre las estructuras regular y estacional
- Estos satélites son útiles para identificar en qué retardos estacionales hay autocorrelaciones no nulas, pero no requieren una parametrización especial

Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Series estacionales
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

Ideas principales (I)

- El análisis univariante de series temporales trata de construir un modelo sencillo, sin variables exógenas, que aproveche la inercia de los datos pasados para predecir los valores futuros de la serie
- Para ello, se parte de tres supuestos: (a) linealidad, (b) estacionariedad (débil) y (c) normalidad
- …y se utilizan diversos instrumentos, incluyendo: (a) gráfico de la serie temporal, (b) función de autocorrelación simple, (c) función de autocorrelación parcial, (d) estadístico de Ljung-Box…
- Si la serie original no es débilmente estacionaria, es necesario inducir esta propiedad mediante una transformación de datos adecuada
- Los procesos lineales tienen pautas de autocorrelación teórica características y reconocibles:
  - En un proceso de ruido blanco, las funciones de autocorrelación y autocorrelación parcial son nulas
  - En un proceso AR (autorregresivo) la función de autocorrelación es infinita y la de autocorrelación parcial es finita (el número de coeficientes significativos indica el orden)
  - En un proceso MA (de media móvil) la función de autocorrelación parcial es finita (el número de coeficientes significativos indica el orden) y la de autocorrelación parcial es infinita
  - En un proceso ARMA mixto ambas funciones son infinitas

Ideas principales (II): Metodología

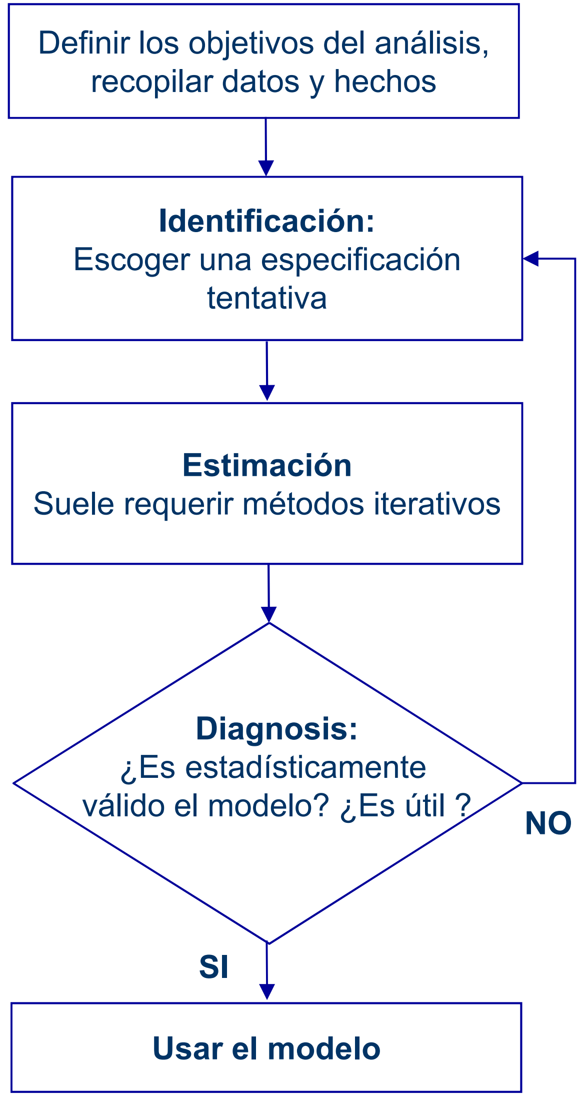
El análisis de series temporales utiliza diversas herramientas estadísticas para formar una metodología de modelización que se estructura en tres fases:

Identificación: Elija una especificación provisional para el DGP en base a las características de datos medibles: "dejar que los datos hablen"

Estimación: Suele requerir métodos iterativos no lineales

Diagnosis: Controla la calidad estadística del modelo ajustado. Algunos controles estándar son:

- Significatividad de los parámetros estimados
- Estacionariedad y homocedasticidad de los residuos
- ¿Existe un patrón de autocorrelación residual que podría ser modelado?



Índice

- Introducción
- Procesos estocásticos elementales
- Instrumentos de identificación
- Identificación y diagnosis
- Notación de retardos
- Ideas principales
- Apéndice. Estudio de los procesos más comunes

A.1. Estudio de los procesos más comunes: AR(1)

$$z_t = c + \phi z_{t-1} + a_t \;\; ; \;\; |\phi| < 1$$

$$E(z_t) = \mu_z = \frac{c}{1-\phi}$$

$$E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = \frac{\sigma_a^2}{1-\phi^2}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k \;\; ; \;\; k \geq 0$$

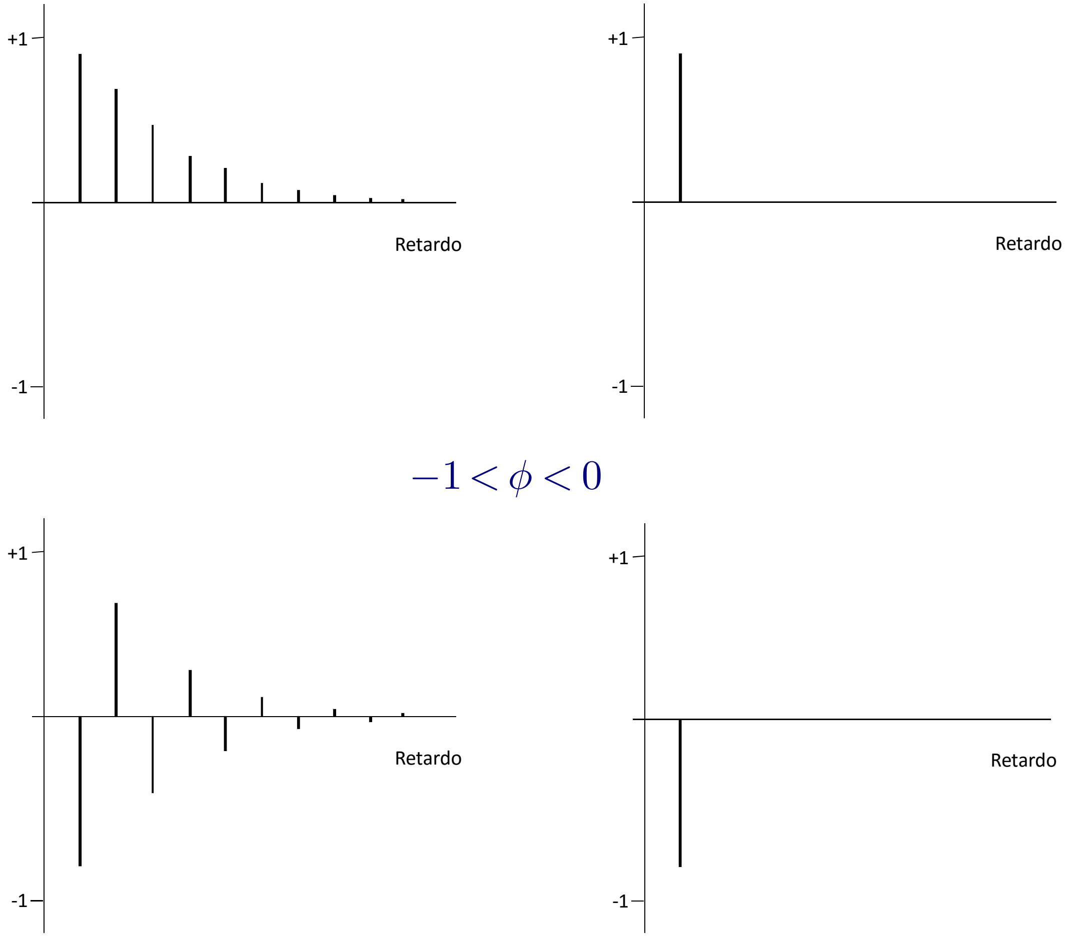
Los procesos AR(1) se reconocen por una ACF infinita y una PACF que se anula a partir del segundo retardo

Si los datos tienen media, es necesario especificar un término constante

ACF PACF

$0 < \phi < 1$

$-1 < \phi < 0$



A.2. Estudio de los procesos más comunes: MA(1)

$$z_t = \mu_z + a_t - \theta a_{t-1} \; ; \; |\theta| < 1$$

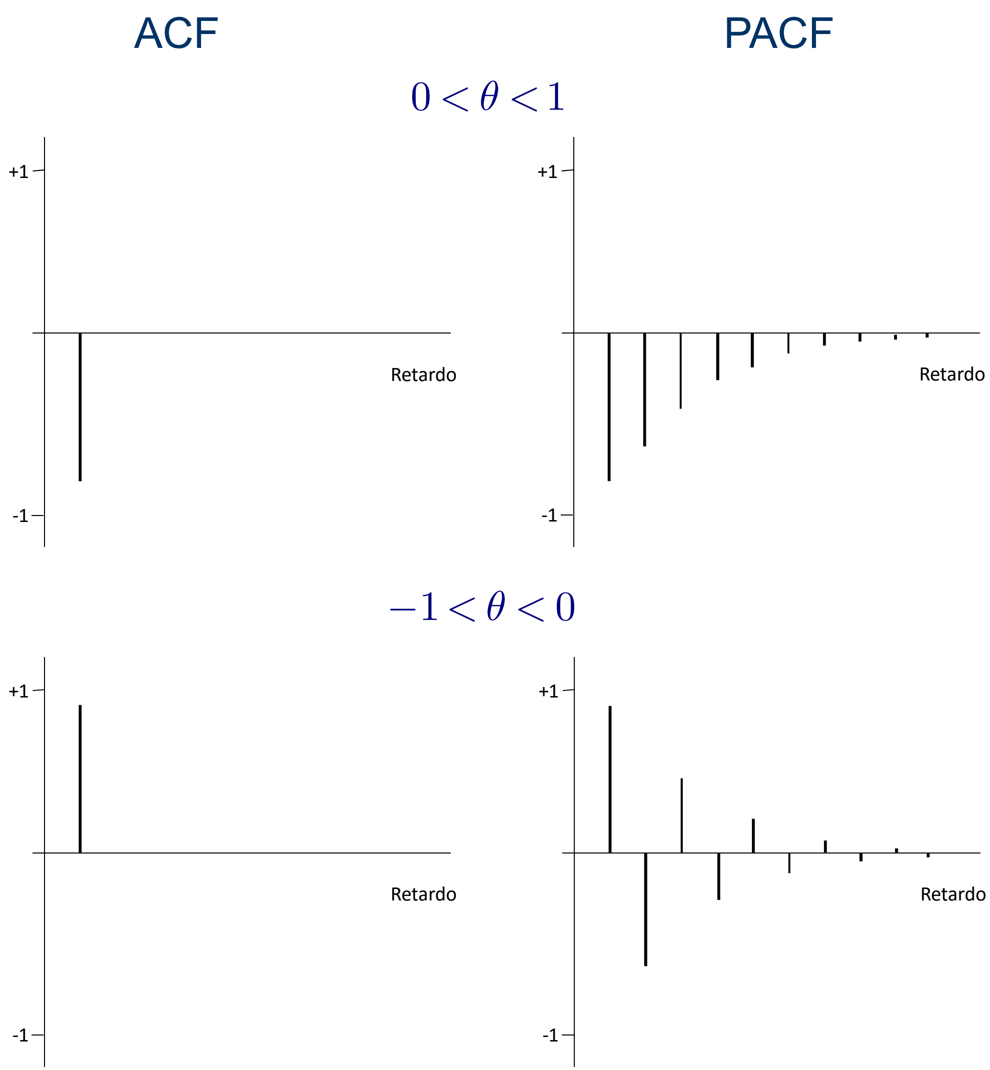
$$E(z_t) = \mu_z$$

$$E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = (1+\theta^2)\sigma_a^2$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \dfrac{-\theta}{1+\theta^2} & si \; k = 1 \\ 0 & si \; k > 1 \end{cases}$$

Los procesos MA(1) se reconocen por una PACF infinita y una ACF que se anula a partir del segundo retardo

Si los datos tienen media, es necesario especificar un término constante



A.3. Estudio de los procesos más comunes: AR(2)

$$z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

$$\phi_2 + \phi_1 < 1\,;\ \phi_2 - \phi_1 < 1\,;\ |\phi_2| < 1$$

$$E(z_t) = \mu_z = \frac{c}{1-\phi_1-\phi_2}$$

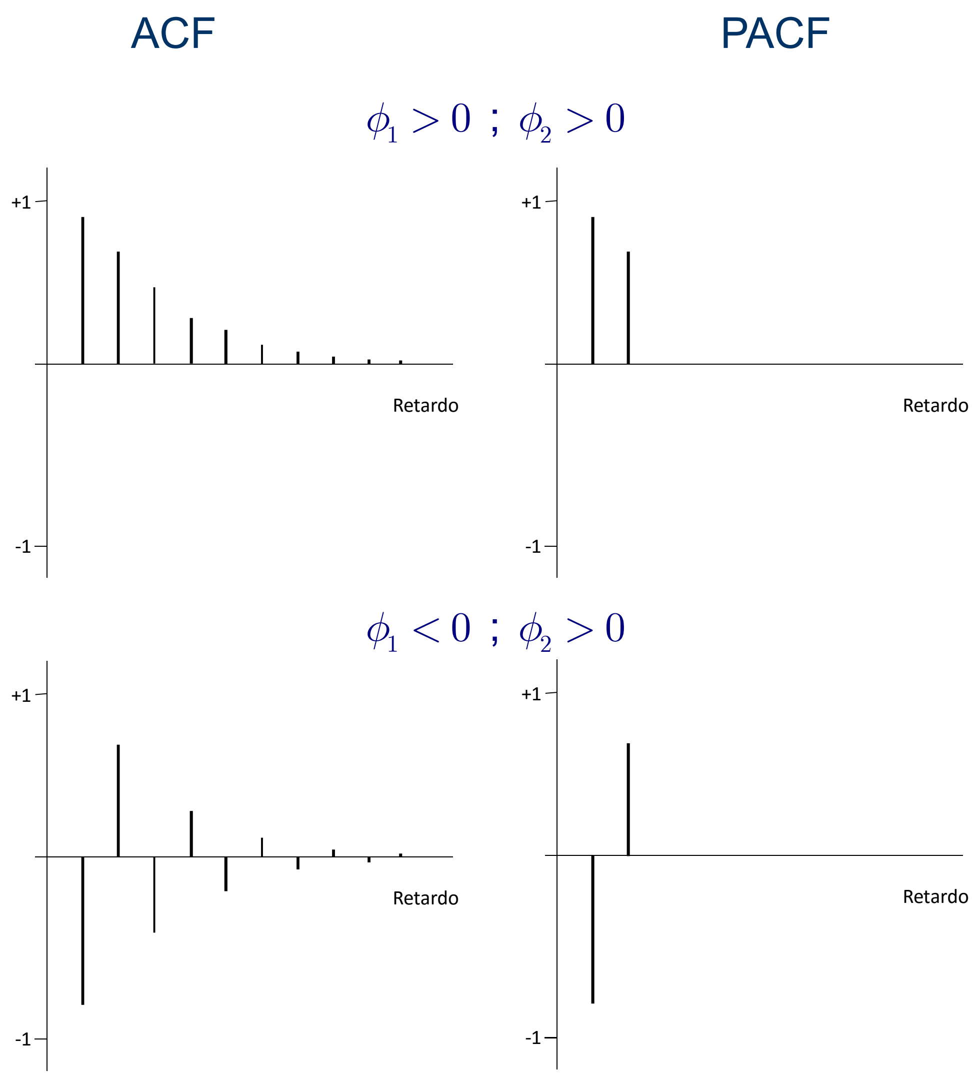
$$\rho_k = \phi_1\,\rho_{k-1} + \phi_2\,\rho_{k-2};\ k \geq 1$$

con:

$$\rho_1 = \frac{\phi_1}{1-\phi_2}\ \ ;\ \rho_2 = \phi_2 + \frac{\phi_1^2}{1-\phi_2}$$

Los procesos AR(2) se reconocen por una ACF infinita y una PACF que se anula a partir del tercer retardo

Si los datos tienen media, es necesario especificar un término constante



A.3. Estudio de los procesos más comunes: AR(2)

$$z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

$$\phi_2 + \phi_1 < 1\,;\ \phi_2 - \phi_1 < 1\,;\ |\phi_2| < 1$$

$$E(z_t) = \mu_z = \frac{c}{1 - \phi_1 - \phi_2}$$

$$\rho_k = \phi_1\, \rho_{k-1} + \phi_2\, \rho_{k-2};\ k \geq 1$$

con :

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}\ \ ;\ \rho_2 = \phi_2 + \frac{\phi_1^2}{1 - \phi_2}$$

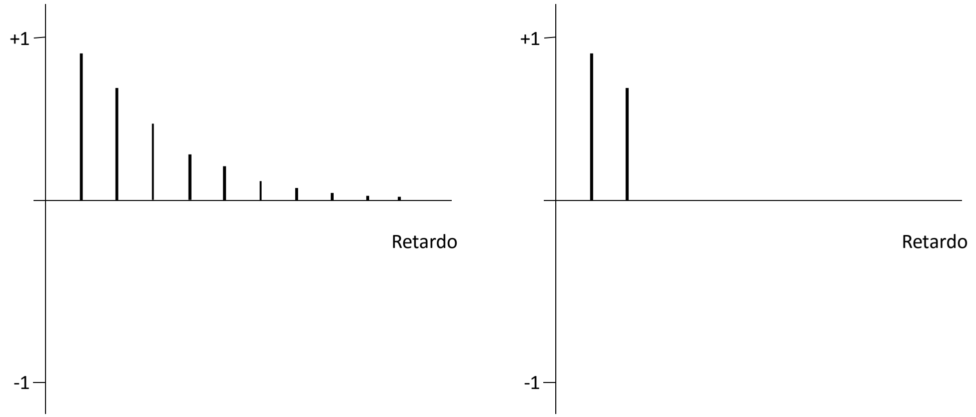
Los procesos AR(2) se reconocen por una ACF infinita y una PACF que se anula a partir del tercer retardo

Si los datos tienen media, es necesario especificar un término constante

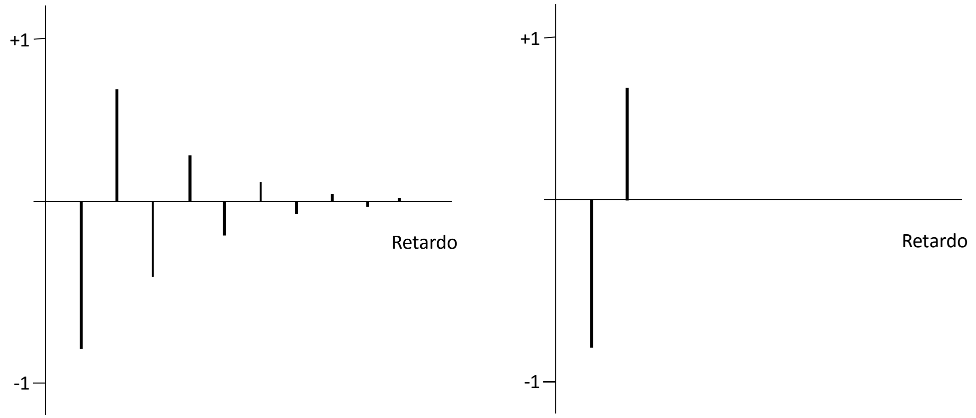
ACF

PACF

$\phi_1 > 0\ ;\ \phi_2 > 0$



$\phi_1 < 0\ ;\ \phi_2 > 0$



A.5. Estudio de los procesos más comunes: MA(2)

$$z_t = \mu_z + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

$$\theta_2 + \theta_1 < 1\,;\, \theta_2 - \theta_1 < 1\,;\, |\theta_2| < 1$$

$$E(z_t) = \mu_z$$

$$E(\tilde{z}_t^2) = \sigma_z^2 = \gamma_0 = (1 + \theta_1^{\,2} + \theta_2^{\,2})\sigma_a^2$$

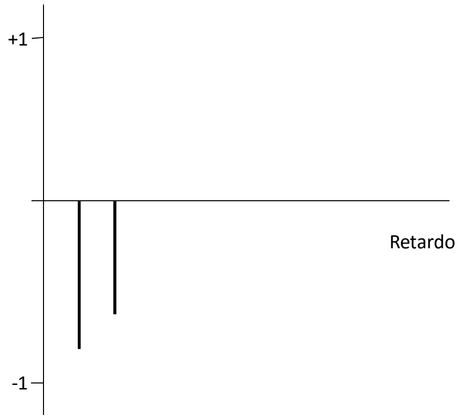
$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \dfrac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & si\ k = 1 \\[2ex] \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2} & si\ k = 2 \\[2ex] 0 & si\ k > 2 \end{cases}$$

Los procesos MA(2) se reconocen por una PACF infinita (no se muestra aquí) y una ACF que se anula a partir del tercer retardo

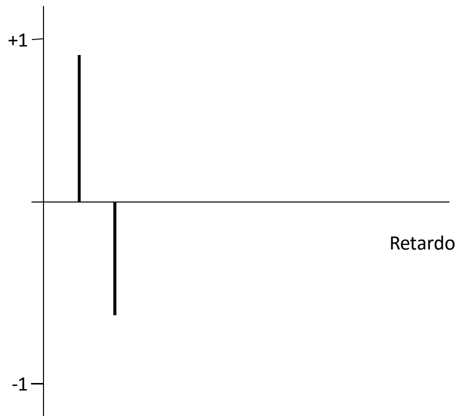
Si los datos tienen media, es necesario especificar un término constante

ACF

$\theta_1 > 0\ ;\ \theta_2 > 0$

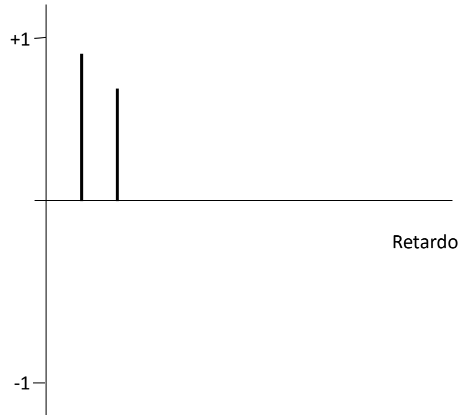


$\theta_1 < 0\ ;\ \theta_2 > 0$

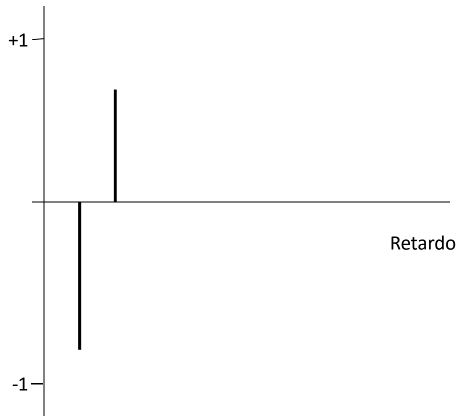


ACF

$\theta_1 < 0\ ;\ \theta_2 < 0$



$\theta_1 > 0\ ;\ \theta_2 < 0$



Miguel Jerez (mjerez@ccee.ucm.es)

Departamento de Fundamentos del Análisis Económico II
(Economía Cuantitativa)

Facultad de Ciencias Económicas, UCM