

# Índice

<b>1. Series temporales vs datos de sección cruzada</b>	<b>2</b>
1.1. Correlación serial vs muestreo aleatorio simple . . . . .	5
1.2. Simplificación del escenario . . . . .	6
<b>2. Procesos estocásticos de segundo orden</b>	<b>6</b>
2.1. Un poco de geometría . . . . .	6
2.2. Primeros momentos de procesos estocásticos de segundo orden . . . . .	7
2.3. Procesos estocásticos (débilmente) estacionarios y la ACF . . . . .	7
<b>3. Notación: convolución y el operador retardo</b>	<b>8</b>
<b>4. Ejemplos de procesos (débilmente) estacionarios</b>	<b>9</b>
4.1. Proceso de ruido blanco . . . . .	9
4.2. Procesos lineales . . . . .	9
4.2.1. Media móvil infinita. $MA(\infty)$ . . . . .	10
4.2.2. Proceso de media móvil de orden $q$ . $MA(q)$ . . . . .	10
4.2.3. Proceso autorregresivo de orden $p$ . $AR(p)$ . . . . .	10
4.2.4. Proceso autorregresivo de media móvil. $ARMA(p, q)$ . . . . .	11
4.2.5. Proceso autorregresivo de media móvil con media no nula . . . . .	12
<b>5. Primeros momentos de procesos lineales causales</b>	<b>12</b>
5.1. Esperanza y autocovarianzas de un proceso lineal causal . . . . .	12
5.2. Covarianza cruzada entre dos procesos lineales causales . . . . .	13
5.3. Las Ecuaciones de Yule-Walker para un $AR(p)$ estacionario . . . . .	13
5.4. Función de autocovarianzas para un $ARMA(p, q)$ . . . . .	14

# Econometría Aplicada. Lección 5

Marcos Bujosa

1 de octubre de 2024

## Resumen

Esta lección veremos las dificultades que ocasiona la correlación serial y algunos tipos de procesos débilmente estacionarios que nos permitirán lidiar con ella. En particular veremos los procesos lineales, su valor esperado y su función de autocovarianzas, la función de covarianzas cruzadas entre dos procesos lineales, y las ecuaciones de Yule-Walker.

- [lección en html](#)
- [lección en mybinder](#)

## Carga de algunas librerías de R

Primero cargamos la librería `tfarima` (Repositorio Cran: <https://cran.r-project.org/web/packages/tfarima/index.html>; repositorio GitHub: <https://github.com/gallegoj/tfarima>)

---

```
library(tfarima) # librería de José Luis Gallego para Time Series
library(readr)  # para leer ficheros CSV
library(ggplot2) # para el scatterplot (alternativamente library(tidyverse))
library(ggfortify) # para pintar series temporales
library(jttools) # para representación resultados estimación
library(zoo)     # para generar objetos ts (time series)
```

---

y además fijamos los parámetros por defecto para las figuras en `png` del notebook

---

```
# fijamos el tamaño de las figuras que se generan en el notebook
options(repr.plot.width = 12, repr.plot.height = 4, repr.plot.res = 200)
```

---

## 1. Series temporales vs datos de sección cruzada

Corresponden a observaciones de un mismo objeto a lo largo del tiempo. El índice indica el instante de cada medición. *El orden cronológico puede ser crucial* al modelar los datos.

- El motivo es que frecuentemente el valor medido en un instante de tiempo está relacionado con otras mediciones próximas en el tiempo (*correlación serial*).
- Si es así, ya no deberíamos asumir que las variables aleatorias del proceso estocástico subyacente,  $\mathbf{X} = (X_t \mid t \in \mathbb{Z})$ , son independientes entre sí.

Esto tiene importantes implicaciones en las técnicas de análisis y los modelos a utilizar.

Veamos algunos ejemplos de series temporales...

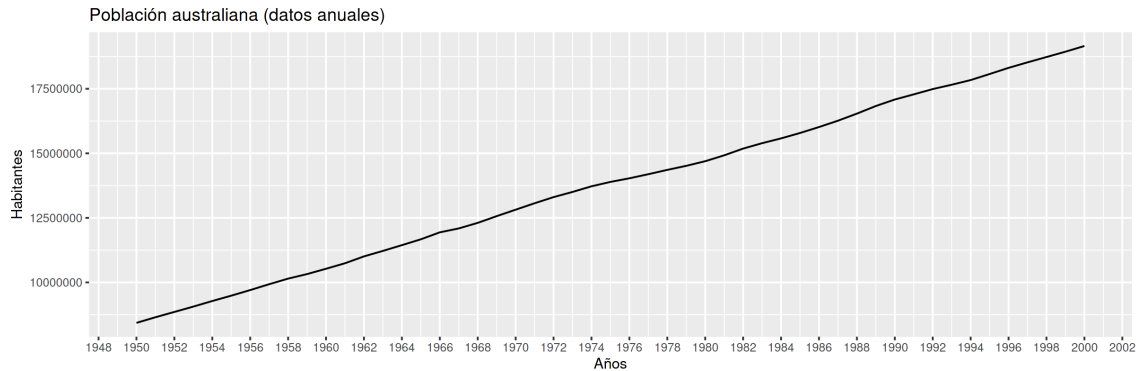
## 1. Población en Australia

---

```
PoblacionAustralia_ts = as.ts( read.zoo('datos/PoblacionAustralia.csv',
                                     header=TRUE,
                                     index.column = 1,
                                     sep=",",
                                     FUN = as.yearmon))

p <- autoplot(PoblacionAustralia_ts)
p <- p + labs(y = "Habitantes", x = "Años") + ggtitle("Población australiana (datos anuales)")
p <- p + scale_x_continuous(breaks = scales::pretty_breaks(n = 20))
p
```

---



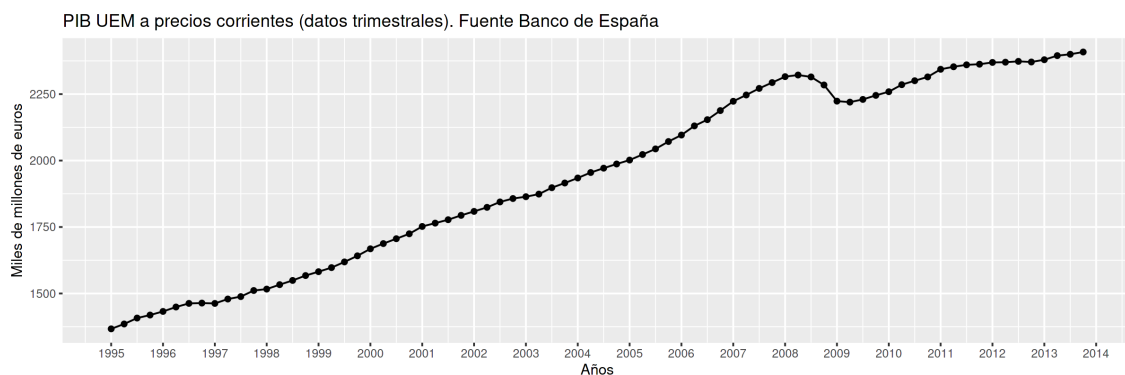
## 2. PIB UEM

---

```
PIB_UEM_df <- read_csv("datos/PIB_UEM.csv",
                       show_col_types = FALSE)

fmt <- "%YQ%q"
PIB_UEM_df$Time <- as.yearqtr(PIB_UEM_df$obs, format = fmt)
# head(PIB_UEM_df, 3)
P <- ggplot(PIB_UEM_df, aes(Time, PIB))
P <- P + geom_point() + geom_line()
P <- P + scale_x_continuous(breaks = scales::pretty_breaks(n = 15))
P <- P + labs(y = "Miles de millones de euros", x = "Años")
P <- P + ggtitle("PIB UEM a precios corrientes (datos trimestrales). Fuente Banco de España")
P
```

---



## 3. Temperatura media en el Parque del Retiro. Madrid

---

```
TemperaturaRetiro_df <- read_csv("datos/Retiro.txt", show_col_types = FALSE)
# Añadimos fechas
```

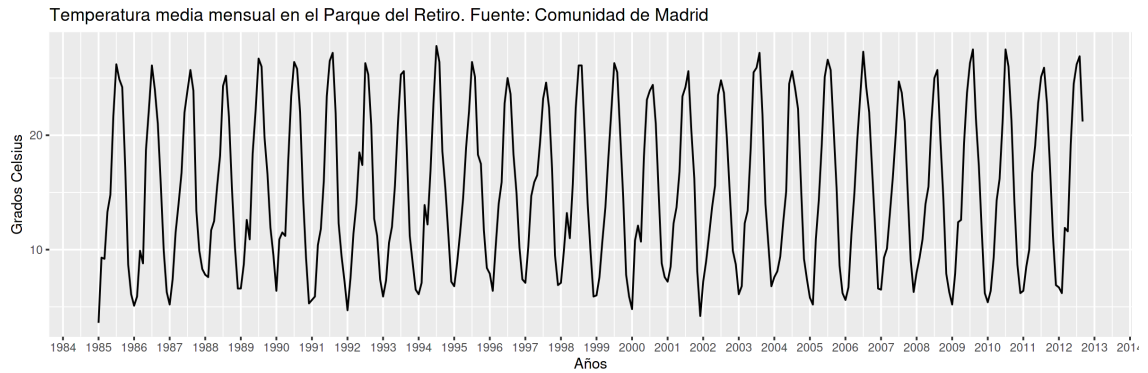
---

```

TemperaturaRetiro_df$Time <- as.yearmon(1985 + seq(0, nrow(TemperaturaRetiro_df)-1)/12)

P <- ggplot(TemperaturaRetiro_df, aes(Time, TemperaturaMedia))
P <- P + geom_line() # + geom_point()
P <- P + scale_x_continuous(breaks = scales::pretty_breaks(n = 25))
P <- P + labs(y = "Grados Celsius", x = "Años")
P <- P + ggtitle("Temperatura media mensual en el Parque del Retiro. Fuente: Comunidad de Madrid")
P

```

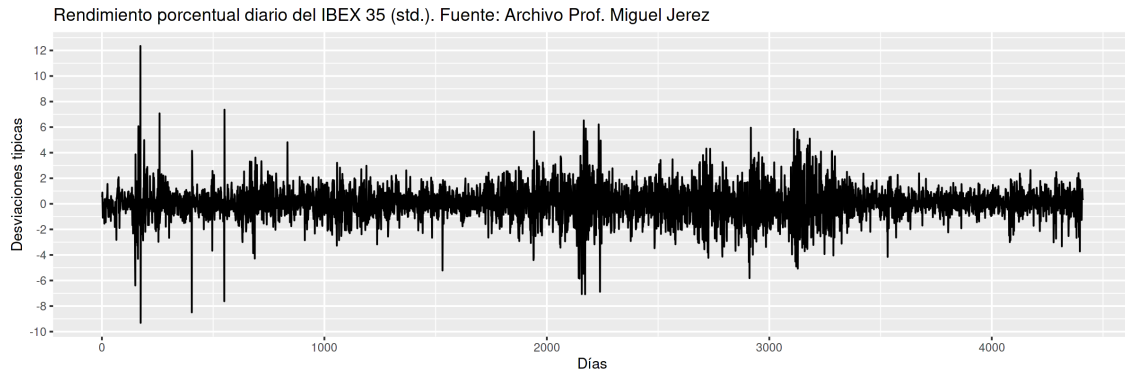


#### 4. Rendimiento porcentual diario del IBEX 35 (std)

```

IBEX35_ts = as.ts( read.csv.zoo("datos/IBEX35.csv",
                                strip.white = TRUE))
P <- autoplot(IBEX35_ts) + scale_y_continuous(breaks = scales::pretty_breaks(n = 12))
p <- P + labs(y = "Desviaciones típicas", x = "Días")
p <- P + ggtitle("Rendimiento porcentual diario del IBEX 35 (std.). Fuente: Archivo Prof. Miguel Jerez")
p

```



- Datos centrados y estandarizados, i.e. el eje vertical está en desviaciones típicas.
- Los *volatility clustering* son característicos de series financieras de alta frecuencia.

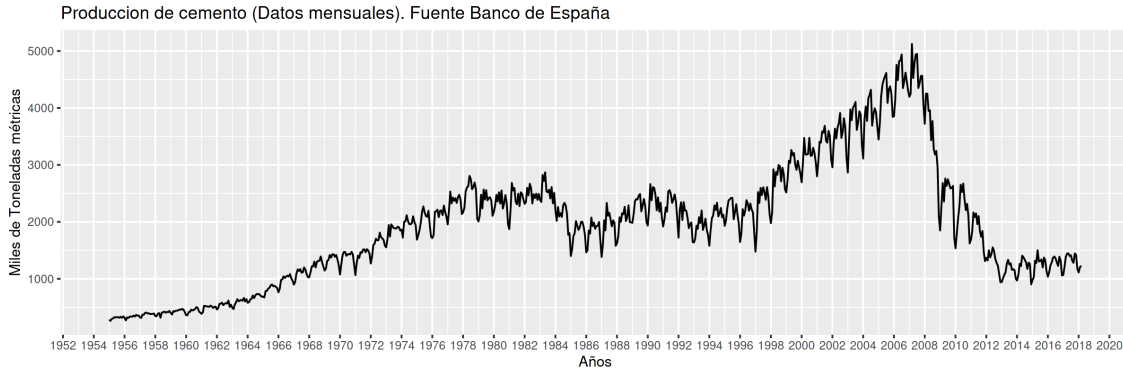
#### 5. Producción de cemento

```

ProduccionCemento_df <- read_csv("datos/ProduccionCemento.csv",
                                show_col_types = FALSE)
fmt <- "%Y%m"
ProduccionCemento_df$Time <- as.yearmon(ProduccionCemento_df$obs, format = fmt)
# head(ProduccionCemento_df, 3)
P <- ggplot(ProduccionCemento_df, aes(Time, ProduccionCemento))

```

```
P <- P + geom_line() # + geom_point()
P <- P + scale_x_continuous(breaks = scales::pretty_breaks(n = 25))
P <- P + labs(y = "Miles de Toneladas métricas", x = "Años")
P <- P + ggtitle("Producción de cemento (Datos mensuales). Fuente Banco de España")
P
```



### 1.1. Correlación serial vs muestreo aleatorio simple

Con datos de

**sección cruzada** solemos asumir que el muestreo es aleatorio simple

- i.e., los datos son realizaciones de variables aleatorias i.i.d.

**series temporales** dicha asunción resulta generalmente errónea

- con frecuencia el nivel esperado (o la volatilidad) parece cambiar con  $t$
- con frecuencia hay dependencia temporal (correlación serial).

**Ejemplo:** no parece aceptable asumir que  $ProdCemento_{1960M01}$  se distribuye igual que  $ProdCemento_{2000M04}$  (ni que sea independiente de  $ProdCemento_{1959M01}$ ).

Veamos por qué esto genera dificultades...

Consideremos el proceso estocástico

$$\mathbf{X} = (X_t \mid t = 0, \pm 1, \pm 2, \dots).$$

Caracterizar su distribución conjunta (todos los momentos) es demasiado ambicioso.

Así que, tentativamente, vamos a fijarnos *solo* en los dos primeros momentos:

$$E(X_t) = \mu_t \quad \text{y} \quad Cov(X_t, X_k) = E[(X_t - \mu_t)(X_k - \mu_k)] = \gamma_{t,k}; \quad t, k \in \mathbb{Z}$$

(si  $k = t$  entonces  $\gamma_{t,t} = Var(X_t) = \sigma_t^2$ ).

Si el proceso  $\mathbf{X}$  fuera gaussiano, conocer estos *parámetros* bastaría para caracterizar la distribución conjunta. Pero aún así...

- necesitaríamos para cada  $X_t$  una muestra suficiente para estimar los parámetros
  - pero en una serie temporal tenemos una sola realización de cada  $X_t$ .
- Además... para cada variable aleatoria  $X_t$  hay infinitos parámetros.

## 1.2. Simplificación del escenario

Si  $\mathbf{X}$  es **débilmente estacionario** se reduce drásticamente el número de parámetros:

$$E(X_t) = \mu \quad (1)$$

$$Cov(X_t, X_{t-k}) = \gamma_k \quad (2)$$

El desafío para el analista es (y nótese el abuso de lenguaje)

**primero** transformar los datos para lograr que sean "**estacionarios**".

- (Algo vimos en la lección 1))

**después** transformar los datos estacionarios en "**ruido blanco**".

- (Es lo que iniciaremos en esta lección y las siguientes)

Todo este proceso constituye la especificación y ajuste de un modelo a la serie temporal.

Antes de atacar los temas de especificación y ajuste de modelos, debemos estudiar un poco los procesos estocásticos débilmente estacionarios que vamos a utilizar.

## 2. Procesos estocásticos de segundo orden

El ambiente natural para estudiar las propiedades de segundo orden de una colección de variables aleatorias es el espacio de variables aleatorias  $X$  definidas en un espacio de probabilidad tales que

$$E(X) = 0 \quad \text{y} \quad E(X^2) < \infty$$

donde  $E$  es el operador esperanza. Denotaremos este espacio con  $H$ .

### 2.1. Un poco de geometría

El espacio, dotado de producto escalar y norma

$$\langle X | Y \rangle = E(XY), \quad \|X\| = \sqrt{E(X^2)}, \quad X, Y \in H,$$

es un espacio de Hilbert,

Nótese que como las variables de  $H$  tienen esperanza cero, el producto escalar entre  $X, Y \in H$  también es

$$\langle X | Y \rangle = Cov(X, Y).$$

Por tanto, en este espacio  $H$  la noción geométrica de ortogonalidad coincide con la noción estadística de *no correlación*. Por tanto, en este contexto los términos producto escalar, covarianza y esperanza del producto serán intercambiables.

Una colección de variables aleatorias pertenecientes a  $H$

$$\mathbf{X} = (X_t | t \in \mathbb{Z}) \quad \text{con} \quad X_t \in H$$

se denomina *proceso estocástico de segundo orden*.

Si  $\mathbf{Y} = (Y_t | t \in \mathbb{Z})$  es tal que  $E(Y_t) = \mu \neq 0$ , entonces  $\mathbf{Y}$  no es de segundo orden.

Pero basta restar  $\mu$  de cada  $Y_t$  para tener un proceso  $(\mathbf{Y} - \mu \mathbf{1})$  de segundo orden.

*Por ello siempre asumiremos (sin pérdida de generalidad) que las variables aleatorias de los procesos estocásticos de esta lección (y la siguiente) tienen esperanza cero.*

## 2.2. Primeros momentos de procesos estocásticos de segundo orden

Si  $E(X_t) < \infty$  para  $t \in \mathbb{Z}$ , entonces  $E(\mathbf{X})$  es la secuencia

$$E(\mathbf{X}) = (E(X_t) \mid t \in \mathbb{Z}) = \sum_{t \in \mathbb{Z}} E(X_t) z^t = (\dots, E(X_{-1}), E(X_0), E(X_1), \dots)$$

Si  $\mathbf{X}$  tiene segundos momentos finitos, la secuencia de autocovarianzas de orden  $k$  es

$$\begin{aligned} \left( Cov(X_t, X_{t-k}) \mid t \in \mathbb{Z} \right) &= (\gamma_{k,t} \mid t \in \mathbb{Z}) \\ &= (\dots, \gamma_{k,-1}, \gamma_{k,0}, \gamma_{k,1}, \gamma_{k,2}, \dots); \quad k \in \mathbb{Z}. \end{aligned}$$

(nótese que la secuencia solo contiene covarianzas de orden  $k$ )

Así, para cada par  $(k, t)$ , tenemos la covarianza  $\gamma_{k,t}$  entre  $X_t$  y  $X_{t-k}$ . Por tanto, en general, tenemos una esperanza para cada  $t$  y una covarianza de orden  $k$  para cada  $t$ . Dado que  $t$  recorre todos los números enteros, ¡esto son muchos momentos! Por eso necesitamos reducir el número de parámetros restringiéndonos a procesos estocásticos débilmente estacionarios.

## 2.3. Procesos estocásticos (débilmente) estacionarios y la ACF

Un proceso estocástico de segundo orden  $\mathbf{X}$  se dice que es *débilmente estacionario* (estacionario en covarianza o, sencillamente, estacionario) si  $E(X_t) = \mu$  para todo  $t \in \mathbb{Z}$  y la covarianza entre  $X_s$  y  $X_t$  solo depende de la diferencia  $s - t$  para todo  $s, t \in \mathbb{Z}$ .

En tal caso, definimos la función de autocovarianzas como:

$$\gamma = (\gamma_k \mid k \in \mathbb{Z}) = (\dots, \gamma_{-1}, \gamma_0, \gamma_1, \gamma_2, \dots) = \sum_{-\infty}^{\infty} \gamma_k z^k.$$

(nótese que esta secuencia sí incluye todas las covarianzas).

Y se denomina matriz de autocovarianzas de  $\mathbf{X}$  a la matriz simétrica

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Tanto la secuencia  $\gamma$  como la matriz  $\mathbf{\Gamma}$  son *\*definidas positivas\**; es decir, para todos los enteros  $n \geq 1$  y escalares  $c_1, c_2, \dots, c_n$

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j \gamma_{i-j} \geq 0$$

ya que

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j Cov(X_i, X_j) = Cov \left( \sum_{i=1}^n c_i X_i, \sum_{j=1}^n c_j X_j \right) = \left\langle \sum_{i=1}^n c_i X_i \mid \sum_{i=1}^n c_i X_i \right\rangle = \left\| \sum_{i=1}^n c_i X_i \right\|^2 \geq 0.$$

Esto es equivalente a que las submatrices principales de  $\mathbf{\Gamma}$  son definidas positivas.

Es más, una secuencia  $\gamma$  es definida positiva si y solo si existe un espacio de Hilbert  $H$  y un proceso estocástico estacionario  $\mathbf{X}$  con  $X_t \in H$  tales que  $\gamma_k = Cov(X_t, X_{t-k})$  para todo  $t, k \in \mathbb{Z}$  (Kolmogorov, 1941).

**Propiedades** de la función de autocovarianzas  $\gamma$  (ACF):

- $\gamma_0 \geq 0$
- $\gamma$  es definida positiva; y por tanto,
  - $\gamma$  es simétrica:  $\gamma_k = \gamma_{-k}$
  - $\gamma$  es acotada:  $|\gamma_k| \leq \gamma_0$

Y, si  $\gamma_0 > 0$ , llamamos *función de autocorrelación* (ACF) a la secuencia:  $\rho = \frac{1}{\gamma_0}(\gamma) = \sum_{k \in \mathbb{Z}} \frac{\gamma_k}{\gamma_0} z^k$ .

### 3. Notación: convolución y el operador retardo

Los procesos estocásticos se pueden sumar elemento a elemento y se pueden multiplicar por escalares. Si  $\mathbf{X}$  e  $\mathbf{Y}$  son dos procesos estocásticos y  $a \in \mathbb{R}$ , entonces

$$\mathbf{X} + \mathbf{Y} = (X_t + Y_t \mid t \in \mathbb{Z}) \quad \text{y} \quad a\mathbf{X} = (a(X_t) \mid t \in \mathbb{Z}).$$

El conjunto de procesos estocásticos con la suma y el producto por escalares es un espacio vectorial.

Sea  $\mathbf{a}$  una secuencia de números y sea  $\mathbf{X}$  un proceso estocástico tales que la suma

$$\sum_{k=-\infty}^{\infty} a_k X_{t-k}$$

converge para todo  $t$ . Entonces:

Definimos el producto convolución  $(*)$  de  $\mathbf{a}$  con  $\mathbf{X}$  como el proceso estocástico:

$$\mathbf{a} * \mathbf{X} = \left( \sum_{r+s=t} a_r X_s \mid t \in \mathbb{Z} \right)$$

es decir

$$(\mathbf{a} * \mathbf{X})_t = \sum_{r+s=t} a_r X_s, \quad \text{para } t \in \mathbb{Z}.$$

Por tanto, cada elemento de  $(\mathbf{a} * \mathbf{X})$  es una combinación de variables aleatorias de  $\mathbf{X}$

Podemos aplicar el operador B sobre los elementos de un proceso estocástico  $\mathbf{X}$ .

$$B\mathbf{X}_t = X_{t1}, \quad \text{para } t \in \mathbb{Z}.$$

Aplicando el operador B repetidamente tenemos

$$B^k \mathbf{X}_t = X_{tk}, \quad \text{para } t, k \in \mathbb{Z}$$

Así, para el polinomio  $\mathbf{a}(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3$ , y el proceso estocástico  $\mathbf{Y}$

$$\begin{aligned} \mathbf{a}(B)\mathbf{Y}_t &= (a_0 + a_1 B + a_2 B^2 + a_3 B^3)\mathbf{Y}_t \\ &= a_0 \mathbf{Y}_t + a_1 \mathbf{Y}_{t-1} + a_2 \mathbf{Y}_{t-2} + a_3 \mathbf{Y}_{t-3} \\ &= (\mathbf{a} * \mathbf{Y})_t, \quad \text{para } t \in \mathbb{Z} \end{aligned}$$

Y en general, si la suma  $\sum_{k=-\infty}^{\infty} a_k \mathbf{Y}_{t-k}$  converge para todo  $t$ , entonces

$$\begin{aligned} \mathbf{a}(B)\mathbf{Y}_t &= (\cdots + a_{-2} B^{-2} + a_{-1} B^{-1} + a_0 + a_1 B + a_2 B^2 + \cdots) \mathbf{Y}_t \\ &= \cdots + a_{-2} \mathbf{Y}_{t+2} + a_{-1} \mathbf{Y}_{t+1} + a_0 \mathbf{Y}_t + a_1 \mathbf{Y}_{t-1} + a_2 \mathbf{Y}_{t-2} + \cdots \\ &= (\mathbf{a} * \mathbf{Y})_t, \quad \text{para } t \in \mathbb{Z} \end{aligned}$$



## 4. Ejemplos de procesos (débilmente) estacionarios

### 4.1. Proceso de ruido blanco

Una secuencia  $\mathbf{U} = (U_t \mid t \in \mathbb{Z})$  de variables aleatorias **incorreladas** y tales que

$$E(U_t) = 0 \quad \text{y} \quad \text{Var}(U_t) = E(U_t^2) = \sigma^2$$

para  $t \in \mathbb{Z}$  y  $0 < \sigma^2 < \infty$  se llama *proceso de ruido blanco*.  $\mathbf{U} \sim WN(0, \sigma^2)$ .

Al ser variables aleatorias incorreladas, su función de autocovarianzas es

$$\gamma(z) = \sigma^2 z^0 = (\dots, 0, 0, \sigma^2, 0, 0, \dots)$$

- Es el proceso estacionario (no trivial) más sencillo.
- Este proceso es el pilar sobre el que definiremos el resto de ejemplos.

### 4.2. Procesos lineales

Sea  $\mathbf{U} \sim WN(0, \sigma^2)$  y sea  $\mathbf{b} \in \ell^2$ ; es decir, una secuencia de cuadrado sumable  $\sum_{j \in \mathbb{Z}} b_j^2 < \infty$ .

Denominamos *proceso lineal* al proceso estocástico  $\mathbf{X} = \mathbf{b} * \mathbf{U}$  cuyos elementos son

$$X_t = (\mathbf{b} * \mathbf{U})_t = \mathbf{b}(B)U_t = \sum_{j=-\infty}^{\infty} b_j U_{t-j}; \quad t \in \mathbb{Z}.$$

$\mathbf{b}(B)$  se denomina *función de transferencia* del filtro lineal que relaciona  $X_t$  con  $U_t$ .

El proceso está bien definido puesto que la serie infinita converge en norma por el Teorema de Riesz-Fisher (Pourahmadi, M. 2001, Teorema 9.7). Y el proceso es estacionario porque, usando la continuidad de los productos escalares (Pourahmadi, M. 2001, Teorema 9.2),

$$\begin{aligned} \gamma_k = \text{Cov}(X_{t+k}, X_t) &= \langle X_{t+k} \mid X_t \rangle = \lim_{m, n \rightarrow \infty} \left\langle \sum_{i=-m}^m b_i U_{t+k-i} \mid \sum_{j=-n}^n b_j U_{t-j} \right\rangle \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} b_i b_j \langle U_{t+k-i} \mid U_{t-j} \rangle \\ &= \sigma^2 \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} b_i b_j \delta_{t+k-i, t-j} \\ &= \sigma^2 \sum_{i=-\infty}^{\infty} b_i b_{i+k} = \sigma^2 (\mathbf{b}(z) * \mathbf{b}(z^{-1}))_k \end{aligned}$$

que solo depende de  $k$  (en la tercera línea,  $\delta_{p,q}$  es la delta de Kronecker; y en la cuarta hemos usado la última ecuación de la Lección 4).

El proceso lineal es “causal” si además  $\mathbf{b}$  es una serie formal (i.e.,  $\text{cogrado}(\mathbf{b}) \geq 0$ )

$$X_t = \sum_{j=0}^{\infty} b_j U_{t-j}; \quad t \in \mathbb{Z}$$

(pues cada  $X_t$  es una suma de variables “del presente y/o el pasado”).

La clase de **procesos lineales causales** incluye muchas e importantes subclases de procesos, algunas de las cuales son objeto principal de estudio de este curso.

#### 4.2.1. Media móvil infinita. $MA(\infty)$

Sea  $\mathbf{U} \sim WN(0, \sigma^2)$  y sea  $\boldsymbol{\psi} \in \ell^2$  una serie formal con infinitos términos NO nulos; entonces el proceso estocástico  $\boldsymbol{\psi} * \mathbf{U}$ , cuyos elementos son

$$X_t = (\boldsymbol{\psi} * \mathbf{U})_t = \boldsymbol{\psi}(B)U_t = \sum_{j=0}^{\infty} \psi_j U_{t-j}; \quad t \in \mathbb{Z}$$

se denomina proceso de *media móvil infinita*  $MA(\infty)$ .

Algunas clases de procesos lineales causales tienen una representación parsimoniosa, pues basta un número finito de parámetros para representarlos completamente. Por ejemplo, cuando  $\boldsymbol{\psi}$  tiene un número finito de términos no nulos...

#### 4.2.2. Proceso de media móvil de orden $q$ . $MA(q)$

Sea  $\mathbf{U} \sim WN(0, \sigma^2)$  y sea  $\boldsymbol{\theta}$  un polinomio de grado  $q$  con  $\theta_0 = 1$ ; entonces el proceso estocástico  $\boldsymbol{\theta} * \mathbf{U}$ , cuyos elementos son

$$X_t = (\boldsymbol{\theta} * \mathbf{U})_t = \boldsymbol{\theta}(B)U_t = \sum_{j=0}^q \theta_j U_{t-j}; \quad t \in \mathbb{Z}$$

se denomina proceso de *media móvil*  $MA(q)$ .

Es decir, si  $\boldsymbol{\theta} = (1 - \theta_1 z - \dots - \theta_q z^q)$ :

$$X_t = U_t - \theta_1 U_{t-1} - \dots - \theta_q U_{t-q}.$$

Hay otros procesos lineales con representación parsimoniosa.

#### 4.2.3. Proceso autorregresivo de orden $p$ . $AR(p)$

Sea  $\mathbf{U} \sim WN(0, \sigma^2)$ , se denomina *proceso autorregresivo de orden  $p$*  a aquel proceso estocástico *estacionario*  $\mathbf{X}$  que es la solución de la siguiente ecuación en diferencias

$$\boldsymbol{\phi} * \mathbf{X} = \mathbf{U}$$

donde  $\boldsymbol{\phi}$  un polinomio de grado  $p$  con  $\phi_0 = 1$ ;

Por tanto,

$$(\boldsymbol{\phi} * \mathbf{X})_t = \boldsymbol{\phi}(B)X_t = \sum_{j=0}^p \phi_j X_{t-j} = U_t.$$

Si  $\boldsymbol{\phi} = (1 - \phi_1 z - \dots - \phi_p z^p)$  entonces  $\mathbf{X} = (X_t \mid t \in \mathbb{Z})$  es solución de la ecuación:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = U_t.$$

El problema con la anterior definición es que la ecuación  $\boldsymbol{\phi} * \mathbf{X} = \mathbf{U}$  no tiene solución única (y en algunos casos ninguna solución es estacionaria). Despejemos  $\mathbf{X}$  para verlo.

Multiplicando ambos lados de la ecuación por una inversa de  $\boldsymbol{\phi}$  tenemos

$$\mathbf{X} = \text{inversa}(\boldsymbol{\phi}) * \mathbf{U}.$$

Y si denotamos la secuencia *inversa*( $\phi$ ) con  $\mathbf{a}$  entonces

$$X_t = \mathbf{a}(\mathbf{B})U_t = \sum_{j \in \mathbb{Z}} a_j U_{t-j}.$$

Pero... ¿Qué secuencia  $\mathbf{a}$  usamos como inversa de  $\phi$ ? Recuerde que hay infinitas y la mayoría no son sumables (si el polinomio  $\phi$  tiene raíces unitarias ninguna lo es).

En tal caso la expresión  $\mathbf{a}(\mathbf{B})U_t = \sum_{j=-\infty}^{\infty} a_j U_{t-j}$  carece de sentido (pues no converge).

**Requisitos** sobre el polinomio autorregresivo  $\phi$ . Para que el proceso AR exista y sea:

1. lineal y estacionario, exigiremos que  $\phi$  no tenga raíces de módulo 1.

Entonces existe una única inversa absolutamente sumable:  $\phi^{-1} \in \ell^1 \subset \ell^2$ .

La inversa  $\mathbf{a} = \phi^{-1}$  corresponde a la única solución *estacionaria* de  $\phi * \mathbf{X} = \mathbf{U}$ . (Si  $\phi$  tuviera raíces de módulo 1 no existiría ni  $\phi^{-1}$ , ni la solución estacionaria).

$$X_t = \phi^{-1}(\mathbf{B})U_t = \sum_{j=-\infty}^{\infty} a_j U_{t-j}$$

2. causal exigiremos que las raíces de  $\phi$  sean mayores que 1 en valor absoluto (raíces fuera del círculo unidad):  $\phi^{-1} = \phi^{-\triangleright}$  (**serie formal**  $\in \ell^1 \subset \ell^2$ ).

$$X_t = \phi^{-1}(\mathbf{B})U_t = \sum_{j=0}^{\infty} a_j U_{t-j}$$

(¡de nuevo un proceso lineal causal!)

El siguiente modelo lineal es una combinación (o generalización) de los dos anteriores.

#### 4.2.4. Proceso autorregresivo de media móvil. ARMA( $p, q$ )

Sea  $\mathbf{U} \sim WN(0, \sigma^2)$ , se denomina *proceso autorregresivo de media móvil* ( $p, q$ ) al proceso estocástico estacionario  $\mathbf{X}$  que es la solución de la ecuación en diferencias:

$$\phi * \mathbf{X} = \theta * \mathbf{U}$$

donde el polinomio *autorregresivo*  $\phi$  tiene grado  $p$  con  $\phi_0 = 1$  y con todas sus raíces fuera del círculo unidad (*por los motivos anteriormente vistos*); y el polinomio *de media móvil*  $\theta$  es de grado  $q$  con  $\theta_0 = 1$ ;

$$\text{es decir, } \mathbf{X} = \frac{\theta}{\phi} * \mathbf{U}; \quad \text{donde } \frac{\theta}{\phi} \equiv \phi^{-1} * \theta$$

Tanto  $\phi^{-1}$  como  $\theta$  son series formales absolutamente sumables y como  $\ell^1$  y las series formales son anillos,  $\phi^{-1} * \theta \equiv \frac{\theta}{\phi} \in \ell^1$  también es una serie formal absolutamente sumable (y por tanto de cuadrado sumable). Consecuentemente el proceso estocástico es un proceso lineal causal.

$$X_t = \frac{\theta}{\phi}(\mathbf{B})U_t = \sum_{j=0}^{\infty} a_j U_{t-j}$$

donde  $\mathbf{a} = \phi^{-1} * \theta$ .

#### 4.2.5. Proceso autorregresivo de media móvil con media no nula

Consideremos un proceso  $\mathbf{Y}$  con media distinta de cero, es decir,

$$E(Y_t) = \mu \neq 0$$

y definamos la secuencia constante  $\boldsymbol{\mu} = \sum_{j \in \mathbb{Z}} \mu z^j = (\dots, \mu, \mu, \mu, \dots)$ .

Decimos que  $\mathbf{Y}$  es un proceso ARMA( $p, q$ ) con media distinta de cero si  $\mathbf{X}$  es ARMA( $p, q$ )

$$\phi * \mathbf{X} = \boldsymbol{\theta} * \mathbf{U}$$

donde  $\mathbf{X} = \mathbf{Y} - \boldsymbol{\mu}$  es evidentemente un proceso de media cero. Por tanto

$$\begin{aligned} \phi * (\mathbf{Y} - \boldsymbol{\mu}) &= \boldsymbol{\theta} * \mathbf{U} \\ \phi * \mathbf{Y} - \phi * \boldsymbol{\mu} &= \boldsymbol{\theta} * \mathbf{U} \\ \phi * \mathbf{Y} &= \phi * \boldsymbol{\mu} + \boldsymbol{\theta} * \mathbf{U} \end{aligned}$$

Es decir, si  $\phi(B)$  es  $1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ , entonces

$$\phi(B)Y_t = c + \boldsymbol{\theta}(B)U_t$$

donde

$$c = (1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu$$

y donde  $\mu = E(Y_t)$ , es un proceso autorregresivo de media móvil ARMA( $p, q$ ) *con media no nula*.

## 5. Primeros momentos de procesos lineales causales

### 5.1. Esperanza y autocovarianzas de un proceso lineal causal

Sea  $\mathbf{X} = \boldsymbol{\psi} * \mathbf{U}$ , donde  $\boldsymbol{\psi}$  es una serie formal de cuadrado sumable ( $\boldsymbol{\psi} \in \ell^2$ ) y donde  $\mathbf{U} \sim WN(0, \sigma^2)$ . Recordando que la convolución es una operación lineal:

$$E(\mathbf{X}) = E(\boldsymbol{\psi} * \mathbf{U}) = \boldsymbol{\psi} * E(\mathbf{U}) = \boldsymbol{\psi} * \mathbf{0} = \mathbf{0}.$$

Consecuentemente, la covarianza de orden  $k$  para cada  $X_t$  es

$$\begin{aligned} \gamma_{k,t} &= E\left[(\boldsymbol{\psi}(B)X_t) \cdot (\boldsymbol{\psi}(B)X_{t-k})\right] \\ &= E\left[(\psi_0 U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} \dots)(\psi_0 U_{t-k} + \psi_1 U_{t-k-1} + \psi_2 U_{t-k-2} \dots)\right] \\ &= \sigma^2 \sum_{j \in \mathbb{Z}} \psi_{j+k} \psi_j \quad \text{ya que } E(U_h U_j) = 0 \text{ si } j \neq h, \end{aligned}$$

que no depende de  $t$  ( $\mathbf{X}$  es estacionario). Es más, por la última ecuación de la lección 4

$$\gamma_k = \sigma^2 \sum_{j \in \mathbb{Z}} \psi_{j+k} \psi_j = \sigma^2 (\boldsymbol{\psi}(z) * \boldsymbol{\psi}(z^{-1}))_k \quad \text{para } k \in \mathbb{Z}.$$

Y, por tanto

$$\boldsymbol{\gamma} = \sigma^2 \boldsymbol{\psi}(z) * \boldsymbol{\psi}(z^{-1}) \tag{3}$$

con grado igual al grado de  $\boldsymbol{\psi}$  y cogrado igual a menos el grado de  $\boldsymbol{\psi}$ .

## 5.2. Covarianza cruzada entre dos procesos lineales causales

Sean  $\mathbf{W} = \boldsymbol{\theta} * \mathbf{U}$  e  $\mathbf{Y} = \boldsymbol{\psi} * \mathbf{U}$ , donde  $\boldsymbol{\theta}$  y  $\boldsymbol{\psi}$  son series formales de cuadrado sumable y donde  $\mathbf{U} \sim WN(0, \sigma^2)$ .

Entonces la covarianza cruzada (de orden  $k \in \mathbb{Z}$ ) entre  $W_t$  e  $Y_{t-k}$  es

$$\begin{aligned} E[W_t \cdot Y_{t-k}] &= E[(\boldsymbol{\theta}(\mathbf{B})U_t) \cdot (\boldsymbol{\psi}(\mathbf{B})U_{t-k})] \\ &= E[(\theta_0 U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} \cdots)(\psi_0 U_{t-k} + \psi_1 U_{t-k-1} + \psi_2 U_{t-k-2} \cdots)] \\ &= \sigma^2 \sum_{j \in \mathbb{Z}} \theta_{j+k} \psi_j \quad \text{ya que } E(U_h U_j) = 0 \text{ si } j \neq h \end{aligned}$$

que tampoco depende de  $t$ . Es más, por la última ecuación de la lección 4

$$\gamma_{\mathbf{W}, \mathbf{Y}}(k) = \sigma^2 \sum_{j \in \mathbb{Z}} \theta_{j+k} \psi_j = \sigma^2 (\boldsymbol{\theta}(z) * \boldsymbol{\psi}(z^{-1}))_k \quad \text{para todo } k \in \mathbb{Z}.$$

Es decir...

Repitiendo los mismos pasos que en el caso de la autocovarianza, llegamos a que la función de covarianzas cruzadas es la secuencia

$$\gamma_{\mathbf{W}, \mathbf{Y}} = \sigma^2 \boldsymbol{\theta}(z) * \boldsymbol{\psi}(z^{-1}) \quad (4)$$

con grado igual al grado de  $\boldsymbol{\theta}$  y cogrado igual a menos el grado de  $\boldsymbol{\psi}$ .

## 5.3. Las Ecuaciones de Yule-Walker para un AR( $p$ ) estacionario

*Por una parte* (lado izquierdo):

Si  $\mathbf{X}$  es un proceso (débilmente) estacionario con  $E(\mathbf{X}) = \mathbf{0}$  y  $\boldsymbol{\phi}$  es una serie formal absolutamente sumable; entonces para  $t, k \in \mathbb{Z}$

$$E[(\boldsymbol{\phi}(\mathbf{B})X_t) \cdot X_{t-k}] = \boldsymbol{\phi}(\mathbf{B})E(X_t \cdot X_{t-k}) = \boldsymbol{\phi}(\mathbf{B})\gamma_k \quad (5)$$

que no depende de  $t$ , por ser  $\mathbf{X}$  es un proceso (débilmente) estacionario.

*Por otra parte* (lado derecho):

Si  $\mathbf{X}$  tiene representación  $\mathbf{X} = \boldsymbol{\psi} * \mathbf{U}$  donde  $\mathbf{U} \sim WN(0, \sigma^2)$  y  $\boldsymbol{\psi} \in \ell^2$  es una serie formal con  $\psi_0 = 1$ ; es decir, si es un proceso lineal causal

$$X_t = U_t + \sum_{j=1}^{\infty} \psi_j U_{t-j},$$

entonces para  $t, k \in \mathbb{Z}$

$$E[U_t \cdot X_{t-k}] = E\left[U_t \left(U_{t-k} + \sum_{j=1}^{\infty} \psi_j U_{t-k-j}\right)\right] = \begin{cases} \sigma^2 & \text{cuando } k = 0 \\ 0 & \text{cuando } k \neq 0 \end{cases} \quad (6)$$

Sea un AR( $p$ ) estacionario:  $\boldsymbol{\phi}(\mathbf{B})X_t = U_t$  donde  $\boldsymbol{\phi}(z) = 1 - \phi_1 z^1 - \cdots - \phi_p z^p$ . Multiplicando por  $X_{t-k}$  y tomando esperanzas:

$$E[(\boldsymbol{\phi}(\mathbf{B})X_t) \cdot X_{t-k}] = E[U_t \cdot X_{t-k}]$$

**para  $k = 0$ :** (por 5 y 6)

$$\boxed{\boldsymbol{\phi}(\mathbf{B})\gamma_0 = \sigma^2} \quad \Rightarrow \quad \gamma_0 - \phi_1 \gamma_1 - \cdots - \phi_p \gamma_p = \sigma^2 \quad \Rightarrow \quad \sigma^2 = \gamma_0 - \sum_{j=1}^p \phi_j \gamma_j.$$

Dividiendo por  $\gamma_0$  (y recordando que  $\rho_0 = 1$ ):

$$\phi(\mathbf{B})\rho_0 = \frac{\sigma^2}{\gamma_0} \Rightarrow \boxed{\gamma_0 = \frac{\sigma^2}{\phi(\mathbf{B})\rho_0}} \Rightarrow \gamma_0 = \frac{\sigma^2}{1 - \sum_{j=1}^p \phi_j \rho_j}.$$

para  $k > 0$ : (por 5 y 6)

$$\boxed{\phi(\mathbf{B})\gamma_k = 0} \Rightarrow \gamma_k - \phi_1\gamma_{k-1} - \cdots - \phi_p\gamma_{k-p} = 0 \Rightarrow \gamma_k = \sum_{j=1}^p \phi_j\gamma_{k-j}.$$

Dividiendo por  $\gamma_0$ :

$$\boxed{\phi(\mathbf{B})\rho_k = 0} \Rightarrow \rho_k - \phi_1\rho_{k-1} - \cdots - \phi_p\rho_{k-p} = 0 \Rightarrow \rho_k = \sum_{j=1}^p \phi_j\rho_{k-j}.$$

Por tanto, la estructura autorregresiva del proceso impone que las autocovarianzas (y las autocorrelaciones) verifiquen las ecuaciones de Yule-Walker.

#### 5.4. Función de autocovarianzas para un ARMA( $p, q$ )

Sea un ARMA( $p, q$ ) estacionario:  $\phi(\mathbf{B})X_t = \theta(\mathbf{B})U_t$  donde  $\phi$  y  $\theta$  no tienen raíces comunes. Multiplicando por  $X_{t-k}$ , tomando esperanzas y sustituyendo  $X_{t-k}$  por su representación MA( $\infty$ ), donde  $\psi = \frac{\theta}{\phi}$ :

$$\underbrace{E[(\phi(\mathbf{B})X_t) \cdot X_{t-k}]}_{\phi(\mathbf{B})\gamma_k \text{ (por 5)}} = E[(\theta(\mathbf{B})U_t) \cdot X_{t-k}] = \underbrace{E[(\theta(\mathbf{B})U_t) \cdot (\psi(\mathbf{B})U_{t-k})]}_{\gamma_{\mathbf{W}, \mathbf{Y}}(k)}$$

Donde hemos usando (5) y renombrando  $\theta(\mathbf{B})U_t = \mathbf{W}$  y  $\psi(\mathbf{B})U_t = \mathbf{Y}$ . Así:

$$\begin{aligned} \phi(\mathbf{B})\gamma_k &= \gamma_{\mathbf{W}, \mathbf{Y}}(k) \\ &= \sigma^2 (\theta(z) * \psi(z^{-1}))_k \end{aligned} \quad \text{por (4)}$$

Y como  $\theta(z) * \psi(z^{-1})$  tiene grado  $q$  y cogrado  $-\infty$

$$\phi(\mathbf{B})\gamma_k = \begin{cases} 0 & k > q \quad (\text{como en un AR}) \\ \sigma^2 (\theta(z) * \psi(z^{-1}))_k & k \leq q \quad (\text{que depende de } \theta \text{ y } \phi) \end{cases} \quad (7)$$