CAPíTULO 8

Análisis de series temporales

Los datos estadísticos y, en particular, los datos económicos se recopilan a menudo en forma de series temporales. Una **serie temporal** es un conjunto ordenado de observaciones $\{z_1, \ldots, z_t, \ldots, z_n\}$ obtenidas en intervalos regulares de tiempo, en donde z_t denota la observación de la variable de interés en el instante o intervalo temporal t. El instante temporal t suele ser un año, trimestre, mes, semana, etc., y determina la frecuencia de observación: anual, trimestral, mensual, semanal, etc. Suponemos, por tanto, que todas las observaciones de la serie se obtienen en instantes equidistantes de tiempo y descartamos así las series temporales compuestas, por ejemplo, de observaciones anuales, trimestrales y mensuales.

La característica distintiva de una serie temporal es la dependencia observacional: el valor de una variable en una determinada fecha depende de los valores de la propia variable en fechas previas. Esta idea subyace tras la especificación de los **procesos estocásticos univariantes** que estudiamos en este tema, los cuales explican la evolución temporal de una variable a partir de su comportamiento pasado. Vamos a caracterizar estos modelos dinámicos por medio de las funciones de autocorrelación simple y parcial.

Análogamente a la interpretación del modelo lineal general como proceso generador de datos, vamos a contemplar una **serie temporal** como una *realización particular de un proceso estocástico*. Uno de los principales propósitos del análisis de series temporales es inferir las propiedades del proceso (población) a partir de una realización particular (muestra), para lo cual limitamos nuestro interés a una clase de procesos que se encuentran en un estado de equilibrio estadístico: los procesos ARMA estacionarios, que tienen momentos estables. Dentro de esta clase, el proceso autorregresivo de primer orden y el proceso de medias móviles de primer orden son los dos procesos más usados.

Los métodos que estudiamos en este capítulo se tratan extensamente en el libro *Time Series Analysis: Forecasting and Control* de Box y Jenkins (1970, 1976, 1994) y a menudo se denominan métodos Box-Jenkins. Estos autores desarrollaron un metodología para sistematizar la construcción de una clase de modelos de series temporales que se ha mostrado muy útil en predicción.

8.1. Procesos estocásticos estacionarios

DEFINICIÓN 53. *Un* **proceso estocástico** *es una colección o secuencia de variables aleatorias* $\{z_t(\omega);\ \omega \in \Omega,\ t \in \mathrm{T}\}$*, en donde* Ω *es el conjunto de todos los sucesos elementales y* T *es un conjunto de índices.*

Observación 39. *Los procesos estocásticos pueden ser discretos o continuos, dependiendo de si el conjunto* T *es contable (números naturales) o incontable (números reales). En este tema estudiamos procesos estocásticos discretos, a los que llamamos abreviadamente procesos.*

Suponemos que las variables aleatorias $z_t(\omega)$ o, simplemente, z_t tienen distribución continua con función de densidad $p(z_t)$ que satisface la condición

$$\int_{-\infty}^{\infty} p(z_t)dz_t = 1$$

De aquí, los momentos de orden r de la variable aleatoria z_t

$$E(z_t^r) = \int_{-\infty}^{\infty} z_t^r p(z_t)dz_t$$

existirán si

$$\int_{-\infty}^{\infty} |z_t|^r p(z_t)dz_t < \infty$$

Además, las variables aleatorias bidimensionales (z_t, z_s) tienen una función de densidad conjunta $p(z_t, z_\tau)$ que satisface la condición

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(z_t, z_s)dz_t dz_s = 1$$

y que nos permite definir momentos de orden (r, s) como

$$E(z_t^r z_\tau^s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_t^r z_t^s p(z_t, z_\tau)dz_t dz_\tau$$

que existirán sí y sólo sí

$$E(z_t^r z_\tau^s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |z_t^r z_t^s| p(z_t, z_\tau)dz_t dz_\tau < \infty$$

En general, el subconjunto de variables aleatorias $(z_{t_1}, \ldots, z_{t_m})$ tiene función de densidad conjunta $p(z_{t_1}, \ldots, z_{t_m})$.

Definición 54. *Una* **serie temporal***, $z_1, \ldots, z_n$ es una realización particular de un proceso estocástico $z_1(\omega), \ldots, z_n(\omega)$.*

En esta definición cada observación z_t de la serie temporal se interpreta como un valor particular de una variable aleatoria $z_t(\omega)$. Con un sólo dato no podemos pretender estimar los momentos de $z_t(\omega)$. Por tanto, para inferir la distribución del proceso estocástico a partir de una serie temporal es necesario restringir nuestro estudio a una clase particular de procesos que tengan momentos estables.

Definición 55. *Un proceso es estacionario de orden r si todos sus momentos hasta el orden r existen y son estables.*

Definición 56. *Un proceso es estacionario de segundo orden si*

1. $E(z_t) = \mu < \infty \quad \forall t \in \mathrm{T}$
2. $E(z_t - \mu)^2 = \sigma^2 < \infty \quad \forall t \in \mathrm{T}$
3. $E(z_t - \mu)(z_s - \mu) = \gamma_{|t-s|} < \infty \quad \forall t, s \in \mathrm{T}$

La notación $E(z_t) = \mu$ y $E(z_t - \mu)^2 = \sigma^2$ indica que la media y la varianza de z_t no depende de t; en otras palabras, todas las variables aleatorias tienen la misma media y la misma varianza (estacionariedad en media y varianza). Análogamente, la notación $E(z_t - \mu)(z_s - \mu) = \gamma_{|t-s|}$ indica que la covarianza entre dos variables aleatorias z_t y z_s depende de la distancia entre sus índices $|t-s|$, pero no depende de t ni de s. Así, todas la variables aleatorias bidimensionales separadas un periodo (z_1, z_2), (z_2, z_3), $\ldots$, (z_{n-1}, z_n) tendrán las misma covarianza γ_1. Del mismo modo, todas la variables aleatorias bidimensionales separadas dos periodos (z_1, z_3), (z_2, z_4), $\ldots$, (z_{n-2}, z_n) tendrán las misma covarianza γ_2. En general, todas las variables aleatorias separadas k periodos (z_t, z_{t-k}) tendrán covarianza γ_k. Por otro lado, con la relación menor que infinito indicamos la existencia correspondiente del momento.

Proposición 63. *Bajo estacionariedad de segundo orden, podemos estimar los dos primeros momentos poblaciones del proceso mediante los correspondientes momentos muestrales de la serie temporal:*

1. *Media muestral*
$$\hat{\mu} = \bar{z} = \frac{\sum_{t=1}^{n} z_t}{n}$$
2. *Varianza muestral*
$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} (z_t - \bar{z})^2}{n}$$
3. *Covarianza muestral en el retardo k*
$$\hat{\gamma}_k = \frac{\sum_{t=k+1}^{n} (z_t - \bar{z})(z_{t-k} - \bar{z})}{n}$$

Definición 57. *Un proceso es estrictamente estacionario si la función de densidad de un subconjunto de m variables aleatorias cualesquiera no se ve afectada por un desplazamiento temporal*

$$p(z_{t_1}, z_{t_2}, \ldots, z_{t_m}) = p(z_{t_1+k}, z_{t_2+k}, \ldots, z_{t_m+k})$$

en donde $t_1, \ldots, t_m$ son m índices no necesariamente consecutivos y k es el desplazamiento temporal.

Observación 40. *Para $m = 1$, la estacionariedad estricta implica que todas las variables aleatorias tienen la misma distribución de probabilidad.*

Definición 58. *Un proceso se dice Gaussiano cuando la distribución de probabilidad de cualquier subconjunto de variables aleatorias es Normal.*

Definición 59. *Si un proceso es Gaussiano y débilmente estacionario, entonces también será estrictamente estacionario.*

Definición 60. *Un* **proceso de ruido blanco** *o puramente aleatorio es una secuencia de variables aleatorias $\{a_t\}$ mutuamente ortogonales con media cero y varianza constante: $E(a_t) = 0$, $E(a_t^2) = \sigma_a^2$ y $E(a_t a_s) = 0$ para $t \neq s$.*

8.2. Funciones de autocorrelación simple y parcial

El coeficiente de correlación simple entre dos variables aleatorias X e Y, denotado por ρ_{XY}, se define como

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(X - EX)(Y - EY)}{\sqrt{E(X - EX)^2 E(Y - EY)^2}}$$

Definición 61. *El coeficiente de autocorrelación simple en el retardo k, denotado por ρ_k, es el coeficiente de correlación simple entre las variables aleatorias z_t y z_{t-k}*

$$\rho_k = \frac{Cov(z_t, z_{t-k})}{\sqrt{Var(z_t)Var(z_{t-k})}} = \frac{E(z_t - \mu)(z_{t-k} - \mu)}{\sqrt{E(z_t - \mu)^2 E(z_{t-k} - \mu)^2}} = \frac{\gamma_k}{\gamma_0}$$

Proposición 64. *El coeficiente de autocorrelación simple en el retardo k es la pendiente en la regresión lineal simple de z_t sobre z_{t-k}.*

Demostración. La regresión simple de z_t sobre z_{t-k} es

$$z_t = \beta_1 + \beta_2 z_{t-k} + u_t$$

en donde u_t es un proceso de ruido blanco. Tomando esperanza matemática tenemos

$$\mu = \beta_1 + \beta_2 \mu$$

Por tanto, la ecuación de regresión simple en desviaciones respecto a las *medias poblaciones* es

$$\tilde{z}_t = \beta_2 \tilde{z}_{t-k} + u_t$$

Multiplicando la ecuación por $\tilde{z}_{t-k}$ y tomando esperanza matemática obtenemos

$$E(\tilde{z}_t \tilde{z}_{t-k}) = \beta_2 E(\tilde{z}_{t-k}^2) + E(u_t z_{t-k})$$

en donde $E(u_t z_{t-k}) = 0$ para $k > 1$. De aquí,

$$\gamma_k = \beta_2 \gamma_0 \quad \Rightarrow \quad \beta_2 = \frac{\gamma_k}{\gamma_0} = \rho_k$$

□

El coeficiente de autocorrelación en el retardo k puede estimarse a partir de los datos de una serie temporal como

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^{n}(z_t - \bar{z})(z_{t-k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}$$

que, en grandes muestras, puede aproximarse por la pendiente estimada en la regresión lineal simple de z_t sobre z_{t-k}

$$\hat{\beta}_2 = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^{n}(z_t - \bar{z})(z_{t-k} - \bar{z})}{\sum_{t=k+1}^{n}(z_{t-k} - \bar{z})^2} \simeq \hat{\gamma}_k$$

El coeficiente de correlación parcial entre dos variables aleatorias X e Y dada Z, denotado por $\rho_{XY.Z}$, se define como el coeficiente de correlación simple entre X e Y después de extraer la influencia de Z. La extensión de esta medida a un proceso estocástico es como sigue.

Definición 62. *El coeficiente de autocorrelación parcial en el retardo k, denotado por ϕ_{kk}, es el la correlación simple entre z_t y z_{t-k} después de extraer la influencia de los retardos intermedios.*

El cálculo de las autocorrelaciones parciales puede basarse en el modelo de regresión múltiple en desviaciones respecto a las *medias poblacionales*

$$\tilde{z}_t = \phi_{1,k}\tilde{z}_{t-1} + \cdots + \phi_{k,k}\tilde{z}_{t-k} + u_t$$

Multiplicando el modelo por $\tilde{z}_{t-k}$ y tomando esperanza matemática

$$E(\tilde{z}_t \tilde{z}_{t-k}) = \phi_{1,k} E(\tilde{z}_{t-1}\tilde{z}_{t-k}) + \cdots + \phi_{k,k} E(\tilde{z}_{t-k}^2) + E(u_t \tilde{z}_{t-k})$$

y tomando esperanza matemática obtenemos

$$\gamma_k = \phi_{1,k}\gamma_{k-1} + \cdots + \phi_{k,k}\gamma_{t-k}$$

Dividiendo por γ_0 podemos especificar el sistema de ecuaciones de Yule-Walker

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix} = \begin{pmatrix} \rho_0 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & \rho_0 & \dots & \rho_{k-2} \\ \vdots & & & \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_0 \end{pmatrix} \begin{pmatrix} \phi_{1,k} \\ \phi_{2,k} \\ \vdots \\ \phi_{k,k} \end{pmatrix} \tag{8.1}$$

que nos permite obtener la autocorrelación parcial $\phi_{k,k}$ en términos de las autocorrelaciones simples $\rho_1, \dots, \rho_k$. Aplicando la regla de Cramer, tenemos que

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-1} & \dots & 1 \end{vmatrix}}$$

Definición 63. *El conjunto* $\{\gamma_k;\ k = 0, \pm 1, \dots\}$ *se denomina función de autocovarianzas.*

Definición 64. *El conjunto* $\{\rho_k;\ k = 0, \pm 1, \}$ *se denomina función de autocorrelación simple (ACF); y su gráfico, correlograma.*

Definición 65. *El conjunto* $\{\phi_{k,k};\ k = 0, \pm 1, \}$ *se denomina función de autocorrelación parcial (PACF).*

8.3. El proceso estacionario lineal general

Definición 66. *El* **proceso lineal general** *expresa cada la variable aleatoria* $\tilde{z}_t$ *como una combinación lineal de todas sus retardos pasados más un término de error puramente aleatorio*

$$\begin{aligned} \tilde{z}_t =& \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + \dots + a_t \\ =& \sum_{j=1}^{\infty} \pi_j \tilde{z}_{t-j} + a_t \end{aligned} \tag{8.2}$$

en donde $\tilde{z}_{t-j} = z_{t-j} - \mu$ $(j = 1, 2, \dots)$ *son los retardos de la variable aleatoria* $\tilde{z}_t = z_t - \mu$, μ *es la media del proceso* z_t, π_j $(j = 1, 2, \dots)$ *son los coeficientes asociados a las variables explicativas y* a_t *es un proceso de ruido blanco.*

Observación 41. *El proceso lineal general (8.2) puede contemplarse como un modelo de regresión lineal con infinitas variables explicativas, que son los retardos de la propia variable dependiente. Se utiliza el término* **proceso univariante** *para enfatizar que el modelo incluye información de una única variable.*

Observación 42. *La ecuación (8.2) se denomina autorregresión de orden infinito, denotada por* $AR(\infty)$, *y a veces se conoce también como la forma* π *de un proceso.*

El algebra con procesos lineales se simplifica considerablemente haciendo uso del **operador de retardo** (en inglés, *lag operator*) o cambio hacia atrás (en inglés, *backward shift operator*), denotado por L o B y definido como $Bz_t = z_{t-1}$. La aplicación repetida del operador B a z_t permite expresar cualquier retardo z_{t-k} en términos de z_t. Así, $B^2 z_t = BBz_t = Bz_{t-1} = z_{t-2}$ y $B^k = z_{t-k}$. De aquí, el proceso lineal puede escribirse como

$$\begin{aligned}\tilde{z}_t =& \pi_1 B\tilde{z}_t + \pi_2 B^2 \tilde{z}_t + \cdots + a_t \\ =& \sum_{j=1}^{\infty} \pi_j B^j \tilde{z}_t + a_t\end{aligned} \tag{8.3}$$

o bien

$$\pi(B)\tilde{z}_t = a_t$$

en donde $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \ldots$ es un polinomio en B de orden infinito.

Observación 43. El operador adelanto o cambio hacia adelante (en inglés, forward shift operator) se denota por F y se define como $Fz_t = z_{t+1}$, cumpliéndose que $F = B^{-1}$.

PROPOSICIÓN 65. **Representación de Wold** *(1938). El proceso lineal general puede expresarse como una combinación lineal de los retardos de un proceso puramente aleatorio*

$$\begin{aligned}\tilde{z}_t =& a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots \\ =& a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} \\ =& \psi(B) a_t\end{aligned} \tag{8.4}$$

en donde $\psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \ldots$ es un polinomio en B de orden infinito con $\psi_0 = 1$.

DEMOSTRACIÓN. Partiendo de $\pi(B)\tilde{z}_t = a_t$, podemos escribir $\tilde{z}_t = (1/\pi(B))a_t = \psi(B)a_t$, en donde $\psi(B) = \pi^{-1}(B)$. Vemos que los polinomios $\psi(B)$ y $\pi(B)$ cumplen la relación $\psi(B)\pi(B) = 1$, de modo que podemos obtener los pesos ψ a partir de los pesos π, y viceversa. □

Observación 44. La ecuación (8.2) es una media móvil de orden infinito, denotada por $MA(\infty)$, y a veces se conoce también como la forma ψ de un proceso.

PROPOSICIÓN 66. *El proceso lineal general es estacionario de segundo orden si $\psi_0, \psi_1, \ldots$ es una serie convergente, esto es, si la suma de los valores absolutos de los pesos ψ_j es finita, $\sum_{j=0}^{\infty} |\psi_j| < \infty$.*

DEMOSTRACIÓN.

- Media: $E(\tilde{z}_t) = 0 \Rightarrow E(z_t) = \mu$.
- Varianza

$$\begin{aligned}E(\tilde{z}_t)^2 =& E(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots)^2 \\ =& E(a_t^2 + \psi_1^2 a_{t-1}^2 + \psi_2^2 a_{t-2}^2 + \cdots + 2\psi_1 a_t a_{t-1} + \ldots) \\ =& \sigma_a^2 (1 + \psi_1^2 + \psi_2^2 + \ldots)\end{aligned}$$

- Covarianzas

$$
\begin{aligned}
E(\tilde{z}_t\tilde{z}_{t-k}) =& E[(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots + \psi_k a_{t-k} + \psi_{k+1} a_{t-k-1} + \psi_{k+2} a_{t-k-2} + \dots) \\
&\times (a_{t-k} + \psi_1 a_{t-k-1} + \psi_2 a_{t-k-2} + \dots)] \\
=& \sigma_a^2 \psi_k + \sigma_a^2 \psi_{k+1}\psi_1 + ... = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}
\end{aligned}
$$

□

Definición 67. *Un proceso es invertible si $\pi_0, \pi_1, \pi_2 \dots$ es una serie convergente, esto es, $\sum_{j=1}^{\infty} |\pi_j| < \infty$. En este caso, el pasado muy remoto es irrelevante en la explicación del presente.*

Observación 45. *La condiciones de estacionariedad e invertivilidad se cumplicarán cuando el proceso tenga una representación MA y una representación AR finitas, respectivamente.*

A pesar de su interés teórico, el proceso lineal general no tiene ninguna relevancia práctica porque incluye un número infinito de parámetros, que no podemos pretender estimar usando muestras finitas. De ahí que sea conveniente buscar **representaciones parsiomoniosas** o escuetas que usen un número finito de parámetros y que sean buenas aproximaciones al proceso lineal general. Tales aproximaciones pueden obtenerse reemplazando el polinomio $\pi(B)$ por un polinomio racional

$$
\pi(B) \simeq \frac{\phi(B)}{\theta(B)} = \frac{1 - \phi_1 B - \dots - \phi_p B^p}{1 - \theta_1 B - \dots - \theta_q B^q}
$$

en donde $\phi(B)$ y $\theta(B)$ son dos polinomios en B de orden finito p y q, respectivamente.

Definición 68. *El proceso mixto autorregresivo-de medias móviles de orden (p, q), denotado por $ARMA(p, q)$ se define como*

$$
z_t = \delta + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}
$$

o, en términos del operador de retardo,

$$
(1 - \phi_1 B - \dots - \phi_p B^p) z_t = \delta + (1 - \theta_1 B - \dots - \theta_q B^q) a_t
$$

en donde p es el orden del polinomio autorregresivo y q el del polinomio de medias móviles.

Proposición 67. *El proceso $ARMA(p, q)$ es estacionario si las raíces B del polinomio autorregresivo $1 - \phi_1 B - \dots - \phi_p B^p = 0$ caen fuera del círculo unitario, es decir, son mayores que la unidad en valor absoluto.*

Proposición 68. *El proceso $ARMA(p, q)$ es invertible si las raíces B del polinomio de medias móviles $1 - \theta_1 B - \dots - \theta_q B^q = 0$ caen fuera del círculo unitario.*

8.4. Proceso autorregresivo de primer orden

Definición 69. *El proceso autorregresivo de primer orden, denotado por $AR(1)$, es*

(8.5) $$z_t = \delta + \phi z_{t-1} + a_t \quad \Leftrightarrow \quad (1 - \phi B) z_t = \delta + a_t$$

en donde δ, ϕ y $\sigma_a^2 = E(a_t^2)$ son los parámetros del modelo y a_t es un proceso de ruido blanco.

Observación 46. *El proceso $AR(1)$ se obtiene como caso especial del proceso lineal general cuando $\pi_1 = \phi_1$ y $\pi_j = 0$ para $j > 1$*

Observación 47. *El proceso $AR(1)$ puede contemplarse como un modelo de regresión simple*

$$y_t = \beta_1 + \beta_2 x_t + u_t \tag{8.6}$$

en donde la variable explicativa es la propia variable dependiente retardada un periodo. El término autorregresivo significa que la variable dependiente se explica por sí misma, sin auxilio de otras variables.

Proposición 69. *El proceso $AR(1)$ tendrá una media estable si $\phi \neq 1$.*

Demostración. El proceso lineal general tiene media estable si $\sum_{j=1}^{\infty} \pi_j \neq 1$. En nuestro caso, $\sum_{j=1}^{\infty} \pi_j = \phi$. □

Proposición 70. *Bajo estacionariedad de primer orden, podemos escribir el proceso $AR(1)$ en desviaciones respecto a la media*

$$\tilde{z}_t = \phi \tilde{z}_{t-1} + a_t \quad \Leftrightarrow \quad (1 - \phi B)\tilde{z}_t = a_t \tag{8.7}$$

Proposición 71. *La representación del Wold del proceso $AR(1)$ es*

$$\tilde{z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \cdots = a_t + \sum_{j=1}^{\infty} \phi^j a_{t-j}$$

Demostración. La forma ψ puede obtenerse siguiendo dos aproximaciones alternativas.

1. Sustitución reiterada de retardos:

$$\begin{aligned}
\tilde{z}_t &= \phi \tilde{z}_{t-1} + a_t \\
\tilde{z}_t &= \phi[\phi \tilde{z}_{t-2} + a_{t-1}] + a_t = \phi^2 \tilde{z}_{t-2} + \phi a_{t-1} + a_t \\
\tilde{z}_t &= \phi^2[\phi \tilde{z}_{t-3} + a_{t-2}] + \phi a_{t-1} + a_t = \phi^3 \tilde{z}_{t-3} + \phi^2 a_{t-2} + \phi a_{t-1} + a_t \\
&\vdots \\
\tilde{z}_t &= \phi^k \tilde{z}_{t-k} + \phi^{k-1} a_{t-(k-1)} + \cdots + \phi^2 a_{t-2} + \phi a_{t-1} + a_t
\end{aligned} \tag{8.8}$$

   Si $|\phi| < 1$, entonces el término $\phi^k \tilde{z}_{t-k}$ es despreciable y tenemos el resultado buscado.

2. Inversión del polinomio autorregresivo:

$$(1 - \phi B)\tilde{z}_t = a_t \quad \Rightarrow \quad \tilde{z}_t = \frac{1}{1 - \phi B} a_t = \psi(B) a_t$$

   en donde $\psi(B)$ es un polinomio en B de orden infinito

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$$

cuyos coeficientes pueden encontrarse de la relación $\psi(B)(1-\phi B)=1$. El polinomio producto

$$\begin{aligned}\psi(B)(1-\phi B) =&1+\psi_1 B+\psi_2 B^2+\cdots+\phi B+\phi\psi_1 B^2+\phi\psi_2 B^2\dots\\ =&1+(\psi_1-\phi)B+(\psi_2-\phi\psi_1)B^2+\cdots+(\psi_j-\phi\psi_{j-1})B^j+\dots\end{aligned}$$

será igual a 1 si sus coeficientes $\psi_j-\phi\psi_{j-1}$ son nulos. De aquí, encontramos que $\psi_j=\phi^j$ y podemos escribir

$$\tilde{z}_t=\psi(B)a_t=a_t+\phi a_{t-1}+\phi^2 a_{t-2}+\dots$$

Esta segunda aproximación a veces se obtiene directamente como una aplicación de la suma de una serie geométrica

$$1+x+x^2+x^3+\cdots=\frac{1}{1-x}\quad\text{cuando}\quad -1<x<1$$

De manera que

$$(1-\phi B)^{-1}=1+\phi B+\phi^2B^2+\dots$$

□

Proposición 72. *El proceso $AR(1)$ será estacionario de segundo orden si $-1<\phi<1$.*

Demostración. El proceso lineal general es estacionario de segundo orden si $\sum_{j=1}^{\infty}\psi_j<\infty$. En nuestro caso, como $\psi_j=\phi^j$, la condición de estacionariedad requiere que $|\phi|<1$. □

Proposición 73. *El proceso $AR(1)$ siempre es invertible.*

Demostración. Es claro que el proceso $AR(1)$ no depende del pasado remoto: $\sum_{j=1}^{\infty}\pi_j=\phi<\infty$. □

Proposición 74. *La función de autocovarianzas $\{\gamma_k\}$ de un proceso $AR(1)$ es*

$$\gamma_k=\phi^k\gamma_0$$

en donde $\gamma_0=\sigma_a^2/(1-\phi^2)$.

Demostración. Para obtener la covarianza en el retardo k, $\gamma_k=E[(z_t-\mu)(z_{t-k}-\mu)]$, multiplicamos la ecuación (8.7) por $\tilde{z}_{t-k}$ y tomamos esperanza matemática

$$\begin{aligned}E(\tilde{z}_t\tilde{z}_{t-k}) =&E[(\phi\tilde{z}_{t-1}+a_t)\tilde{z}_{t-k}]\\ =&\phi E(\tilde{z}_{t-1}\tilde{z}_{t-k})+E(a_t\tilde{z}_{t-k})\end{aligned}$$

en donde $E(a_t\tilde{z}_{t-k})=\sigma_a^2$ si $k=0$ y $E(a_t\tilde{z}_{t-k})=0$ si $k>0$. Vemos que para $k=0$, $\gamma_0=\sigma_a^2/(1-\phi^2)$; y para $k>0$, γ_k cumple una ecuación en diferencias de primer orden $\gamma_k=\phi\gamma_{k-1}$. □

Observación 48. *La varianza γ_0 del proceso puede obtenerse directamente de la forma ψ. En efecto,*

$$\begin{aligned}\gamma_0 = E(\tilde{z}_t^2) =& E[(\sum_{j=0}^{\infty} \phi^j a_{t-j})^2] = E[\sum_{j=0}^{\infty} \phi^{2j} a_{t-j}^2 + \sum_{j=0}^{\infty}\sum_{h\neq j}^{\infty} \phi^j \phi^h a_{t-j}a_{t-h}] \\ =& \sigma_a^2 \sum_{j=0}^{\infty} \phi^{2j}\end{aligned}$$

que es la suma de una progresión geométrica de razón ϕ^2. De aquí, $\gamma_0 = \sigma_a^2/(1-\phi^2)$.

Proposición 75. *La función de autocorrelación $\{\rho_k\}$ de un proceso $AR(1)$ es*

$$\rho_k = \phi^k$$

Observación 49. *Se dice que el proceso $AR(1)$ tiene una memoria infinita para indicar que z_t está correlacionado con cualquier retardo z_{t-k}.*

Proposición 76. *La función de autocorrelación parcial $\{\phi_{k,k}\}$ se anula o corta para retardos k mayores que 1, siendo $\phi_{1,1} = \rho_0 = \phi$.*

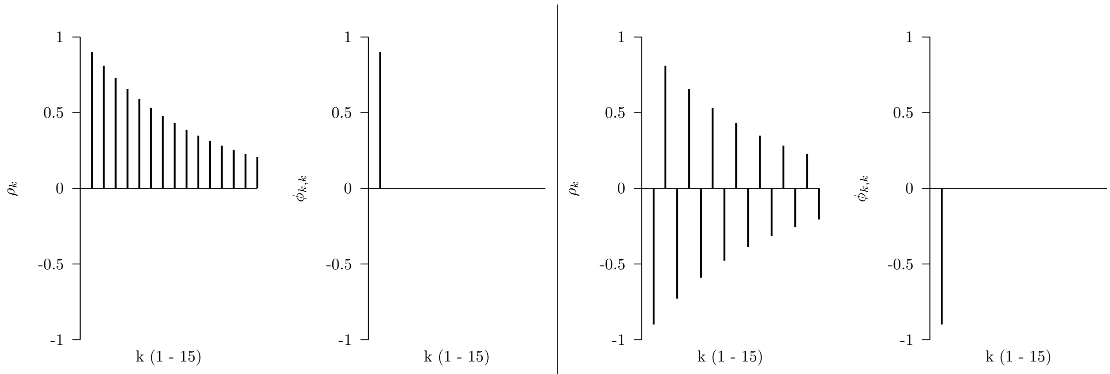
Demostración. La función de autocorrelación parcial puede calcularse a partir de las ecuaciones de Yule-Walker (8.1), resultando que

$$\begin{aligned}\phi_{11} =& \rho_1 = \phi \\ \phi_{22} =& \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \phi \\ \phi & \phi^2 \end{vmatrix}}{\begin{vmatrix} 1 & \phi \\ \phi & 1 \end{vmatrix}} = 0 \\ \vdots&\end{aligned}$$

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-1} & \dots & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \phi & \dots & \phi \\ \phi & 1 & \dots & \phi^2 \\ \vdots & \vdots & \dots & \vdots \\ \phi^{k-1} & \phi^{k-2} & \dots & \phi^k \end{vmatrix}}{\begin{vmatrix} 1 & \phi & \dots & \phi^{k-1} \\ \phi & 1 & \dots & \phi^{k-2} \\ \vdots & \vdots & \dots & \vdots \\ \phi^{k-1} & \phi^{k-2} & \dots & 1 \end{vmatrix}} = 0 \tag{8.9}$$

Vemos que el determinante del numerador para ϕ_{kk} ($k > 1$) se anula porque la última columna es ϕ veces la primera. □

El cuadro (2) muestra las funciones de autocorrelación simple y parcial para dos modelos $AR(1)$. Si $\phi = 0{,}9 > 0$, $\{\rho_k\}$ decrece exponencialmente al aumentar el retardo k, mientras que $\{\phi_{kk}\}$ sólo tiene un coeficiente distinto de cero en el primer retardo. Si $\phi = -0{,}9 < 0$, la función de autocorrelación simple decrece alternando en signo, y la función de autocorrelación parcial toma un valor negativo en el primer retardo.



(a) $\phi = 0{,}9$ (b) $\phi = -0{,}9$

Figura 1: Funciones de autocorrelación simple y parcial para dos procesos $AR(1)$

Los resultados anteriores se extienden fácilmente al proceso autorregresivo estacional de primer orden, denotado por $AR(1)_s$,

$$z_t = \delta + \phi z_{t-s} + a_t \quad \Leftrightarrow \quad (1 - \phi B^s) z_t = \delta + a_t$$

que es útil en la descripción de series trimestrales ($s = 4$) y mensuales ($s = 12$). Por ejemplo, cuando pensamos que las ventas de una empresa en un mes determinado dependen de las ventas en el mismo mes del año anterior.

8.5. Proceso de medias móviles de primer orden

Definición 70. *La ecuación de un proceso de medias móviles de primer orden, denotado por $MA(1)$, es*

$$z_t = \mu + a_t - \theta a_{t-1} \quad o \quad z_t = \mu + (1 - \theta B) a_t \tag{8.10}$$

en donde μ, θ y $\sigma_a^2 = E(a_t^2)$ son los parámetros del modelo y a_t es un proceso de ruido blanco.

Observación 50. *El proceso $MA(1)$ es un caso especial de la forma ψ del proceso lineal general que se obtiene fijando $\psi_1 = \theta$ y $\psi_j = 0\ \forall j > 1$.*

Proposición 77. *El proceso $MA(1)$ en desviaciones respecto a la media es*

$$\tilde{z}_t = a_t - \theta a_{t-1} \quad o \quad \tilde{z}_t = (1 - \theta B) a_t$$

en donde $\tilde{z}_t = z_t - \mu$

Demostración. Es inmediato comprobar que el término constante μ es la media del proceso

$$E(z_t) = \mu + E(a_t) - \theta E(a_{t-1}) = \mu$$

□

Proposición 78. *La forma π del proceso $MA(1)$ es*

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \cdots + a_t = -\sum_{j=1}^{\infty} \theta^j z_{t-j} + a_t$$

Demostración. Seguimos las dos aproximaciones descritas para el proceso $AR(1)$.

1. Sustitución reiterada de errores retardos:

$$\begin{aligned}
\tilde{z}_t =& a_t - \theta a_{t-1}\\
\tilde{z}_t =& a_t - \theta(\tilde{z}_{t-1} + \theta a_{t-2}) = a_t - \theta\tilde{z}_{t-1} - \theta^2 a_{t-2}\\
\tilde{z}_t =& a_t - \theta\tilde{z}_{t-1} - \theta^2(\tilde{z}_{t-2} + \theta a_{t-3}) = a_t - \theta\tilde{z}_{t-1} - \theta^2\tilde{z}_{t-2} - \theta^3 a_{t-3}\\
&\vdots\\
\tilde{z}_t =& -\theta\tilde{z}_{t-1} - \theta^2\tilde{z}_{t-2} - \cdots - \theta^{k-1}\tilde{z}_{t-k+1} + a_t - \theta^k a_{t-k}
\end{aligned} \tag{8.11}$$

   en donde el término $\theta^k a_{t-k}$ tenderá a cero cuando $k \to \infty$ si $|\theta| < 1$.

2. Inversión del polinomio de medias móviles:

$$\tilde{z}_t = (1-\theta B)a_t \quad \Rightarrow \quad \frac{1}{1-\theta B}\tilde{z}_t = a_t \quad \Rightarrow \quad \pi(B)z_t = a_t$$

   en donde $\pi(B)$ es un polinomio en B de orden infinito

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$$

   cuyos coeficientes pueden encontrarse de la relación $\pi(B)(1-\theta B) = 1$. El polinomio producto

$$\begin{aligned}
\pi(B)(1-\theta B) =& 1 - \pi_1 B - \pi_2 B^2 - \cdots - \theta B + \theta\pi_1 B^2 + \theta\pi_2 B^2 \dots\\
=& 1 - (\pi_1 + \theta)B - (\pi_2 + \theta\pi_1)B^2 - \cdots - (\pi_j + \theta\pi_{j-1})B^{j-1} + \dots
\end{aligned}$$

   será igual a 1 si sus coeficientes $\pi_j + \theta\pi_{j-1}$ son nulos. De aquí, encontramos que los pesos $\pi_j = -\theta^j$.

□

Proposición 79. *Un proceso $MA(1)$ siempre es estacionario.*

Demostración. Se cumple que $\sum_{j=1}^{\infty} |\psi_j| = |\theta| < \infty$. □

Proposición 80. *Un proceso $MA(1)$ es invertible si $-1 < \theta < 1$.*

Demostración. Se cumplirá que $\sum_{j=1}^{\infty} |\pi_j| = \sum_{j=1}^{\infty} |\theta^j| < \infty$ cuando $|\theta| < 1$. □

Proposición 81. *La función de autocovarianzas de un proceso $MA(1)$ es*

$$\gamma_k \begin{cases} (1-\theta^2)\sigma_a^2 & k = 0\\ -\theta\sigma_a^2 & k = 1\\ 0 & k > 1 \end{cases}$$

Demostración. Es claro que

$$E(\tilde{z}_t\tilde{z}_{t-k}) = E[(a_t - \theta a_{t-1})(a_{t-k} - \theta a_{t-k-1})]$$

□

Proposición 82. *La función de autocorrelación simple de un proceso $MA(1)$ es*

$$\rho_k \begin{cases} -\dfrac{\theta}{1-\theta^2} & k = 1\\ 0 & k > 1 \end{cases}$$

Observación 51. Se dice que la memoria del proceso $MA(1)$ es de un periodo porque $\rho_k = 0$ para $k > 1$.

Observación 52. *Si* $|\theta| < 1$, *entonces* $|\rho_1| < 0{,}5$.

PROPOSICIÓN 83. *La función de autocorrelación parcial de un proceso* $MA(1)$ *es*

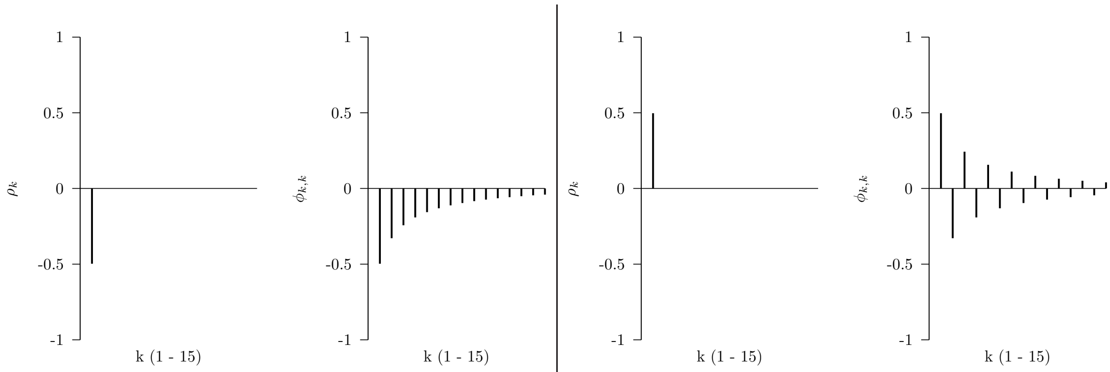
$$\phi_{kk} = \frac{-\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \quad \text{para} \quad k > 0$$

DEMOSTRACIÓN. Resolviendo las ecuaciones de Yule-Walker para distintos retardos, obtenemos

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = -\frac{\rho_1^2}{1-\rho_1^2}, \quad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1^3}{1-2\rho_1^2}, \ldots$$

□

El cuadro (2) muestra las funciones de autocorrelación simple y parcial para dos modelos MA(1). Cuando el parámetro MA es positivo, la ACF tiene un coeficiente negativo en el primer retardo y la PACF se amortigua alternando en signo. Por el contrario, cuando el parámetro MA es negativo, la ACF tiene un coeficiente positivo en el primer retardo, y la PACF decrece exponencialmente por debajo de cero.



(a) $\theta = 0{,}9$ (b) $\theta = -0{,}9$

Figura 2: Funciones de autocorrelación simple y parcial para dos procesos $MA(1)$

Los resultados anteriores se extienden fácilmente al proceso de medias móviles estacional de primer orden, denotado por $MA(1)_s$,

$$z_t = \delta + a_t - \theta a_{t-s} \quad \Leftrightarrow \quad z_t = \delta + (1-\theta B^s)a_t$$

8.6. Procesos no estacionarios

Las series temporales que observamos en economía suelen ser no estacionarias en media. Por ejemplo, en el gráfico temporal (3) de la serie mensual Indices de Precios Industriales en España vemos que la media local (la media de un subconjunto de observaciones) aumenta en el tiempo. Las series temporales con estas características se denominan series no estacionarias y no pueden ser descritas directamente mediante procesos estacionarios.

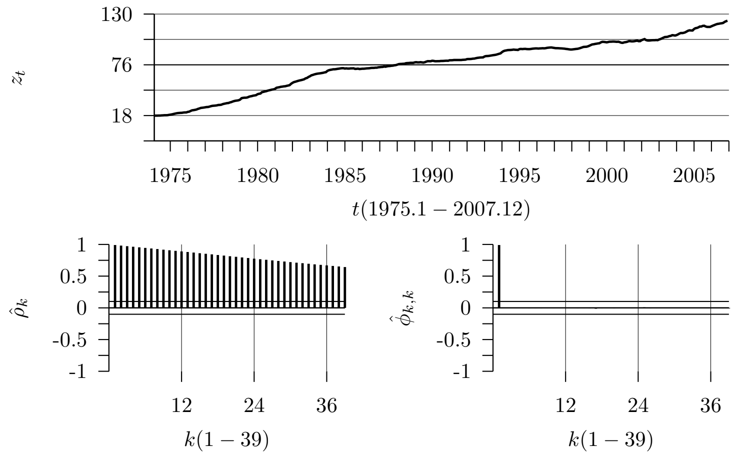


Figura 3: Gráfico temporal y funciones de autocorrelación para la serie mensual Indices de Precios Industriales

El modelo de regresión con tendencia lineal

$$y_t = \beta_0 + \beta_1 t + u_t, \quad t = 1, \dots, n$$

es un candidato razonable para describir series que fluctúan alrededor de una tendencia lineal. El modelo puede ampliarse, en caso necesario, especificando un proceso ARMA para el término de error u_t. Por ejemplo, suponiendo que $u_t = \phi u_{t-1} + a_t$, en donde a_t es un proceso de ruido blanco.

Una aproximación alternativa consiste en ajustar un proceso autorregresivo de primer orden

$$(1 - \phi B) y_t = \delta + u_t$$

en donde u_t es un proceso de ruido blanco. Esta especificación puede justificarse por la forma de las funciones de autocorrelación simple y parcial mostradas en el gráfico (3). Cuando $\phi = 1$, el proceso y_t no es estacionario. Sin embargo, su primera diferencia sí es estacionaria, $E(y_t - y_{t-1}) = \delta$.

Definición 71. *La diferencia primera de un proceso y_t es $y_t - y_{t-1}$ o $(1 - B)y_t$.*

Definición 72. *La diferencia segunda de un proceso y_t es la diferencia primera de la diferencia primera $(y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$ o $(1 - B)^2 y_t$.*

Definición 73. *La diferencia de orden d de un proceso y_t es $(1 - B)^d y_t$ o $\nabla^d y_t$, en donde $\nabla = 1 - B$ se denomina operador diferencia.*

Muchas series temporales no estacionarias en media pueden transformarse en estacionarias o bien ajustando polinomios de tendencias o bien tomando sucesivas diferencias. La serie resultante puede ser entonces descrita por un modelo ARMA.

Definición 74. *El modelo de regresión con tendencia determinista y autocorrelación es*

$$\begin{aligned} y_t =& \beta_0 + \beta_1 t + \cdots + \beta_r t^r + u_t \\ u_t =& \phi_1 u_{t-1} + \cdots + \phi_p u_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - a_{t-q} \end{aligned}$$

Definición 75. *Un proceso y_t es integrado de orden d si al diferenciarlo d veces obtenemos un proceso $z_t = (1 - B)^d y_t$ estacionario.*

Definición 76. *El proceso ARIMA(p,d,q) es*

$$(1-\phi_1 B - \cdots - \phi_p B^p)(1-B)^d y_t = \delta + (1-\theta_1 B - \cdots - \theta_q B^q)a_t$$

Definición 77. *El proceso $ARIMA(0,1,0)$ sin término constante*

$$(1-B)y_t = a_t$$

se denomina paseo aleatorio, también conocido como paseo del borracho. Si el modelo incluye término constante, se denomina paseo aleatorio con deriva.

Observación 53. *Al diferenciar una serie cuando no es necesario obtenemos modelos MA no invertibles. Por ejemplo, si una serie temporal ha sido generada por un modelo de tendencia lineal, la diferencia primera conduce a un proceso $MA(1)$ no invertible. Para verlo, escribimos el modelo de regresión en dos instantes consecutivos*

$$\begin{aligned} y_t =& \beta_0 + \beta_1 t + u_t \\ y_{t-1} =& \beta_0 + \beta_1 (t-1) + u_{t-1} \end{aligned}$$

y restando obtenemos

$$y_t - y_{t-1} = \beta_1 + u_t - u_{t-1}, \quad t = 2, \ldots, n$$

en donde ∇y_t es un proceso $MA(1)$ no invertible ($\theta = 1$) si u_t es un proceso de ruido blanco.

Definición 78. *La diferencia estacional $z_t - z_{t-s}$ o $(1-B^s)z_t$ se emplea para eliminar la estacionalidad.*

En el análisis de series temporales reales la decisión de si una serie es estacionaria o no estacionaria puede basarse en la inspección de su gráfico temporal y en la función de autocorrelación simple. Las series no estacionarias tienen medias locales inestables y la función de autocorrelación decrece muy lentamente. Un procedimiento estadístico más formal es el contraste de Dickey-Fuller

Definición 79. *En el modelo $y_t = \phi y_{t-1} + u_t$, se rechaza la hipótesis nula de no estacionariedad $H_0 : \phi = 1$ frente a la alternativa de estacionariedad $H_1 : \phi < 1$ cuando*

$$DF = \frac{\hat{\phi} - 1}{se(\hat{\phi})} < c_\alpha$$

en donde $\hat{\phi} = \sum_{t=2}^n y_t y_{t-1} / \sum_{t=2}^n y_{t-1}^2$ es el estimador de mínimos cuadrados de ϕ, $se(\hat{\phi})$ es la raíz de $\hat{V}(\hat{\phi}) = \hat{\sigma}_u^2 / \sum_{t=2}^n y_{t-1}^2$ y c_α es el valor crítico para el nivel de significación α en una distribución no estándar tabulada por Dickey y Fuller.

Algunas series económicas, además de ser no estacionarias en media, tienen varianzas locales inestables. La no estacionariedad en varianza puede corregirse tomando logaritmos, que es un tipo de transformación Box-Cox.

Definición 80. *La transformación potencia Box-Cox es una familia de transformaciones que inducen linealidad, homocedastidad (varianza estable) y normalidad*

$$y_t^{(\lambda)} = \frac{z_t^\lambda - 1}{\lambda} = \begin{cases} y_t & \lambda = 1 \\ \sqrt{y_t} & \lambda = 0{,}5 \\ \ln(y_t) & \lambda = 0 \\ \dfrac{1}{\sqrt{y_t}} & \lambda = -0{,}5 \\ \dfrac{1}{y_t} & \lambda = -1{,}0 \end{cases}$$

8.7. Predicción con modelos ARIMA

La predicción lineal general del valor futuro z_{n+h} desde el origen n es una combinación lineal de las observaciones pasadas z_n, z_{n-1}, z_{n-2}, ...

$$\hat{z}_n(h) = \alpha_0 z_n + \alpha_1 z_{n-1} + \alpha_2 z_{n-2} + \dots$$

y, mediante sustituciones sucesivas, puede expresarse en términos de los errores pasados

$$\hat{z}_n(h) = \beta_0 a_n + \beta_1 a_{n-1} + \beta_2 a_{n-2} + \dots$$

Nos interesa elegir los pesos α_i o β_i de manera que el error de predicción $e_n(h) = z_{n+h} - \hat{z}_n(h)$ tenga error cuadrático medio mínimo.

Proposición 84. *La predicción de error cuadrático medio mínimo de z_{n+h} en el origen n y a horizonte h es la esperanza de z_{n+h} condicionada a todas las observaciones disponibles hasta el origen n*

$$\hat{z}_t(h) = E(z_{n+h}|z_n, z_{n-1}, z_{n-2}, \dots) = \psi_h a_n + \psi_{h+1} a_{n-1} + \dots$$

Demostración. Si el modelo es estable, el valor futuro z_{n+h} vendrá generado por

$$z_{n+h} = a_{n+h} + \psi_1 a_{n+h-1} + \dots + \psi_h a_n + \psi_{h+1} a_{n-1} + \dots$$

El error de predicción

$$e_n(h) = z_{n+h} - \hat{z}_n(h) = a_{n+h} + \psi_1 a_{n+h-1} + \dots + (\psi_h - \beta_0) a_n + (\psi_{h+1} - \beta_1) a_{n-1} + \dots$$

es insesgado y tiene varianza

$$V(e_n(h)) = \sigma_a^2 (1 + \psi_1^2 + \dots + \psi_{h-1}) + \sigma_a^2 [(\psi_h - \beta_0)^2 + (\psi_{h+1} - \beta_1)^2 + \dots]$$

La varianza, error cuadrático medio $E(z_{n+h} - \hat{z}_n(h))$, será mínima cuando $\beta_i = \psi_{h+i}$ ($i = 0, 1, \dots$). De aquí, la predicción de error cuadrático medio mínimo

$$\hat{z}_n(h) = \psi_h a_n + \psi_{h+1} a_{n-1} + \psi_{h+2} a_{n-2} + \dots$$

es la esperanza de z_{n+h} condicionada a todas las observaciones pasadas. □

Observación 54. *En la predicción con modelos ARIMA hacemos uso de los siguientes resultados:*

1. $E(z_t|z_n, z_{n-1}, \dots) = \hat{z}_n(t-n)$ *cuando* $t > n$.
2. $E(z_t|z_n, z_{n-1}, \dots) = z_t$ *cuando* $t \leq n$
3. $E(a_t|z_n, z_{n-1}, \dots) = 0$ *cuando* $t > n$.

4. $E(a_t|z_n, z_{n-1}, \dots) = a_t$ *cuando* $t \leq n$

Ejemplo 16. *Predicción con el modelo* $AR(1)$.

Predicción: $\hat{z}_n(h) = \phi\hat{z}_n(h-1)$ *con* $z_n(1) = z_n$.
Error de predicción: $e_n(h) = a_{n+h} + \phi a_{n+h-1} + \cdots + \phi^{h-1} a_{n+1}$
Varianza: $V(e_n(h) = \sigma^2(1 + \phi^2 + \cdots + \phi^{2(h-1)})$
Predicción por intervalo: $\hat{z}_n(h) \pm c \times \sqrt{\sigma_a^2(1 + \phi^2 + \cdots + \phi^{2(h-1)})}$ *en donde* c *es el valor crítico tal que* $Prob(|N(0,1)| < c) = 1 - \alpha$.

Ejemplo 17. *Predicción con el modelo* $MA(1)$.

Predicción: $\hat{z}_n(1) = -\theta a_n$ *y* $\hat{z}_n(h) = 0$ *para* $h > 1$.
Error de predicción: $e_n(1) = a_{n+1}$ *y* $e_n(h) = a_{n+h} + \theta a_{n+h-1}$ *para* $h > 1$.
Varianza: $V(e_n(1) = \sigma^2$ *y* $V(e_n(h)) = (1 + \theta^2)$ *para* $h > 1$.
Predicción por intervalo: $\hat{z}_n(1) \pm c \times \sqrt{\sigma^2}_a$ *y* $\hat{z}_n(h) \pm c \times \sqrt{\sigma_a^2(1 + \theta^2)}$ *para* $h > 1$.

8.8. Resumen

1. Una serie temporal es una realización particular de proceso estocástico.
2. El proceso estocástico lineal general expresa cada observación de una serie temporal como una combinación lineal de las observaciones pasadas.
3. Un proceso es débilmente estacionario si sus dos primeros momentos existen y son estables.
4. Los procesos ARMA son una aproximación al proceso lineal general que incluye un número finito de parámetros.
5. Los procesos ARMA están caracterizados por las funciones de autocorrelación simple y parcial, que nos permiten distinguir unos procesos de otros.
6. Las series temporales no estacionarias en media puede convertirse en estacionarias extrayendo tendencias deterministas o diferenciando.
7. Las series temporales no estacionarias en varianza pueden convertirse en estacionarias tomando logaritmos o cualquier otra transformación Box-Cox.
8. La dinámica de los modelos ARIMA es conveniente para calcular predicciones de forma recursiva.

Palabras clave

Proceso estocástico
Serie temporal
Proceso lineal general
Estacionariedad
Invertibilidad
Función de autocorrelación
Procesos integrados
Procesos ARIMA
Predicción

8.9. Ejercicios

1. Suponga que $z_t = \phi_1 z_{t-1} + u_t$ y $u_t = \phi_2 u_{t-1} + a_t$, en donde a_t es un proceso de ruido blanco. Demuestre que z_t sigue un proceso $AR(2)$.
2. Suponga que $z_t = u_t - \theta_1 u_{t-1}$ y $u_t = a_t - \theta_2 a_{t-1}$, en donde a_t es un proceso de ruido blanco. Demuestre que z_t sigue un proceso $MA(2)$.
3. Suponga que $z_t = \phi_1 z_{t-1} + u_t$ y $u_t = a_t - \theta_2 a_{t-1}$, en donde a_t es un proceso de ruido blanco. Demuestre que z_t sigue un proceso $ARMA(1,1)$.

4. Explique detalladamente las propiedades de un proceso $ARMA(1,1)$ y su uso en predicción.

8.10. Ejercicios resueltos

1. Escribimos los dos modelos en notación retardos

$$(1-\phi_1 B)z_t = u_t$$
$$(1-\phi_2 B)u_t = a_t$$

Premultiplicando la primera ecuación por $(1-\phi_2 B)$ obtenemos

$$(1-\phi_2 B)(1-\phi_1 B)z_t = (1-\phi_2 B)u_t = a_t$$

que podemos escribir como

$$(1-\delta_1 B+\delta_2 B^2)z_t = a_t$$

en donde $\delta_1 = \phi_1 + \phi_2$ y $\delta_2 = -\phi_1\phi_2$.