ASPECTOS TEORICOS DE SERIES DE TIEMPO ESTACIONARIAS

JESUS VIRGILIO ALEMAN VALERIO

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS
EN ESTADISTICA EXPERIMENTAL

Universidad Autonoma Agraria
Antonio Narro
PROGRAMA DE GRADUADOS
Buenavista, Saltillo, Coah.
DICIEMBRE DE 1992

Tesis elaborada bajo la supervisión del comité particular de asesoría y aprobada como requisito parcial para optar al grado de

MAESTRO EN CIENCIAS EN

ESTADISTICA EXPERIMENTAL

COMITE PARTICULAR

Asesor principal: ______________________

Dr. Rolando Cavazos Cadena

Asesor: ______________________

M.C. Félix de Jesús Sánchez Pérez

Asesor: ______________________

M.C. Mario Cantú Sifuentes

____

Dr. José Manuel Fernández Brondo

Subdirector de Asuntos de Postgrado

Buenavista, Saltillo, Coah., diciembre 1992

AGRADECIMIENTOS

Agradezco sinceramente al Dr. Rolando Cavazos Cadena por la ayuda y orientación brindada incondicionalmente durante mis estudios y especialmente durante la elaboración de este trabajo.

a mis padres

a mis hermanos

a mis maestros

a mis amigos

y a todas aquellas personas, que de manera directa o indirecta contribuyeron en la realización de este trabajo.

COMPENDIO

ASPECTOS TEORICOS DE SERIES DE TIEMPO ESTACIONARIAS

POR

JESUS VIRGILIO ALEMAN VALERIO

MAESTRIA

ESTADISTICA EXPERIMENTAL

UNIVERSIDAD AUTONOMÁ AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA. DICIEMBRE 1992

Dr. Rolando Cavazos Cadena -Asesor-

Palabras Clave: Función de autocovarianza, polinomio causal, causalidad; filtro, inversibilidad,solución única de sistema de ecuaciones de Yule-Walker (Y-W), procesos AR(p), MA(q) y ARMA(p,q).

Este trabajo, trata sobre algunos aspectos relevantes de la teoría de series de tiempo univariadas. Aunque la exposición gira alrededor de nociones clásicas, también se

abordan resultados recientes, los cuales se relacionan, principalmente, con el sistema de ecuaciones de Yule-Walker (Y-W). Dicho sistema surge de manera natural al cosiderar el problema de determinar la función de autocovarianza de un proceso autorregresivo, o más generalmente, de un proceso ARMA. En esta sección, se estudia el problema de la unicidad de las soluciones del sistema (Y-W), y se presenta el siguiente resultado: El sistema de Yule-Walker asociado a un polinomio causal posee solución **única.** Además se estudia el problema de diseño de filtros lineales, proporcionándose las ecuaciones que los coeficientes de un filtro deben satisfacer para dejar invariante a una función de tendencia polinomial.

ABSTRACT

THEORIC ASPECTS OF STATIONARY TIME SERIES

BY

JESUS VIRGILIO ALEMAN VALERIO

MASTER OF SCIENCE

EXPERIMENTAL STATISTIC

UNIVERSIDAD AUTONOMA AGRARIA ANTONIO NARRO

BUENAVISTA SALTILLO COAHUILA. MEXICO. DECEMBER 1992

Dr Rolando Cavazos Cadena -Advisor-

Key words: Autocovariance function, causal polynomial, causality; filter, inversibility, unique solution of Yule-Walker equations system (Y-W), AR(p), MA(q) and ARMA(p,q) proceses.

This work concerns basic notions in the theory of univariate time series. Our exposition referes, mainly, to classical notions, but we also present some recent development concerning the Yule-Walker system of linear equations, which asise when considering the problem of finding the autocovariance function of an autorregresive proceses or, more generally, of an ARMA processes. At this point, our main results can be summarized as follow: The Yule-Walker equations associated to causal polynomial have a unique solution. On the other hand, we also study the problem of designing a linear filter $\mathcal{L}$, and we give the system equations that the coefficients of the filter must satisfy so that a given polynomial function is invariant under action of $\mathcal{L}$.

INDICE DE CONTENIDO

I INTRODUCCION

Las series de tiempo constituyen, sin duda alguna, una herramienta de gran utilidad en la teoría y práctica de la estadística. El presente trabajo, trata sobre algunos aspectos teóricos de series de tiempo univariadas, poniendo especial énfasis en la unicidad de la solución del sistema de ecuaciones de Yule-Walker.

Se inicia el trabajo con el capítulo II, en el que se proporciona un breve resumen sobre **procesos estocásticos**, ya que constituyen el marco de referencia para el tratamiento de las series de tiempo que se considerarán en este trabajo. A continuación, se introduce la noción de serie estacionaria y se presentan varios ejemplos, incluyendo el bien conocido proceso de "ruido blanco" que aparecerá repetidamente en nuestra discusión general. posteriormente, su función de autocovarianza y de autocorrelación, las cuales, para el caso de procesos estacionarios, sólo dependen de la diferencia que exista entre dos tiempos dados; dichas funciones constituyen la herramienta básica en el análisis de una serie de tiempo. En otra parte de este capítulo se introduce el concepto de "filtro" el cual es una transformación de un proceso estocástico a otro mediante ciertas reglas que pueden expresarse generalmente a través de fórmulas recursivas; se considera, en particular, la clase de filtros sumables, que tienen la propiedad de transformar un proceso estacionario en otro de la misma clase. Como ejemplo, se obtiene el modelo de medias móviles que se estudia brevemente en el capítulo III

Con el propósito de aumentar la clase de modelos que se utilizarán en la descripción de fenómenos aleatorios que requiera de las series de tiempo, en el capítulo IV, se habla con más detalle de los procesos autorregresivos. La exposisión se inicia con la presentación de la ecuación del modelo, la cual depende de "p" parámetros de ponderaciones para las observaciones pasadas, así como de un parámetro de dispersión para las perturbaciones.

A continuación consideramos el problema de determinar la función de autocovarianza de un proceso autorregresivo asociado a un polinomio causal $\varphi(z)$. En este punto surge de manera natural el sistema de ecuaciones de Yule-Walker y consideramos a profundidad el problema de establecer la unicidad de las soluciones de dicho sistema.

Habiendo considerado los procesos autorregresivos y de medias móviles, se procede en el capítulo V al estudio de series de tiempo que pueden ser modeladas utilizando una combinación tanto autorregresiva como de medias móviles. Tales procesos constituyen el tipo mixto y permiten ampliar de manera considerable la clase de modelos para series de tiempo estacionarias. Se presenta la ecuación para tales modelos, incluyendo "p" parámetros que constituyen las ponderaciones en la parte autorregresiva, "q" parámetros para las correspondientes ponderaciones en la parte de medias móviles.

II CONCEPTOS BASICOS

1. INTRODUCCION

Desde un punto de vista práctico, una serie de tiempo consiste de un conjunto de observaciones x_t registradas en ciertos "instantes" $t \in T$, donde T es un conjunto dado, usualmente contenido en $\mathbb{R}$. El análisis de la serie tiene como objetivo principal la construcción de un modelo que describa adecuadamente el comportamiento observado de $\{x_t\} \equiv \{x_t | \ t \in T\}$ para, posteriormente, realizar las tareas de pronóstico y control. Con este propósito seguiremos el enfoque usual, el cual postula que cada valor observado x_t es el valor tomado por una variable aleatoria X_t, y que x_t se conoce en el instante $t \in T$. Esto conduce de manera natural a considerar familias de variables aleatorias, las cuales se denominan procesos estocásticos o series de tiempo.

2.2 SERIES DE TIEMPO

Se inicia esta sección con una definición formal de serie de tiempo.

Definición 2.1. Una serie de tiempo o proceso estocástico con conjunto de índices (o tiempos) T, es una colección

$$\{X_t | t \in T \} \qquad (2.1)$$

de variables aleatorias definidas en un mismo espacio de probabilidad $(\Omega, \mathcal{F}, P)$.

Si el conjunto de índices T está contenido en la clase de los enteros $\mathbb{Z} := \{\cdots -2, -1, 0, 1, 2, \cdots\}$, el proceso

$$\{X_t\} \equiv \{X_t | t \in T\}$$

se dice que se desarrolla en *tiempo discreto*, mientras que si T es un intervalo de $\mathbb{R}$ el proceso se realiza en *tiempo continuo*.

A cada proceso estocástico $\{X_t\}$ como en (2.1) le corresponde una distribución probabilística; ésta queda especificada por la familia de distribuciones conjuntas correspondientes a cada subcolección finita

$$(X_{t(1)}, X_{t(2)}, \cdots, X_{t(n)}),$$

donde $t(1), \cdots, t(n) \in T$ son arbitrarios. Esto significa que para determinar el comportamiento (conjunto) de $\{X_t\}$ es necesario encontrar

(a) la distribución de cada variable X_t, $t \in T$;

(b) la distribución de cada pareja de variables aleatorias

$$(X_t, X_s), \ s, t \in T,$$

y en general, debe determinarse

(c) la distribución de cada n-ada

$$(X_{t(1)}, X_{t(2)}, \cdots, X_{t(n)})$$

para cualquier $t(1),\cdots,t(n) \in T$

Por otro lado, usualmente no se conoce la distribución de una serie de interés $\{X_t\}$, y entonces sus valores observados $\{x_t\}$ deben utilizarse para estimar la distribución de la serie generadora $\{X_t\}$. Como es posible imaginar a partir de los comentarios anteriores, *en general* este es un trabajo formidable. Así, en un problema práctico, es necesario tratar de obtener un modelo simple que permita realizar la tarea de estimación de manera relativamente sencilla. En los próximos capítulos estudiaremos algunos de los modelos que se han utilizado con más éxito en las aplicaciones.

A continuación introducimos dos importantes funciones asociadas con una serie de tiempo $\{X_t\}$.

Definición 2.2. Suponga que la serie $\{X_t\}$ es tal que

$$E[X_t^2] < \infty,\ t \in T.$$

En esta situación,

(*i*) la *función de medias* $m:T \longrightarrow \mathbb{R}$ se define por

$$m(t) := E[X_t],\ t \in T,$$

mientras que

(*ii*) la *función de autocovarianza* $G:T \times T \longrightarrow \mathbb{R}$ está dada por

$$G(s,t) := \mathrm{Cov}[X_s, X_t] = E[(X_s - m(s))(X_t - m(t))],\ s,t \in T.$$

Las funciones que se acaban de definir son extremadamente importantes. El valor $m(t)$ especifica el valor esperado de X_t, mientras que $G(\cdot,\cdot)$ determina la asociación lineal entre las variables X_s y X_t; vea, por ejemplo, Dudewicz y Mishtra (1988), Mood y Graybill (1971) o Cavazos Cadena (1990).

La matriz de covarianza $\Sigma(t(1),t(2),\cdots, t(N)) \equiv \Sigma$

del vector

$$(X_{t(1)}, X_{t(2)},\cdots,X_{t(N)})$$

puede expresarse en términos de la función de autocovarianza G en la Definición 2.2 como sigue:

$\Sigma =$

$$\begin{bmatrix} G(t(1),t(1)) & G(t(1),t(2))\cdots & G(t(1),t(N)) \\ G(t(2),t(1)) & G(t(2),t(2))\cdots & G(t(2),t(N)) \\ \vdots & \vdots & \vdots \\ G(t(N),t(1)) & G(t(N),t(2))\cdots & G(t(N),t(N)) \end{bmatrix}. \qquad (2.2)$$

Es claro que esta matriz es simétrica. Por otro lado Σ es también *no negativa definida*. Para ver esto, seleccione α

$$\begin{aligned} \alpha'\Sigma\alpha &= \sum_{i,j=1} a_i G(t(i),t(j))\, a_j \\ &= \sum_{i,j=1}^{N} a_i \mathrm{Cov}[X_{t(i)},X_{t(j)}] a_j \\ &= \mathrm{Cov}[\sum_{i=1}^{N} a_i X_{t(i)}, \sum_{j=1}^{N} a_j X_{t(j)}] \\ &= \mathrm{Var}\, [\sum_{i=1}^{N} a_i X_{t(i)}] \geq 0. \end{aligned}$$

Puesto que $\alpha \in \mathbb{R}^N$ es arbitrario, los cálculos anteriores muestran que Σ es, efectivamente, una matriz no negativa definida.

2.3 SERIES ESTACIONARIAS

Como se mencionó en la sección anterior, el problema de estimar las funciones de medias y de autocovarianza es de gran importancia. Sin embargo, para poder abordar este problema, es necesario restringir la clase de series de tiempo bajo consideración. Para entender esto, suponga que se desea estimar $m(1)$ (= $E[X_1]$), esto es, la media de la distribución de X_1, y que se dispone de los valores observados de $X_i = x_i$ para ciertos índices i contenidos en T. Note ahora que, en general, entre estas cantidades sólo el valor $X_1 = x_1$ proviene de una variable con la distribución de interés (la cual es la distribución de X_1). Así, aunque se tienen los valores de X_i para varios índices i, a lo más uno de estos es útil en lo que al problema de estimar $m(1)$ se refiere, y entonces no es posible pretender obtener algún estimador "razonablemente bueno" de $m(1)$. La situación cambiaría si todas las variables aleatorias X_i tuvieran la misma distribución, pues en este caso, todas los valores observados x_i corresponderían a variables aleatorias con la misma distribución de interés. Similarmente, si se desea estimar

$G(1,2) = Cov[X_1, X_2]$ (= covarianza de la distribución del vector (X_1,X_2)), sólo la pareja (X_1,X_2) tiene la distribución de interés, y por lo tanto tampoco será posible obtener un buen estimador de $G(1,2)$; desde luego, la situación sería diferente si otras parejas (X_i,X_j) tuvieran la misma distribución de (X_1,X_2). A continuación, se introduce una clase de procesos para las cuales, en principio, es posible obtener buenos estimadores de $m(\cdot)$ y $G(\cdot,\cdot)$.

Definición 2.3. Una serie de tiempo $\{X_t | t \in T\}$ (con $T \subset \mathbb{Z}$) es *estrictamente estacionaria* si, para toda sucesión $t(1), t(2), \cdots, t(n) \in T$ y cada *desplazamiento* $h \in \mathbb{Z}$,

$$\left.\begin{array}{l} \textit{los vectores} \\ \qquad\qquad (X_{t(1)}, X_{t(2)}, \cdots, X_{t(n)}) \\ \qquad\qquad\qquad y \\ \qquad\qquad (X_{t(1)+h}, X_{t(2)+h}, \cdots, X_{t(n)+h}) \\ \text{tienen } \textit{la misma distribución.} \end{array}\right\} \quad (2.4)$$

Observación 2.1. No es difícil ver que $\{X_t | t \in \mathbb{Z}\}$ es una serie estrictamente estacionaria si y sólo si lo siguiente ocurre: Para todo $N = 1,2,3,\cdots$ y $h \in \mathbb{Z}$,

los vectores $(X_1, X_2, \cdots, X_N)$ y $(X_{1+h}, X_{2+h}, \cdots, X_{N+h})$

tienen la misma distribución; vea por ejemplo, Brockwell y

Davis (1987).

Suponga ahora que $\{X_t\}$ es una serie estrictamente estacionaria con $E[X_1^2] < \infty$. En este caso, para todo $t \in \mathbb{Z}$, X_t y X_1 tienen la misma distribución; esto se desprende de (2.4) con $n = 1$, $t(1) = t$ y $h = t - 1$. De este modo, la función de medias $m(\cdot)$ en la Definición 2.2 se reduce a una constante,

$$m(t) = m(1) \equiv \mu, \quad t \in \mathbb{Z}, \tag{2.5}$$

y si se desea estimar μ, se puede disponer de todos los valores observados x_i para obtener un buen estimador. Por otro lado, la pareja (X_r, X_s) tiene la misma distribución que (X_1, X_{1+s-r}), lo cual se sigue usando (2.4) con $n = 2$, $t(1) = 1$, $t(2) = 1 + s - r$, y $h = r-1$. Esto implica que

$$G(r,s) = \text{Cov}\ [X_r, X_s] = \text{Cov}\ [X_1,\ X_{1+s+r}]$$

esto es,

$$G(r,s) = G(1,\ 1 + s-r). \tag{2.6}$$

A partir de esta igualdad se concluye que Cov $[X_r, X_s]$ depende sólo de la diferencia s - r. Esto permite describir la estructura de covarianzas de una serie estacionaria mediante una función más simple que $G(\cdot,\cdot)$.

Definición 2.4. Suponga que la serie $\{X_t | t \in \mathbb{Z}\}$ es estrictamente estacionaria. La función $\gamma : \mathbb{Z} \longrightarrow \mathbb{R}$ se define mediante

$$\gamma(h) := \mathrm{Cov}[X_1, X_{1+h}], \quad h \in \mathbb{Z}. \tag{2.7}$$

Observe ahora que (2.5) y (2.7) combinadas implican que

$$\begin{aligned} G(r,s) &= \mathrm{Cov}\ [X_r, X_s] \\ &= \mathrm{Cov}\ [X_1, X_{1+s-r}] = \gamma(r-s),\ r,s, \in \mathbb{Z}, \end{aligned} \tag{2.8}$$

de manera que la estructura de covarianzas de $\{X_t\}$ está completamente contenida en la función $\gamma(\cdot)$ recién definida en (2.7). Por esta razón llamaremos a $\gamma(\cdot)$ *la* función de autocovarianza de la serie $\{X_t\}$. Este nombre también fue aplicado a la función $G(\cdot,\cdot)$ en la Definición 2.2, pero, debido a (2.7), esta duplicidad no debe causar confusión alguna.

A continuación se ven algunos ejemplos de series estrictamente estacionarias.

Ejemplo 2.1. Suponiendo que $\{Z_t|\ t \in \mathbb{Z}\}$ es una sucesión de *variables aleatorias independientes e idénticamente distribuidas con media* 0 *y varianza* $\sigma^2 > 0$, lo cual se abreviará escribiendo

$$\{Z_t\} \sim \mathrm{IID}(0, \sigma^2). \tag{2.9}$$

En este caso $\{Z_t\}$ es una serie estrictamente estacionaria. Para ver esto, se selecciona un entero positivo N y $h \in \mathbb{Z}$. Entonces, para intervalos arbitrarios $J_1, J_2, \cdots, J_N \subset \mathbb{R}$

$$P[Z_1 \in J_1, Z_2 \in J_2, \cdots, Z_N \in J_N]$$

$$= P[Z_1 \in J_1]\cdot P[Z_2 \in J_2]\cdots P[Z_N \in J_N]$$

$$= P[Z_1 \in J_1]\cdot P[Z_1 \in J_2]\cdots P[Z_1 \in J_N]$$

donde las igualdades son consecuencia inmediata de (2.9). Similarmente,

$$P[Z_{1+h} \in J_1, Z_{2+h} \in J_2, \cdots, Z_{N+h} \in J_N]$$

$$= P[Z_{1+h} \in J_1]\cdot P[Z_{2+h} \in J_2]\cdots P[Z_{N+h} \in J_N]$$

$$= P[Z_1 \in J_1]\cdot P[Z_1 \in J_2]\cdots P[Z_1 \in J_N],$$

donde se han usado el supuesto que las Z_is tienen la misma distribución para obtener la segunda igualdad. Por lo tanto,

$$P[Z_1 \in J_1, Z_2 \in J_2, \cdots, Z_N \in J_N]$$

$$= P[Z_{1+h} \in J_1, Z_{2+h} \in J_2, \cdots, Z_{N+h} \in J_N],$$

y como J_1, $J_2, \cdots, J_N$ son intervalos arbitrarios, se concluye que $(Z_1, Z_2, \cdots, Z_N)$ y $(Z_{1+h}, Z_{2+h}, \cdots, Z_{N+h})$ tienen la misma distribución, y donde se desprende que $\{Z_t\}$ es una serie estrictamente estacionaria; vea la Observación 2.1.

Por otro lado, es claro que $m(t) = E[Z_t] = 0$, $t \in \mathbb{Z}$, y que $\gamma(0) = E[Z_t^2] = \sigma^2$(por (2.9)). Además, debido a que Z_1 y Z_{1+h} son independientes para $h \neq 0$, se tiene que $\gamma(h) = \mathrm{cov}[Z_1, Z_{1+h}] = 0$.

En resumen: Si $\{Z_t\} \sim \mathrm{IID}(0, \sigma\ )$, entonces

$$m(t) = 0, \quad t \in \mathbb{Z},$$

$$\gamma(0) = \sigma^2 \quad \text{y} \quad \gamma(h) = 0 \text{ para } h \neq 0. \qquad \square$$

Ejemplo 2.2 En este ejemplo se verá cómo a partir de una serie estrictamente estacionaria pueden construirse otras.
(i) Suponga ahora que $\{W_t\}$ es estrictamente estacionaria, y sea g: $\mathbb{R}^k \longrightarrow \mathbb{R}$ una función arbitraria. Se define una nueva serie $\{X_t\}$ mediante

$$X_t = g(W_t,\ W_{t+1},\ \cdots, W_{t+k-1}),\quad t \in \mathbb{Z}.$$

En este caso $\{X_t\}$ es una serie estrictamente estacionaria. Se demostrará esto en el caso en que $k = 2$, ya que el caso general se establece de manera similar, salvo que la notación se complica excesivamente. Como en el ejemplo anterior, Se selecciona un entero positivo N y $h \in \mathbb{Z}$. Entonces, para intervalos arbitrarios $J_1, J_2, \cdots, J_N \subset \mathbb{R}$

$$\begin{aligned} &P[X_1 \in J_1, X_2 \in J_2, \cdots, X_N \in J_N] \\ &\qquad = P[g(W_1,W_2) \in J_1,\ g(W_2,W_3) \in J_2, \cdots, g(W_N,W_{N+1}) \in J_N] \\ &\qquad = P[(W_1,W_2) \in H_1, (W_2,W_3) \in H_2, \cdots, (W_N,W_{N+1}) \in H_N], \quad (2.10) \end{aligned}$$

donde $H_i = g^{-1}(J_i)$. Ahora defina el conjunto $\mathscr{C}$ *en* $\mathbb{R}^{N+1}$ mediante

$$\mathscr{C} := \{(w_1, \cdots, w_{N+1}) \mid (w_i, w_{i+1}) \in H_i,\ 1 \le i \le N\}.$$

Con esta notación, (2.10) es equivalente a

$$P[X_1 \in J_1, X_2 \in J_2, \cdots, X_N \in J_N]$$

$$= P[(W_1, W_2, \cdots, W_{N+1}) \in \mathscr{C}] \ . \qquad (2.11)$$

De la misma manera puede mostrarse que

$$P[X_{1+h} \in J_1, X_{2+h} \in J_2, \cdots, X_{N+h} \in J_N]$$

$$= P[(W_{1+h}, W_{2+h}, \cdots, W_{N+1+h}) \in \mathscr{C}],$$

$$= P[(W_1, W_2, \cdots, W_{N+1}) \in \mathscr{C}],$$

donde la segunda igualdad se desprende del hecho de que $\{W_t\}$ es estrictamente estacionaria. Combinando este resultado con (2.11) concluimos que

$$P[X_1 \in J_1, X_2 \in J_2, \cdots, X_N \in J_N]$$

$$= P[X_{1+h} \in J_1, X_{2+h} \in J_2, \cdots, X_{N+h} \in J_N];$$

como los intervalos J_i son arbitrarios, se desprende que

$$(X_1, X_2, \cdots, X_N) \text{ y } (X_{1+h}, X_{2+h}, \cdots, X_{N+h})$$

tienen la misma distribución. A partir de este hecho, la Observación 2.1 implica que $\{X_t\}$ es una serie estrictamente estacionaria.

(ii) Veamos ahora algunos casos particulares del resultado que acabamos de establecer:

(a) Si $g(x,y) = \mu + x + y$,

$$X_t = \mu + W_t + W_{t+1}.$$

Si $\{W_t\}$ tiene función de autocovarianza γ_W entonces la correspondiente función γ_X de $\{X_t\}$ esta dada por

$$\begin{aligned}\gamma_X(h) &= \mathrm{Cov}[X_1, X_{1+h}] \\ &= \mathrm{Cov}[\mu + W_1 + W_{1+1},\ \mu + W_{1+h} + W_{(1+h)+1}] \\ &= \mathrm{Cov}[W_1 + W_2,\ W_{1+h} + W_{2+h}] \\ &= \mathrm{Cov}[W_1, W_{1+h}] + \mathrm{Cov}[W_1, W_{2+h}] \\ &\qquad + \mathrm{Cov}[W_2, W_{1+h}] + \mathrm{Cov}[W_2, W_{2+h}] \\ &= \gamma_W(h) + \gamma_W(h+1) + \gamma_W(h-1) + \gamma_W(h),\end{aligned}$$

y finalmente,

$$\gamma_X(h) = 2\cdot\gamma_W(h) + \gamma_W(h+1) + \gamma_W(h-1),\ h \in \mathbb{Z}.$$

(**b**) Suponga que $g(x,y) = x/y$ si $y \neq 0$, $g(x,y) = 1$ si $y = 0$. En este caso

$X_t = g(W_t, W_{t+1}) = W_t/W_{t+1}$ si $W_{t+1} \neq 0$;

$X_t = 1$ si $W_{t+1} = 0$.

Es importante notar que, aunque $\{X_t\}$ es una serie estrictamente estacionaria, en general la función de medias o de autocovarianza de esta serie no está definida, a pesar de que $E[W_t^2] < \infty$. Por ejemplo, se supone que las variables W_t tienen distribución normal con media cero y varianza 1 y, desde luego, las W_t son independientes. En este caso, X_t

tiene densidad de Cauchy, la cual está dada por

$$f(x) = [\pi(1 + x^2)]^{-1}, \quad x \in \mathbb{R},$$

de donde se desprende que ningún momento de X_t es finito; vea Dudewicz y Mishira (1988) o Mood, et. al., (1974) para detalles. □

A continuación introduciremos una noción más general de proceso estacionario.

Definición 2.5. Una serie de tiempo $\{X_t | t \in \mathbb{Z}\}$ es *estacionaria (en el sentido de segundo orden o débil)* si satisface las siguientes condiciones:

(*i*) Para todo $t \in \mathbb{Z}$, $\mathrm{E}[X_t^2] < \infty$;

(*ii*) Para todo $t \in \mathbb{Z}$, $\mathrm{E}[X_t] = \mathrm{E}[X_1]$, es decir, la función de medias de la serie $\{X_t\}$ es constante, y

(*iii*) Para cada $r,s \in \mathbb{Z}$,

$$\mathrm{Cov}[X_r, X_s] = \mathrm{Cov}[X_1, X_{1+s-r}], \tag{2.12}$$

esto es, la covarianza entre X_r y X_s depende solo de la diferencia entre los índices r y s.

Observación 2.2. (*i*) Si $\mathrm{E}[X_t^2] < \infty$, entonces, usando la desigualdad de Cauchy-Schwarz, se sigue que

$$\mathrm{E}[|X_t|] \le \{\mathrm{E}[X_t^2]\}^{1/2} < \infty,$$

esto es, X_t tiene esperanza finita y la condición *(ii)* en la Definición 2.5 tiene sentido.

(ii) La condición *(iii)* en la definición anterior puede establecerse de la siguiente manera:

$$G(r,s) = G(1,\ 1 + r - s),\ r,s \in \mathbb{Z}.$$

Como en el caso estrictamente estacionario, definimos la función de autocovarianza de una serie estacionaria $\{X_t\}$ (en el sentido de la Definición 2.5) mediante

$$\gamma(h) = \mathrm{Cov}[X_1, X_{1+h}],\ h \in \mathbb{Z};$$

note que en este caso se tiene

$$\mathrm{Cov}[X_r, X_s] = \gamma(r-s). \tag{2.13}$$

(ii) Si $\{X_t\}$ es una serie estacionaria su función de autocovarianza tiene las siguientes propiedades:

Para todo $t \in \mathbb{Z}$,

(a) $\gamma(0) = \mathrm{Cov}[X_t, X_t] = \mathrm{Var}[X_t] \geq 0$; vea (2.12).

(b) $|\gamma(t)| = \mathrm{Cov}[X_1, X_{1+t}] \leq \{\mathrm{Var}(X_1)\cdot\mathrm{Var}(X_{1+t})\}^{1/2}$, y usando (a) se desprende que

$$|\gamma(t)| \leq \gamma(0).$$

(c) $\gamma(t) = \mathrm{Cov}[X_1, X_{1+t}] = \mathrm{Cov}[X_{1+t}, X_1] = \gamma(-t)$, esto es, $\gamma(\cdot)$ es una función simétrica.

En el siguiente ejemplo se discutirá la diferencia entre la noción de estacionariedad estricta en la Definición 2.3 y la idea de estacionariedad en la Definición 2.5.

Ejemplo 2.3. (*i*) Suponga que la serie $\{X_t\}$ es estrictamente estacionaria.

(a) De acuerdo a la Definición 2.5, para que $\{X_t\}$ sea una serie estacionaria es necesario que $E[X_t^2] < \infty$. Como se observó en el Ejemplo 2.2(*ii-b*) es posible que $E[X_t^2]$ no sea finita, y entonces $\{X_t\}$ no es una serie estacionaria (en el sentido de segundo orden); vea la parte (*i*) en la Definición 2.5.

(b) Si $E[X_t^2] < \infty$, entonces $\{X_t\}$ es estacionaria en el sentido de segundo orden. Esto se desprende del hecho de que en este caso la esperanza de X_t esta bien definida, y no depende t (es decir, la función de medias es constante), lo cual es consecuencia de que las X_t's tienen la misma distribución. El hecho de que $G(r,s)$ depende sólo de la diferencia $r - s$ fue verificado anteriormente.

Podemos resumir esta discusión de la siguiente manera:

> Si $\{X_t\}$ es una serie estrictamente estacionaria, entonces $\{X_t\}$ es estacionaria en el sentido de segundo orden si y sólo si $E[X_t^2] < \infty$.

(ii) Suponga ahora que $\{X_t\}$ es estacionario en el sentido de segundo orden. A continuación veremos mediante un ejemplo que $\{X_t\}$ no necesita ser estrictamente estacionaria. Con este fin, considere dos sucesiones $\{V_t\}$ y $\{B_t\}$ donde

(a) Las variables aleatorias V_t son *independientes e idénticamente distribuidas* (*iid*) con distr*ibución común* N(0,1);

(b) Las variables aleatorias B_t son iid con distribución común *Ber(1/2), y*

(c) Las sucesiones $\{V_t\}$ y $\{B_t\}$ son independientes.

Ahora defina

$$X_t := V_t \text{ si } t \text{ es par;}$$

$$X_t := 2\cdot B_t - 1 \text{ si } t \text{ es impar.}$$

En este caso no es díficil ver que para todo $t \in \mathbb{Z}$,

$$E[X_t] = 0 \tag{2.14}$$

y

$$E[X_t^2] = 1. \tag{2.15}$$

Mas aún, para $r \neq s$ se tiene lo siguiente:

Si r y s son pares,

$$\text{Cov}[X_r, X_s] = \text{Cov}[V_r, V_s] = 0,$$

pues las variables aleatorias V_t son independientes.
Similarmente, si r y s son impares,

$$\mathrm{Cov}[X_r, X_s] = \mathrm{Cov}[B_r, B_s] = 0.$$

Por último, si r es par y s es impar (o viceversa)

$$\mathrm{Cov}[X_r, X_s] = \mathrm{Cov}[V_r, B_s] = 0,$$

ya que la sucesión $\{V_r\}$ es independiente de $\{B_s\}$.
Esta discusión implica que $G(r,s) = \mathrm{Cov}[X_r, X_s] = 0$ si $r \neq s$. Combinando esto con (2.15) se obtiene que G(r,s) depende solo de r - s, y por lo tanto el proceso $\{X_r\}$ es una serie estacionaria (en el sentido de segundo orden). Sin embargo, es claro que $\{X_t\}$ no es estacionaria en el sentido estricto, pues X_1 y X_2 *no* tienen la misma distribución. □

La motivación para introducir las series de tiempo estacionarias fue el tener la posibilidad de obtener estimadores razonables de las funciones de medias y de autocovarianza. Para un proceso estacionario, la función de medias es constante, digamos μ, y un candidato obvio para ser considerado como estimador de μ es la media muestral de las variables observadas, las cuales supondremos que son $X_1, X_2, \cdots, X_n$. Se Concluye esta sección con un resultado que establece condiciones bajo las cuales la media muestral de las variables observadas es un estimador consistente de μ.

Teorema 2.1. Sea $\{X_t\}$ una serie de tiempo estacionaria con

$$\mu = \mathrm{E}[X_t],$$

y sea $\overline{X}_n$ la media muestral basada en las observaciones X_1,

$X_2, \cdots X_n$, esto es,

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n, \qquad n \geq 1$$

Entonces,

$$(i) \quad n \cdot \mathrm{Var}(\bar{X}_n) = \sum_{h=-n+1}^{h=n+1} \gamma(h) \cdot [1 - |h|/n],$$

y

$$(ii) \quad \text{Si } S := \sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \text{ entonces para todo } \varepsilon > 0,$$

$$P[|\bar{X}_n - \mu| \geq \varepsilon] \longrightarrow 0 \quad \text{cuando } n \rightarrow \infty,$$

esto es, $\{\bar{X}_n\}$ es una sucesión consistente de estimadores de μ; vea Dudewicz y Mishira (1988), Mood, et. al., (1974) o Cavazos Cadena (1990).

Demostración. *(i)* Si se observa que

$$\mathrm{Var}[\sum_{t=1}^{n} X_t] = \mathrm{Cov}[\sum_{t=1}^{n} X_t, \sum_{s=1}^{n} X_s]$$

$$= \sum_{t,s=1}^{n} \mathrm{Cov}[X_t, X_s]$$

$$= \sum_{t,s=1}^{n} \gamma(t-s). \tag{2.16}$$

Ahora es fácil ver que cuando s y t varian en $\{1,2,\cdots,n\}$ la diferencia $t-s$ toma valores en $\{z | z \in \mathbb{Z}, |z| < n\}$. Más aún,

(a) para $h = 1, 2, \cdots, n-1$, $t-s = h$ sólo cuando

$$(t,s) = (h+1,1), (h+2,2), \cdots, (n,n-h),$$

esto es, $t-s = h$ para $n - h = n - |h|$ parejas (t,s).

(b) Para $h = -1, -2, \cdots, -(n-1)$, $t-s = h$ sólo cuando

$$(t,s) = (1,-h+1),\ (2,-h+2),\ \cdots,\ (h+n,\ n),$$

y por lo tanto $t-s = h$ para $h + n = n - |h|$ parejas (t,s) (observe que $|h| = -h$ pues $h < 0$).

(c) Para $h = 0$ $t - s = 0$ cuando y sólo cuando

$$(t,s) = (1,1),\ (2,2),\ \cdots, (n,n),$$

es decir, $t-s = 0$ para $n = n - 0$ parejas.

Usando (a), (b) y (c) vemos que, en (2,16), el término $\gamma(h)$ aparece en $n - |h|$ ocasiones, y que esto ocurre para cada h con $|h| < n$. En consecuencia,

$$\mathrm{Var}[\sum_{t=1}^{n} X_t] = \sum_{h=-n+1}^{h=n+1} \gamma(h)\cdot[n - |h|],$$

y entonces,

$$\mathrm{Var}(\overline{X}_n) = n^{-2}\ \mathrm{Var}[\sum_{t=1}^{n} X_t]$$

$$= n^{-2} \sum_{h=-n+1}^{h=n+1} \gamma(h)\cdot[n - |h|],$$

$$= n^{-1} \sum_{h=-n+1}^{h=n+1} \gamma(h)\cdot[1 - |h|/n],$$

de donde se desprende que

$$n\cdot\mathrm{Var}(\overline{X}_n) = \sum_{h=-n+1}^{h=n+1} \gamma(h)\cdot[1 - |h|/n].$$

(ii) Note que

$$n\cdot \mathrm{Var}(\overline{X}_n) = \sum_{h=-n+1}^{h=n+1} \gamma(h)\cdot[1 - |h|/n].$$

$$\leq \sum_{h=-n+1}^{h=n+1} |\gamma(h)|\cdot[1 - |h|/n].$$

$$\leq \sum_{h=-\infty}^{h=\infty} |\gamma(h)| = S < \infty.$$

Luego, $\mathrm{Var}(\overline{X}_n) \leq S/n$, y usando la desigualdad de Chevichev vemos que para todo $\varepsilon > 0$,

$$P[|\overline{X}_n - \mu| \geq \varepsilon] \leq S/[n\varepsilon^2] \longrightarrow 0 \quad \text{conforme } n \rightarrow \infty.$$

Esto concluye la demostración del teorema. ∎

Observación 2.3. Bajo ciertas condiciones que no se discutirán aquí, puede demostrarse que la función de autocovarianza $\gamma(\cdot)$ de un proceso estacionario $\{X_t\}$ puede ser estimada consistentemente por la función de autocovarianza muestral $\hat{\gamma}$, la cual se define a continuación.

Se considera que se han observado $X_1, X_2, \cdots X_n$. Entonces,

$$\hat{\gamma}(h) := \sum_{t=1}^{n-h} (X_t - \overline{X}_n)^2\,(X_{t+h} - \overline{X})/\,n, \qquad \text{si } |h| < n;$$

$$:= 0 \quad \text{cuando } |h| \geq n.$$

El lector interesado puede consultar Box y Jenkins (1976), Anderson(1976) o Brockwell and Davis (1987)

2.4 FILTROS

Otro concepto de utilidad en el estudio de las series de tiempo es el de *filtro*. En general, se entiende por filtro a la operación que transforma una serie dada $\{X_t | t \in T\}$ en otra serie $\{Y_t | t \in T\}$, de modo que ya hemos visto algunos casos de filtros en el Ejemplo 2.2. Hay muchas maneras en que una serie puede ser transformada, pero las transformaciones lineales son las más importantes.

Definición 2.6. Supongase que $\{X_t | t \in \mathbb{Z}\}$ es una serie estacionaria (en el sentido de segundo orden) y que $\{a_k | k \in \mathbb{Z}\}$ es una sucesión sumable, esto es,

$$\sum_{k=-\infty}^{\infty} |a_k| < \infty. \tag{2.17}$$

Defina la serie $\{ Y_t | t \in \mathbb{Z}\}$ mediante

$$Y_t = \sum_{k=-\infty}^{\infty} a_k X_{t-k}, \quad t \in \mathbb{Z}. \tag{2.18}$$

En este caso, se dice que $\{Y_t | t \in \mathbb{Z}\}$ se obtiene a partir de $\{X_t | t \in \mathbb{Z}\}$ mediante el filtro (sumable) $\{a_k | k \in \mathbb{Z}\}$, mientras que $\{Y_t | t \in \mathbb{Z}\}$ es el proceso filtrado.

Note ahora que (2.18) implica que $|Y_t| \le \sum_{k=-\infty}^{k=\infty} a_k |X_{t-k}|$, y entonces

$$E[|Y_t|] \le \sum_{k=-\infty}^{k=\infty} a_k E[|X_{t-k}|]$$

$$\le \sum_{k=-\infty}^{k=\infty} a_k E[|X_1|] < \infty, \qquad (2.19)$$

donde las dos últimas desigualdades dependen de que $\{X_t\}$ sea estacionaria. Esto implica que la serie que define a Y_t en (2.18) es convergente con probabilidad 1. Por otro lado ya hemos visto un caso de la acción de un filtro lineal; en el Ejemplo 2.2(*iia*), el proceso $\{X_t | t \in \mathbb{Z}\}$ se obtuvo a partir de la serie estacionaria $\{W_t | t \in \mathbb{Z}\}$ mediante la aplicación del filtro $\{a_k\}$ dado por $a_0 = a_{-1} = 1$ y $a_k = 0$ para $k \neq 0, -1$.

A continuación se verá que al aplicar un filtro sumable a una serie estacionaria $\{X_t\}$, el resultado es de nuevo una serie estacionaria.

Teorema 2.2. Suponga que

(i) $\{X_t | t \in \mathbb{Z}\}$ es un proceso estacionario,

y

(ii) El filtro $a = \{a_k\}$ es sumable, esto es satisface (2.17).

Sea $\{Y_t\}$ el proceso obtenido a partir de $\{X_t\}$ aplicando el filtro a, es decir,

$$Y_t = \sum_{k=-\infty}^{k=\infty} a_k X_{t-k}, \quad t \in \mathbb{Z}.$$

Entonces $\{Y_t\}$ es una serie estacionaria.

Demostración. Primero observe que, por (2.19), Y_t tiene esperanza bien definida, y que

$$E[Y_t] = \sum_{k=-\infty}^{k=\infty} a_k E[X_{t-k}] = \sum_{k=-\infty}^{k=\infty} a_k E[X_1],$$

por lo que $E[Y_t]$ es constante. Para concluir, es necesario ver que $E[Y_t^2] < \infty$, y que $\mathrm{Cov}[Y_r, Y_s]$ depende sólo de $s - r$. Con este fin obvserve que para todo $r, s, n \in \mathbb{Z}$, se tiene que

$$\mathrm{Cov}\Big[\sum_{k=-n}^{n} a_k X_{r-k}, \sum_{j=-n}^{n} a_j X_{s-j}\Big]$$

$$= \sum_{k,j=-n}^{n} a_k \gamma_X(s - r + k - j) a_j. \tag{2.20}$$

Haciendo $s = r$ se obtiene que

$$\begin{aligned} \mathrm{Var}\Big[\sum_{k=-n}^{n} a_k X_{r-k}\Big] &= \mathrm{Cov}\Big[\sum_{k=-n}^{n} a_k X_{r-k}, \sum_{j=-n}^{n} a_j X_{r-j}\Big] \\ &= \sum_{k,j=-n}^{n} a_k \gamma_X(k - j) a_j. \\ &\le \sum_{k,j=-n}^{n} |a_k \gamma_X(0) a_j| \\ &= \gamma_X(0) \cdot \sum_{k,j=-n}^{n} |a_k a_j| \\ &= \gamma_X(0) \cdot \Big\{ \sum_{k=-n}^{n} |a_k| \Big\}^2, \end{aligned}$$

y por lo tanto

$$\mathrm{Var}\Big[\sum_{k=-n}^{n} a_k X_{r-k}\Big] \le \gamma_X(0) \cdot \Big\{\sum_{k=-\infty}^{\infty} |a_k|\Big\}^2, \quad n, r \in \mathbb{Z}.$$

Tomando límite cuando $n \longrightarrow \infty$ en el lado izquierdo de esta última desigualdad se obtiene que

$$\mathrm{Var}\,[Y_r] \le \gamma(0)\cdot\{\sum_{k=-\infty}^{\infty} |a_k|\}^2, \quad r \in \mathbb{Z},$$

lo cual significa que Y_r tiene varianza finita para todo r. En este caso tomando límite cuando $n \longrightarrow \infty$ en ambos lados de (2.20) se obtiene que

$$\mathrm{Cov}[Y_r,Y_s] = \sum_{k,j=-\infty}^{\infty} a_k\gamma_X(s - r + k - j)a_j, \; r,s, \in \mathbb{Z}.$$

Por lo tanto $\mathrm{Cov}[Y_r,Y_s]$ depende sólo de $s - r$ y se concluye que $\{Y_t\}$ es estacionario. ∎

Las series de tiempo estacionarias han sido estudiadas intensamente y existe una amplia teoría acerca de ellas. Sin embargo, muchas series de tiempo que surgen en las aplicaciones *no* son estacionarias. Veremos en la siguiente sección como un filtro lineal puede ser extremadamente útil para transformar una serie no estacionaria en una nueva serie que si es estacionaria, y para la cual existen métodos poderosos de estudio. Concluimos esta parte de nuestra exposición introduciendo una notación especial para filtros lineales.

Notación. Considere el filtro asociado con la sucesión $\{a_k\}$ dada por $a_1 = 1$, $a_k = 0$ para $k \ne 1$. En este caso (2.18) se transforma en

$$Y_t = X_{t-1},$$

y denotamos este filtro por B, esto es,

$$BX_t := X_{t-1}, \quad t \in \mathbb{Z}. \tag{2.21}$$

Observe ahora que $B(BX_t) = BY_t = Y_{t-1} = X_{t-2}$, y entonces

$$B^2X_t = X_{t-2}.$$

En general, definimos

$$B^kX_t := X_{t-k}, \quad t,k \in \mathbb{Z}.$$

Con esta notación, (2.18) puede escribirse como

$$Y_t = \sum_{k=-\infty}^{\infty} a_k B^k X_t,$$

o en notación más compacta,

$$Y_t = a(B)X_t,$$

donde

$$a(B) := \sum_{k=-\infty}^{k=\infty} a_k B^k;$$

recuerde que $a = \{a_k\}$.

2.5. DISEÑO DE UN FILTRO

Con frecuencia las series que surgen en las aplicaciones pueden descomponerse como

$$X_t = m_t + s_t + Y_t,\ t \in \mathbb{Z}, \tag{2.21}$$

donde

(*a*) $\{Y_t\}$ es una serie estacionaria con media cero,

(*b*) m_t es una función que representa la tendencia del proceso; usualmente m_t es (o puede aproximarse por) una función polinomial,

$$m_t = c_0 + c_1 t + c_2 t^2 + \ldots + c_n t^n, \tag{2.22}$$

como en el caso en que $\{X_t\}$ es la serie de las poblaciones de algunos paises; vea, por ejemplo, Brockwell y Davis (1987) pp. 14-25.

(c) La componente s_t representa una componente periódica o repetitiva en $\{X_t\}$, esto es, para todo t se tiene que $s_t = s_{t+d}$, donde d es un entero positivo el cual se denomina el *período* de s.

Note que si X_t se descompone como en (2.21), entonces $E[X_t] = m_t + s_t$, la cual (en general) depende de t y en este caso $\{X_t\}$ *no* es una serie estacionaria. Dado que la serie de interés se descompone como en (2.21), un problema importante es estimar la componente de tendecia m_t, la componente periódica s_t así como construir una serie estacionaria a partir de X_t. Por lo pronto, veamos como la aplicación de filtros lineales puede transformar $\{X_t\}$ en una serie estacionaria.

Ejemplo 2.4. (*i*) Suponga que $X_t = a + b\cdot t + s_t + Y_t$ donde s_t tiene período 12 y la serie $\{Y_t\}$ es estacionaria con media cero. En este caso

$$(1 - B^{12})X_t = X_t - X_{t-12}$$

$$= (a + b\cdot t + s_t + Y_t) - (a + b\cdot(t-12) + s_{t-12} + Y_{t-12})$$

$$= 12\cdot b + Y_t - Y_{t-12},$$

donde hemos usado que $s_t = s_{t-12}$. Usando el Teorema 2.2 vemos que $(1 - B^{12})X_t$ es un proceso estacionario con media $12\cdot b$. Además,

$$(1 - B)(1 - B^{12})X_t = (1 - B^{12})X_t - B(1 - B^{12})X_t$$

$$= Y_t - Y_{t-12} -(Y_{t-1} - Y_{t-13})$$

$$= Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

es un proceso estacionario con media cero.

(*ii*) Suponga que $X_t = (a + b\cdot t)\cdot s_t + Y_t$ donde s_t es otra vez una componente estacional con período 12. En este caso

$$(1 - B^{12})X_t = (1 - B^{12})[(a + b\cdot t)\cdot s_t] + (1 - B^{12})Y_t.$$

Además,

$$(1 - B^{12})[(a + b\cdot t)\cdot s_t]$$

$$= [(a + b\cdot t)\cdot s_t] -[(a + b\cdot(t-12)\cdot s_{t-12}$$

$$= [(a + b\cdot t)\cdot s_t] -[(a + b\cdot(t-12)\cdot s_t$$

$$= 12b\cdot s_t$$

y entonces,

$$(1 - B^{12})X_t = 12b \cdot s_t + (1 - B^{12})Y_t.$$

Como s_t tiene período 12, $(1 - B^{12})s_t = s_t - s_{t-12} = 0$, de donde se desprende que

$$(1 - B^{12})(1 - B^{12})X_t = (1 - B^{12})(1 - B^{12})Y_t$$

$$= Y_t - 2 \cdot Y_{t-12} + Y_{t-24},$$

y usando el Teorema 2.2 vemos que $\{(1 - B^{12})(1 - B^{12})X_t\}$ es una serie estacionaria. □

En el ejemplo anterior se estableció que la aplicación de un filtro lineal puede transformar una serie no estacionaria en una serie que *sí* es estacionaria. A continuación se aplicará un filtro para obtener un estimador insesgado de la función de tendencia m_t cuando ésta es un polinomio.

Ejemplo 2.5. Sea $\{X_t\}$ la serie dada por $X_t = m_t + Y_t$, donde $\{Y_t\}$ es una serie estacionaria con media cero.

(*i*) En este inciso se supondrá que la tendencia m_t es la función cuadrática

$$m_t = c_0 + c_1 t + c_2 t^2,$$

y que el filtro $a(B)$ está dado por

$$a(B) = \sum_{k=-7}^{7} a_k B^k,$$

donde $a_i = a_{-i}$, y

$$[a_0, a_1, \cdots, a_7] = (1/320)\cdot[74,67,46,21,3,-5,-6,-3].$$

Se verá ahora que

$$a(B)X_t = \sum_{k=-7}^{7} a_k B^k X_t = \sum_{k=-7}^{7} a_k X_{t-k}$$

es un estimador insesgado de m_t.

Con este fin se observa que

$$a(B)X_t = \sum_{k=-7}^{7} a_k X_{t-k}$$

$$= \sum_{i=-7}^{7} a_i m_{t-i} + \sum_{i=-7}^{7} a_i Y_{t-i},$$

y como $E[Y_k] = 0$, vemos que

$$E[a(B)X_t] = \sum_{i=-7}^{7} a_i m_{t-i} + \sum_{i=-7}^{7} a_i E[Y_{t-i}] = \sum_{i=-7}^{7} a_i m_{t-i}.$$

Luego, $E[a(B)X_t] = m_t$ si y sólo si

$$m_t = \sum_{i=-7}^{7} a_i m_{t-i}. \qquad (2.23)$$

Ahora se tendrá que verificar esta ecuación para todo $t \in \mathbb{Z}$ y cada función cuadrática m_t, y es claramente suficiente comprobar (2.23) para cada una de las siguientes tres posibilidades (*a*)-(*c*) para m_t.

(*a*) $m_t = 1$ para todo $t \in \mathbb{Z}$.

En este caso (2.23) es equivalente a $1 = \sum_{i=-7}^{7} a_i$, una igualdad que es fácilmente verificable a partir de la definición de los a_i's.

(*b*) $m_t = t$.

En este caso debemos ver que

$$t = \sum_{i=-7}^{7} a_i \cdot (t - i).$$

Para esto, recuerde que $1 = \sum_{i=-7}^{7} a_i$ de donde se desprende que

$$t = \sum_{i=-7}^{7} a_i \cdot (t - i) + \sum_{i=-7}^{7} a_i \cdot i.$$

Además, debido a que $a_i = a_{-i}$ se tiene que $\sum_{i=-7}^{7} a_i \cdot i = 0$ y entonces $t = \sum_{i=-7}^{7} a_i \cdot (t - i)$, que es lo que se deseaba establecer.

(*c*) $m_t = t^2$.

Note que $t = \sum_{i=-7}^{7} a_i \cdot (t - i)$, $t \in \mathbb{Z}$ implica que,

$$\begin{aligned} t^2 &= \sum_{i=-7}^{7} a_i \cdot t \cdot (t - i) \\ &= \sum_{i=-7}^{7} a_i (t - i)^2 + \sum_{i=-7}^{7} a_i \cdot i \cdot (t - i) \\ &= \sum_{i=-7}^{7} a_i (t - i)^2 \\ &\qquad + t \cdot \sum_{i=-7}^{7} a_i \cdot i + \sum_{i=-7}^{7} a_i i^2 \\ &= \sum_{i=-7}^{7} a_i (t - i)^2 + \sum_{i=-7}^{7} a_i i^2. \end{aligned}$$

Ahora notese que $\sum_{i=-7}^{7} a_i i^2 = 2\cdot\sum_{i=1}^{7} a_i i^2 = 0$, donde la última igualdad se obtiene fácilmente a partir de la definición de las a_i's, y entonces

$$t^2 = \sum_{i=-7}^{7} a_i (t - i)^2,$$

que es (2.23) con $m_t = t^2$.

Usando *(a)*-*(c)* se concluye que (2.23) ocurre para toda función cuadrática m_t, y en este caso

$$E[a(B)X_t] = m_t.$$

(ii) Si $a(B)$ es el filtro de la parte *(i)*, se verá ahora que $E[a(B)X_t] = m_t$ para toda función cúbica

$$m_t = c_0 + c_1 t + c_2 t^2 + c_3 t^3.$$

Como antes, $E[a(B)X_t] = m_t$ es equivalente a

$$m_t = \sum_{i=-7}^{7} a_i m_{t-i},$$

lo cual se deberá verificar cuando m_t es un polinomio de grado ≤ 3. Usando la parte *(i)* basta verificar esta igualdad para $m_t = t^3$. Para lograr esto primero notese que

$$t^2 = \sum_{i=-7}^{7} a_i \cdot (t - i)^2,$$

(por la parte *(i)*), de donde se sigue que

$$t^3 = \sum_{i=-7}^{7} a_i \cdot (t - i)^3 - \sum_{i=-7}^{7} a_i (- i)^3.$$

Por otro lado, usando que $a_i = a_{-i}$ se ve que $\sum_{i=-7}^{7} a_i(-i)^3$ $= -\sum_{i=-7}^{7} a_i i^3 = 0$, y entonces

$$t^3 = \sum_{i=-7}^{7} a_i \cdot (t - i)^3,$$

lo cual demuestra que (2.23) ocurre también para polinomios de grado 3.

El filtro que se considera en este ejemplo se conoce como filtro de 15 puntos de Spencer. Para más detalles acerca de filtros, vea, por ejemplo, Kendall y Stuart (1976). □

En el ejemplo anterior hemos visto que es posible estimar la componente de tendencia del proceso $X_t = m_t + Y_t$ por medio de un filtro lineal, siempre y cuando m_t sea un polinomio de grado ≤ 3. En los siguientes Teoremas abordamos el problema de diseñar un filtro que permita estimar una función de tendencia polinomial de grado arbitrario.

Teorema 2.3. Supongase que $\{a_k| \ |k| \le r\}$ es un filtro simétrico, esto es,

$$a_i = a_{-i}, \ i = 1,2,\cdots,r.$$

Entonces, las siguientes condiciones (a) (b) y (c) son equivalentes.

(*a*) Para toda tendencia $m_t = c_0 + c_1 t + \ldots c_s t^s$, $t \in \mathbb{Z}$,

$$m_t = \sum_{k=-r}^{r} a_k m_{t-k} \quad t \in \mathbb{Z}.$$

(*b*) $t^s = \sum_{h=-r}^{r} a_k (t-k)^s$, $\quad t \in \mathbb{Z}$.

(*c*) Para todos los enteros n con $1 \le n \le s/2$,

$$\sum_{k=1}^{r} a_k k^{2n} = 0,$$

y además

$$a_0 + 2 \cdot \sum_{k=1}^{r} a_k = 1.$$

Demostración. Es claro que (*a*) implica (*b*). Ahora supongase que para todo $t \in \mathbb{Z}$

$$t^s = \sum_{k=-r}^{r} a_k (t-k)^s. \tag{2.24}$$

Como ambos lados de esta igualdad son polinomios se sigue que ésta ocurre para todo $t \in \mathbb{R}$. Luego, derivando r veces ambos lados de (2.24) respecto a t se obtiene que

$$t^{s-r} = \sum_{k=-r}^{r} a_k (t-k)^{s-r},$$

y como $0 \le r \le s$ es un entero arbitrario, la condición (*a*) se sigue de inmediato. Por lo tanto se ha demostrado que (*a*) es válido si y sólo si (*b*) lo es. Para concluir, se verá que (*b*) y (*c*) son equivalentes. Observe que (*b*) ocurre si y sólo si

$$t^s = \sum_{k=-r}^{r} a_k (t-k)^s$$

$$= a_0 t^s + \sum_{k=1}^{r} a_k [(t+k)^s + (t-k)^s] \qquad (\text{pues } a_k = a_{-k})$$

$$= a_0 t^s + \sum_{k=1}^{r} a_k \left[\sum_{y=0}^{s} \binom{s}{y} [t^y (k)^{s-y} + t^y (-k)^{s-y}] \right]$$

$$= a_0 t^s + \sum_{k=1}^{r} a_k \sum_{\substack{0 \le y \le s \\ s-y \text{ par}}}^{s} \binom{s}{y} [2 \cdot t^y \cdot (k)^{s-y}]$$

$$= a_0 t^s + \sum_{k=1}^{r} a_k \sum_{\substack{0 \le y \le s \\ s-y \text{ par}}}^{s} \binom{s}{s-y} [2 \cdot t^y \cdot (k)^{s-y}]$$

$$= a_0 t^s + \sum_{k=1}^{r} a_k \cdot 2 \sum_{0 \le 2n \le s} \binom{s}{s-2n} t^{s-2n} (k)^{2n}$$

$$= a_0 t^s + 2 \cdot \sum_{0 \le 2n \le s} \sum_{k=1}^{r} a_k (k)^{2n} \binom{s}{2n} t^{s-2n}$$

y por lo tanto, (*b*) es equivalente a

$$t^s = (a_0 + 2 \sum_{k=1}^{r} a_k) \cdot t^s + 2 \cdot \sum_{1 \le 2n \le s} [\sum_{k=1}^{r} a_k k^{2n}] \binom{s}{2n} t^{s-2n}$$

donde la igualdad ocurre para todo $t \in \mathbb{Z}$. Como ambos lados de esta ecuación son polinomios, ésta es equivalente a

$$a_0 + 2 \cdot \sum_{k=1}^{r} a_k = 1, \text{ y } \sum_{k=1}^{r} a_k k^{2n} = 0, \quad 1 \le n \le s/2,$$

y esto termina la demostración del teorema. ∎

Es importante notar que las ecuaciones en la condición (*c*) del Teorema anterior tienen solución única si y sólo si $r \le [s/2]$ = parte entera de $s/2$; si $r < [s/2]$ las ecuaciones en (c) tienen la solución (única) $a_0 = 1$ y $a_k = 0$ para $k \ne 0$. Por otro lado, para $r > [s/2]$ el sistema tiene múltiples soluciones.

La primera parte del siguiente resultado es una consecuencia del teorema anterior, mientras que la parte (*ii*) se obtiene usando la técnica de multiplicadores de Lagrange..

Teorema 2.4. Suponga que $X_t = m_t + Y_t$, donde m_t es un polinomio de grado $\le s$ y $\{Y_t\}$ es una serie estacionaria. Entonces, si $\{ a_k |\ |k| \le r\}$ es un filtro simétrico,

(*i*) $\hat{m}_t := \sum_{k=-r}^{r} a_k X_{t-k}$ es un estimador insesgado de m_t si y sólo si

$$a_0 + 2\cdot\sum_{k=1}^{r} a_k = 1, \text{ y } \sum_{k=1}^{r} a_k k^{2n} = 0, \quad 1 \le n \le s/2. \qquad (2.25)$$

(*ii*) Si $\{Y_t\} \sim \mathrm{IID}(0,\sigma^2)$, entonces $\mathrm{Var}(\hat{m}_t)$ es mínima si y sólo si existen $\lambda_0, \lambda_1, \ldots \lambda_{[s/2]} \in \mathbb{R}$ tales que

$$(a_0, 2a_1, \cdots, 2a_r)$$

$$= \lambda_0(1,2,\cdots,2) + \sum_{n=1}^{[s/2]} \lambda_n (0,1^{2n},2^{2n},\cdots,r^{2n}), \quad (2.26)$$

donde $(a_0, a_1, \ldots a_r)$ satisface (2.25).

Ejemplo 2.5. Se usará el Teorema 2.4(*ii*) para obtener el filtro simétrico $\{a_k |\ |k| \le r\}$ que produce una varianza mínima para $\hat{m}_t$.

(*i*) Se Toma, por ejemplo, $s = 2$, lo cual significa que m_t es una función polinomial de grado 2. En este caso, el sistema (2.25) tiene soluciones múltiples sólo cuando $r > s/2 = 1$. Cuando $r = 2$, la ecuación (2.26) es

$$(a_0, 2a_1, 2a_2) = \lambda_0(1,2,2) + \lambda_1(0,1^2,2^2), \qquad (2.27)$$

donde (a_0, a_1, a_2) satisface (2.25). Para resolver (2.27) observemos que este sistema es equivalente a

$$a_0 = \lambda_0, \qquad 2a_1 = 2\lambda_0 + \lambda_1, \qquad \text{y} \qquad 2a_2 = 2\lambda_0 + 4\lambda_1.$$

Reemplazando en (2.25) se obtiene

$$\lambda_0 + (2\lambda_0 + \lambda_1) + (2\lambda_0 + 4\lambda_1) = 1;$$

$$(2\lambda_0 + \lambda_1)\cdot 1^2 + (2\lambda_0 + 4\lambda_1)\cdot 2^2 = 0.$$

Estas ecuaciones son equivalentes a

$$5\lambda_0 + 5\lambda_1 = 1 \qquad \text{y} \qquad 10\lambda_0 + 17\lambda_1 = 0,$$

cuya solución es

$$\lambda_0 = 17/35 \qquad \text{y} \qquad \lambda_1 = -10/35.$$

Luego

$$a_0 = \lambda_0 = 17/35;$$

$$a_{-1} = a_1 = [2\lambda_0 + \lambda_1]/2 = [2\cdot 17/35 + (-10/35)]/2 = 12/35;$$

$$a_{-2} = a_2 = [2\lambda_0 + 4\lambda_1]/2 = [2\cdot 17/35 + 4\cdot(-10/35)]/2 = -3/35.$$

(*ii*) Cuando $r = 3$ el mejor filtro $\{b \mid |k| \leq 3\}$, satisface

$$(b_0, 2\cdot b_1, 2\cdot b_2, 2\cdot b_3) = \lambda_0\cdot(1,2,2,2) + \lambda_1\cdot(0,1^2,2^2,3^2),$$

es decir

$$b_0 = \lambda_0, \quad 2\cdot b_1 = 2\lambda_0 + \lambda_1, \quad 2\cdot b_2 = 2\lambda_0 + 4\lambda_1, \quad 2\cdot b_3 = 2\lambda_0 + 9\lambda_1.$$

Reemplazando en (2.25) se obtiene

$$1 = (b_0 + 2\cdot b_1 + 2\cdot b_2 + 2\cdot b_3) = 7\lambda_0 + 14\lambda_1;$$

$$0 = 2\cdot b_1\cdot 1^2 + 2\cdot b_2\cdot 2^2 + 2\cdot b_3\cdot 3^2$$

$$= (2\lambda_0 + \lambda_1) + (2\lambda_0 + 4\lambda_1)\cdot 4 + (2\lambda_0 + 9\lambda_1)\cdot 9$$

$$= 28\lambda_0 + 98\lambda_1.$$

Así, debemos resolver el sistema

$$7\lambda_0 + 14\lambda_1 = 1;$$

$$28\lambda_0 + 98\lambda_1 = 0.$$

La solución es

$$\lambda_0 = 1/3, \; \lambda_1 = -\,2/21$$

y entonces

$$b_0 = \lambda_0 = 1/3$$

$$b_{-1} = b_1 = [2\lambda_0 + \lambda_1]/2 = 6/21$$

$$b_{-2} = b_2 = [2\lambda_0 + 4\lambda_1]/2 = [2/3 + 4(-2/21)]/2 = 3/21;$$

$$b_{-3} = b_3 = [2\lambda_0 + 9\lambda_1]/2 = [2/3 + 9(-2/21)]/2 = -2/21.$$

□

III PROCESOS DE MEDIAS MOVILES

3.1 INTRODUCCION

En este capítulo, se estudiarán una clase de procesos estacionarios de gran importancia en series de tiempo, los cuales se conocen como procesos de medias móviles, o procesos MA. Este modelo se obtiene aplicando un filtro lineal a una sucesión de variables aleatorias no correlacionadas con varianza finita, obteniendose como resultado un proceso estacionario.

La organización del capítulo es la siguiente:

En la sección 3.2 se introduce una noción de proceso de medias móviles. A continuación, en la sección 3.3 se define la noción de inversibilidad y se dan las condiciones necesarias y suficientes para que un proceso de medias móviles sea inversible. Finalmente, en la sección 3.4 se estudia el problema de determinar la función de autocovarianza en procesos MA y se presentan dos ejemplos concretos donde se obtiene tal función.

3.2 EL MODELO DE MEDIAS MOVILES.

Los procesos MA que estudiamos a continuación contienen un número finito de "ponderaciones" diferentes de cero. Como antes, denotaremos mediante $\{ Z_t | t \in \mathbb{Z} \}$ a un conjunto de variables aleatorias no correlacionadas con varianza común. De manera informal, un proceso $\{X_t\}$ es de medias móviles si X_t es un "promedio" ponderado de q+1 perturbaciones Z_i, desde la actual Z_t, hasta una en el pasado q períodos atrás ; además los coeficientes del promedio ponderado no cambian con el tiempo. Todo proceso MA resulta ser estacionario, ya que es el resultado de filtrar la serie estacionaria de perturbaciones $\{ Z_t\}$ mediante un filtro sumable, y tal serie siempre es estacionaria.

Definición 3.1. Una serie $\{ X_t | t \in \mathbb{Z} \}$ es un proceso de medias móviles de orden q (MA(q)), si satisface la siguiente ecuación de diferencias:

$$X_t = Z_t - \theta_1 Z_{t-1} - \ . \ . \ . - \theta_q Z_{t-q}, \ t \in \mathbb{Z}, \tag{3.1}$$

donde las Z_t' s son no correlacionadas con varianza común; para más detalles vea Box y Jenkins (1976) pp. 67-72 y Brokwell y Davis (1987) pp. 89-90

Note que (3.1) puede escribirse de la forma

$$X_t = \ (1-\theta_1 B - \theta_2 B^2 - \ . \ . \ . \ - \theta \ B_q^q) Z_t \tag{3.2}$$

o bien de la forma reducida

$$X_t = \theta(B)Z_t. \qquad (3.3)$$

A continuación se introducirá la noción de inversibilidad, la cual es particularmente simple para el el caso de procesos MA.

3.3 PROCESOS INVERSIBLES.

Iniciamos esta sección definiendo de manera precisa el concepto de inversibilidad

Definición 3.2. Sea $\{X_t\}$ un proceso MA(q) como en (3.1). El proceso $\{X_t\}$ se dice inversible si

$$\theta(z) := 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q \neq 0,$$

para todos los números complejos z que satisfagan $|z| \leq 1$.

La noción de inversibilidad introducida en la definición anterior es importante y está ligada a la siguiente pregunta:

$$\text{si } X_t = \theta(B)Z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)Z_t, \; t \in \mathbb{Z},$$

¿ es posible expresar $\{Z_t\}$ en la forma

$$Z_t = \sum_{k=0}^{\infty} a_k X_{t-k}, \; t \in \mathbb{Z} \text{ ?}$$

La respuesta es afirmativa si y solo si $\{X_t\}$ es **inversible.**

Tener la posibilidad de "despejar" Z_t de la ecuación $Z_t = \theta(B)\ X_t$ es útil para implementar procedimientos de estimación de los parámetros $\theta_1 \cdots \theta_q$ en (3.1). El lector interesado puede encontrar más información sobre este tema en Brockwell y Davis (1988 y 1989), así como en Melard (1984) o Berk (1974).

Ejemplo 3.1 Sea $\{X_t\}$ el proceso MA(1) dado por

$$X_t = Z_t - \theta\ Z_{t-1}.$$

En este caso $\theta(z) = 1 - \theta_1 z$ y la única raíz de $\theta(z)$ es $z = 1/\theta_1$, la cual tiene valor absoluto mayor a uno si y solo si $|\theta_1| < 1$. Luego $\{X_t\}$ es inversible cuando y sólo cuando $|\theta_1| < 1$. En este caso, se puede despejar Z_t como sigue:

Notese que $X_t = (1 - \theta_1 B)\ Z_t$ y entonces, procediendo formalmente, se tiene que

$$Z = (1 - \theta_1 B)^{-1}\ X$$

$$= \sum_{k=0}^{\infty} (\theta_1 B)^k\ X_t,$$

esto es,

$$Z_t = \sum_{k=0}^{\infty} \theta_1^k\ X_{t-k}\ ;$$

Notese que la serie converge cuando $|\theta_1| < 1$ (esto es, cuando $\{X_t\}$ es inversible) y diverge en otro caso.

Ejemplo 3.2 Considerese el proceso MA(2) dado por

$$X_t = 1 - 0.4\ Z_{t-1} + 0.04\ Z_{t-2}.$$

observese que el polinomio $\theta(z)$ correspondiente a esta serie es

$$\theta(z) = 1 - 0.4\ Z + 0.04\ Z^2$$

$$= (1 - 0.2\ z)^2$$

por lo que $\theta(z)$ tiene la única raíz doble $z = 1/0.2 = 5$ la cual cae fuera del disco unitario y se concluye que el proceso $\{X_t\}$ es inversible. □

3.4 FUNCION DE AUTOCOVARIANZA.

En esta sección, se dará una expresión para la función de autocovarianza de un proceso (la función de autocovarianza ya se definió en el cap. 2). MA como en (3.1). notese que para k > 0

$$\gamma(k) = E[\ X_t,\ X_{t+k}]$$

$$= E[\sum_{s=0}^{q} \theta_s Z_{t-s}, \sum_{r=0}^{q} \theta_r Z_{t+k-r}]$$

$$= \sum_{s,r=0}^{q} \theta_s \theta_r E[Z_{t-s}, Z_{t+k-r}]$$

$$= \sum_{s=0}^{q-r} \theta_s \theta_{s+k},$$

donde $\theta_0 = 1$.

Finalmente, usando que $\gamma(\cdot)$ es una función simetrica se concluye que

$$\gamma(k) = \sum_{s=0}^{q-k} \theta_s \theta_{s+|k|}, \quad k \in \mathbb{Z}.$$

IV PROCESOS AUTORREGRESIVOS

4.1. INTRODUCCION

En este capítulo se estudiará una clase muy importante de series estacionarias, a saber, los procesos autorregresivos. Las razones por las que estos procesos son útiles en las aplicaciones son diversas, y aquí se mencionan sólo dos de ellas:

(*i*) Los cálculos para determinar las funciones importantes del proceso, por ejemplo la función de autocovarianza pueden realizarse mediante algoritmos eficientes, y

(*ii*) Una serie de tiempo arbitraria puede aproximarse por un proceso autorregresivo en el siguiente sentido: si se conoce la función de covarianza γ_Y de una serie arbitraria $\{Y_t\}$, puede encontrarse un proceso autorregresivo $\{X_t\}$ cuya función de autocovarianza γ_X coincide con γ_Y para todos los retardos h con $|h| < M$, donde M es un entero positivo arbitrario.

La organización del capítulo es la siguiente: En la Sección 2 se introducen los procesos que estudiaremos en este capítulo así como la noción de polinomio causal. Luego, en la Sección 3 se describe un método para calcular la función de autocovarianza el cual involucra la solución de un sistema especial de ecuaciones lineales conocido como el sistema de Yule-Walker. En la sección 4 se establece un resultado acerca de la unicidad de las soluciones de ese sistema, y éste se demuestra en la Sección 6, después se presentan algunos resultados preliminares en la sección 5.

4.2. EL MODELO AUTORREGRESIVO

Empezamos esta sección con una definición formal de la noción de proceso autorregresivo.

Definición 4.1. Una serie $\{X_t | t \in \mathbb{Z}\}$ estacionaria con media cero es un proceso autorregresivo de orden p (AR(p)) si satisface una ecuación (de diferencias) de la forma

$$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + Z_t, \quad t \in \mathbb{Z}, \tag{4.1}$$

donde $\{ Z_t\}$ es una sucesión de variables aleatorias no correlacionadas con varianza común $\sigma^2 > 0$, y $\varphi_1, \varphi_2, \cdots, \varphi_p \in \mathbb{R}$ son números dados.

Note que (4.1) puede escribirse como

$$(1- \varphi_1 B - \phi_2 B^2 - \cdots - \varphi_p B^p)\ X_t = Z_t,$$

donde $B^k \ X_t = \ X_{t-k}$; vea la sección 2.3 del capítulo 2. Podemos abreviar aún más esta ecuación escribiendo

$$\varphi(B)X_t = Z_t, \tag{4.2}$$

donde el polinomio $\varphi(z)$ está dado por

$$\varphi(z) = 1 + \varphi_1 z + \varphi_2 z^2 + \cdots + \varphi_p z^p, \tag{4.3}$$

al cual denominamos el *polinomio autorregresivo* del proceso $\{ X_t \}$.

La primera pregunta que surge en relación a la definición anterior es la siguiente:

Dados los coeficientes $\varphi_1, \varphi_2, \ldots, \varphi_p$ así como el ruido blanco $\{Z_t\}$,

¿bajo qué condiciones existe un proceso $\{X_t\}$

que satisfaga (4.1)?

En el ejemplo siguiente se estudia esta pregunta en el caso especial p = 1.

Ejemplo 4.1. Supongase que $\varphi(z) = 1 - \varphi_1 z$. El objetivo es determinar para que valores de $\varphi_1 \in \mathbb{R}$ existe un proceso $\{X_t\}$ que satisfaga

$$X_t - \varphi_1 X_{t-1} = Z_t. \tag{4.4}$$

Se verá este problema en cada uno de los siguientes casos a) -c).

a) $|\varphi_1| = 1$.

En esta situación se verá que *no* existe $\{X_t\}$ que satisfaga (4.4). Esto es, demostrar que no existe solución estacionaria $\{X_t\}$ de la ecuación en diferencias (4.4). Con este fin, suponga que $\{X_t\}$ satisface (4.4). Iterando esta ecuación se desprende inmediatamente que

$$X_t = \sum_{k=0}^{r-1} \varphi_1^k Z_{t-k} + \varphi_1^r X_{t-r}. \tag{4.5}$$

Por otro lado, como $\{X_t\}$ es un proceso estacionario, se tiene que $\|X_t\| = \|\varphi^r X_{t-r}\| = M$ $(< \infty)$ para todo t,r. Entonces (4.5) implica que

$$(2M)^2 \geq || X_t - \varphi_1^r X_{t-r} ||^2 = || \sum_{k=0}^{r-1} \varphi_1^k Z_{t-k}||^2$$

$$= \sum_{k=0}^{r-1} |\varphi_1|^{2k} \cdot ||Z_{t-k}||^2,$$

donde hemos usado la condición de que las variables Z_t son no correlacionadas para obtener la segunda igualdad. Como $|\varphi_1| = 1$ y $\mathrm{Var}(Z_t) = \sigma^2$, la ecuación anterior implica que

$$(2M)^2 \geq ||X_t - \varphi^r X_{t-r}||^2 = \sum_{k=0}^{r-1} 1 \cdot \sigma^2 = r\ \sigma^2,$$

y entonces $\sigma^2 \leq (2M)^2/r$. Como *r* es un entero positivo arbitrario, esta última desigualdad implica que $\sigma^2 = 0$, lo cual es una contradicción. Por lo tanto, concluimos que, si

$|\varphi_1| = 1$, entonces, no existe un proceso $\{X_t\}$ que satisfaga (4.4).

b) $|\varphi_1| > 1$. En estas circunstancias se verá que *sí* existe solución estacionaria $\{X_t\}$ de la ecuación (4.4) y que ésta es única.

Suponga que $\{X_t\}$ satisface (4.4). Despejando X_{t-1} se tiene

$$X_{t-1} = (1/\varphi_1)X_t - (1/\varphi_1)Z_t,$$

o equivalentemente,

$$X_t = (1/\varphi_1)X_{t+1} - (1/\varphi_1)Z_{t+1}.$$

Iterando esta ecuación se obtiene

$$\begin{aligned} X_t &= (1/\varphi_1)[(1/\varphi_1)X_{t+2} - (1/\varphi_1)Z_{t+2}] - (1/\varphi_1)Z_{t+1} \\ &= (1/\varphi_1)^2X_{t+2} - (1/\varphi_1)^2Z_{t+2} - (1/\varphi_1)Z_{t+1}^2, \end{aligned}$$

y en general,

$$X_t = -\sum_{k=1}^{r}(1/\varphi_1)^k Z_{t+k} + (1/\varphi_1)^r X_{t+r}, \quad r > 0. \qquad (4.6)$$

Como $|\varphi_1| > 1$, $\|(1/\varphi_1)^r X_{t+r}\| = |(1/\varphi_1)|^r \|X_{t+r}\| \longrightarrow 0$ cuando $r \longrightarrow \infty$, y entonces, tomando límite conforme $r \longrightarrow \infty$ en (4.6), se concluye que

$$X_t = -\sum_{k=1}^{\infty}(1/\varphi_1)^k Z_{t+k}. \qquad (4.7)$$

Recíprocamente, si *se define* X_t mediante (4.7), se tiene que

$$X_t = -\sum_{k=1}^{\infty}(1/\varphi_1)^k Z_{t+k}$$

$$= (-1/\varphi_1)Z_{t+1} - \sum_{k=2}^{\infty}(1/\varphi_1)^k Z_{t+k}$$

$$= (-1/\varphi_1)Z_{t+1} - \sum_{k=1}^{\infty}(1/\varphi_1)^{k+1} Z_{t+1+k}$$

$$= (-1/\varphi_1)Z_{t+1} + (1/\varphi_1)\cdot\{-\sum_{k=1}^{\infty}(1/\varphi_1)^k Z_{t+1+k}\}$$

$$= (-1/\varphi_1)Z_{t+1} + (1/\varphi_1)X_{t+1}.$$

Luego, para todo t,

$$X_t = (-1/\varphi_1)Z_{t+1} + (1/\varphi_1)X_{t+1},$$

y entonces

$$X_{t+1} = \varphi_1 X_t + Z_{t+1},\ t \in \mathbb{Z},$$

y esto significa que $\{X_t\}$ satisface (4.4).

c) $|\varphi_1| < 1$. Se notará que para esta condición también existea una única solución estacionaria de (4.4).

Supongase que $\{X_t\}$ es solución de (4.4), esto es,

$$X_t = \varphi_1 X_{t-1} = Z_t,$$

e iterando esta ecuación, es fácil ver que

$$X_t = Z_t + \varphi_1 Z_{t-1} + \varphi_1^2 X_{t-2},$$

y en general, para todo $k > 0$,

$$X_t = Z_t + \varphi_1 Z_{t-1} + \ldots + \varphi_1^k Z_{t-k} + \varphi_1^{k+1} Z_{t-k-1}.$$

Como $|\varphi_1| < 1$, $\|\varphi_1^{k+1} Z_{t-k-1}\| = |\varphi_1|^{k+1} \|Z_{t-k-1}\| = |\varphi_1|^{k+1}\sigma^2 \longrightarrow 0$ conforme $k \longrightarrow \infty$, y la ecuación anterior implica que

$$X_t = \sum_{k=0}^{\infty} \varphi_1^k Z_{t-k}.$$

Recíprocamente, si X_t está *definido* por la ecuación anterior, entonces

$$\begin{aligned} X_t &= Z_t + \sum_{k=1}^{\infty} \varphi_1^k Z_{t-k} \\ &= Z_t + \sum_{k=0}^{\infty} \varphi_1^{k+1} Z_{t-1-k} \\ &= Z_t + \varphi_1 \cdot \sum_{k=0}^{\infty} \varphi_1^k Z_{t-1-k}, \end{aligned}$$

esto es,

$$X_t = Z_t + \varphi_1 X_{t-1},$$

y entonces $\{X_t\}$ satisface la ecuación (4.4).

Se puede resumir la discusión en este ejemplo como sigue:

Existe un proceso $\{X_t\}$ que satisfaga (4.4)

si y sólo si $|\varphi_1| \neq 1$.

Observese que $\varphi(z) = 1 - \varphi_1 z$ tiene una sola raíz, a saber $r = 1/\varphi_1$, y que la condición $|\varphi_1| \neq 1$ equivale a $|r| \neq 1$.
Por lo tanto la conclusión anterior puede establecerse equivalentemente de la siguiente manera:

Existe $\{ X_t\}$ que satisfaga (4.4) si y sólo si

el polinomio autorregresivo $\varphi(z) = 1 - \varphi_1 z$

no se anula en $\{ z \mid z \in \mathbb{C}, |z| = 1 \}$. □

Lo importante del ejemplo anterior es que su conclusión se generaliza al caso de procesos autorregresivos de orden arbitrario.

Teorema 4.1. Sea $\varphi(z) = 1 - \varphi_1 z - \varphi_2 z^2 - \cdots - \varphi_p z^p$ un polinomio con coeficientes reales. Entonces, para que exista un proceso $\{ X_t\}$ que satisfaga

$$\varphi(B)X_t = Z_t, \ t \in \mathbb{Z},$$

es necesario y suficiente que el polinomio autorregresivo $\varphi(z)$ no se anule en $\{z \mid z \in \mathbb{C}, |z| = 1\}$.
Demostración. Aquí sólo se establecerá la suficiencia; una demostración completa puede verse en Brockwell y Davis (1987).

Suponga que $\varphi(z) \neq 0$ en $\{z \mid |z| = 1\}$, y defina $\psi(z)$ mediante

$$\psi(z) := 1/\varphi(z), \ |z| \neq 1.$$

Entonces ψ se expresa en la forma

$$\psi(z) = \sum_{k=-\infty}^{\infty} \psi_k z^k,$$

donde la serie converge absolutamente en una región anular $\mathcal{R} = \{z| \ z \in \mathbb{C} \text{ y } r_1 < |z| < r_2\}$ donde $r_1 < 1 < r_2$.
Ahora definase X_t mediante

$$X_t = \psi(B)Z_t = \sum_{k=-\infty}^{\infty} \psi_k Z_{t-k}, \ t \in \mathbb{Z}. \tag{4.8}$$

En este caso,

$$\varphi(B)X_t = \sum_{i=0}^{p} \varphi_i X_{t-i}$$

$$= \sum_{i=0}^{p} \varphi_i \sum_{k=0}^{\infty} \psi_k Z_{t-i-k}$$

$$= \sum_{i=0}^{p} \sum_{k=0}^{\infty} \varphi_i \cdot \psi_k Z_{t-i-k},$$

esto es,

$$\varphi(B)X_t = \sum_r \left(\sum_{i+k=r}^{\infty} \varphi_i \cdot \psi_k\right) Z_{t-r}. \tag{4.9}$$

Por otro lado, observese que $\psi(z) = 1/\varphi(z)$ equivale a $\varphi(z)\psi(z) = 1$ y usando que $\psi(z) = \sum_k \psi_k z^k$ y $\varphi(z) = \sum_{s=0}^{p} \varphi_s z^s$ se desprende que

$$1 = \psi(z)\varphi(z)$$

$$= \{\sum_k \psi_k z^k\} \cdot \{\sum_s \varphi_s z^s\}$$

y entonces,

$$1 = \sum_r \{\sum_{i+k=r} \varphi_i \psi_k\} z^r .$$

Comparando los coeficientes de z^r en ambos lados de esta ecuación vemos que

$$\sum_{i+k=r} \varphi_i \psi_k = 0 \quad \text{si } r \neq 0,$$

y además

$$\sum_{i+k=r} \varphi_i \psi_k = 1 \quad \text{si } r = 0.$$

Combinando estos hechos con (4.9) concluimos que

$$\varphi(B) X_t = \sum_r (\sum_{i+k=r} \varphi_i \psi_k) Z_{t-r} = Z_t ,$$

pues el coeficiente de Z_{t-r} se anula cuando $r \neq 0$. □

El Teorema anterior establece que existe un proceso AR(p) con polinomio autorregresivo $\varphi(z)$ si y sólo si $\varphi(z) \neq 0$ para todos los números complejos z con $|z| = 1$. Los polinomios que satisfacen $\varphi(z) \neq 0$ cuando $|z| \leq 1$ reciben un nombre especial.

Definición 4.2. El polinomio $\varphi(z)$ se dice *causal* si $\varphi(z) \neq 0$ para todos los números complejos z que satisfacen $|z| \leq 1$; se supone que $\varphi(z)$ tiene coeficientes reales.

Cuando $\varphi(z)$ es un polinomio causal la función $\psi(z) = 1/\varphi(z)$, se expresa como una serie de potencias

$$\psi(z) = \sum_{k=0}^{\infty} \psi_k z^k$$

y entonces la única solución de $\varphi(B)X_t = Z_t$ está dada por

$$X_t = \sum_{k=0}^{\infty} \psi_k Z_{t-k}; \tag{4.10}$$

vea (4.9). Esto significa que X_t es una función de perturbaciones Z_s con $s \leq t$. La idea es que las perturbaciones Z_s con $s > t$, las cuales se interpretan como ocurriendo en tiempos posteriores a t, no aparecen en la expresión (4.10) para X_t y, por lo tanto, cuando $\varphi(z)$ es un polinomio causal, la unica solución de $\varphi(B)X_t = Z_t$ no depende de perturbaciones futuras. Es importante mencionar que un proceso AR(p) arbitrario puede expresarse como en (4.1) para algún polinomio causal $\varphi(z)$ y un ruido blanco adecuado.; vea, por ejemplo, Brockwell y Davis (1987).

Concluimos esta sección con un ejemplo en el que determinaremos condiciones necesarias y suficientes para que un polinomio de grado 2 sea causal.

Ejemplo 4.1. Considere el polinomio $\varphi(z) = 1 - \varphi_1 z - \varphi_2 z^2$. Veremos que $\varphi(z)$ es causal si y solo si los parámetros (φ_1, φ_2) pertenece a la región triangular determinada por

$$\left.\begin{array}{l}\varphi_1 + \varphi_2 < 1 \\ \varphi_1 - \varphi_2 < 1 \\ |\varphi_2| < 1.\end{array}\right\} \qquad (4.11)$$

Primero supongase que φ es causal.

En este caso $\varphi(z)$ no se anula en $[-1,1]$. Como $\varphi(0) = 1$, debemos tener que $\varphi(1) > 0$, es decir $\varphi(1) = 1 - \varphi_1 - \varphi_2 > 0$, y similarmente $\varphi(-1) = 1 + \varphi_1 - \varphi_2 > 0$, lo cual da las primeras dos desigualdades en (4.11). Para terminar, sean r_1 y r_2 las raices de $\varphi(z)$. Entonces $|r_1| > 1$, $|r_2| > 1$ y $\varphi(z)$ se escribe como

$$\varphi(z) = (1- r_1^{-1}z)\ (1- r_2^{-1}z),$$

de donde se desprende que $\varphi_2 = (r_1 r_2)^{-1}$, y entonces $|\varphi_2| = |r_1 r_2|^{-1} < 1$, lo cual establece la tercera desigualdad en (4.11).

Recíprocamente supongase que φ_1 y φ_2 satisfacen (4.11).

Tenemos que ver que $\varphi(z)$ es causal, es decir, que las raices r_1 y r_2 de $\varphi(z)$ tienen valor absoluto mayor a 1. Con este fin, note que las primeras dos desigualdades en (4.11) significan que $\varphi(1) > 0$ y $\varphi(-1) > 0$.

Si $r_1 \in (-1,1)$ se tienen la dos posibilidades (a) y (b) descritas a continuación.

(a) $r_1 = r_2$.

En este caso $|r_1 r_2| < 1$ y entonces $|\varphi_2| = |r_1 r_2|^{-1} > 1$, lo cual contradice la tercera desigualdad en (4.11).

(b) $r_1 \in (-1,1)$ es una raíz *simple* de $\varphi(z)$; note que $r_1 \neq 0$, pues $\varphi(0) = 1$.

Si $r_1 \in (-1,0)$ entonces φ toma necesariamente valores negativos en el intervalo $(r_1,0)$ y, ya que $\varphi(1) > 0$, φ debe anularse también en algún punto $r_2 \in (r_1,1)$. Por lo tanto la otra raíz r_2 tiene valor absoluto menor que 1 y entonces $|r_1 r_2| < 1$, lo cual implica que $|\varphi_2| > 1$, en contradicción con la tercera desigualdad en (4.11). De manera similar se prueba que $r_1 \in (0,1)$ conduce a una contradicción.

Esto demuestra que $\varphi(z)$ no posee raíces *reales* en el conjunto $\{z| \ |z| \leq 1\}$. Sin embargo esto no demuestra todavía que φ es causal, pues debemos ver que φ no se anula para *cualquier* número *complejo* z con $|z| \leq 1$. Para ver esto suponga que $c \in \mathbb{C}$ *no* es real y que $\varphi(c) = 0$.

En este caso, la otra raíz de φ es $\bar{c}$ y φ se escribe como

$$\varphi(z) = (1 - z/c)\cdot(1 - z/\bar{c}),$$

de donde se desprende que $\varphi_2 = (c\cdot\bar{c})^{-1} = |c|^{-2}$. Usando la tercera desigualdad en (4.11) vemos que $|c|^{-2} < 1$, y entonces $|c| > 1$, esto es c no pertenece al conjunto. $\{z| \ |z| \leq 1\}$. En resumen, hemos visto que la condición (4.11) implica que $\varphi(z)$ no posee raíces z con $|z| \leq 1$, y por lo tanto φ es causal. $\square$

4.3. DETERMINACION DE LA FUNCION DE AUTOCOVARIANZA.

En esta sección verá como encontrar la función de autocovarianza $\gamma(\cdot)$ de un proceso AR(p) en (4.1); se supondrá que el polinomio autorregresivo $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ es causal. En este caso X_t se expresa en términos de las perturbaciones $\{ Z_s | \ s \leq t \}$; vea (4.10).

El siguiente **método de los dos pasos**, descrito por Brokwell y Davis (1987), es un método de cálculo muy conveniente para determinar $\gamma(\cdot)$.

Paso 1. Encontrar $\gamma(0), \gamma(1), \cdots, \gamma(p)$ resolviendo el sistema

$$\left.\begin{aligned} \sum_{k=0}^{p} \gamma(k)\varphi_k &= \sigma^2 \\ \sum_{k=0}^{p} \gamma(|i-k|)\varphi_k &= 0 \quad i = 1,2, \ . \ . \ ., p. \end{aligned}\right\} \tag{4.12}$$

el cual es el sistema de ecuaciones de Yule-Walker (Y-W) asociado a φ.

Paso 2. Usar $\gamma(i) = \sum_{k=1}^{p} \gamma(i-k)\varphi_k$, $i > p$, para determinar $\gamma(p+1)$, $\gamma(p+2)$, . . ., en una manera *recursiva.*

Para que el método de dos pasos descrito al inicio de la sección esté bien justificado es necesario demostrar que el sistema de Yule-Walker tiene solución única cuando el polinomio autorregresivo del proceso es causal. En la siguientes secciones se abordará este problema.

4.4. UNICIDAD DE LA SOLUCION DEL SISTEMA DE YULE-WALKER

En esta sección se establece un resultado acerca de la unicidad de las soluciones del sistema de Yule-Walker (4.12) en el caso en que el polinomio autorregresivo es causal. Este resultado de unicidad puede ser verificado fácilmente cuando el grado de φ es pequeño, digamos $p = 1$ o $p = 2$, pero cuando φ es de grado arbitrario, la situación es distinta. En Achilles (1987) o Lütkepohi y Maschke (1988) puede encontrarse una demostración de que el sistema (4.12) tiene solución única cuando φ es un polinomio causal. El objetivo que se tiene en el resto del capítulo es calcular el determinante de la matriz del sistema de Yule-Walker, y la presentación está basada en Cavazos-Cadena (1992). Se Verá que cuando φ es causal, la matriz de (4.12) posee un determinante no nulo, confirmando la unicidad de la solución del sistema.

El resultado que que se expone tiene como consecuencia principal la siguiente:

El sistema de (Y-W) asociado a un polinomio φ tiene una solución única si y sólo si

$$r_i r_j \neq 1, \quad i,j = 1,2,\cdots,p. \tag{4.13}$$

donde $r_1, r_2, \ldots r_p$ son las raices de φ.

(Vea el Teorema 4.2 lineas abajo.)

Notese que cuando φ es un polinomio causal todas sus raices están fuera del disco unitario, y en este caso (4.13) se satisface claramente.

Antes de enunciar el siguiente teorema, encontraremos la matriz del sistema de Yule-Walker asociado a φ. Empezaremos introduciendo la siguiente notación referente a los coeficientes de $\varphi(z)$:

$$\varphi_k := 0 \quad \text{para} \quad k < 0 \quad \text{ó} \quad k > p. \tag{4.14}$$

Ahora observe que para i > 0,

$$\sum_{k=0}^{p} \gamma(|i-k|)\varphi_k = \sum_{k=0}^{i} \gamma(i-k)\varphi_k + \sum_{k=i+1}^{p} \gamma(k-i)\varphi_k$$

$$= \sum_{j=0}^{i} \gamma(j)\varphi_{i-j} + \sum_{j=1}^{p-i} \gamma(j)\varphi_{i+j}$$

y usando la convención (4.14) obtenemos

$$\sum_{k=0}^{p} \gamma(|i-k|)\varphi_k = \gamma(0)\varphi_i + \sum_{j=1}^{p} \gamma(j)\cdot[\varphi_{i-j} + \varphi_{i+j}].$$

Así el sistema (4.12) puede ser escrito como

$$\left.\begin{array}{r} \sum_{k=0}^{p} \gamma(k)\varphi_k = \sigma^2; \\ \gamma(0)\varphi_i + \sum_{j=1}^{p} \gamma(j)\cdot[\varphi_{i-j} + \varphi_{i+j}] \end{array}\right\} \tag{4.15}$$

La matriz (cuadrada) de coeficientes de este sistema será denotada por $M(\varphi)$. Claramente, $M(\varphi)$ es de orden $(p+1)$ y está dado como sigue:

Para $i = 0,1,2, \ldots, p$

$$
\begin{aligned}
M(\varphi)_{i,0} &:= \varphi_i ; \\
M(\varphi)_{i,j} &:= \varphi_{i-j} + \varphi_{i+j} \ , \ j = 1,2 \ldots, p.
\end{aligned}
\tag{4.16}
$$

El siguiente teorema contiene una fórmula para el determinante de $M(\varphi)$ y, como consecuencia directa, se desprende un criterio para la no singularidad de $M(\varphi)$.

Teorema 4.2. Sea $\varphi(z) = 1+ \varphi_1 z + \ldots + \varphi_p z^p$ un polinomio complejo con grado p. Si las raices de φ son $r_1, \ldots, r_p$ y $M(\varphi)$ es como en (4.16), entonces (i) y (ii) ocurren.

(i) $$\operatorname{Det} M(\varphi) = \prod_{1\leq i<j\leq p} [1-(r_i r_j)^{-1}] \cdot \prod_{i=1}^{p} (1-r_i^{-2}) ; \tag{4.17}$$

para p = 1 el primer producto en (4.17) es 1.

(ii) $M(\varphi)$ es no singular si y sólo si $r_i r_j \neq 1$ para todo $i,j = 0,1,2, \ldots, p$.

La demostración de este resultado será dada más adelante. Por el momento notamos que (ii) se sigue inmediatamente de (i), y que (4.17) es fácilmente verificable para pequeños valores de p. Por ejemplo, suponga que p = 1 y que $\varphi(z) = 1- \varphi_1 z$. En este caso

$$M(\varphi) = \begin{bmatrix} 1 & \varphi_1 \\ \varphi_1 & 1 \end{bmatrix}$$

vea (4.16). Entonces Det $M(\varphi) = 1 - \varphi_1^2$ y esto queda como (4.17) con $p = 1$, pues φ tiene la (única) raiz $r_1 = -1/\varphi_1$. Para $p = 2$ se factoriza $\varphi(z)$ como

$$\varphi(z) = (1 + a_1 z)\cdot(1 + a_2 z),$$

donde las raices de φ son

$$r_i = -1/a_i,\ i = 1,2,$$

y entonces

$$\varphi(z) = 1 + (a_1 + a_2)z + a_1 a_2 z^2.$$

Por lo tanto, $M(\varphi)$ está dada por

$$M(\varphi) = \begin{bmatrix} 1 & a_1 + a_2 & a_1 a_2 \\ a_1 + a_2 & a_1 a_2 + 1 & 0 \\ a_1 a_2 & a_1 + a_2 & 1 \end{bmatrix}.$$

Expandiendo Det $M(\varphi)$ por la tercera columna se obtiene

Det $M(\varphi)$

$$= a_1a_2[(a_1+ a_2)^2-a_1a_2(1+ a_1a_2)] + (1+ a_1a_2) - (a_1+ a_2)^2$$

$$=-(a_1+ a_2)^2\cdot(1- a_1a_2) + (1+ a_1a_2)\cdot[1- (a_1a_2)^2]$$

$$= (1- a_1a_2)\cdot[-(a_1+ a_2)^2+ (1+ a_1a_2)^2]$$

$$= (1- a_1a_2)\cdot(1- a_1^2)\cdot(1- a_2^2),$$

y reemplazando a_i por $-1/ r_i$ se obtiene (4.17) con $p = 2$.

La demostración del Teorema 4.2 en el caso general es por inducción y será presentada en la Sección 5.

4.5 PRELIMINARES BASICOS

Esta sección contiene algunos resultados preliminares que serán utilizados en la demostración del Teorema 4.2.

Se inicia introduciendo una notación especial para algunos objetos de interés asociados a un polinomio φ.

Notación. Recuerdese que $\mathbb{Z}$ denota al conjunto de enteros, $\mathbb{N} = \{0,1,2,\cdots\}$ mientras que $\mathbb{C}$ representa al conjunto de números complejos. Por otro lado, el espacio vectorial complejo ℓ consiste de todos los vectores $V:\mathbb{N} \longrightarrow \mathbb{C}$ con la propiedad de que, para algún $C \in \mathbb{N}$, $V(k) = 0$ si $|k| > C$; ℓ está dotado de las operaciones usuales de adición y multiplicación por escalar. El *operador de desplazamiento* $s:\ell \longrightarrow \ell$ está definido como sigue:

Para $V \in \ell$

$$s(V)(0):=0, \text{ y } s(V)(k):= V(k-1), k = 1,2, \cdots; \tag{4.18}$$

además,

$$s^0(V)=V \quad \text{y} \quad s^n(V)= s^{n-1}(s(V)) \tag{4.19}$$

Por otro lado, los renglones y columnas de una matriz cuadrada M serán numerados a partir de cero y Det M denota el determinante de M. Para vectores $V_0, V_1, \cdots, V_n \in \ell$ definimos la matriz cuadrada $M_{n+1}(V_0, V_1, \cdots, V_n)$ de orden $(n+1)$ por

$$M_{n+1}(V_0, V_1, \cdots, V_n):= v_i(j), \quad i,j = 0,1,2,\cdots,n. \tag{4.20}$$

Por otro lado, span$\{V_0, V_1, \cdots, V_n\}$ denota el espacio vectorial espacio generado por V_i's y Dim span$\{V_0, V_1, \cdots, V_n\}$ es la dimensión correspondiente. Para concluir, con un polinomio $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ de grado p asociamos dos vectores $\vec{\varphi}$ y $\overleftarrow{\varphi}$ en ℓ definidos como sigue:

$$\left.\begin{array}{l} \vec{\varphi}(k):=\varphi(k) \quad y \quad \overleftarrow{\varphi}(k) = \varphi_{p-k}, \quad k=0,1,2,\ldots p; \\ \vec{\varphi}(k):=0 \quad y \quad \overleftarrow{\varphi}(k):=0, \quad k > p. \end{array}\right\} \qquad (4.21)$$

Con cada polinomio φ se asociará una sucesión en $\mathfrak{L}$ definida a continuación.

Definición 4.2. Sea $\varphi(z) = 1+ \varphi_1 Z +\ldots+ \varphi_p z^p$ un polinomio de grado p. La secuencia $V^{\varphi} = \{ V_t^{\varphi} \mid t\in\mathbb{N} \} \subset \mathfrak{L}$ se define como sigue :

(i) para $1 \le n < p$

$$V_n^{\varphi}(0):= \varphi_n, \quad y \quad V_n^{\varphi}(k):= \varphi_{n+k} + \varphi_{n-k} \ ,k = 1,2,\cdots.$$

(ii) Para $n \in \mathbb{N}$, $V_{-n}^{\varphi}:= s^n(\vec{\varphi})$ y $V_{n+p}^{\varphi}:= s^n(\overleftarrow{\varphi})$.

A partir (4.14), (4.20), (4.21) y de la definición anterior, se desprende que la secuencia V^{φ} está relacionada con $M(\varphi)$ através de la siguiente igualdad:

$$M(\varphi) = M_{p+1}(V_0^{\varphi},V_1^{\varphi},\cdots,V_p^{\varphi}) \qquad (4.22)$$

El principal resultado de esta sección es el siguiente.

Teorema 4.3. Supongase que $\varphi(Z) = 1+ \varphi_1 Z +\cdots+ \varphi_p Z^p$ tiene grado p y satisface $\varphi(b) = \varphi(1/b) = 0$ donde $b \in \mathbb{C} -\{0\}$. Entonces (i)-(iii) ocurren.

(i) $\operatorname{Dim} \operatorname{span} \{ V^{\varphi}_{-1}, V^{\varphi}_{1}, \cdots, V^{\varphi}_{p+1}\} \le p + 1$.

(ii) Para cada $a \in \mathbb{C}$,

$$\operatorname{Det} M_{p+2}(V^{\varphi}_{0} + a\cdot V^{\varphi}_{-1},\ V^{\varphi}_{1} + a\cdot V^{\varphi}_{0}, \cdots, V^{\varphi}_{p+1} + a\cdot V^{\varphi}_{p}) = 0.$$

(iii) Para todo $a \in \mathbb{C}$, $\operatorname{Det} M[(1 + aZ)\varphi] = 0$.

La demostración de este resultado es algo técnica y se ha dividido en los Lemas 4.1-4.5 que aparecen a continuación. Estos resultados auxiliares se refieren a las siguientes nociones.

Definición 4.3. Sea $V = \{V_t | t \in \mathbb{Z}\}$ una secuencia en $\mathfrak{L}$.

(i) V tiene propiedad $\mathfrak{D}(p)$ si, para todo $n \in \mathbb{N}$

$$\operatorname{Dim} \operatorname{span} \{ V_t | -n \le t \le p+n \} \le p+n.$$

(ii) Dado $a \in \mathbb{C}$, la secuencia $T_a V = \{ T_a V_t |\ t \in \mathbb{Z} \}$ está definida mediante

$$T_a V_t := V_t + aV_{t-1},\ t \in \mathbb{N}.$$

Ahora se iniciará el camino hacia una demostración del Teorema 4.2. La primera etapa es la siguiente.

Lema 4.1. Sea $\varphi(z) = 1+ \varphi_1 Z +\dots+ \varphi_p z^p$ un polinomio de grado p y $a \in \mathbb{C}-\{ 0 \}$. Si $\theta(z) = (1 - az)\varphi(z)$, entonces

$$V^{\theta} = T_a V^{\varphi}.$$

Demostración. Sea $n \in \{1,2,\cdots,p\}$. A partir de la Definición 4.2 se sigue que

$$V_n^{\theta}(0) = \theta_n = \varphi_n + a\cdot\varphi_{n-1} = V_n^{\varphi}(0) + a\cdot V_{n-1}^{\varphi}(0) = T_a V_n^{\varphi}(0),$$

y para $k = 1,2,\cdots,$

$$\begin{aligned} V_n^{\theta}(k) &= \theta_{n+k} + \theta_{n-k} \\ &= (\varphi_{n+k} + a\cdot\varphi_{n+k-1}) + (\varphi_{n-k} + a\cdot\varphi_{n-k-1}) \\ &= (\varphi_{n+k} + \varphi_{n-k}) + a\cdot[\varphi_{n-1-k} + \varphi_{n-1+k}] \\ &= V_n^{\varphi}(k) + a\cdot V_{n-1}^{\varphi}(k) \\ &= T_a V_n^{\varphi}(k). \end{aligned}$$

Se Concluye que, para $1 \le n \le p$, $V_n^{\varphi} = T_a V_n^{\varphi}$. Para completar la demostración esta igualdad debe ser verificada para $n \le 0$ y $n > p$. Con este fin, primero observe que

$$\vec{\theta} = \vec{\varphi} + a\cdot s(\vec{\varphi}) \quad y \quad \overleftarrow{\theta} = s(\overleftarrow{\varphi}) + a\cdot\overleftarrow{\varphi};$$

esto se desprende de (4.18)-(4.21). Entonces, para $n \geq 0$,

$$V^{\theta}_{-n} = s^{n}(\vec{\theta}) = s^{n}[\vec{\varphi} + a\cdot s(\vec{\varphi})].$$

$$= s^{n}(\vec{\varphi}) + a\cdot s^{n+1}(\vec{\varphi})$$

$$= V^{\varphi}_{-n} + a\cdot V^{\varphi}_{-n-1}$$

$$= T_a V^{\varphi}_{-n}.$$

Similarmente,

$$V^{\theta}_{n+p+1} = s^{n}(\overleftarrow{\theta}) = s^{n}[s(\overleftarrow{\varphi})\ a\cdot\overleftarrow{\varphi}\ ]$$

$$= s^{n+1}(\overleftarrow{\varphi}) + a\cdot s^{n}(\overleftarrow{\varphi})$$

$$= V^{\varphi}_{n+1+p} + a\cdot V^{\varphi}_{n+p}$$

$$= T_a\ V^{\varphi}_{n+p+1}.$$

Así, se ha visto que $V^{\theta}_{-n} = T_a V^{\varphi}_{-n}$ y $V^{\theta}_{n+p+1} = T_a\ V^{\varphi}_{n+p+1}$ para todo $n \in \mathbb{N}$, y como ya se ha notado, ésto completa la demostración. ∎

El siguiente resultado muestra cómo se comporta una transformación T_a respecto a la propiedad $\mathfrak{D}(p)$.

Lema 4.2. Sea $a \in \mathbb{C}$ arbitrario y suponga que

$V = \{V_t | t \in \mathbb{N}\} \subset \mathfrak{L}$ tiene la propiedad $\mathfrak{D}(k)$.

Entonces T_aV tiene la propiedad $\mathfrak{D}(k+1)$.

Demostración. Notese que $T_aV_t \in$ span $\{ V_t, V_{t-1}\}$ y entonces, para $r \in \mathbb{N}$ arbitrario,
$\text{span}\{T_aV_t | -r \le t \le k+1+r\} \subset \text{span}\{V_t | -(r+1) \le t \le k+(r+1)\}$.
Ahora observese que si V tiene la propiedad $\mathfrak{D}(k)$, el espacio al lado derecho de la relación anterior tiene dimensión $\le k + r + 1$ y, consecuentemente,

$$\text{Dim span } \{ T_aV_t | -r \le t \le (k+1)+ r \} \le (k + 1) + r,$$

lo cual demuestra que T_aV tiene propiedad $\mathfrak{D}(k+1)$. ∎

Los siguientes dos lemas relacionan la propiedad $\mathfrak{D}(k)$ con secuencias V^{φ}.

Lema 4.3. Sea $\varphi(z)$ un polinomio de grado $p \ge 1$ y suponga que $\varphi(1)= 0$ ó $\varphi(-1)= 0$. Entonces,

$$V^{\varphi} \text{ tiene la propiedad } \mathfrak{D}(p).$$

Demostración . Primero supongase que $p = 1$. Cuando $\varphi(1) = 0$ se tiene $\varphi(z) = \varphi(0)\ (1-z)$ y, usando (4.21), se ve que $\vec{\varphi} = -\overleftarrow{\varphi}$, el cual implica que para $n \ge 0$, $V^{\varphi}_{-n} = s^n(\vec{\varphi}) = -s^n(\overleftarrow{\varphi}) = -V^{\varphi}_{1+n}$. Si $\varphi(-1) = 0$ se sigue que $\varphi(z) = \varphi(0)\ (1 + z)$ y entonces $\vec{\varphi} = \overleftarrow{\varphi}$ y se concluye que $V^{\varphi}_{-n} = V^{\varphi}_{1+n}$, $n \in \mathbb{N}$. Por lo tanto, en cualquier caso se ve que para todo $r \in \mathbb{N}$,

$$\text{span}\{ V_t^{\varphi} \mid -r \le t \le 1+ r \} = \text{span}\{ V_t^{\varphi} \mid 1 \le t \le 1+ r\},$$

y como el espacio al lado derecho tiene r + 1 generadores concluimos que Dim span{ V_t^{φ} $|-r \le t \le 1+ r\} \le r + 1$, i.e, V^{φ} tiene la propiedad $\mathfrak{D}(1)$; vea la Definición 4.3(i).

Completaremos la demostración por inducción en p. Suponga que el resultado es verdadero para p = k, y sea φ un polinomio de grado k+1 con raices 1 o -1. En este caso es claramente posible factorizar φ como $\varphi(z) = (1+ az)\theta(z)$ donde $a \in \mathbb{C}$, y $\theta(z)$ tiene grado k y satisface $\theta(1) = 0$ ó $\theta(-1) = 0$.
Por la hipótesis de inducción V^{θ} tiene la propiedad $\mathfrak{D}(k)$, y entonces $V^{\varphi} = T_a V^{\theta}$ tiene la propiedad $\mathfrak{D}(k + 1)$, por el lema anterior. ■

Lema 4.4. Sea $\varphi(z)$ un polinomio de grado $p \ge 2$ tal que $\varphi(b) = \varphi(1/b) = 0$ para algún $b \in \mathbb{C} -\{0,-1,1\}$. Entonces, V^{φ} tiene propiedad $\mathfrak{D}(p)$.

Demostración. El argumento es similar al usado en la demostración del Lema 4.3. Primero supongamos que p = 2. En este caso

$$\varphi(z) = \varphi(0)\cdot(1 - z/b)\cdot(12 - bz)$$

$$= \varphi(0)\cdot[1 - (b+b^{-1})z + z^2],$$

y usando (4.2.1) vemos que $\vec{\varphi} = -\overleftarrow{\varphi}$. Por la definición 4.1(i) esto nos dice que $V^{\varphi}_{-n} = V^{\varphi}_{2+n}$, $n \geq 0$, de donde se concluye que para todo $r \in \mathbb{N}$,

$$\operatorname{span}\{V^{\varphi}_{t} \mid -r \leq t \leq 2+r\} = \operatorname{span}\{V^{\varphi}_{t} \mid 1 \leq t \leq 2+r\},$$

y como el lado derecho tiene $r + 2$ generadores concluimos que Dim span $\{V^{\varphi}_{t} \mid -r \leq t \leq 2+ r\} \leq r + 2$, i.e, V^{φ} tiene la propiedad $\mathfrak{D}(2)$. El resultado para p arbitrario se obtiene por un argumento de inducción similar al usado en la demostración del Lema 4.3. ∎

Se Concluye esta sección con un resultado sera muy útil en la demostración del Teorema 4.2.

Lema 4.5. Sea $\varphi(z)$ un polinomio de grado p con $\varphi(0) = 1$. Entonces,

$$\operatorname{Det} M(\varphi) = \operatorname{Det} M_{p+2}(V^{\varphi}_{0}, V^{\varphi}_{1}, \ldots V^{\varphi}_{p+1}).$$

Demostración. Por conveniencia defina

$$L := M_{p+2}(V^{\varphi}_{0}, V^{\varphi}_{1}, \ldots V^{\varphi}_{p+1})$$

y observemos que

La submatriz obtenida eliminando la fila $p + 1$ así como la columna p + 1 de L es $M_{p+1}(V_0^{\varphi}, V_1^{\varphi}, \ldots V_p^{\varphi})$.

Ahora se calculan las componentes de la última columna de L. Primero recuerde la Definición 4.2(i) y note que

$$L_{0,p+1} = V_0^{\varphi}(p+1) = s^0(\vec{\varphi})(p+1) = \vec{\varphi}(p + 1) = 0;$$

vea (4.14), (4.18) y (4.19). Por otro lado, para $1 \leq n \leq p$,

$$L_{n,p+1} = \varphi_{n+p+1} + \varphi_{n-p-1} = 0,$$

donde hemos usado (4.14) en la última igualdad.

Finalmente,

$$L_{p,p+1} = V_p^{\varphi}(p+1) = s^0(\overleftarrow{\varphi})(p+1) = \overleftarrow{\varphi}(p+1) = 0,$$

y

$$L_{p+1,p+1} = s(\overleftarrow{\varphi})(p+1) = \overleftarrow{\varphi}(p) = \varphi_0 = \varphi(0) = 1.$$

Resumiendo:

(b) La ultima columna de L consiste totalmente de ceros excepto por el elemento en el último renglón, el cual es uno.

Para concluir, se expande Det L por la pasada columna. En éste caso, (a) y (b) implican que

$$\text{Det } L = \text{Det } M_{p+1}(V_0^{\varphi}, V_1^{\varphi}, \ldots V_p^{\varphi}) = \text{Det } M(\varphi),$$

donde la última igualdad se sigue de (4.22). ∎

Finalmente ya se tienen los argumentos necesarios para demostrar el Teorema 4.3.

Demostración del Teorema 4.3. Sea $\varphi(z) = 1+ \varphi_1 z +\cdots+ \varphi_p z^p$ un polinomio de grado p con $\varphi(b) = \varphi(1/b) =0$ para algún $b \in \mathbb{C} - \{0\}$.

(i) Cuando $b = 1$ o $b =-1$ el Lema 4.3 dice que V^φ tiene propiedad $\mathcal{D}(p)$ y, por Lema 4.4, lo mismo ocurre cuando $b \neq 1, -1$. Entonces, por la Definición 4.3(i) se concluye que

$$\text{Dim span } \{V^\varphi_{-1}, V^\varphi_0, \ . \ . \ . V^\varphi_{p+1}\} \le p+1.$$

(ii) Observese que

$\text{span}\{V^\varphi_r + a\cdot V^\varphi_{r-1} | r = 0,1,2,\cdots,p+1\} \subset \text{span}\{V^\varphi_t \ | -1 \le t \le p+1\}$, y entonces $\text{Dim span } \{V^\varphi_r + a\cdot V^\varphi_{r-1} | r = 0,1,2,\cdots, p+1\} \le p+1$, por (i) se sigue que los $p + 2$ vectores $V^\varphi_r + a\cdot V^\varphi_{r-1}$, $r = 0,1,2,\cdots,(p+1)$ son linealmente dependientes en $\mathfrak{L}$, y es claro que éste hecho implica la dependencia lineal de los renglones de $M_{p+2}(V^\varphi_0, + a\cdot V^\varphi_{-1},\cdots,V^\varphi_{p+1}+ a\cdot V^\varphi_p)$, y entonces

$$\text{Det } M_{p+2}(V^\varphi_0, +a\cdot V^\varphi_{-1},\cdots,V^\varphi_{p+1}+ a\cdot V^\varphi_p) = 0.$$

(iii) Defina $\psi(z):= (1 + az)\cdot\varphi(z)$. Entonces ψ tiene grado $p+1$, y usando (4.22) con $p+1$ y en lugar de p y ψ en lugar de φ, respectivamente, se ve que

$$M(\psi) = M_{p+2}(V_0^{\psi},\ V_1^{\psi},\ \ldots\ V_{p+1}^{\psi}),$$

$$= M_{p+2}(V_0^{\varphi} + a\cdot V_{-1}^{\varphi},\ \ldots\ V_{p+1}^{\varphi} + a\cdot V_p^{\varphi}),$$

donde se usa el Lema 4.1 para obtener la segunda igualdad. Finalmente, la parte (ii) nos da $M(\psi) = 0$. ■

4.6 DEMOSTRACION DEL TEOREMA 4.2

Después de los preliminares de la sección anterior, finalmente ha llegado la hora de establecer la validez del Teorema 4.2. se puede Iniciar el trabajo

Demostración del teorema 4.2. Como se mencionó anteriormente es suficiente probar la parte (i). Sea φ el polinomio de grado p dado por

$$\varphi(z) = 1 + \varphi_1 Z + \ldots + \varphi_p z^p,$$

y podemos factorizar a φ como

$$\varphi(z) = \prod_{i=1}^{p} (1 + a_i z),$$

donde las raices de φ son $-1/a_i$, $i = 1,2,\cdots,p$. Con esta notación se deberá demostrar que

$$\text{Det } M[\prod_{i=1}^{p}(1+ a_i z)] = \prod_{1 \leq i<j\leq p} [1-a_i a_j] \prod_{i=1}^{p}(1-a_i^2). \quad (4.23)$$

Se verificará esta igualdad mediante un argumento de inducción. Supongase que (4.22) es verdadero para $p = n \geq 2$, y sean $a_1, a_2, \ldots a_{n+1}$ numeros complejos no nulos. Podemos Definir

$$\psi(z) := \prod_{i=1}^{p} (1 + a_i z), \tag{4.23}$$

y para $c \in \mathbb{C}$ defina $F(c)$ por

$$F(c) = \operatorname{Det} M_{n+2}(V_0^{\psi} + c \cdot V_{-1}^{\psi}, \ldots V_{n+1}^{\psi} + c \cdot V_n^{\psi}). \tag{4.24}$$

Combinando los Lemas 4.1 y 4.2 obtenemos

(a) $F(c) = \operatorname{Det} M[(1 + cz) \cdot \psi(z)]$; en particular,

$$F(a_{n+1}) = \operatorname{Det} M[\prod_{i=1}^{n+1} (1 + a_i z)]. \tag{4.25}$$

Por otro lado, usando la multilinealidad de la función determinante (4.24) implica que

(b) $F(c)$ es un polinomio en c con grado $\leq n + 2$;

(c) $F(1) = F(-1) = 0$.

Para verificar (c), sea $\psi^*(z) := (1 + z) \prod_{i=2}^{n} (1 + a_i z)$ y note que $\psi^*(-1) = 0$, y que $(1+z)\ \psi(z) = (1 + a_1 z)\psi^*(z)$; vea (4.23). Entonces el inciso (a) anterior $F(1) = \operatorname{Det} M[(1+z)\ \psi(z)] = \operatorname{Det} M[(1+a_1 z)\psi^*(z)] = 0$, donde la anterior igualdad se justifica por el Teorema 4.3(iii) con ψ y -1 en lugar de φ y b, respectivamente. Similarmente puede verse que $F(-1) = 0$.

(d) $F(1/a_i) = 0$, $i = 1,2,\cdots,n$.

Para demostrar esto, seleccione $k,i \in \{1,2,\cdots,n\}$ con $k \neq i$, y definamos

$$\tilde{\psi}(z) := (1+ z/a_i)\cdot\Pi^{(k)}(1+ a_j z), \tag{4.26}$$

donde $\Pi^{(k)}$ indica el producto sobre todo $j \neq k$, $1 \leq j \leq n$; note que

$$(1+ z/a_i)\cdot\psi(z) = (1+ a_k z)\cdot\tilde{\psi}(z).$$

A partir de (4.26) es claro que $\tilde{\psi}(-a_i)=0$ y, como $k \neq i$, $\tilde{\psi}(z)$ contiene el factor $(1+ a_i z)$, de donde se desprende que $\tilde{\psi}(-1/a_i) = 0$. Por lo tanto, usando (a) se obtiene $F(1/a_i) = \operatorname{Det} M[(1+ z/a_i)\cdot\psi(z)] = \operatorname{Det} M[(1+ a_k z)\cdot\tilde{\psi}(z)]$, y usando el teorema 4.3(iii) con $\tilde{\psi}$ y $-a_i$ en lugar de φ y b, respectivamente, se concluye que $F(1/a_i) = 0$.

Para continuar supongase, *por el momento* que, $a_1,a_2,\cdots,a_n$ son diferentes números en $\mathbb{C}-\{ 0,1,-1 \}$. En este caso, (c) y (d) muestran que el polinomio $F(c)$ tiene $n + 2$ raices, a saber, $1,-1$ y $1/a_i$, $i = 1,2,\cdots,n$. Combinando este hecho con (b) vemos que $F(\cdot)$ tiene grado $n + 2$ y puede ser factorizado como

$F(c) = (1- c)(1+ c)\prod_{i=2}^{n}(1-a_i c)\,F(0)$. Poniendo $c = a_{n+1}$ y usando (a) se sigue que

$$\operatorname{Det} M[\prod_{i=1}^{n+1}(1-a_i z)] = (1- a_{n+1}^2)\cdot\prod_{i=1}^{n}(1- a_i a_{n+1})\cdot F(0). \qquad (4.27)$$

Ahora usando (4.24) se ve que $F(0) = \operatorname{Det}M_{n+2}(V_0^{\psi}, V_1^{\psi}, \ldots, V_{n+1}^{\psi})$ y entonces $F(0) = M(\psi)$, por el Lema 4.5 aplicado a ψ. Ahora, la hipótesis de inducción implica que

$$F(0) = \prod_{i=1}^{n}(1 - a_i^2)\prod_{1 \le i < j \le n}(1 - a_i a_j),$$

y combinando esta igualdad con (4.27) se obtiene

$$\operatorname{Det} M\,[\prod_{i=1}^{n+1}(1- a_i z)] = \prod_{i=1}^{n+1}(1- a_i^2)\prod_{1 \le i < j \le n}(1- a_i a_j), \qquad (4.28)$$

el cual es (4.13) con $p = n+1$. Aunque (4.28) ha sido establecido bajo la hipótesis que $a_1, a_2, \cdots, a_n$ son números diferentes en $\mathbb{C} - \{0\}$, debido a que ambos lados de (4.28) son funciones continuas de las a_i's, la igualdad es válida para $a_1, a_2, \cdots, a_n$ arbitrarios. En resumen, suponiendo que (4.23) se satisface para $p = n$, hemos demostrado que también se cumple para $p = n + 1$. Esto completa la demostración del Teorema 4.2. ∎

V PROCESOS A R M A

INTRODUCCION

En éste capítulo, se expone una combinación de los procesos autorregresivos y procesos de medias móviles. Tales procesos son del tipo mixto y permiten ampliar de manera considerable la clase de modelos para series de tiempo estacionarias.

Los modelos mixtos ocurren frecuentemente en situaciones prácticas dada la flexibilidad que tienen para representar fenómenos aleatorios de tipo estacionario. Los modelos mixtos tienen, además, una propiedad muy importante que puede describirse, de manera algo informal, como sigue: en comparación con los modelos puramente autorregresivos o puramente de medias móviles, un modelo mixto requiere menos parámetros para brindar una buena aproximación de un fenómeno aleatorio.

Este último capítulo se organizó de la siguiente manera: En la sección 5.2 se introducen los procesos ARMA, así como las definiciones de causalidad e inversibilidad. Por último, en la sección 5.3 se describen tres métodos para calcular la función de autocovarianza para procesos ARMA.

5.2. EL MODELO ARMA

Los modelos autorregresivos y de promedios móviles de orden (p,q), usualmente abreviados ARMA(p,q), son con frecuencia utilizados para describir y pronosticar observaciones en series de tiempo, y están definidos de la siguiente manera.

Definición 5.1. Una serie $\{ X_t | \ t \in \mathbb{Z} \}$ estacionaria con media cero se dice que es un proceso ARMA de orden p,q (ARMA (p,q)) si satisface una ecuación (de diferencias) de la forma

$$X_t - \varphi_1 X_{t-1} - \ldots - \varphi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q} \qquad t \in \mathbb{Z} \qquad (5.1)$$

donde, como antes, las Z_t's son variables aleatorias no correlacionadas con varianza común $\sigma^2 > 0$, y $\varphi_1, \varphi_2, \ldots, \ \varphi_p, \ \theta_1, \theta_2, \ldots, \theta_q \ \in \mathbb{R}$. Observe ahora que (5.1) puede escribirse como

$$(1 - \varphi_1 B - \ldots - \varphi_p B^p) X_t = (1 + \theta_1 B + \ldots + \theta_q B^q) Z_t,$$

donde $B^k \ X_t = \ X_{t-k}$ (vea la sección 2.3 del capítulo 2) y podemos abreviar aún más la expresión anterior escribiendo

$$\varphi(B) \ X_t = \theta(B) \ Z_t, \qquad (5.2)$$

donde los polinomios $\varphi(z)$ y $\theta(z)$ están dados por

$$\left.\begin{array}{l}\varphi(z) := 1- \varphi_1 z - \ldots - \varphi_p z^p \\ \theta(z) := 1- \theta_1 z - \ldots - \theta_q z^q,\end{array}\right\} \qquad (5.3)$$

los cuales son el polinomio autorregresivo y de medias móviles del proceso, respectivamente.

Definición 5.2. Un proceso ARMA(p,q) definido por las ecuaciones (5.2) se dice que es causal (o, más precisamente, que es una función causal de $\{Z_t\}$ si existe una sucesión de constantes $\{\psi_j\}$ tal que

$$\left.\begin{array}{l}\text{a)}\ \sum_{j=0}^{\infty} |\psi_j| < \infty, \\ \qquad \text{y} \\ \text{b)}\ X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}\end{array}\right\} \qquad (5.5)$$

Note que, la causalidad no sólo es una propiedad del proceso $\{X_t\}$, sino de la relación entre los dos procesos $\{X_t\}$ y $\{Z_t\}$. También podremos decir que $\{X_t\}$ es causal, si éste se obtiene a partir de $\{Z_t\}$ por aplicación de un filtro lineal causal.

El siguiente teorema establece las condiciones necesarias y suficientes para que un proceso ARMA(p,q) sea causal. También proporciona la representación de X_t en términos de $\{ Z_s, s \leq t \}$.

Teorema 5.1. Sea $\{ X_t\}$ un proceso ARMA(p,q) para el cual los polinomios $\varphi(\cdot)$ y $\theta(\cdot)$ no tienen ceros en común. Enonces $\{ X_t\}$ es causal si y sólo si $\varphi(z) \neq 0$ para todo $z \in \mathbb{C}$ tal que $|z| \leq 1$. Los coeficientes $\{ \psi_t\}$ en (5.5) están determinados por la relación

$$\psi(z) = \sum_{j=0}^{\infty} \varphi_j z^j = \theta(z)/\varphi(z) \qquad |z| \leq 1. \tag{5.6}$$

Para una demostración vea Brockwell y Davis (1987).

En los procesos ARMA(p,q), la causalidad es de gran importancia, pero existe otra condición, la cual está íntimamente relacionada con la causalidad, tal es el caso de la noción de **inversibilidad** que a continuación se define.

Definición 5.3. Un proceso ARMA(p,q) definido por (5.2) se dice que es inversible si existe una secuencia de constantes $\{ \Pi_j \}$ tal que $\sum_{j=0}^{\infty}|\Pi_j| < \infty$, además

$$Z_t = \sum_{j=0}^{\infty} \Pi_j X_{t-j} \quad ,t \in \mathbb{Z}. \tag{5.7}$$

Así como la causalidad, la propiedad de inversibilidad no sólo es una propiedad del proceso $\{ X_t\}$, sino de las relaciones de los dos procesos $\{ X_t\}$ y $\{ Z_t\}$. En el siguiente teorema, cuya demostración puede encontrarse en Brockwell y Davis (1987), Box y Jenkins (1976), o Fuller (1976). A continuación daremos las condiciones necesarias y suficientes para que $\{ X_t\}$ sea inversible.

Teorema 5.2. Sea $\{ X_t\}$ un proceso ARMA(p,q) para el cual los polinomios $\varphi(\cdot)$ y $\theta(\cdot)$ no tienen ceros en común. Entonces $\{ X_t\}$ es inversible si y sólo si $\theta(z) \neq 0$ para todo $z \in \mathbb{C}$ tal que $|z| \leq 1$.

Por otro lado, es interesante notar que los coeficientes $\{ \Pi_j\}$ en (5.7) están determinados por la relación

$$\Pi(z) = \sum_{j=0}^{\infty} \Pi_j z^j = \varphi(z)/ \theta(z), \quad |z| \leq 1. \tag{5.8}$$

Ejemplo 5.1. A continuación determinaremos la causalidad e inversibilidad de varios procesos ARMA.

(i) Sea $X_t + 1.9X_{t-1} + 0.88X_{t-2} = Z_t + 2Z_{t-1} + 1/3Z_{t-2}$. Note que el polinomio autorregresivo es

$$\varphi(z) = 1- 1.9z + 0.88z^2 = (1 + 1.1z)(1 + 0.88z)$$

cuyas raices son

$$r_1 = 1/1.1 \quad y \quad r_2 = 1/0.8$$

Por lo tanto, se verá que $\{X_t\}$ no es causal ya que $|r_1| < 1$, es decir, una raíz no está fuera del círculo unitario.

Se Analiza ahora la parte de promedios móviles cuyo polinomio es

$$\theta(z) = 1 + 2z + 1/3z^2,$$

Las raices de $\theta(z)$ son r_1 y r_2, donde

$$r_i = [-2 \pm (2^2 -4(1/3))^{1/2}] /(2/3)$$

$$= [-2 \pm (8/3)^{1/2}]/(2/3)$$

$$= [-3 \pm (6)^{1/2}],$$

y entonces

$$r_1 = -0.55 \quad y \quad r_2 = -5.45.$$

Notese que, tales raíces **no cumplen** con las condiciones de inversibilidad, es decir, solo una raíz tiene valor absoluto mayor que 1.

(ii) Sea $X_t + 1.6X_{t-1} = Z_t - 0.4Z_{t-1} + 0.04Z_{t-2}$.

Como en el proceso anterior, el polinomio autorregresivo es

$$\varphi(z) = 1 + 1.6z,$$

y su única raíz es $r_1 = -5 / 8$. Por lo tanto el proceso **no es causal** ya que $|r_1| \leq 1$.
Por otro lado,

$$\theta(z) = 1 - 0.4z + 0.004z^2 = (1 + 0.2z)^2,$$

y este proceso tiene la única raíz (doble) $r = 5$, y se desprende que $\{ X_t \}$ es inversible. □

Para concluir esta sección se verá un ejemplo donde se analiza la inversibilidad de un proceso ARMA(p,q).

Ejemplo 5.2. Suponga que $\{ X_t \}$ es un proceso ARMA(p,q) inversible como en (5.1) con

$$Z_t = \sum_{j=0}^{\infty} \Pi_j X_{t-j}$$

se demostará que la secuencia $\{ \Pi_j \}$ está determinada por las ecuaciones

$$\Pi_j + \sum_{k=1}^{\min(q,j)} \theta_k \Pi_{j-k} = - \varphi_j, \; j \in \mathbb{Z}^+,$$

donde definimos $\varphi_0 = -1$, $\theta_k = 0$ para $k > q$ y $\varphi_j = 0$ para $j > p$.

Recuerdese que $\Pi(z)$ esta determinada por

$$\varphi(z) = \Pi(z)\, \theta(z), \tag{5.9}$$

o más explicitamente,

$$-\sum_{k=0}^{p} \varphi_k z^k = \sum_{j=0}^{\infty} \Pi_j z^j \sum_{r=0}^{q} \theta_r z^r$$

$$= \sum_{j=0}^{\infty} \left(\sum_{k=0}^{j} \Pi_{j-k}\, \theta_k \right) z^j,$$

donde $\theta_j = 0$ y $s > q$, por lo tanto para $j \geq 0$,

$$\sum_{k=0}^{j} \Pi_{j-k}\, \theta_k = - \varphi_j \tag{5.10}$$

donde $\varphi_j = 0$ para $j > p$, En esta ecuación, en lugar de sumar desde k=0 a k=j, es suficiente sumar para k=0 a k= min(j,k). Así (5.10) es equivalente a

$$\sum_{k=0}^{\min(q,j)} \Pi_{j-k} \theta_k = - \varphi_j,$$

y, usando que $\theta = 1$, se concluye que

$$\Pi_j + \sum_{k=1}^{\min(j,q)} \Pi_{j-k}\, \theta_k = - \varphi_j. \qquad (5.11)$$

5.3. DETERMINACION DE LA FUNCION DE AUTOCOVARIANZA

Se inicia esta sección, diciendo que $\{ X_t | \ t \in \mathbb{Z}\}$ definida en términos de las ecuaciones lineales en diferencia, implica una familia paramétrica de pocesos estacionarios, los procesos autorregresivos y de promedios móviles o ARMA(p,q). para cualquier función de autocovarianza $\gamma(\cdot)$ tal que lim $h \longrightarrow \infty$ $\gamma(h) = 0$, y para cualquier $k > 0$, es posible encontrar un proceso ARMA(p,q) con función de autocovarianza $\gamma_x(\cdot)$ tal que, $\gamma_x(h) = \gamma(h)$, $h = 0,1,2,\ldots,k$. Es por esto que la función de autocovarianza juega un papel importante en la selección de un modelo ARMA para modelar series de tiempo.

A continuación describiremos tres métodos para la obtención de la función de autocovarianza.

Primer Método. La función de autocovarianza γ de el proceso ARMA(p,q) causal $\varphi(B)X_t = \theta(B)Z_t$ es:

$$\gamma(k) = \sigma^2 \sum_{j=0}^{\infty} \psi_i \psi_{j+|k|} \qquad (5.12)$$

donde

$$\psi(z) = \sum_{j=0}^{\infty} \varphi_j z^j = \theta(z)/\varphi(z) \text{ para } |z| \leq 1. \qquad (5.13)$$

Debido a que $\theta(\cdot)$ y $\varphi(\cdot)$ son los polinomios de grado q y p respectivamente, se puede reescribir (5.13) en la forma $\psi(Z)\varphi(Z) = \theta(Z)$ e igualar los coeficientes de z^j para obtener

$$\psi_j - \sum_{0<k\leq j} \varphi_k\psi_{j-k} = \theta_j \qquad 0\leq j< \max(p,q+1) \qquad (5.14)$$

y

$$\psi_j - \sum_{0<k\leq p} \varphi_k\psi_{j-k} = 0 \qquad j \geq \max(p,q+1) \qquad (5.15)$$

Por lo tanto se puede resolver (5.14) y (5.15) sucesivamente para $\psi_0, \psi_1, \ldots$ Así

$$\left.\begin{aligned} &\psi_0 = \theta_0 = 1 \\ &\psi_1 = \theta_1 + \psi_0\varphi_1 = \theta_1 + \varphi_1 \\ &\psi_2 = \theta_2 + \psi_0\varphi_2 + \psi_1\varphi_1 = \theta_2 + \varphi_2 + \theta_1\varphi_1 + \varphi_1^2 \end{aligned}\right\} \quad (5.16)$$

.

.

.

etc. una vez que se han calculado los números ψ_j, la función de autocovarianza se determina usando (5.12).

Segundo Método. Un procedimiento alternativo para calcular la función de autocovarianza $\gamma(\cdot)$ de un proceso ARMA(p,q) causal está basado en las ecuaciones en diferencia para

$\gamma(k)$, $k=0,1,2,\ldots$, las cuales son obtenidas al multiplicar a ambos lados de (5.4) por X_{t-k} y al tomar esperanzas se obtiene:

$$\gamma(k) - \varphi_1\gamma(k-1) - \ldots \varphi_p\gamma(k-p) = \sigma^2 \sum_{k\leq j\leq q} \theta_j \psi_{j-k} \quad (5.17)$$

para

$$0 \leq k \leq \max(p,q+1) \quad \text{y}$$

$$\gamma(k) = \varphi_1\gamma(k-1) - \ldots \varphi_p\gamma(k-p) \quad (5.18)$$

para

$$k > \max(p,q+1).$$

Tercer metodo.- La determinación numérica de la función de autocovarianza $\gamma(\cdot)$ de las ecuaciones (5.17) y (5.18) puede ser llevada a cabo primero encontrando $\gamma(0),\gamma(1),\ldots,,\gamma(p)$ de las ecuaciones con $k= 0,1,2,\ldots,p$ y entonces usando las ecuaciones subsecuentes para determinar $\gamma(p+1),\gamma(p+2),\ldots$ de manera recursiva.

LITERATURA CITADA

Achilles, M.S. (1987), Unique Solution of the Yule-Walker systems Metrika, 34, 237-251.

Anderson, T.W. (1971), The statistical analysis of time series, John Wiley, New York 370 p.

Berk, K.N. (1974), Consistent Autorregresive Spectral Estimates, Annals of Statistics, 2, 489-502.

Box, G.E.P, and Jenkins G.W.M: (1976), Time series analysis forecasting and control, Holden Day, 46 p.

Box, G.E.P. and Pierce, D.A. (1970), Time Series Analysis; Forecasting and Control, Holden-Day, San Francisco.

Brockwell,P.J., and Davis R.A. (1987), Time series theory and methods, Springer, New York. 77 p.

Brockwell,P.J., and Davis R.A. (1988), Aplication of the Innovation Algorithm to estimate the coefficient of filter, stochastic, 25, 129-143.

Bockwell,P.J., and Davis R.A. (1989), Simple Consistent estimates of the coefficients of a moving average process, Applied probability, 17, 143-151.P

Cavazos-Cadena R. (1990), Fundamentos de estadística, parte I, Departamento de Matemáticas CINVESTAV-IPN, México, D.F. 149 p.

Cavazos-Cadena R. (1992), On the Yule-Walker Ecuations in Time Series Analysis, Report N°, 01-05-92, Departamento de Estadística y Cálculo, UAAAN, México.

Dudewicz E.J.. and Mishira K. (1988), Modern Mathematical Statistics, John Wiley, New York.

Fuller, W.A. (1976), Introduction to Statistical Time Series, John Wiley, New York. 460 p.

Kendall, M.G. and Stuart.A. (1976), The Advanced Theory of Statistic, Vol. 3, Griffin, London. 430 p.

Lütkepohi J., and Maschke R.S. (1988), A Simple proof of the Uniquenese of the solution of the Yule-Walker Ecuations Metrika 35, 287-289.

Melard, G. (1984), A fast algorithm for the exact likelihood of moving average models, Applied statistic, 33, 104-114.

Mood, A.M, Graybill, F.A., and Boes D.C. (1974), Introduction to the theory of statistic, McGraw-Hill, New York. 520 p.