# LAS ECUACIONES DE YULE-WALKER EN EL ANÁLISIS DE SERIES DE TIEMPO

*ANA PAULA ISAIS TORRES*

TESIS

Presentada como Requisito Parcial para
Obtener el Grado de:

MAESTRÍA PROFESIONAL EN
ESTADÍSTICA APLICADA

UNIVERSIDAD AUTÓNOMA AGRARIA
"ANTONIO NARRO"
PROGRAMA DE GRADUADOS
Buenavista, Saltillo, Coahuila, México
Febrero de 2012

Universidad Autónoma Agraria Antonio Narro
Subdirección de Postgrado

LAS ECUACIONES DE YULE-WALKER EN EL ANÁLISIS DE SERIES DE TIEMPO

TESIS

Por:

ANA PAULA ISAIS TORRES

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

MAESTRÍA PROFESIONAL
EN ESTADÍSTICA APLICADA

Comité Particular

Asesor principal: ______________________
Dr. Rolando Cavazos Cadena

Asesor: ______________________
Dr. Mario Cantú Sifuentes

Asesor: ______________________
M. C. Félix de Jesús Sánchez Pérez

______________________
Dr. Fernando Ruiz Zárate
Subdirector de Postgrado
Buenavista, Saltillo, Coahuila, Febrero de 2012

COMPENDIO

LAS ECUACIONES DE YULE-WALKER EN EL ANÁLISIS DE SERIES DE TIEMPO

Por

ANA PAULA ISAIS TORRES

MAESTRÍA PROFESIONAL

EN ESTADÍSTICA APLICADA

UNIVERSIDAD AUTÓNOMA AGRARIA
ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA, Febrero de 2012

Dr. Rolando Cavazos Cadena –Asesor–

**Palabras clave:** Procesos autorregresivos, Ecuaciones de Yule-Walker, Ecuaciones de pronóstico, Solución única, Matriz invertible, Polinomio causal, Determinación de una función de autocovarianza.

Este trabajo trata sobre el sistema de ecuaciones lineales de Yule-Walker en el análisis de procesos autorregresivos. Dado un polinomio complejo $\varphi(z)$ que satisface $\varphi(0) = 1$, se determina la correspondiente matriz de Yule-Walker $M(\varphi)$ del sistema asociado a $\varphi$, y el principal resultado el trabajo consiste en obtener una fórmula explícita para el determinante de $M(\varphi)$ en terminos de las raíces de $\varphi(z)$. Tal resultado produce el siguiente criterio para la no singularidad de $M(\varphi)$: La matriz $M(\varphi)$ es invertible si, y sólo si, el producto de dos raíces $\varphi$ es siempre distinto de 1; esta propiedad implica que el sistema de ecuaciones de Yule-Walker asociado con un polinomio causal tiene una única solución. La forma en que este resultado se utiliza de manera implícita en la litertura se discute brevemente.

ABSTRACT

THE YULE-WAKER EQUATIONS IN
TIME SERIES ANALYSIS

BY

ANA PAULA ISAIS TORRES

MASTER

APPLIED STATISTICS

UNIVERSIDAD AUTÓNOMA AGRARIA
ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA, February, 2012

Dr. Rolando Cavazos Cadena –Advisor–

**Key Words:** Autoregressive processes, Yule-Walker equations, Forecasting equations, Unique solution, Invertible matrix Causal polynomial, Determination of an autocovariance function.

This work concerns the Yule-Walker system of linear equations arising in the study of autoregressive processes. Given a complex polynomial $\varphi(z)$ satisfying $\varphi(0) = 1$, the matrix $M(\varphi)$ of the Yule-Walker system associated with $\varphi$ is determined and the main result of this work establishes an explicit formula for the determinant of $M(\varphi)$ in terms of the roots of $\varphi(z)$. Such a result renders the following criterion for the non-singularity of $M(\varphi)$: The matrix $M(\varphi)$ is invertible if and only if the product of two roots of $\varphi$ is always different form 1, a property that yields that the Yule-Walker system associated with a causal polynomial has a unique solution. The way in which this result is implicitly used in the time series literature is briefly discussed.

# Índice de Contenido

# Capítulo 1

# INTRODUCCIÓN

El presente trabajo se ubica en un área conocida como el "análisis de series de tiempo", y particularmente trata sobre la idea de una *serie estacionaria* de segundo orden, un concepto que se discutirá más adelante. En el estudio de dichas series la noción de función de autocovarianza desempeña un papel central, especialmente con relación al problema de pronósticos, así que determinar dicha función es un problema de extrema importancia. Para una amplia clase de procesos estacionarios, concretamente, la familia de procesos ARMA, la función de autocovarianza se determina mediante un procedimiento recursivo, cuya primera etapa consite en resolver un sistema de ecuaciones lineales, denominado sistema de Yule-Walker; *el propósito central de este trabajo* es establecer que dicho sistema tiene una solución única, esto es que su matriz de coeficientes es no singular. En este capítulo se hace una descripción del contexto básico en que se desarrolla este trabajo, se describe el principal problema que se estudiará, y se presenta una visión general de la organización del material subsecuente.

## 1.1. Series Estacionarias

De manera intuitiva, una series de tiempo es una sucesión $\{X_t\}$ de variables aleatorias definidas sobre un mismo espacio de probabilidad, donde la variable aleatoria $X_t$ se interpreta como una observación que se realiza en el tiempo $t$, el cual, para el caso considerado, puede variar en un subconjunto de los números enteros. El rasgo fundamental de la sucesión $\{X_t\}$ es que, en contraste con el supuesto comúnmente adoptado en la teoría estadística clásica (Borovkov, 1999, Dudewicz y Mishra, 1998, Shao, 2010, Wackerly *et. al.*, 2009), no se supone la independencia de las variables $X_t$, ni que éstas tengan la misma distribución ó características que permiten incluir en el estudio una gran variedad de observaciones que surgen en la práctica, como las ventas diarias de un centro comercial, la asistencia semanal a cines, la población de un país, o la aparición de manchas solares (Brockwell y Davis, 1998, Shumway y Stoffer 2006). Por otro lado, una serie de tiempo $\{X_t\}$ se llama *estacionaria* cuando, para todo entero $h$, las propiedades "fundamentales"de un segmento $(X_1, X_2, \ldots, X_n)$ son las mismas que las del segmento trasladado $(X_{1+h}, X_{h+2}, \ldots, X_{n+h})$. Como no se harán supuestos sobre la distribución de las variables $X_t$, las propiedades "importantes"se referiran a los momentos: concretamente, si la serie es estacionaria satisface los siguientes dos requerimientos:

(i) Todas las variables $X_t$ tienen los mismos momentos de primer y segundo orden, es decir

$$E[X_t] \text{ y } E[X_t^2] \text{ son finitos y no dependen de } t.$$

(ii) El segundo momento de la diferencia $X_s - X_t$, es el mismo que el segundo momento de $X_{s+h} - X_{t+h}$ para todo entero $h$. Estas dos condiciones son equivalentes a las propiedades (a) y (b):

(a) $E[X_t] = \mu$ no depende de $t$;

(b) Para cada $s$ y $t$, $\mathrm{Cov}\,(X_s, X_t)$ está bien definida y depende sólo de la diferencia entre $s$ y $t$.

Estas propiedades permiten definir la función de autocovarianza asociada a la serie estacionaria $\{X_t\}$ mediante

$$\gamma(h) = E[X_{t+h}X_t], \quad h = 0, \pm, \pm 2, \ldots. \tag{1.1.1}$$

Para una serie estacionaria, en el sentido definido anteriormente, no es difícil observar que para cada entero $h$ y cada entero positivo $n$,

$$E[(X_1, X_2 \ldots, X_n)'] = E[(X_{1+h}, X_{2+h} \ldots, X_{n+h})']$$

y

$$\mathrm{Var}[(X_1, X_2 \ldots, X_n)'] = \mathrm{Var}[(X_{1+h}, X_{2+h} \ldots, X_{n+h})']$$

de manera que los segmentos $(X_1, X_2 \ldots, X_n)$ y $(X_{1+h}, X_{2+h} \ldots, X_{n+h})$ tienen el mismo valor esperado y la misma matriz de varianza; en este sentido, ambos segmentos tienen las mismas propiedades "importantes", y el análisis de una serie de tiempo estacionaria estará basado en propiedades de segundo orden. Debido a que si $\{X_t\}$ es una serie estacionaria, entonces $\{X_t - \mu\}$ también lo es, en este trabajo se supondrá, sin pérdida de generalidad, que la esperanza de $X_t$ es nula. La idea de serie estacionaria es bastante restrictiva y, en general, una de tales series, no es un modelo razonable para los datos que se observan comúnmente en la práctica; sin embargo, una serie estacionaria es la componente esencial del denominado *modelo clásico*, el cual captura una amplia gama de situaciones reales (Brockwell y Davis, 1998).

## 1.2. El Problema de Pronóstico

*El problema fundamental* para una serie estacionaria puede describirse como sigue: Dadas las observaciones $X_1, \ldots, X_n$ registradas hasta el tiempo $n$, determinar "la mejor"aproximación para la variable $X_{n+h}$ que se observará en el tiempo futuro $n + h$ en términos de los datos disponibles $X_1, \ldots, X_n$, el cual se denomina el *problema de pronóstico*. Un resultado conocido en la teoría estadística establece que la mejor aproximación a $X_{n+h}$ en el sentido de minimizar el error cuadrático esperado es la esperanza condicional $g(X_1, \ldots, X_n) = E[X_{n+h}|X_1, \ldots, X_n]$; sin embargo, al no conocer la distribución de $\{X_t\}$, la función $g(\cdot)$ no se puede calcular, y en la teoría de series de tiempo, la mejor aproximación que se pretende encontrar es "la mejor aproximación lineal basada en los datos observados hasta el momento actual $n$":

$$\hat{g}(X_1, \ldots, X_n) = a_1 X_1 + \cdots + a_n X_n \tag{1.2.2}$$

donde $a_1, \ldots, a_n$ son constantes y

$$\begin{aligned} \min_{b_1,\ldots,b_n} & E[((X_{n+h} - (b_1 X_1 + \cdots + b_n X_n))^2] \qquad (1.2.3) \\ &= E[((X_{n+h} - (a_1 X_1 + \cdots + a_n X_n))^2]. \qquad (1.2.4) \end{aligned}$$

tales números $a_i$ siempre existen y se expresan solamente en términos de la función de autocovarianza $\gamma(\cdot)$; de hecho, $a_1, \ldots, a_n$ pueden determinarse por medio de diferenciación (Fulks 1980, Khuri 2002).

En un sentido general, el primer problema de interés en este trabajo se refiere a la determinación de la función de autocovarianza de una clase especial de procesos estacionarios; como se analizará en el siguiente capítulo, dicha función de autocovarianza permite determinar los pronósticos de las observaciones futuras.

## 1.3. Proyecciones

Dado que este trabajo involucra la determinación de las constantes $a_i$ en (1.2.2), las cuales tienen la propiedad de optimalidad en (1.2.4), las ideas básicas de *espacio vectorial* y *producto interno* son de importancia fundamental en el desarrollo subsecuente; para un estudio de estos conceptos en espacios de dimensión finita, vea, por ejemplo, Graybill (2001), Harville (2008), Hoffman y Kunze (1975) y, particularmente, Strang (2003). La intuición del caso de dimensión finita es de gran importancia en el estudio de series de tiempo, aunque el entorno de trabajo es, naturalmente, el espacio

$$L^2 = L^2(\Omega, \mathcal{F}, P)$$

de todas las variables aleatorias $Y : \Omega \to R$ con segundo momento finito, y donde $(\Omega, \mathcal{F}, P)$ es el espacio de probabilidad donde la serie $\{X_t\}$ está definida. Este espacio es de dimensión infinita y el producto interno está especificado por

$$\langle Y_1, Y_2 \rangle = E[Y_1 Y_2], \quad Y_1, Y_2 \in L^2,$$

y la norma correspondiente está dada por

$$\|Y\| = \sqrt{\langle Y, Y \rangle} = E[Y^2]^{1/2}.$$

Con respecto a esta norma, $L^2$ es un espacio métrico completo, es decir, toda sucesión de Cauchy es convergente (Apostol 1980, Rudin 1984, Royden 2003). Considere ahora, una sucesión finita $W_1, \ldots, W_k$ de variables en $L^2$, y sea $\mathcal{W}$ el espacio vectorial generado por estas variables:

$$\mathcal{W} = \{c_1 W_1 + \cdots + c_n W_n \mid c_1, c_2, \ldots, c_n \in \mathbb{R}\}$$

el cual es un subespacio cerrado de $L^2$ (Rudin 1984, Royden 2003). Dado $Y \in L^2$, existe una única variable aleatoria $P_{\mathcal{W}}Y$ que satisface las siguientes propiedades:

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \|Y - P_{\mathcal{W}}Y\| \leq \|Y - \mathbf{w}\| \quad \text{para todo } \mathbf{w} \in \mathcal{W}. \tag{1.3.5}$$

Estas condiciones son equivalentes a

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \langle P_{\mathcal{W}}Y, \mathbf{w}\rangle = \langle Y, \mathbf{w}\rangle, \quad \mathbf{w} \in \mathcal{W}. \tag{1.3.6}$$

(Brockwell y Davis 1998, Strang 2003); debido a la (bi-)linealidad del producto interno, estas condiciones pueden expresarse como

$$P_{\mathcal{W}}Y \in \mathcal{W} \quad \text{y} \quad \langle P_{\mathcal{W}}Y, W_i\rangle = \langle Y, W_i\rangle, \quad i = 1, 2, \ldots, k. \tag{1.3.7}$$

La variable $P_{\mathcal{W}}Y$ se llaman la proyección (ortogonal) de $Y$ sobre $\mathcal{W}$ y es una transformación lineal; la equivalencia de (1.3.5–1.3.7) se conoce *teorema de proyección*. La siguiente propiedad será útil: Si $\mathcal{W}_0$ y $\mathcal{W}_1$ son dos subespacios de $L^2$, entonces

$$\mathcal{W}_0 \subset \mathcal{W}_1 \Rightarrow P_{\mathcal{W}_0}P_{\mathcal{W}_1} = P_{\mathcal{W}_0} \tag{1.3.8}$$

A partir de esta discusión es claro que el mejor pronóstico $g(X_1, \ldots, X_n)$ para $X_{n+1}$ dadas las observaciones $X_1, \ldots, X_n$, es la proyección de $X_{n+1}$ sobre el espacio generado por $X_1, \ldots, X_n$; compare (1.2.2) y (1.2.4) con (1.3.5).

## 1.4. El Objetivo Principal

Como se verá en el Capítulo 2, la determinación de prónosticos para una serie estacionaria involucra directamente a la correspondiente función de autocovarianza. Para una clase de modelos denomindados autorregresivos y de promédios móviles, el cálculo de la función de autocovarianza requiere resolver una sistema de ecuaciones lineales conocido como el sistema de Yule-Walker, y el objetivo central de este trabajo es demostrar que dicho sistema tiene solución única, esto es, que la matriz de coeficientes es invertible. Para alcanzar esta meta, se determinará una fórmula para el determinante de una matriz de Yule-Walker correspondiente a un proceso ARMA, y se verá que dicho determinante es diferente de cero.

## 1.5. La Organización

La presentación de este trabajo ha sido organizada de la siguiente manera:

En el Capítulo 2, dada una serie de tiempo estacionaria se formulan las ecuaciones que determinan el pronóstico $\hat{X}_{n+h}$ en términos de los datos actuales $X_1, \ldots, X_n$ y se encuentra una expresión simple para el correspondiente error de pronóstico $v_n^h$, analizando sus propiedades de monotonicidad y su comportamiento conforme $h$ se incrementa sin límite.

En el Capítulo 3 se introduce una clase importante de procesos estacionarios que surgen resolviendo ecuaciones en diferencias, los cuales se denomina procesos autorregresivos y de promedios móviles o, simplemente, procesos ARMA. La variablilidad de esos procesos se debe a la presencia de una serie denominada ruido blanco, la cual consiste de variables aleatorias no correlacionadas de varianza común, y que que influye en los datos observables de una manera simple, generando un modelo que es sencillo de analizar, pero lo suficientemente versátil para ser utilizado, ya que la función de autocovarianza de cualquier serie estacionaria $\{Y_t\}$ puede aproximarse, tanto como se desee, por la función de autocovarianza de un proceso ARMA adecuadamente seleccionado. La presentación incluye una discusión sobre la existencia y unicidad de procesos ARMA, y finaliza con la introducción del concepto de *causalidad* de un proceso respecto a un ruido blanco.

En el Capítulo 4 se describe el procedimento de dos fases para determinar la función de autocovarianza de un proceso ARMA , se formulan las ecuaciones de Yule-Walker, y se establece la fórmula sobre el determinate de la matriz de dicho sistema en el Teorema 4.4.1.

# Capítulo 2

# ECUACIONES DE PRONÓSTICO

Dada una serie de tiempo estacionaria, en este capítulo se formulan las ecuaciones que determinan la mejor aproximación lineal para una observación futura en términos de los datos disponibles, y se encuentra una expresión simple para el correspondiente error de pronóstico. Se analizan dos ideas sobre el comportamiento de los errores que, desde el punto de vista intuitivo, parece natural esperar que sean válidas. El estudio que se realiza muestra que dichas ideas no son, en general, correctas, pero un análisis más profundo muestra que, bajo condiciones adecuadas o después de una reformulación apropiada, el comportamiento de los errores está de acuerdo con la intuición.

## 2.1. Formulación de las Ecuaciones

Dada una serie estacionaria $\{X_t\}$ con función de autocovarianza $\gamma(\cdot)$, en esta sección se establecen las ecuaciones para determinar la proyección ortogonal de una variable $X_{n+h}$, donde $h > 0$, sobre el espacio generado por las variables $X_1, \ldots, X_n$ donde, por el momento, el entero $n$ es fijo, y $\mathcal{L}_n$ es el subespacio de $L^2$ generado por las variables $X_t$ con $1 \leq t \leq n$:

$$\mathcal{L}_n(X_1, \ldots, X_n) \equiv \mathcal{L}_n : = \{a_1 X_1 + \cdots + a_n X_n \mid (a_1, \ldots, a_n) \in \mathrm{I\!R}^n\}. \qquad (2.1.1)$$

Para cualquier variable aleatoria $Y$ con varianza finita, la proyección ortogonal de $Y$ sobre $\mathcal{L}_n$ se denotará mediante $\hat{Y}$:

$$\hat{Y} = P_{\mathcal{L}_n} Y, \tag{2.1.2}$$

de manera que $\hat{Y}$ está caracterizada por los siguientes requerimientos:

$$\hat{Y} \in \mathcal{L}_n, \quad \text{y} \quad \langle Y, X_k \rangle = \left\langle \hat{Y}, X_k \right\rangle, \quad k = 1, 2, \ldots, n; \tag{2.1.3}$$

vea el Teorema de proyección en el capítulo 2 de Brockwell y Davis (1998). En palabras, $\hat{Y}$ es una combinación lineal de $X_1, \ldots, X_n$, y los productos internos de $\hat{Y}$ y $Y$ con $X_i$ coinciden, $i = 1, 2, \ldots, n$. Ahora se determinará la forma que estas condiciones adoptan para el caso en que $Y$ es la variable $X_{n+h}$ que se observará $h$ unidades después de haber registrado el valor de $X_n$; primero, se estudiará el caso $h = 1$. Observe que la inclusión $\hat{X}_{n+h} \in \mathcal{L}_n$ significa que $\hat{X}_{n+h}$ es una combinación lineal de $X_1, \ldots, X_n$, de manera que existen constantes $\phi_{n\,1}, \phi_{n\,2}, \ldots, \phi_{n\,n}$ tales que

$$\hat{X}_{n+1} = \phi_{n\,1} X_n + \phi_{n\,2} X_{n-1} + \cdots + \phi_{n\,n} X_1,$$

o de forma más compacta,

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}, \tag{2.1.4}$$

A continuación, note que la condición de que $X_{n+1}$ y $\hat{X}_{n+1}$ tengan los mismos productos internos con las variables $X_r$, $r = 1, 2, \ldots, n$ puede escribirse en términos de la función de autocovarianza:

$$\begin{aligned} \langle X_{n+1}, X_r \rangle &= \left\langle \hat{X}_{n+1}, X_r \right\rangle \\ &= \left\langle \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}, X_r \right\rangle \\ &= \sum_{j=1}^{n} \phi_{n\,j} \langle X_{n+1-j}, X_r \rangle, \quad r = 1, 2, \ldots, n, \end{aligned} \tag{2.1.5}$$

igualdades que equivalen a

$$\gamma(n+1-r) = \sum_{j=1}^{n} \phi_{n\,j} \gamma(n+1-j-r), \quad r = 1, 2, \ldots, n.$$

Una forma más conveniente para este sistema se obtiene escribiendo

$$i = n + 1 - r;$$

con esta notación, cuando $r$ toma cualquier valor entre 1 y $n$, $i$ también lo hace y el anterior sistema de ecuaciones se escribe como

$$\gamma(i) = \sum_{j=1}^{n} \phi_{n\,j}\gamma(i-j), \quad i = 1, 2, \ldots, n.$$

Se definirá ahora los vectores $\boldsymbol{\gamma}_n, \boldsymbol{\varphi}_n \in \mathbb{R}^n$ y la matriz $\boldsymbol{\Gamma}_n$ de orden $n \times n$ mediante

$$\boldsymbol{\gamma}_n = \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n) \end{bmatrix}, \quad \boldsymbol{\varphi}_n = \begin{bmatrix} \phi_{n\,1} \\ \phi_{n\,2} \\ \vdots \\ \phi_{n,\,n} \end{bmatrix}, \quad \boldsymbol{\Gamma}_n = [\gamma(i-j)]_{i,j=1,2,\ldots,n}; \tag{2.1.6}$$

con esta notación, las ecuaciones anteriores equivalen a

$$\boldsymbol{\gamma}_n = \boldsymbol{\Gamma}_n \boldsymbol{\varphi}_n, \tag{2.1.7}$$

sistema que siempre es consistente, por el teorema de proyección. Después de encontrar una solución $\boldsymbol{\varphi}_n$ para esta ecuación vectorial, el error cuadrático de pronóstico, denotado por $v_n$ y definido mediante

$$v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2, \tag{2.1.8}$$

puede determinarse de forma muy simple. En efecto, observando que

$$\left\langle X_{n+1} - \hat{X}_{n+1}, X_r \right\rangle = 0$$

para todo $r = 1, 2, \ldots, n$ (vea la primera igualdad en (2.1.5)), como $\hat{X}_{n+1}$ es una combinación lineal de $X_1, \ldots, X_n$ se obtiene que

$$\left\langle X_{n+1} - \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle = 0, \tag{2.1.9}$$

lo que equivale a

$$\left\langle \hat{X}_{n+1}, X_{n+1} \right\rangle = \left\langle \hat{X}_{n+1}, \hat{X}_{n+1} \right\rangle = \|\hat{X}_{n+1}\|^2.$$

Con esto en mente,

$$\begin{aligned}
\|X_{n+1}-\hat{X}_{n+1}\|^2 &= \left\langle X_{n+1}-\hat{X}_{n+1}, X_{n+1}-\hat{X}_{n+1}\right\rangle \\
&= \left\langle X_{n+1}-\hat{X}_{n+1}, X_{n+1}\right\rangle - \left\langle X_{n+1}-\hat{X}_{n+1}, \hat{X}_{n+1}\right\rangle \\
&= \left\langle X_{n+1}-\hat{X}_{n+1}, X_{n+1}\right\rangle \text{ (por la ecuación (2.1.9))} \\
&= \langle X_{n+1}, X_{n+1}\rangle - \left\langle \hat{X}_{n+1}, X_{n+1}\right\rangle \\
&= \gamma(0) - \left\langle \hat{X}_{n+1}, \hat{X}_{n+1}\right\rangle
\end{aligned}$$

donde, de nueva cuenta, se utilizó la ecuación (2.1.9) para establecer la última igualdad. Por lo tanto,

$$\|X_{n+1}-\hat{X}_{n+1}\|^2 = \gamma(0) - \left\langle \hat{X}_{n+1}, \hat{X}_{n+1}\right\rangle. \tag{2.1.10}$$

Ahora se calculará la norma cuadrática de $\hat{X}_{n+1}$. Usando (2.1.4) se tiene que

$$\begin{aligned}
\left\langle \hat{X}_{n+1}, \hat{X}_{n+1}\right\rangle &= \left\langle \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}, \sum_{i=1}^{n} \phi_{n\,i} X_{n+1-i}\right\rangle \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} \phi_{n\,i}\langle X_{n+1-j}, X_{n+1-i}\rangle \phi_{n\,j} \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} \phi_{n\,i}\gamma(i-j)\phi_{n\,j} \\
&= \boldsymbol{\varphi}_n' \boldsymbol{\Gamma}_n \boldsymbol{\varphi}_n
\end{aligned}$$

y recordando que $\boldsymbol{\Gamma}_n\boldsymbol{\varphi}_n = \boldsymbol{\gamma}_n$, se obtiene que $\left\langle \hat{X}_{n+1}, \hat{X}_{n+1}\right\rangle = \boldsymbol{\varphi}_n'\boldsymbol{\gamma}_n$; combinando esta relación con (2.1.10), se llega a la expresión

$$v_n = \|X_{n+1}-\hat{X}_{n+1}\|^2 = \gamma(0) - \boldsymbol{\varphi}_n'\boldsymbol{\gamma}_n.$$

De esta manera, al resolver la ecuación (2.1.7) es posible determinar, tanto la proyección $\hat{X}_{n+1}$ por medio de (2.1.4), como el error de pronóstico $v_n$ a través de la expresión anterior desplegada. Esta discusión se resume en el siguiente teorema.

**Teorema 2.1.1** *Suponga que $\{X_t\}$ es una serie estacionaria con función de autocovarianza $\gamma(\cdot)$. En este caso, la proyección $\hat{X}_{n+1}$ de $X_{n+1}$ sobre el espacio $\mathcal{L}_n$ generado por $X_1, \ldots, X_n$ está dada por*

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{n\,j} X_{n+1-j}$$

*donde $\boldsymbol{\varphi}_n = (\phi_{n\,1}, \ldots, \phi_{n\,n})'$ es cualquier vector que satisface el sistema de ecuaciones*

$$\boldsymbol{\gamma}_n = \boldsymbol{\Gamma}_n \boldsymbol{\varphi}_n,$$

*y la notación es como en (2.1.6). Más aún, el error de pronóstico $v_n$ está dado por*

$$v_n = \|\hat{X}_{n+1} - X_{n+1}\|^2 = \gamma(0) - \boldsymbol{\varphi}_n' \boldsymbol{\gamma}_n.$$

Para concluir esta sección, ahora se discutirá brevemente la determinación de la proyección $\hat{X}_{n+h}$ de $X_{n+h}$ sobre $\mathcal{L}_n$. Como $\hat{X}_{n+h} \in \mathcal{L}_n$ existen constantes $\phi_{n\,1}^h, \phi_{n\,2}^h, \ldots, \phi_{n\,n}^h$ tales que

$$\hat{X}_{n+h} = \phi_{n\,1}^h X_n + \phi_{n\,2}^h X_{n-1} + \cdots + \phi_{n\,n}^h X_1, )$$

y entonces las condiciones (2.1.3) implican que, para cada $i = 1, 2, \ldots, n$,

$$\begin{aligned} \gamma(h+i-1) &= \langle X_{n+h}, X_{n+1-i} \rangle \\ &= \left\langle \hat{X}_{n+h}, X_{n+1-i} \right\rangle \\ &= \left\langle \sum_{j=1}^{n} \phi_{n\,j}^h X_{n+1-j}, X_{n+1-i} \right\rangle \\ &= \sum_{j=1}^{n} \phi_{n\,j}^h \langle X_{n+1-j}, X_{n+1-i} \rangle \\ &= \sum_{j=1}^{n} \phi_{n\,j}^h \gamma(i-j); \end{aligned}$$

Definiendo los vectores $\boldsymbol{\gamma}_n^h, \boldsymbol{\varphi}_n^h \in \mathbb{R}^n$ mediante

$$\boldsymbol{\gamma}_n^h = \begin{bmatrix} \gamma(h) \\ \gamma(h+1) \\ \vdots \\ \gamma(h+n-1) \end{bmatrix} \quad \text{y} \quad \boldsymbol{\varphi}_n^h = \begin{bmatrix} \phi_{n\,1}^h \\ \phi_{n\,2}^h \\ \vdots \\ \phi_{n,\,n}^h \end{bmatrix} \tag{2.1.11}$$

se desprende que

$$\gamma_n^h = \mathbf{\Gamma}_n \boldsymbol{\varphi}_n^h \tag{2.1.12}$$

donde $\mathbf{\Gamma}_n$ es la matriz $n \times n$ definida en (2.1.6). Más aún, siguiendo el mismo procedimiento que para en el caso $h = 1$, es posible obtener una fórmula simple para el error de pronóstico $\|X_{n+h} - \hat{X}_{n+h}\|^2$:

$$v_n^h \colon = \|X_{n+h} - \hat{X}_{n+h}\|^2 = \|X_{n+h}\|^2 - \|\hat{X}_{n+h}\|^2 = \gamma(0) - \boldsymbol{\varphi}_n^{h\prime} \boldsymbol{\gamma}_n^h. \tag{2.1.13}$$

## 2.2. Comportamiento de los Errores

En esta sección se ejemplifica el cálculo de proyecciones mediante la solución del sistema (2.1.7) o (2.1.12), y se analizan dos aspectos simples que, a primera vista, parece natural esperar sobre los errores de pronóstico:

(i) En el tiempo $n$, se han observado $X_1, X_2, \ldots, X_n$, y a medida que $h > 0$ aumenta, el tiempo de observación $n+h$ de la variable $X_{n+h}$ se aleja más del tiempo actual $n$. Esto puede generar la impresión de que cada vez será más difícil pronósticar $X_{n+h}$ usando los datos $X_1, \ldots, X_n$, lo cual se reflejará en el incremento del error de pronóstico $v_n^h = \|X_{n+h} - \hat{X}_{n+h}\|^2$ al crecer $h$. La expresión formal de esta primera impresión es

$$v_n^h \leq v_n^{h+1}, \quad h = 1, 2, 3, \ldots$$

(ii) Llevando al extremo la idea anterior, conforme $h$ crece sin límite, parecería natural esperar que "la influencia"de $X_1, \ldots, X_n$ sobre $X_{n+h}$ tiende a desvanecerse conforme $h$ se incrementa, en el sentido de que $\lim_{h \to 0} \hat{X}_{n+h} = 0$. En este caso, usando que $v_n^h = \gamma(0) - \|\hat{X}_{n+h}\|^2$, parece natural esperar que

$$\lim_{h \to \infty} v_n^h = \gamma(0).$$

A partir de los siguientes ejemplos, se verá que *estas dos impresiones no son correctas en general*; sin embargo, se mostrará que la intuición va en el camino correcto bajo condiciones adecuadas, o después de una reformulación apropiada de las ideas.

**Ejemplo 2.2.1** Considere el proceso de promedios móviles $X_t = Z_t + Z_{t-1} + Z_{t-5}$, donde la serie $\{Z_t\}$ es un ruido blanco con varianza unitaria, en este caso, la función de autocovarianza está dada por

$$\gamma(h) = \begin{cases} 3, & \text{si } h = 0 \\ 1, & \text{si } h = \pm 1 \text{ o } h = \pm 5, \\ 0, & \text{de otra forma;} \end{cases}$$

vea, por ejemplo, Brockwell y Davis (1998), Fuller (1998), o Shumway y Stoffer (2006). Dado $n = 2$, se determinarán, para cada $h > 0$, las proyecciones $\hat{X}_{2+h}$ de $X_{h+2}$ sobre $\mathcal{L}_2$.

(i) Para encontrar $\hat{X}_3 = \hat{X}_{2+1}$ se resuelve el sistema (2.1.7) (que equivale al sistema (2.1.12) con $h = 1$):

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \gamma(2) \end{bmatrix} = \begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} \begin{bmatrix} \phi_{2\,1} \\ \phi_{2\,2} \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \phi_{2\,1} \\ \phi_{2\,2} \end{bmatrix}.$$

Este sistema tiene matriz invertible, y su única solución es $\boldsymbol{\varphi}_2 = (3/8, -1/8)'$. Por lo tanto, $\hat{X}_3 = (3/8)X_2 - (1/8)X_1$, y el error de pronóstico es

$$v_2^1 = v_2 = \|X_3 - P_{\mathcal{L}_2} X_3\|^2 = 3 - (3/8, -1/8) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 21/8.$$

(ii) Las proyecciones $\hat{X}_4 = \hat{X}_{2+2}$ y $\hat{X}_5 = \hat{X}_{2+3}$ se determinan resolviendo (2.1.12) con $h = 2$ y $h = 3$, respectivamente. Observe ahora que $\boldsymbol{\gamma}_2^h = (\gamma(h), \gamma(h+1))' = (0, 0)'$ para $h = 2, 3$, de modo que el sistema $\boldsymbol{\gamma}_h = \boldsymbol{\Gamma}_2 \boldsymbol{\varphi}_2^h$ tiene la única solución $\boldsymbol{\varphi}_2^h = (0, 0)'$, y entonces $\hat{X}_4 = \hat{X}_5 = 0$ y los correspondientes errores de pronóstico son

$$\begin{aligned} v_2^2 &= \|X_4 - \hat{X}_4\|^2 = \gamma(0) - \boldsymbol{\varphi}_2^{2\prime} \boldsymbol{\gamma}_2^2 = 3 \\ v_2^3 &= \|X_5 - \hat{X}_5\|^2 = \gamma(0) - \boldsymbol{\varphi}_2^{3\prime} \boldsymbol{\gamma}_2^3 = 3 \end{aligned}$$

(iii) La proyección $\hat{X}_6 = \hat{X}_{2+4} = P_{\mathcal{L}_2} X_6$ se encuentra observando que $\gamma_2^4 = (\gamma(4), \gamma(5))' = (0, 1)'$. En este caso, (2.1.12) con $h = 4$ se reduce a

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma(4) \\ \gamma(5) \end{bmatrix} = \begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} \begin{bmatrix} \phi_{2\,1}^4 \\ \phi_{2\,2}^4 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \phi_{2\,1}^4 \\ \phi_{2\,2}^4 \end{bmatrix}.$$

cuya única solución es $\boldsymbol{\varphi}_2^4 = (-1/8, 3/8)'$. La proyección $\hat{X}_6$ es $X_1/8 - 3X_2/8$ y el correspondiente error de pronóstico es

$$v_2^4 = \|X_6 - P_{\mathcal{L}_2}X_6\|^2 = \gamma(0) - \boldsymbol{\varphi}_2^{4\prime}\gamma_2^4 = 3 - 3/8 = 21/8$$

(iv) Para encontrar la proyección $\hat{X}_7 = \hat{X}_{2+5} = P_{\mathcal{L}_2}X_7$ observe que $\gamma_2^5 = (\gamma(5), \gamma(6)' = (1, 0)'$, de tal forma que (2.1.12) con $h = 5$ equivale a

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma(5) \\ \gamma(6) \end{bmatrix} = \begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} \begin{bmatrix} \phi_{2\,1}^5 \\ \phi_{2\,2}^5 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \phi_{2\,1}^5 \\ \phi_{2\,2}^5 \end{bmatrix}.$$

cuya única solución es $\boldsymbol{\varphi}_2^5 = (3/8, -1/8)'$; se tiene que $\hat{X}_7 = (3X_1 - X_2)/8$ y el error de pronóstico es

$$v_2^5 = \|X_7 - \hat{X}_7\|^2 = \gamma(0) - \boldsymbol{\varphi}_2^{5\prime}\gamma_2^5 = 3 - 3/8 = 21/8$$

(v) Para evaluar $\hat{X}_{2+h} = P_{\mathcal{L}_2}X_{2+h}$ cuando $h \geq 6$, note que la especificación de $\gamma(\cdot)$ implica que $\boldsymbol{\gamma}_2^h = (\gamma(h), \gamma(h+1))' = (0, 0)'$. Por lo tanto, la solución de $\boldsymbol{\gamma}_2^h = \boldsymbol{\Gamma}_2\boldsymbol{\varphi}_2^h$ es $\boldsymbol{\varphi}_2^h = (0, 0)'$, y $\hat{X}_{2+h} = 0$ para $h \geq 6$; el correspondiente error de pronóstico es $v_2^h = \|X_{2+h} - \hat{X}_{2+h}\|^2 = \|X_{2+h}\|^2 = \gamma(0) = 3$.

En el ejemplo anterior $n = 2$ y la sucesión de errores de pronóstico es

$$(v_2^1, v_2^2, v_2^3, v_2^4, v_2^5, v_2^6, v_2^7, \dots) = \left(\frac{21}{8}, 3, 3, \frac{21}{8}, \frac{21}{8}, 3, 3, \dots\right)$$

y es claro que $v_2^h$ no es una función monótona creciente de $h$, de tal forma que la propiedad (2.2) no se satisface. Sin embargo, como $v_2^h = 3 = \gamma(0)$ para $h \geq 6$, en este ejemplo la convergencia en (2.2) *si ocurre*. Sin embargo, el siguiente ejemplo muestra que la validez de (2.2) tampoco puede garantizarse en el caso general.

**Ejemplo 2.2.2** Considere la serie $\{X_t\}$ definida mediante

$$X_t = A\cos(\pi t/3) + B\sin(\pi t/3),$$

donde $A$ y $B$ son variables aleatorias no correlacionadas con varianza unitaria. En este caso, la correspondiente función de autocovarianza es

$$\gamma(h) = \cos(\pi h/3);$$

vea Brockwell y Davis (1998) para una deducción detallada de este hecho. Ahora tome $n = 1$ y considere el correspondiente error de pronóstico

$$v_1^h = \|X_{1+h} - P_{\mathcal{L}_1} X_{1+h}\|^2 = \gamma(0) - \|P_{\mathcal{L}_1} X_{1+h}\|^2 = 1 - \|P_{\mathcal{L}_1} X_{1+h}\|^2$$

En este caso $P_{\mathcal{L}_1} X_{1+h} = \phi_{1\,1}^h X_1$ y las ecuaciones (2.1.12) se reducen a $\gamma(h) = \gamma(0)\phi_{1\,1}^h$, de donde $\phi_{1\,1}^h = \gamma(h)/\gamma(0) = \cos(\pi h/3)$, de manera que

$$\|P_{\mathcal{L}_1} X_{1+h}\|^2 = \|\cos(\pi h/3) X_1\|^2 = \cos^2(\pi h/3)$$

y entonces $v_1^h = 1 - \cos^2(\pi h/3) = \sin^2(\pi h/3)$. Esta expresión es periódica respecto a $h$, y al variar $h$ toma los valores $1/2, 1/2, 0, 1/2, 1/2, 0$ una y otra vez. Por lo tanto, $v_1^h$ no converge a $\gamma(0) = 1$ conforme $h$ tiene a $\infty$.

## 2.3. Una Condición de Invertibilidad

La determinación del vector de coeficientes $\boldsymbol{\varphi}_n^h$ requiere resolver el sistema (2.1.12) cuya matriz es $\boldsymbol{\Gamma}_n = [\gamma(i-j)]_{i,.j=1,\ldots,n}$. En esta sección se presenta una condición sobre la función de autocovarianza $\gamma(\cdot)$ que asegura que dicha matriz $\boldsymbol{\Gamma}_n$ es invertible para cada $n$.

**Teorema 2.3.1** *Suponga que $\gamma(\cdot)$ es la función de autocovarianza de un proceso estacionario $\{X_t\}$ para la cual las siguientes condiciones (i) y (ii) se satisface:*

*(i)* $\gamma(0) > 0$;

*(ii)* $\lim_{k\to\infty} \gamma(k) = 0$.

*En estas circunstancias, para cada $n = 1, 2, 3, \ldots$ la matriz $\boldsymbol{\Gamma}_n$ en (2.1.6) es invertible*

**Demostración.** Note que $\boldsymbol{\Gamma}_1 = [\gamma(0)]$ es una matriz invertible, pues $\gamma(0) > 0$. El argumento ahora procede por contradicción. Suponga que $\boldsymbol{\Gamma}_r$ es una matriz singular para algún $r > 1$. En este caso, debido a que $\boldsymbol{\Gamma}_1$ es invertible, existe un entero $n \geq 1$ tal que

$$\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_n \text{ son inveritbles, pero } \boldsymbol{\Gamma}_{n+1} \text{ es singular.}$$

Luego, existe un vector $\mathbf{b} = (b_1, \ldots, b_n, b_{n+1})' \in \mathbb{R}^{n+1}$ tal que

$$\mathbf{\Gamma}_{n+1}\mathbf{b} = 0, \quad \mathbf{b} \neq 0. \tag{2.3.14}$$

Observando que $\mathbf{\Gamma}_{n+1}$ se representa en bloques como

$$\mathbf{\Gamma}_{n+1} = \begin{bmatrix} \mathbf{\Gamma}_n & \tilde{\gamma} \\ \tilde{\gamma}' & \gamma(0) \end{bmatrix},$$

la igualdad en (2.3.14) se escribe como

$$\begin{bmatrix} \mathbf{\Gamma}_n & \tilde{\gamma} \\ \tilde{\gamma}' & \gamma(0) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{b}} \\ b_{n+1} \end{bmatrix} = 0, \quad \text{donde } \tilde{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Ahora se verá que $b_{n+1} \neq 0$. En efecto, si $b_{n+1} = 0$, la ecuación anterior implica que $\mathbf{\Gamma}\tilde{\mathbf{b}} = 0$, y entonces, por la invertibilidad de $\mathbf{\Gamma}_n$, se obtiene que $\tilde{\mathbf{b}} = 0$, lo que combinado con el supuesto de que $b_{n+1} = 0$ implica que $\mathbf{b} = 0$, en contradicción con (2.3.14). Sustituyendo a $\mathbf{b}$ por $b_{n+1}^{-1}\mathbf{b}$ en (2.3.14) se puede suponer que

$$\mathbf{b} = (b_1, b_2, \ldots, b_n, 1)'$$

Con lo anterior, note que

$$\left\| \sum_{s=1}^{n+1} b_s X_s \right\|^2 = \sum_{i,j=1}^{n+1} b_i \gamma(i-j) b_j = \mathbf{b}'\mathbf{\Gamma}_{n+1}\mathbf{b} = 0,$$

donde se utilizó la condición $\mathbf{\Gamma}_{n+1}\mathbf{b} = 0$ para establecer la última igualdad. Por lo tanto

$$\left\| \sum_{s=1}^{n} b_s X_s + X_{n+1} \right\| = 0 \tag{2.3.15}$$

lo cual significa que $X_{n+1} = -\sum_{s=1}^{n} b_s X_s$, esto, es

$$X_{n+1} \in \mathcal{L}(X_1, X_2, \ldots, X_n).$$

Por otro lado, la estacionaridad del proceso $\{X_t\}$ implica que

$$\left\| \sum_{s=1}^{n} b_s X_{s+h} + X_{n+h+1} \right\| = 0 \text{ para todo } h,$$

y entonces $\sum_{s=1}^n b_s X_{s+h} + X_{n+h+1} = 0$, es decir $X_{n+h+1} = -\sum_{s=1}^n b_s X_{s+h}$, de donde se obtiene que

$$X_{n+h+1} \in \mathcal{L}(X_{h+1}, X_{h+2}, \ldots, X_{h+n}), \quad h = 1, 2, 3, \ldots$$

Combinando esta relación con (2.3), se obtiene que

$$X_{n+h} \in \mathcal{L}(X_1, X_2, \ldots, X_n), \quad h = 1, 2, 3, \ldots$$

y entonces, para cada $h > 0$, existe un vector $\mathbf{b}^h = (b_1^h, \ldots, b_n^h) \in \mathbb{R}^n$ tal que

$$X_{n+h} = \sum_{i=1}^n b_i^h X_i \tag{2.3.16}$$

A partir de esta igualdad se obtendrá una contradicción. Primero observe que debido a que $\mathbf{\Gamma}_n$ es invertible, existen constantes positivas $m$ y $M$ tales

$$m\|\mathbf{x}\|^2 \leq \mathbf{x}'\mathbf{\Gamma}_n\mathbf{x} \leq M\|\mathbf{x}\|^2, \quad \mathbf{x} \in \mathbb{R}^n;$$

vea, por ejemplo, Hoffman y Kunze (1980), Graybill (2001), o Lipschutz (1995). Por otro lado, a partir de (2.3.16), cálculos similares a los realizados en la sección precedente arrojan que $\gamma(0) = \|X_{n+h}\|^2 = \mathbf{b}^{h\prime}\mathbf{\Gamma}_n\mathbf{b}^h$, relación que al combinarse con las anteriores desigualdades desplegadas implica que

$$m\|\mathbf{b}^h\|^2 \leq \mathbf{b}^{h\prime}\mathbf{\Gamma}_n\mathbf{b}^h = \gamma(0) \quad \text{y} \quad \gamma(0) = \mathbf{b}^{h\prime}\mathbf{\Gamma}_n\mathbf{b}^h \leq \|\mathbf{b}^h\|^2.$$

Por lo tanto

$$\frac{\gamma(0)}{M} \leq \|\mathbf{b}^h\|^2 \leq \frac{\gamma(0)}{m}, \tag{2.3.17}$$

esto es, $\{\mathbf{b}^h,\ h = 1, 2, 3, \ldots\}$ es una sucesión acotada. Para concluir, observe que

$$\begin{aligned}
\gamma(0) &= \|X_{n+h}\|^2 \\
&= \operatorname{Cov}(X_{n+h}, X_{n+h}) \\
&= \operatorname{Cov}\left(X_{n+h}, \sum_{i=1}^n b_i^h X_i\right) \\
&= \sum_{i=1}^n b_i^h \operatorname{Cov}(X_{n+h}, X_i) \\
&= \sum_{i=1}^n b_i^h \gamma(n+h-i)
\end{aligned}$$

y entonces (2.3.17) permite establecer que

$$\gamma(0) \leq \sqrt{\frac{\gamma(0)}{m}} \sum_{i=1}^{n} |\gamma(n+h-i)|.$$

Tomando el límite conforme $h$ tiende a $\infty$, la condición $\lim_{k\to\infty} \gamma(k) = 0$ implica que $\gamma(0) \leq 0$, lo cual contradice el supuesto $\gamma(0) > 0$. Esto muestra que $\mathbf{\Gamma}_n$ es invertible para cada $n$, finalizando el argumento. $\square$

## 2.4. Consecuencia de la Condición de Invertibilidad

En esta sección se estudia una implicación de la invertibilidad de las matrices $\mathbf{\Gamma}_n$ sobre el comportamiento de los errores de pronóstico $v_n^h$. Como se mencionó en la Sección 2.2, desde un punto de vista intuitivo, parece natural esperar que $v_n^h = \|X_{n+h} - P_{\mathcal{L}_n} X_{n+h}\|$ converja a $\gamma(0)$ conforme $h$ se incrementa, lo cual es un reflejo de la idea de que, a medida que $h$ aumenta, la variable $X_{n+h}$ que se observará en el tiempo $n+h$ tendrá cada vez "menos relación"con las variables $X_1, \ldots, X_n$ que se han observado desde el tiempo 1 hasta el tiempo $n$. Sin embargo, los ejemplos en la Sección 2.2 muestran que en general dicha impresión no es correcta en general. A continuación se muestra que la convergencia (2.2) es válida bajo las condiciones del Teorema 2.3.1.

**Teorema 2.4.1** *Suponga que $\{X_t\}$ es un proceso estacionario cuya función de autocovarianza $\gamma(\cdot)$ satisface las condiciones*

$$\gamma(0) > 0 \quad y \quad \lim_{k\to\infty} \gamma(k) = 0. \tag{2.4.18}$$

*En este caso,*

$$v_n^h = \|X_{n+h} - P_{\mathcal{L}_n} X_{n+h}\| \to \gamma(0) \quad \textit{conforme } h \to \infty.$$

**Demostración.** Recuerde que $P_{\mathcal{L}_n} X_{n+h} = \sum_{k=1}^{n} \phi_{n\,k}^h X_{n+1-k}$, donde

$$\boldsymbol{\varphi}_n^h = (\phi_{n\,1}^h, \phi_{n\,2}^h, \ldots, \phi_{n\,n}^h)$$

satisface $\gamma_n^h = \mathbf{\Gamma}_n \boldsymbol{\varphi}_n^h$, y el vector $\boldsymbol{\gamma}_n^h$ es como en (2.1.11). Como la matriz $\mathbf{\Gamma}_n$ es invertible, por el Teorema 2.3.1, se desprende que $\boldsymbol{\varphi}_n^h = \mathbf{\Gamma}_n^{-1}\boldsymbol{\gamma}_n^h$ y entonces la fórmula (2.1.13) puede escribirse como

$$v_n^h = \gamma(0) - \boldsymbol{\gamma}_n^{h\prime}\mathbf{\Gamma}_n^{-1}\boldsymbol{\gamma}_n^h;$$

puesto que $\boldsymbol{\gamma}_n^h = (\gamma(h), \gamma(h+1), \ldots, \gamma(h+n-1))' \to 0$ conforme $h$ tiende a $\infty$, por (2.4.18), tomando el límite cuando $h \to \infty$ en la igualdad anterior, se obtiene que $\lim_{h\to\infty} v_n^h = \gamma(0)$. $\square$

Por otro lado, el otro resultado que intuitivamente pareció natural en la Sección 2.2 fue la relación de monotonicidad (2.2). Observando que la función de autocovarianza en el Ejemplo 2.2.1 satisface la condición (2.4.18), se desprende que en ese contexto las matrices $\mathbf{\Gamma}_n$ son invertibles, y entonces el análisis de dicho ejemplo muestra que (2.2) no puede garantizarse aún bajo condiciones que garantizan la invertibilidad de $\mathbf{\Gamma}_n$. Para concluir esta sección, se mostrará que la validez de la relación de monotonicidad depende *de la información disponible* en el tiempo $n$. Suponga que al momento $n$ se han observado no sólo $X_1, X_2, \ldots, X_n$, sino todas las variables $X_t$ con $t \leq n$ y defina $\mathcal{H}_n$ como la cerradura del espacio generado por las variables $X_t$ con $-\infty < t \leq n$:

$$\mathcal{H}_n = \overline{\mathcal{L}(X_t,\, t \leq n)}. \tag{2.4.19}$$

Denotando por $\tilde{X}_{n+h}$ a la proyección de $X_{n+h}$ sobre $\mathcal{H}_n$, esto es,

$$\tilde{X}_{n+h} = P_{\mathcal{H}_n} X_{n+h}$$

se tiene que $\tilde{X}_{n+h}$ es el mejor pronóstico lineal de la observación futura $X_{n+h}$ en términos de la información disponible en el tiempo $n$ representada por $\mathcal{H}_n$. El correspondiente error cuadrático de pronóstico es

$$\tilde{v}_n^h = \|X_{n+h} - \tilde{X}_{n+h}\|^2 = \|X_{n+h} - P_{\mathcal{H}_n} X_{n+h}\|^2.$$

**Teorema 2.4.2** *Con la notación precedente, se tiene que*

$$\tilde{v}_n^h \leq \tilde{v}_n^{h+1}, \quad h = 1, 2, 3, \ldots$$

**Demostración.** Se presentará un bosquejo del argumento: Observe que por la definición del operador de proyección y la especificación de la cerradura de un subespacio,

$$\begin{aligned} \tilde{v}_n^h &= \|X_{n+h} - \tilde{X}_{n+h}\|^2 \\ &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_{n+h} - \sum_{i=1}^{k} a_i X_{n+1-i} \right\|^2, \end{aligned}$$

y aplicando la estacionaridad de $\{X_n\}$ se obtiene que

$$\begin{aligned} \tilde{v}_n^h &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_{n+h} - \sum_{i=1}^{k} a_i X_{n+1-i} \right\|^2 \\ &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_h - \sum_{i=1}^{k} a_i X_{1-i} \right\|^2 \\ &= \|X_h - P_{\mathcal{L}(X_s,\, s\leq 0)} X_h\|^2 \\ &= \|X_h - P_{\mathcal{H}_0} X_h\|^2; \end{aligned}$$

por lo tanto,

$$\tilde{v}_n^h = \|X_h - P_{\mathcal{H}_0} X_h\|^2 = \|X_h\|^2 - \|P_{\mathcal{H}_0} X_h\|^2. \tag{2.4.20}$$

Procediendo de forma similar,

$$\begin{aligned} \tilde{v}_n^{h+1} &= \|X_{n+h+1} - \tilde{X}_{n+h+1}\|^2 \\ &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_{n+h+1} - \sum_{i=1}^{k} a_i X_{n+1-i} \right\|^2, \end{aligned}$$

y entonces

$$\begin{aligned} \tilde{v}_n^{h+1} &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_{n+h+1} - \sum_{i=1}^{k} a_i X_{n+1-i} \right\|^2 \\ &= \min_k \min_{a_1,a_2\dots a_k} \left\| X_h - \sum_{i=1}^{k} a_i X_{0-i} \right\|^2 \\ &= \|X_h - P_{\mathcal{L}(X_s,\, s\leq -1)} X_h\|^2 \\ &= \|X_h - P_{\mathcal{H}_{-1}} X_h\|^2. \end{aligned}$$

En consecuencia,

$$\tilde{v}_n^{h+1} = \|X_h - P_{\mathcal{H}_{-1}} X_h\|^2 = \|X_h\|^2 - \|P_{\mathcal{H}_{-1}} X_h\|^2. \tag{2.4.21}$$

Debido a que $\mathcal{H}_{-1} \subset \mathcal{H}_0$, por (2.4.19), se tiene la desigualdad $\|P_{\mathcal{H}_{-1}} X_h\|^2 \leq \|P_{\mathcal{H}_0} X_h\|^2$, relación que al combinarse con (2.4.20) y (2.4.21) implica que $\tilde{v}_n^{h+1} \geq \tilde{v}_n^h$, la cual es la conclusión deseada. $\square$

Los teoremas de esta sección muestran que, aunque las impresiones intuitivas mencionadas en la Sección 2.2 no son correctas en general, si son válidas bajo condiciones apropiadas o cuando se reformulan adecuadamente.

# Capítulo 3

# SERIES AUTORREGRESIVAS Y DE PROMEDIOS MÓVILES

En este capítulo se introduce una clase importante de procesos estacionarios que surgen resolviendo ecuaciones en diferencias, los cuales se denomina procesos ARMA. La variablilidad de esos procesos se debe a la presencia de un ruido blanco que influye en los datos observables de una manera simple, generando un modelo que es sencillo de analizar, pero lo suficientemente versátil para ser utilizado, ya que la función de autocovarianza de cualquier serie estacionaria $\{Y_t\}$ puede aproximarse, tanto como se deseé, por la función de autocovarianza de un proceso ARMA adecuadamente seleccionado. La presentación incluye una discusión sobre la existencia y unicidad de processos ARMA, y finaliza con la introducción del concepto de *causalidad* de un proceso respecto a un ruido blanco.

## 3.1. Procesos ARMA

En esta sección se introduce un clase muy importante de procesos estacionaros, a saber, la familia de procesos autorregresivos y de promedios móviles, comúnmente referida como la clase ARMA de series estacionarias.

**Definición 3.1.1** *Considere un ruido blanco $\{Z_t\} \sim WN(0, \sigma^2)$ y sea $\{X_t\}$ una serie estacionaria con media nula.*

*(i) La serie $\{X_t\}$ es un proceso de promedios móviles de orden q (MA(q)) si existen números $\theta_1, \ldots, \theta_q \in \mathbb{R}$ tales que*

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

*(ii) El proceso $\{X_t\}$ es autorregresivo de orden p (AR(p)) si*

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t$$

*para números $\phi_1, \ldots, \phi_p \in \mathbb{R}$; (iii) La serie $\{X_t\}$ es un proceso autorregresivo y de promedios móviles de orden $(p, q)$ (ARMA$(p, q)$) si existen números $\phi_1, \ldots, \phi_p$ y $\theta_1, \ldots, \theta_q$ tales que*

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}. \tag{3.1.1}$$

Son varias las razones por las que los procesos ARMA desempeñan un papel importante en la teoría y aplicaciones de las series de tiempo: Primeramente, el problema de prónostico puede analizarse de manera sencilla para esos procesos, y algoritmos generales –como el algoritmo de innovaciones– se implementan de manera simple y eficiente para estos procesos. Por otro lado, es claro que ningún proceso real obedecerá de manera exacta a un modelo teórico, pero para cualquier proceso estacionario $\{Y_t\}$, es posible seleccionar un proceso ARMA $\{X_t\}$ cuya función de autocovarianza esta arbitrariamente cercana a la de $\{Y_t\}$, en el sentido de que $\max_h |\gamma_X(h) - \gamma_Y(h)|$ es tan pequeña como se deseé si el orden $(p, q)$ y los coeficientes del proceso ARMA se eligen adecuadamente (Brockwell y Davis, 1998, Fuller, 1988).

Si $\phi_1, \ldots, \phi_p$ y $\theta_1, \ldots, \theta_q$ son como en la Definición 3.1.1, los polinomios $\phi(z)$ y $\theta(z)$ se especifican mediante

$$\theta(z) = 1 + \theta_1 z_1 + \cdots + \theta_q z^q \tag{3.1.2}$$

y

$$\phi(z) = 1 - \phi_1 z_1 - \cdots - \phi_p z^p, \tag{3.1.3}$$

los cuales se denominan polinomio de promedios móviles y autorregresivo, respectivamente. Evaluando estos polinomios en el operador de retardo $B$, se obtiene que

$$\theta(B)Z_t = (1 + \theta_1 B + \cdots + \theta_q B^q)Z_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

y

$$\phi(B)X_t = (1 - \phi_1 B - \cdots - \phi_p B^p)X_t = X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}$$

de tal manera que las ecuaciones que satisface un proceso $\{X_t\}$ cuando es MA($q$), AR($p$) o ARMA($p, q$) pueden escribirse como

$$X_t = \theta(B)Z_t, \quad \phi(B)X_t = Z_t, \quad \text{y} \quad \phi(B)X_t = \theta(B)Z_t,$$

respectivamente. Las siguientes secciones se ocupan de los problemas de existencia y unicidad de soluciones a la ecuación (3.1.1). Más explícitamente, se analizarán las siguientes preguntas: Dado un ruido blanco $\{Z_t\}$ y los polinomios $\theta(z)$ y $\phi(z)$ ¿Existe un proceso estacionario $\{X_t\}$ que satisfaga (3.1.1)? y en caso afirmativo ? es única la serie $\{X_t\}$?

## 3.2. Existencia

En esta sección se establece un criterio suficiente para garantizar que las ecuaciones (3.1.1) tengan una solución $\{X_t\}$. Las condiciones involucran a las raíces del polinomio autorregresivo $\phi(z)$. Primeramente, note que si $\phi(z) = 1$, las ecuaciones (3.1.1) se reducen a $X_t = \theta(B)Z_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$, $t = 0, \pm 1, \pm 2, \ldots$, ecuaciones que claramente tienen solución única. En adelante, se supondrá que $\phi(z)$ tiene grado $p > 1$, de manera que $\phi_p \neq 0$. En este caso, $\phi(z)$ posee $p$ raíces, las cuales (pueden repetirse y) se denotarán mediante $\xi_1, \ldots, \xi_p$:

$$\phi(\xi_i) = 0, \quad i = 1, 2, \ldots, p. \tag{3.2.4}$$

Debido a que $\phi(0) = 1$ (vea (3.1.3)), el polinomio autorregresivo se factoriza como

$$\phi(z) = (1 - z/\xi_1)(1 - z/\xi_2)\cdots(1 - z/\xi_p) = \prod_{i=1}^{p}(1 - z\xi_i^{-1}). \tag{3.2.5}$$

**Teorema 3.2.1** *Si las raíces del polinomio autorregresivo $\phi(z)$ satisfacen*

$$|\xi_i| \neq 1, \quad i = 1, 2, \ldots, p,$$

*entonces las siguientes afirmaciones (i)–(iii) son válidas:*

*(i) La función* $1/\phi(z)$ *se expande en serie de Laurent alrededor del círculo unitaro* $|z| = 1$ *( Alfhors 1980, Rudin 1984, Apostol 1980). Más precisamente, existen constantes* $R_0$ *y* $R_1$ *con* $R_0 < 1 < R_1$ *tales que*

$$\frac{1}{\phi(z)} = \sum_{k=-\infty}^{\infty} \psi_k z^k, \quad R_0 < |z| < R_1 \tag{3.2.6}$$

*donde la serie converge absolutamente en la región indicada:*

$$\sum_{k=-\infty}^{\infty} |\psi_k||z|^k < \infty, \quad R_0 < |z| < R_1$$

*(ii) Considere la sucesión* $\psi = \{\psi_k,\, k = 0, \pm 1, \pm 2, \ldots\}$ *y el correspondiente filtro* $\psi(B)$. *Si la serie estacionaria* $\{X_t\}$ *se define como*

$$X_t = \psi(B)\theta(B)Z_t, \quad t = 0, \pm 1, \pm 2, \ldots,$$

*entonces* $\{X_t\}$ *satisface las ecuaciones ARMA en (3.1.1).*

*(iii) Si* $\{X_t\}$ *y* $\{\tilde{X}_t\}$ *son dos soluciones a las ecuaciones (3.1.1), entonces* $X_t = \tilde{X}_t$ *para todo* $t$.

De acuerdo a este resultado, cuando las raíces de $\phi(\cdot)$ no tienen módulo 1, entonces existe un proceso estacionario que satisface las ecuaciones ARMA, y dicho proceso es único. El argumento para establecer este teorema utiliza los resultados preliminares que se establecen en el siguiente Lema.

**Lema 3.2.1** *Sea* $a \in \mathbb{C}$ *tal que* $|a| \neq 1$. *En este caso, las siguientes expansiones en serie son válidas para* $1/(1 - z/a)$.

*(i) Si* $|a| > 1$, *entonces, para cada entero positivo* $m$,

$$\frac{1}{(1 - z/a)^m} = \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{z^{k-m+1}}{a^{k-m+1}} \quad si\ |z| < |a|.$$

*Similarmente,*

*(ii) Cuando* $|a| < 1$ *y* $m = 1, 2, 3, \ldots$

$$\frac{1}{(1 - z/a)^m} = (-1)^m \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{a^{k+1}}{z^{k+1}}, \quad cuando\ |a| < |z|.$$

**Demostración.** Recuerde que para cualquier número complejo $r$ con $|r| < 1$, se tiene la sigiuente expansión en serie :

$$\frac{1}{1-r} = \lim_{n\to\infty} \sum_{k=0}^{n} r^k = \sum_{k=0}^{\infty} r^k. \tag{3.2.7}$$

Usando esta relación con $|r|$ en lugar de $r$, se tiene que

$$\sum_{k=0}^{\infty} |r|^k = 1/(1-|r|) < \infty,$$

de manera que la convergencia de la serie anterior es absoluta. Puesto que una serie puede derivarse término a término en su región de convergencia (Fulks, 1982, Khuri, 2002), a partir de (3.2.7) se desprende que

$$\frac{d}{dr}\frac{1}{1-r} = \frac{1}{(1-r)^2} = \sum_{k=1}^{\infty} kr^{k-1}$$

Derivando sucesivamente,

$$\begin{aligned}
\frac{d}{dr}\frac{1}{(1-r)^2} &= \frac{2\cdot 1}{(1-r)^3} = \sum_{k=2}^{\infty} k(k-1)r^{k-2} \\
\frac{d}{dr}\frac{1}{(1-r)^3} &= \frac{3\cdot 2\cdot 1}{(1-r)^4} = \sum_{k=3}^{\infty} k(k-1)(k-2)r^{k-3} \\
\frac{d}{dr}\frac{1}{(1-r)^4} &= \frac{4\cdot 3\cdot 2\cdot 1}{(1-r)^5} = \sum_{k=4}^{\infty} k(k-1)(k-2)(k-3)r^{k-4}
\end{aligned}$$

En general, aplicando esta idea, se llega a la siguiente expresión para cualquier entero $m \geq 1$:

$$\frac{(m-1)!}{(1-r)^m} = \sum_{k=m-1}^{\infty} (k)_{m-1} r^{k-m+1},$$

donde $(k)_s = k(k-1)\cdots(k-s+1)$ es el número de permutaciones de tamaño $s$ de una conjunto de tamaño $k$. Recordando que

$$\frac{(k)_s}{s!} = \binom{k}{s}$$

se desprende que

$$\frac{1}{(1-r)^m} = \sum_{k=m-1}^{\infty} \binom{k}{m-1} r^{k-m+1}, \quad |r| < 1, \quad m = 1, 2, \ldots \tag{3.2.8}$$

donde la serie converge absolutamente.

(i) Suponga que $|a| > 1$ y sea $z$ tal que $|z/a| < 1$. Sustituyendo $r$ por $z/a$ en (3.2.8) se desprende que

$$\frac{1}{(1-z/a)^m} = \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{z^{k-m+1}}{a^{k-m+1}}, \quad |z| < |a|, \quad m = 1, 2, \ldots \tag{3.2.9}$$

y la serie es absolutamente convergente.

(ii) Suponga ahora que $|a| < 1$ y sea $z$ tal que $|a/z| < 1$, esto es $|z| > |a|$. Usando (3.2.8) con $a/z$ en lugar de $r$, se obtiene

$$\frac{1}{(1-a/z)^m} = \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{a^{k-m+1}}{z^{k-m+1}};$$

multiplicando ambos lados por $1/(z/a)^m = (a/z)^m$ se desprende que

$$\frac{1}{(z/a-1)^m} = \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{a^{k+1}}{z^{k+1}},$$

y entonces, para todo $z$ tal que $|z| > |a|$

$$\frac{1}{(1-z/a)^m} = (-1)^m \sum_{k=m-1}^{\infty} \binom{k}{m-1} \frac{a^{k+1}}{z^{k+1}}, \quad m = 0, \pm 1, \pm 2, \ldots$$

□

**Demostración del Teorema 3.2.1.** Sean $a_1, \ldots, a_d$ las raíces diferentes de $\phi(\cdot)$. Con esta notación, cada $a_i$ es igual a algún $\xi_j$ en (3.2.5); si $m_i$ es el número de veces que $a_i$ se repite en la sucesión $\xi_1, \ldots, \xi_p$, la factorización (3.2.5) puede escribirse como

$$\phi(z) = \prod_{i=1}^{d} (1 - z/a_i)^{m_i}$$

de manera que

$$\frac{1}{\phi(z)} = \frac{1}{\prod_{i=1}^{d} (1 - z/a_i)^{m_i}} \tag{3.2.10}$$

Utilizando la descomposición en fracciones parciales para una función racional (Fulks 1982, Khuri 2002), la anterior expresión puede escribirse como

$$\frac{1}{\phi(z)} = \sum_{i=1}^{d} \sum_{j=1}^{m_i} \frac{A_{ji}}{(1 - z/a_i)^j}$$

Por el Lema 3.2.1, cada término en la suma precedente puede expresarse como

$$\frac{A_{ji}}{(1 - z/a_i)^j} = \sum_{k=-\infty}^{\infty} \psi_k^{(i\,j)} z^k$$

donde la serie converge absolutamente en una región de la forma

$$R_{i\,j} < |z| < R^{i\,j}$$

donde $R_{i\,j} < 1 < R^{i\,j}$. Combinando las tres relaciones desplegadas se deduce que $1/\phi(z)$ puede expresarse como en (3.2.6), donde la serie converge absolutamente en la región $R_0 < |z| < R_1$, y las constantes están dadas por $R_0 = \text{máx}\{R_{i\,j}\} < 1$ y $R_1 = \text{mín}\{R^{i\,j}\} > 1$. (ii) Como la serie en (3.2.6) converge absolutamente en una región que incluye al círculo unitario, se desprende que la $\sum_k |\psi_k| < \infty$. Por lo tanto, el filtro $\psi(B)$ está bien definido y la serie $X_t = \psi(B)\theta(B)Z_t$ es estacionaria. Más aún,

$$\phi(B)X_t = \phi(B)\psi(B)\theta(B)Z_t = \theta(B)Z_t \tag{3.2.11}$$

pues, recordando que $\phi(z)\psi(z) = \phi(z)[1/\phi(z)] = 1$ obtiene que

$$W_t = \phi(B)\psi(B)W_t$$

para cada serie estacionaria, hecho que fue utilizado con $W_t = \theta(B)Z_t$ para establecer la segunda igualdad en (3.2.11). Por lo tanto, $X_t$ satisface las ecuaciones ARMA.

(iii) Si

$$\phi(B)X_t = \theta(B)Z_t, \quad \text{y} \quad \phi(B)\tilde{X}_t = \theta(B)Z_t,$$

entonces $\phi(B)X_t = \phi(B)\tilde{X}_t$; aplicando el filtro $\psi(B)$ en ambos lados de esta igualdad se concluye que $X_t = \psi(B)\phi(B)X_t = \psi(B)\phi(B)\tilde{X}_t = \tilde{X}_t$, completando la demostración del teorema. □

El argumento usado para establecer el Teorema 3.2.1 revela que en la expansión (3.2.6) aparecen potencias negativas de $z$ sólo cuando alguna de las raíces $\xi_i$ de $\phi(z)$ tiene módulo menor a 1. Este hecho es importante, y para propósitos de referencia futura se resalta a continuación.

**Corolario 3.2.1** *Si todas las raíces de $\phi(z)$ tienen módulo mayor que 1, entonces los coeficientes $\psi_k$ en (3.2.6) son tales que $\psi_k = 0$ para $k < 0$, de manera que*

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

## 3.3. Raíces Sobre el Círculo Unitario

El criterio para la existencia de un proceso ARMA establece que si el polinomio autorregresivo no se anula sobre el círculo unitario, entonces tal proceso existe. Luego, es interesante analizar que sucede cuando $\phi(z) = 0$ para algún $z$ con módulo 1; como se muestra en el siguiente teorema, en ese caso las ecuaciones ARMA no necesariamente tienen solución.

**Teorema 3.3.1** *Suponga que los polinomios $\phi(z)$ y $\theta(z)$ no tienen raíces comunes. En este caso, si $\phi(\alpha) = 0$ para algún número complejo $\alpha$ con $|\alpha| = 1$, entonces, no existe proceso estacionario $\{X_t\}$ que satisfaga las euaciones $\phi(B)X_t = \theta(B)Z_t$, $t = 0, \pm 1, \pm 2, \ldots$*

**Demostración.** Como $\phi(z)$ y $\theta(z)$ no tienen raíces comunes, la condición $\phi(\alpha) = 0$ implica que

$$\theta(\alpha) \neq 0.$$

Por otro lado, debido a que $\alpha$ es una raíz de $\phi(z)$, es posible factorizar $\phi(z)$ como

$$\phi(z) = (1 - z\alpha^{-1})\tilde{\phi}(z) \tag{3.3.12}$$

donde el polinomio $\tilde{\phi}(z)$ tiene grado $p-1$. A partir de este punto el argumento es por contradicción. Suponga que para algún proceso estacionario $\{X_t\}$ se satisface que

$$\phi(B)X_t = \theta(B)Z_t$$

y considere el nuevo porceso estacionario $\{Y_t\}$ especificado mediante

$$Y_t = \tilde{\phi}(B)X_t$$

En este caso

$$\begin{aligned} Y_t - \alpha^{-1}Y_{t-1} &= (1 - B\alpha^{-1})Y_t \\ &= (1 - B\alpha^{-1})\tilde{\phi}(B)X_t \\ &= \phi(B)X_t \\ &= \theta(B)Z_t, \end{aligned}$$

donde se utilizó la factorización (3.3.12) para establecer la tercera igualdad. A continuación, a partir de la igualdad $Y_t - \alpha^{-1}Y_{t-1} = \theta(B)Z_t$ se desprende que

$$Y_t = \theta(B)Z_t + \alpha^{-1}Y_{t-1} = \theta(B)Z_t + \alpha^{-1}\theta(B)Z_{t-1} + \alpha^{-2}Y_{t-2},$$

y en general, para cada entero positivo $N$,

$$Y_t = \sum_{k=0}^{N} \alpha^{-k}\theta(B)Z_{t-k} + \alpha^{-N-1}Y_{t-N-1}$$

Ahora se estudiará la suma en esta expresión:

$$\begin{aligned} \sum_{k=0}^{N} \alpha^{-k}\theta(B)Z_{t-k} &= \sum_{k=0}^{N} \alpha^{-k} \sum_{s=0}^{q} \theta_s Z_{t-k-s} \\ &= \sum_{k=0}^{N} \sum_{s=0}^{q} \alpha^{-k}\theta_s Z_{t-k-s} \end{aligned}$$

de donde se desprende que

$$Y_t - \alpha^{-N-1}Y_{t-N-1} = \sum_{k=0}^{N} \sum_{s=0}^{q} \alpha^{-k}\theta_s Z_{t-k-s}. \tag{3.3.13}$$

En esta última sumatoria, se agruparán los términos de acuerdo al valor de $m = k+s$, y lugo se sumara desde $m = 0$ hasta su valor máximo, que es $m = N + q$:

$$\begin{aligned} \sum_{k=0}^{N} \sum_{s=0}^{q} \alpha^{-k}\theta_s Z_{t-k-s} &= \sum_{m=0}^{N+q} \sum_{s,k:\ s+k=m,\ 0\leq s\leq q,\ 0\leq k\leq N} \alpha^{-k}\theta_s Z_{t-k-s} \\ &= \sum_{m=0}^{N+q} \sum_{s,k:\ s+k=m,\ 0\leq s\leq q,\ 0\leq k\leq N} \alpha^{-k}\theta_s Z_{t-m} \\ &= \sum_{m=0}^{N+q} Z_{t-m}c_m \end{aligned}$$

donde

$$c_m = \sum_{\substack{s,k:\ s+k=m \\ 0\le s\le q,\ 0\le k\le N}} \alpha^{-k}\theta_s; \tag{3.3.14}$$

combinado este desarrollo con (3.3.13) se obtiene

$$Y_t - \alpha^{-N-1}Y_{t-N-1} = \sum_{m=0}^{N+q} Z_{t-m}c_m \tag{3.3.15}$$

Ahora se analizará el valor de $c_m$. Suponga que $q \le m \le N$ y note que en la suma que define a $c_m$, la igualdad $k = m - s$ ocurre y $0 \le s \le q$, de manera que la condición $0 \le k \le N$ se satisface automáticamente. Por lo tanto,

$$\begin{aligned} c_m &= \sum_{\substack{s,k:\ s+k=m \\ 0\le s\le q,\ 0\le k\le N}} \alpha^{-k}\theta_s \\ &= \sum_{0\le s\le q,} \alpha^{-(m-s)}\theta_s \\ &= \alpha^{-m}\sum_{s=0}^{q} \alpha^{s}\theta_s \\ &= \alpha^{-m}\theta(\alpha) \end{aligned}$$

y puesto que $|\alpha| = 1$, se desprende que

$$|c_m| = |\theta(\alpha)| \quad \text{si } q \le m \le N. \tag{3.3.16}$$

Tomando ahora la norma cuadrática en ambos lados de (3.3.15) y recordando que $\{Z_t\} \sim WN(0, \sigma^2)$, se obtiene que (Graybill, 2001, Lipchutz 1995)

$$\|Y_t - \alpha^{-N-1}Y_{t-N-1}\|^2 = \sum_{m=0}^{N+q} \|Z_{t-m}\|^2|c_m| \ge \sum_{m=q}^{N} \|Z_{t-m}\|^2|c_m|^2$$

y entonces

$$\|Y_t - \alpha^{-N-1}Y_{t-N-1}\|^2 \ge \sum_{m=q}^{N} \sigma^2|\theta(\alpha)|^2 = (N-q)\sigma^2\|\theta(\alpha)|^2.$$

Finalmente, la desigualdad del triángulo y la estacionaridad de la serie $\{Y_t\}$ implican que

$$\|Y_t - \alpha^{-N-1}Y_{t-N-1}\| \le \|Y_t\| + |\alpha|^{-N-1}\|Y_{t-N-1}\| \le 2\gamma_Y(0)$$

y junto con la precedente desigualdad desplegada, esta relación implica que $4\gamma_Y(0) \geq (N-q)\sigma^2\|\theta(\alpha)|^2$, y entonces

$$\frac{4\gamma_Y(0)}{N-q} \geq \sigma^2\|\theta(\alpha)|^2.$$

Tomando el límite cuando $N$ tiende a infinito, se llega a $\theta(\alpha) = 0$, lo cual contradice (3.3). El origen de esta contradicción es el supuesto de que existe una serie estacionaria $\{X_t\}$ que satisface $\phi(B)X_t = \theta(B)Z_t$, y entonces tal serie no existe. □

## 3.4. Causalidad

De acuerdo al Teorema 3.2.1, cuando el polinomio autorregresivo $\phi(z)$ no tiene raíces sobre el círculo unitario, las ecuaciones $\phi(B)X_t = \theta(B)Z_t$ tienen la única solución

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k Z_{t-k} \tag{3.4.17}$$

donde $\psi(z) = \sum_{k=-\infty}^{\infty} \psi_k z^k = \theta(z)/\phi(z)$. La perturbación $Z_t$ se manifiesta en el tiempo $t$, así que

$$\text{si } \psi_k = 0 \text{ para } k < 0, \text{ entonces } X_t = \sum_{k=0}^{\infty} \psi_k Z_{t-k}$$

y la observación $X_t$ es una función de las perturbaciones $Z_r$ que han ocurrido antes de $t$, o en el tiempo $t$. En contraste, si $\psi_k \neq 0$ para algún $k < 0$, entonces la suma en (3.4.17) contiene el término $\psi_k Z_{t-k}$, en el cual $t - k > t$, indicando que $Z_{t-k}$ se mainfestará en un tiempo posterior a $t$, de modo que $X_t$ depende de perturbacions que surgirán en el futuro, situación que no se antoja razonable. Cuando $\psi_k = 0$ para $k < 0$, se dice que la *serie* $\{X_t\}$ *es una función causal de* $\{Z_t\}$. A partir de la demostración del Teorema 3.2.1, es claro que para que un coeficiente $\psi_k$ con $k < 0$ sea no nulo es necesario y suficiente que el polinomio autorregresivo $\phi(z)$ tenga raíces dentro del círculo unitario. Por lo tanto, la causalidad del proceso $\{X_t\}$ depende sólo del polinomio autorregresivo.

**Definición 3.4.1** *Un polinomio $\phi(z)$ de grado $p$ mayor a cero se denomina causal si y sólo si todas sus raíces $\xi_1, \xi_2, \ldots, \xi_p$ se ubican fuera del disco unitario, esto es,*

$$|\xi_i| > 1, \quad i = 1, 2, \ldots, p.$$

Debido a que determinar las raíces de un polinomio $\phi(z)$ no es, en general, una tarea sencilla, es conveniente disponer de un criterio que permita decidir si el polinomio es causal o no, *sin determinar sus raíces*. En la literatura, es posible encontrar enunciado uno de tales criterios para polinomios de grado dos, el cual se discute a continuación. Como se verá después del siguiente análisis, el problema de construir un criterio para la causalidad de un polinomio es realmente interesante.

**Ejemplo 3.4.1** En la literatura se propone el siguiente criterio para la causalidad de un polinomio

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2$$

de grado dos con coeficientes reales. Para que $\phi(z)$ sea un polinomio causal es necesario (Brockwell and Davis, 1998) que los coeficientes $(\phi_1, \phi_2)$ satisfagan

$$\begin{aligned} \phi_1 + \phi_2 &< 1, \\ \phi_1 - \phi_2 &< 1, \\ |\phi_2| &< 1. \end{aligned} \tag{3.4.18}$$

Estas condiciones se proponen como necesarias y suficientes en Shumway y Stoffer (2006).

A continuación se analizará la necesidad de las condiciones (3.4.18) para la causalidad de $\phi(z)$. Como antes, las raíces de $\phi(z)$ se denotarán mediante $\xi_1$ and $\xi_2$. Utilizando que $\phi(z) = (1 - \xi_1^{-1} z)(1 - \xi_2^{-1} z)$ se desprenden las siguientes expresiones para los coeficiente $\phi_1$ y $\phi_2$:

$$\phi_1 = \frac{1}{\xi_1} + \frac{1}{\xi_2}, \quad \phi_2 = -\frac{1}{\xi_2 \xi_1}. \tag{3.4.19}$$

Para verificar la necesidad de las condiciones (3.4.18) debe mostrarse que

$$|\xi_1| > 1 \quad \text{and} \quad |\xi_2| > 1 \Rightarrow \text{ (3.4.18)}. \tag{3.4.20}$$

Considere los siguiented dos casos:

**Caso 1:** Las raíces de $\phi$ son genuinamente complejas, en el sentido de que su parte imaginaria es no nula:

$$\xi_1 = a + ib, \quad \text{y} \quad \xi_2 = a - ib, \quad b \neq 0.$$

En estas circunstancias

$$\phi_1 = \frac{1}{a+ib} + \frac{1}{a-ib} = \frac{2a}{a^2+b^2}, \quad \phi_2 = -\frac{1}{(a+ib)(a-ib)} = -\frac{1}{a^2+b^2}$$

así que

$$\phi_2 - \phi_2 = \frac{2a+1}{a^2+b^2}.$$

Observe ahora que

$$\begin{aligned}\phi_2 - \phi_2 < 1 \quad &\Longleftrightarrow \quad \frac{2a+1}{a^2+b^2} < 1 \\ &\Longleftrightarrow \quad 2a+1 < a^2+b^2 \\ &\Longleftrightarrow \quad 2 < (a-1)^2 + b^2\end{aligned}$$

Así, si (3.4.20) ocurre, entonces debe tenerse la siguiente implicación:

$$a^2 + b^2 > 1 \Rightarrow 2 < (a-1)^2 + b^2.$$

Sin embargo, si se asigna $a = 1.01$ y $b = 0.01$ se sigue que $a^2+b^2 > 1$ y $(a-1)^2 + b^2 = 0.0002 < 2$, de manera que la implicación anterior no se satisface, y por lo tanto en el caso actual, la afirmación (3.4.20) *no es correcta*.

**Caso 2:** Las raíces $\xi_1$ y $\xi_2$ son reales.

En este contexto se analizarán tres caso exahustivos.

(i) $\xi_1 > 0$ y $\xi_2 > 0$.

En esta situación, si (3.4.20) es válida, entonces la siguiente afirmación es correcta :

$$\xi_1 > 1 \quad \text{and} \quad \xi_2 > 1 \Rightarrow \phi_1 - \phi_2 < 1$$

Por medio de (3.4.19) esto es equivalente a

$$\xi_1 > 1 \quad \text{and} \quad \xi_2 > 1 \Rightarrow \frac{\xi_2 + \xi_1 + 1}{\xi_1 \xi_2} < 1$$

Sin embargo, si se tiene $\xi_1 = 2 = \xi_2$, no es difícil ver que que la implicación anterior es falsa. En consecuencia, (3.4.20) falla también en el cotexto actual.

(ii) $\xi_1 < 0$ y $\xi_2 < 0$.

En estas condiciones, $\xi_1\xi_2 > 0$ y

$$\left.\begin{array}{c}\xi_1 < -1\\ \text{y}\\ \xi_2 < -1\end{array}\right\} \Rightarrow \begin{cases}\phi_1 - \phi_2 = \dfrac{\xi_1 + \xi_2 + 1}{\xi_1\xi_2} < \dfrac{\xi_1}{\xi_1\xi_2} = \dfrac{1}{\xi_2} < 0 < 1\\ \phi_1 + \phi_2 = \dfrac{\xi_1 + \xi_2 - 1}{\xi_1\xi_2} < \dfrac{\xi_1}{\xi_1\xi_2} = \dfrac{1}{\xi_2} < 0 < 1\\ |\phi_2| = \dfrac{1}{|\xi_1\xi_2|} = \dfrac{1}{\xi_1\xi_2} < 1\end{cases}$$

Así, (3.4.20) ocurre en el presente contexto.

(iii) Las raíces son reales con diferente signo. En este caso $\xi_1\xi_2 < 0$ y, sin pérdida de generalidad, puede suponerse que $\xi_1 < 0$ y $\xi_2 > 0$. Note ahora que

$$\xi_1 < -1 \quad \text{y} \quad \xi_2 > 1 \Rightarrow \phi_2 < 0 \quad \text{y} \quad |\phi_2| = \frac{1}{|\xi_1|\xi_2} < 1$$

$$\begin{aligned}\xi_1 < -1 \quad \text{y} \quad \xi_2 > 1 \quad &\Rightarrow \quad \xi_1\xi_2 < 0, \quad \xi_2 > 1 \quad \text{y} \quad \xi_1 + \xi_2 + 1 > \xi_1\\ &\Rightarrow \quad \xi_2 > 1 \quad \text{and} \quad \frac{\xi_1 + \xi_2 + 1}{\xi_1\xi_2} < \frac{1}{\xi_2}\\ &\Rightarrow \quad \frac{\xi_1 + \xi_2 + 1}{\xi_1\xi_2} < 1\\ &\Rightarrow \quad \phi_1 - \phi_2 < 1\\ &\Rightarrow \quad \phi_1 + \phi_2 < 1\end{aligned}$$

donde la penúltima implicación se debe a (3.4.19), y la desigualdad $\phi_2 < 0$ fue utilizada en el último paso. Las dos últimas relaciones desplegadas implican que (3.4.20) ocurre en las presentes circunstancias. □

El siguiente resultado establece un criterio para la causalidad de un polinomio, y una demostración puede encontrarse en Chacón Hernández (2010).

**Teorema 3.4.1** *Si $\phi(z)$ es un polinomio con raíces distintas $a_1, a_2 \ldots, a_d$, cada una con multiplicidad $m_1, m_2, \ldots, m_d$, respectivamente, y si $\phi(z) \neq 0$ para todo $z$ tal que $|z| = 1$, entonces*

*(i) El número de raíces de $\varphi(z)$ dentro del círculo unitario está dado por*

$$\sum_{i:\ |a_i|<1} m_i = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{\phi'(z)}{\phi(z)}\, dz,$$

*donde la integral en el lado derecho es una integral compleja sobre el círculo unitario* $\mathcal{C} = \{z \mid |z| = 1\}$. *Por lo tanto,*

*(ii) El polinomio* $\phi(z)$ *es causal si, y sólo si,*

$$\int_{\mathcal{C}} \frac{\phi'(z)}{\phi(z)}\, dz = 0.$$

Como se establece en el siguiente resultado, si una serie estacionaria $\{X_t\}$ satisface la ecuación autorregresiva $\varphi(B)X_t = Z_t$, donde $\{Z_t\}$ es un ruido blanco, entonces no hay pérdida de generalidad, si se supone que $\varphi(z)$ es un polinomio causal, esto es, que todas sus raíces tienen módulo menor a 1.

**Teorema 3.4.2** *Suponga que la serie estacionaria* $\{X_t\}$ *satisface la ecuación* $\varphi(B)X_t = \theta(B)Z_t$, *donde* $\{Z_t\}$ *es un ruido blanco y* $\varphi(z)$ *y* $\theta(z)$ *son polinomios. En este caso, existe un polinomio causal* $\tilde{\varphi}(z)$ *y un ruido blanco* $\{\tilde{Z}_t\}$ *tales que el proceso* $\{X_t\}$ *satisface la ecuación*

$$\tilde{\varphi}(B)X_t = \theta(B)\tilde{Z}_t, \quad t \in \mathbb{N}.$$

Este teorema puede describirse verbalmente como sigue: Cuando se considera un proceso ARMA, siempre puede suponerse que el correspondiente polinomio autorregresivo tiene sus raíces dentro del círculo unitario. Una demostración de este resultado puede encontrarse en Brockwell y Davis (1998), Fuller (1998) o Shumway y Stopoffer (2006).

# Capítulo 4

# ECUACIONES DE YULE-WALKER

## 4.1. Autocovarianza de un Proceso ARMA

En esta sección se describe el sistema de ecuaciones que permiten determinar recursivamente la función de autocovarianza de un proceso ARMA$(p, q)$. Como punto de partida, se analizará el caso de un proceso autorregresivo $\{X_t\}$ que satisface una ecuación de diferencias de la forma

$$X_t + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} = Z_t \tag{4.1.1}$$

donde las variables aleatorias $Z_t$ son no-correlacionadas de media cero y varianza común $\sigma^2 > 0$, de manera que el proceso $\{Z_t\}$ es un ruido blanco; vea, por ejemplo, Anderson (1971) pp. 166–176, o Box y Jenkins (1976), pp. 53–65. Como ya se ha mencionado, tal proceso existe si y sólo si el polinomio

$$\varphi(z) := 1 + \varphi_1 z + \cdots + \varphi_p z^p$$

es tal que $\varphi(z) \neq 0$ para todos los números complejos $z$ con $|z| = 1$, y en este caso no hay pérdida de generalidad al suponer que el polinomio $\varphi(z)$ es causal, esto es, que $\varphi(z) \neq 0$ para todo $z$ con $|z| \leq 1$, condición que se supondrá en la siguiente discusión. Considere el problema de determinar la función de autocovarianza $\gamma(\cdot)$ de $\{X_t\}$, la cual está se denota por

$$\gamma(h) := \mathrm{Cov}\left(X_{t+h}, X_t\right), \quad h = 0, \pm 1, \ldots \tag{4.1.2}$$

Como punto de partida, sea $i$ un entero positivo y multiplique ambos lados de (4.1.1) por $X_{t-i}$ para obtener

$$X_t X_{t-i} + \varphi_1 X_{t-1} X_{t-i} + \cdots + \varphi_p X_{t-p} X_{t-i} = Z_t X_{t-i}$$

si se toma el valor esperado en ambos lados de esta expresión se desprende que

$$E[X_t X_{t-i}] + \varphi_1 E[X_{t-1} X_{t-i}] + \cdots + \varphi_p E[X_{t-p} X_{t-i}] = E[Z_t X_{t-i}],$$

relación que al combinarse con (4.1.2) conduce a

$$\gamma(i) + \varphi_1 \gamma(i-1) + \cdots + \varphi_p \gamma(i-p) = E[Z_t X_{t-i}], \tag{4.1.3}$$

Recuerde ahora que la causalidad del polinomio $\varphi(z)$ implica que la variable $X_{t-i}$ se expresa como una combinación lineal (infinita) de las variables $Z_{t-i}, Z_{t-i-1}, Z_{t-i-2}, \ldots,$ esto es,

$$X_{t-i} = \sum_{r=0}^{\infty} \psi_r Z_{t-i-r} \tag{4.1.4}$$

donde $\sum_{r=0}^{\infty} \psi_r z^r = 1/\varphi(z)$ para $|z| < 1$. Recordando que $i > 0$, se tiene que $E[Z_t Z_{t-i-r}] = 0$ para todo $r \geq 0$ (recordando que $\{Z_t\}$ es un ruido blanco), y entonces la expresión anterior implica que

$$E[Z_t X_{t-i}] = \sum_{r=0}^{\infty} \psi_r E[Z_t Z_{t-i-r}] = 0,$$

y combinando esta igualdad con (4.1.3) se desprende que

$$\gamma(i) + \varphi_1 \gamma(i-1) + \cdots + \varphi_p \gamma(i-p) = 0, \quad i = 1, 2, 3, \ldots \tag{4.1.5}$$

Por otro lado, multiplicando ambos lados de (4.1.1) por $X_t$ se desprende que

$$X_t X_t + \varphi_1 X_{t-1} X_t + \cdots + \varphi_p X_{t-p} X_t = Z_t X_t$$

de manera que al aplicar el operador de esperanza en ambos lados se llega a

$$E[X_t X_t] + \varphi_1 E[X_{t-1} X_t] + \cdots + \varphi_p E[X_{t-p} X_t] = E[Z_t X_t],$$

relación que al combinarse con (4.1.2) implica que

$$\gamma(0) + \varphi_1\gamma(-1) + \cdots + \varphi_p\gamma(-p) = E[Z_t X_t]$$

y aplicando la simetría de la función de autocovarianza, se concluye que

$$\gamma(0) + \varphi_1\gamma(1) + \cdots + \varphi_p\gamma(p) = E[Z_t X_t], \tag{4.1.6}$$

Usando ahora la expresión (4.1.4) con $i = 0$ se obtiene

$$X_t = \sum_{r=0}^{\infty} \psi_r Z_{t-r}$$

de modo que

$$E[Z_t X_t] = \sum_{r=0}^{\infty} \psi_r E[Z_t Z_{t-r}] = \psi_0 E[Z_t^2] = \psi_0 \sigma^2$$

como $\psi_0 = \psi(0) = 1/\varphi(0) = 1$, a partir de (4.1.6) se concluye que

$$\gamma(0) + \varphi_1\gamma(1) + \cdots + \varphi_p\gamma(p) = \sigma^2 \tag{4.1.7}$$

esta discusión se resume en el siguiente teorema.

**Teorema 4.1.1** *Considere el proceso autorregresivo* $X_t + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} = Z_t$ *donde el polinomio* $\varphi(z) = 1 + \varphi_1 z + \cdots \varphi_p z^p$ *es causal. En este caso, la función de autocovarianza* $\gamma(\cdot)$ *de* $\{X_t\}$ *satisface las ecuaciones*

$$\begin{aligned} \gamma(0) + \varphi_1\gamma(1) + \cdots + \varphi_p\gamma(p) &= \sigma^2 \\ \gamma(i) + \varphi_1\gamma(i-1) + \cdots + \varphi_p\gamma(i-p) &= 0, \quad i = 1, 2, 3, \ldots\ , \end{aligned}$$

**Observación 4.1.1** *Un sistema de ecuaciones análogo al anterior es satisfecho por la función de autocovarianza de un proceso* $ARMA(p, q)$*; la diferencia con las ecuaciones precedentes se ubica en los lados derechos de las ecuaciones.*

## 4.2. Procedimiento Recursivo

El Teorema 4.1.1 permite establecer un procedimiento de computo para la función de autocovarianza de un proceso autorregresivo, el cual se describe en Brockwell and Davis (1987), p. 97.

**Fase 1.** Determine $\gamma(0), \gamma(1), \cdots, \gamma(p)$ resolviendo el sistema

$$\begin{aligned} \textstyle\sum_{k=0}^{p} \gamma(k)\varphi_k &= \sigma^2 \\ \textstyle\sum_{k=0}^{p} \gamma(|i-k|)\varphi_k &= 0, \end{aligned} \tag{4.2.8}$$

el cual es el sistema de ecuaciones de Yule-Walker asociado al polinomio $\varphi(\cdot)$

**Fase 2.** Usando la relación

$$\gamma(i) = -\sum_{k=1}^{p} \gamma(i-k)\varphi_k, \quad i > p,$$

determine $\gamma(p+1)$, $\gamma(p+2), \gamma(p+3), \ldots$, recursivamente.

Para que este procedimiento quede bien establecido, es necesario mostrar que el sistema (4.2.8) tiene una única solución, un hecho que puede verificarse fácilmente para valores "pequeños"del grado de $\varphi(z)$, por ejemplo $p = 1$ o $p = 2$, pero el problema se vuelve realmente interesante para polinomios de grado mayor. Como ya se mencionó, el principal objetivo de este trabajo es mostrar que el sistema de ecuaciones de Yule-Walker tiene solución única para un polinomio causal de grado arbitrario. Esta conclusión fue obtenida por Luthkephol (2000), pero el argumento que se presenta en este trabajo es nuevo, basandose en ideas básicas de álgebra lineal. Así, el problema fundamental que se analiza a continuación se establece como sigue: Dado un polinomio $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ , determine una condición necesaria y suficiente para que el sistema (4.2.8) tenga una única solución $\{\gamma(h) \,|\, h = 0, 1, \cdots, p\}$. La solución a este problema se presenta más adelante en el Teorema 4.4.1, y arroja el siguiente criterio: El sistema de Yule-Walker asociado a un polinomo $\varphi(z)$ tiene sólución única si, y sólo si,

$$r_i r_j \neq 1, \quad i, j = 1, 2, \cdots, p$$

donde $r_1, \ldots, r_p$ son la raíces del polinomio $\varphi(z)$. Note que cuando $\varphi(z)$ es un polinomio causal, todas sus raíces $r_i$ se ubican dentro del disco unitario, y en este caso es claro que la condición (4.2) se satisface claramente. La demostración de este resultado utiliza el método de inducción, y será presentada después de establecer las herramientas técnicas necesarias. La organización del material subsecuente es como sigue: En la siguiente sección se introduce la notación y terminología necesarias, para formular posteriormente el principal resultado como Teorema 4.4.1. A continuación, en la Sección 5 se establecen las herramientas técnicas necesarias para demostrar el teorema principal, y el argumento se presenta en la Sección 6, y la presentación concluye en la Sección 7 con algunas observaciones finales.

## 4.3. Notación y Terminología

En el desarrollo del capítulo $\mathbb{Z}$ and $\mathbb{N}$ denotan a los conjuntos de todos los números enteros y enteros no negativos, respectivamente, y $\mathbb{C}$ representa al conjunto de números complejos. El espacio vectorial complejo $\mathcal{L}$ consiste de todos los vectores $\mathbf{v} : \mathbb{N} \to \mathbb{C}$ con la propiedad de que $\mathbf{v}(k) = 0$ para todos los índices $k$ suficientemente grandes y $\mathcal{L}$ esta dotado con las operaciones usuales de adición y multiplicación por escalar. En este contexto, el operador de traslado $s : \mathcal{L} \to \mathcal{L}$ se define como sigue: Para $\mathbf{v} \in \mathcal{L}$

$$s(\mathbf{v})(0) := 0, \quad \text{y} \quad s(\mathbf{v})(k) := \mathbf{v}(k-1), \quad k = 1, 2, \ldots; \tag{4.3.9}$$

adicionalmente, la $n$-ésima composición del operador $s$ consigo mismo se denota mediante $s^n$, así que

$$s^0(\mathbf{v}) = \mathbf{v}, \quad \text{y} \quad s^n(\mathbf{v}) = s^{n-1}(s(\mathbf{v})). \tag{4.3.10}$$

Por otro lado, las filas y columnas de una matriz cuadrada $M$ se numeran iniciando desde cero, y $\text{Det}M$ representa el determinante de $M$. Para vectores $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{L}$, se define una matriz cuadrada $M_{n+1}(\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n)$ de orden $(n+1)$ mediante

$$M_{n+1}(\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n) := \mathbf{v}_i(j), \quad i, j = 0, 1, \ldots, n, \tag{4.3.11}$$

mientras que $\text{span}\{\mathbf{v}_0, \ldots, \mathbf{v}_n\}$ denota al espacio vectorial generado por $\mathbf{v}_0, \ldots, \mathbf{v}_n$; la dimensión de este espacio se denota mediante $\dim \text{span}\{\mathbf{v}_0, \ldots, \mathbf{v}_n\}$. Para finalizar, dado un polinomio $\varphi(z) = \varphi_0 + \varphi_1 z + \cdots + \varphi_p z^p$ de grado $p$, los vectores $\overrightarrow{\varphi}$, $\overleftarrow{\varphi} \in \mathcal{L}$ se definen como sigue:

$$\begin{aligned} &\text{Para } k = 0, 1, \ldots, p \qquad \overrightarrow{\varphi}(k) :\ = \varphi_k \text{ y } \overleftarrow{\varphi}(k) :\ = \varphi_{p-k}, \\ &\text{y, para } k > p, \qquad \overrightarrow{\varphi}(k) = \overleftarrow{\varphi}(k) = 0. \end{aligned} \tag{4.3.12}$$

mientras que la siguiente convención notacional referente a los coeficientes del polinomio $\varphi(z)$ se utilizará:

$$\varphi_k := 0 \quad \text{si } k < 0 \quad \text{o} \quad k > p. \tag{4.3.13}$$

## 4.4. Resultado Principal

Sea $\varphi(z)$ un polinomio de grado $p$. El objetivo de esta sección es enunicar una fórmula para el determinante de de la matriz correspondiente al sistema Yule-Walker (4.2.8). Como punto de partida note que para $i > 0$,

$$\begin{aligned} \sum_{k=0}^{p} \gamma(|i-k|)\varphi_k &= \sum_{k=0}^{i} \gamma(i-k)\varphi_k + \sum_{k=i+1}^{p} \gamma(k-i)\varphi_k \\ &= \sum_{j=0}^{i} \gamma(j)\varphi_{i-j} + \sum_{j=1}^{p-i} \gamma(j)\varphi_{i+j} \end{aligned}$$

y usando la convención (4.3.13) se desprende que

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

Así, el sistema de Yule-Walker (4.2.8) se puede escribir como

$$\begin{aligned} \sum_{k=0}^{p} \gamma(k)\varphi_k &= \sigma^2 \\ \gamma(0)\varphi_i + \sum_{j=1}^{p} \gamma(j)[\varphi_{i-j} + \varphi_{i+j}] &= 0, \quad i = 1, 2, \ldots, p. \end{aligned} \tag{4.4.14}$$

La matriz cuadrada de este sistema se denotará mediante $M(\varphi)$. Con esta notación, no es díficil ver que $M(\varphi)$ es de orden $(p+1)$ y está dada por la siguiente expresión: Para $i = 0, 1, 2, \ldots, p$,

$$M(\varphi)_{i\,0} := \varphi_i, \text{ and } M(\varphi)_{i\,j} := \varphi_{i-j} + \varphi_{i+j}, \quad j = 1, 2, \cdots, p \tag{4.4.15}$$

por ejemplo, $M(\varphi)_{0\,0} = \varphi_0$ y, si $j > 0$, $M(\varphi)_{0\,j} = \varphi_{0-j} + \varphi_{0+j} = \varphi_j$, en concordancia con la primera ecuación en (4.4.14). El siguiente teorema presenta una fórmula para el determinante de $M(\varphi)$ y, como un subproducto, se obtiene un criterio para la no singularidad $M(\varphi)$.

**Teorema 4.4.1** *Sea $\varphi(z) = 1+\varphi_1 z+\cdots+\varphi_p z^p$ un polinomio complejo de grado $p$. Si las raíces de $\varphi(\cdot)$ son $r_1, \cdots, r_p$ y $M(\varphi)$ es como en (4.4.15), entonces las siguientes afirmaciones (i) and (ii) son válidas:*

*(i) El determinante de $M(\varphi)$ está dado por*

$$\text{Det}M(\varphi) = \prod_{1\leq i<j\leq p} [1-(r_i r_j)^{-1}] \prod_{i=1}^{p} (1-r_i^{-2}) \tag{4.4.16}$$

*donde, siguiendo la convención usual, para $p = 1$ el primer producto en esta relación desplegada es 1.*

*Consecuentemente,*

*(ii) $M(\varphi)$ es invertible si, y sólo si, $r_i r_j \neq 1$ para todo $i, j = 1, \ldots, p$.*

Este resultado se establecerá en la Section 6; por el momento, es conveniente notar que la parte (ii) se desprende como consecuencia directa de la fórmula en la parte (i). Por otro lado, (4.4.16) se verifica fácilmente para valores pequeños de $p$. Por ejemplo, para $p = 1$, el polinomio $\varphi(z)$ está dado por $\varphi(z) = 1 + \varphi_1 z$ y la fórmula (4.4.15) establece que

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

así que $\text{Det}M(\varphi) = 1 - \varphi_1^2$ , expresión que coincide con (4.4.16) para $p = 1$, ya que $\varphi(\cdot)$ tiene la única raíz $r_1 = -1/\varphi_1$. Cuando $p = 2$ factorice $\varphi(z)$ como

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

donde las raíces de $\varphi(z)$ son $r_i = -1/a_i$, $i = 1, 2$. En este caso se desprende que

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

y la matriz $M(\varphi)$ está dada por

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

Entonces, expandiendo Det$M(\varphi)$ por la tercera columna, se obtiene la siguiente expresión:

$$\begin{aligned}
\mathrm{Det}M(\varphi) &= a_1a_2[(a_1+a_2) - a_1a_2(1+a_1a_2)] + (1+a_1a_2) - (a_1+a_2)^2 \\
&= -(a_1+a_2)^2(1-a_1a_2) + (1+a_1a_2)[1-(a_1a_2)^2] \\
&= (1-a_1a_2)[-(a_1+a_2)^2 + (1+a_1a_2)^2] \\
&= (1-a_1a_2)(1-a_1^2)(1-a_2^2),
\end{aligned}$$

y al sustituire $a_i$ por $-1/r_i$ se obtiene la fórmula (4.4.16)para el caso $p = 2$. La demostración de (4.4.16) en el caso general se realizará por el método de indución y será presentada en la Sección 6.

## 4.5. Resultados Auxiliares

Esta sección contiene las herramientas técnicas que se utilizarán en la demostración del Teorema 4.4.1. El punto de partida es la siguiente idea.

**Definición 4.5.1** . *Sea $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ un polinomio de grado p. La sucesión $\mathbf{V}^\varphi = \{V_t^\varphi \,|\, t \in \mathbb{Z}\} \subset \mathcal{L}$ se define como sigue:*

*(i) Para $0 \le n < p$,*

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

*(ii) Si $n \in \mathbb{N}$*

$$V_{-n}^\varphi : = s^n(\overrightarrow{\varphi}) \qquad y \qquad V_{n+p}^\varphi : = s^n(\overleftarrow{\varphi});$$

*en este punto es conveniente recordar la notación introducida en (4.3.9)–(4.3.12).*

Comparando la definición anterior con (4.3.9)–(4.3.13) se desprende que la sucesión $\mathbf{V}^\varphi$ está relacionada con la matriz $M(\varphi)$ a través de la siguiente igualdad:

$$M(\varphi) = M_{p+1}(V_0^\varphi, V_1^\varphi, \ldots, V_p^\varphi). \tag{4.5.17}$$

El siguiente teorema es el principal resultado técnico de esta sección.

**Teorema 4.5.1** *Suponga que* $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ *tiene grado* $p$ *y satisface* $\varphi(b) = \varphi(1/b) = 0$ *para algún* $b \in \mathbb{C} \setminus \{0\}$. *En este caso las siguientes afirmaciones (i)–(iii) son válidas:*

*(i)* $\dim \operatorname{span}\{V_{-1}^{\varphi}, V_0^{\varphi}, \cdots, V_{p+1}^{\varphi}\} \leq p + 1.$

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

$$\mathrm{Det} M_{p+2}(V_0^{\varphi} + aV_{-1}^{\varphi}, V_1^{\varphi} + aV_0^{\varphi}, \cdots, V_{p+1}^{\varphi} + aV_p^{\varphi}) = 0.$$

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

La demostración del Teorema 4.5.1 se ha dividido en cuatro fases establecidas en los siguientes Lemas. La siguiente notación será útil.

**Definición 4.5.2** *Sea* $\mathbf{V} = \{V_t | t \in \mathbb{Z}\}$ *una sucesión en* $\mathcal{L}$. *(i)* $\mathbf{V}$ *tiene la propiedad* $\mathcal{D}(p)$ *si, para cada* $n \in \mathbb{N}$,

$$\dim \operatorname{span}\{V_t | -n \leq t \leq p + n\} \leq p + n.$$

*(ii) Dado* $a \in \mathbb{C}$, *la sucesión* $T_a\mathbf{V} = \{T_a V_t \,|\, t \in \mathbb{Z}\}$ *se define mediante*

$$[T_a\mathbf{V}]_t \equiv T_a V_t := V_t + aV_{t-1}, \quad t \in \mathbb{Z}.$$

El inicio de la demostración del Teorema 4.5.1 es el siguiente resultado.

**Lema 4.5.1** *Sea* $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ *un polinomio de grado* $p$ *y* $a \in \mathbb{C} \setminus \{0\}$. *Si* $\theta(z) = (1 + az)\varphi(z)$, *entonces* $\mathbf{V}^{\theta} = T_a\mathbf{V}^{\varphi}$.

**Demostración.** Sea $n \in \{1, 2, \ldots, p\}$ un entero arbitrario. a partir de la Definición 4.5.1 se desprende que

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

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

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

Estas dos últimas relaciones desplegadas muestran que, para $1 \leq n \leq p$, $V_n^\theta = T_a V_n^\varphi$, y para completar el argumento esta igualdad debe verificarse $n < 0$ y $n > p$. Para alcanzar este objetivo, note primero que $\overrightarrow{\theta} = \overrightarrow{\varphi} + as(\overrightarrow{\varphi})$ y $\overleftarrow{\theta} = s(\overleftarrow{\varphi}) + a\overleftarrow{\varphi}$, relaciones que se obtienen de (4.3.9) y (4.3.12). entonces, para $n \geq 0$,

$$\begin{aligned}
V_{-n}^\theta &= s^n(\overrightarrow{\theta}) \\
&= s^n[\overrightarrow{\varphi} + as(\overrightarrow{\varphi})] \\
&= s^n(\overrightarrow{\varphi}) + as^{n+1}(\overrightarrow{\varphi}) \\
&= V_{-n}^\varphi + aV_{-n-1}^\varphi \\
&= T_a V_{-n}^\varphi
\end{aligned}$$

Similarmente,

$$\begin{aligned}
V_{n+p+1}^\theta &= s^n(\overleftarrow{\theta}) \\
&= s^n[s(\overleftarrow{\varphi}) + a\overleftarrow{\varphi}] \\
&= s^{n+1}(\overleftarrow{\varphi}) + as^n(\overleftarrow{\varphi}) \\
&= V_{n+1+p}^\varphi + aV_{n+p}^\varphi \\
&= T_a V_{n+1+p}^\varphi.
\end{aligned}$$

Luego, se ha establecido que $V_{-n}^\theta = T_a V_{-n}^\varphi$ y $V_{n+1+p}^\theta = T_a V_{n+1+p}^\varphi$ para todo $n \in \mathbb{N}$, como ya se ha mencionado, esto completa la demostración. □

En el siguiente lema se analiza la relación entre la propiedad $\mathcal{D}(k)$ con la transformación $T_a$; vea la Definición 4.5.2.

**Lema 4.5.2** *Sea $a \in \mathbb{C}$ un número arbitrario, y suponga que $\mathbf{V} = \{V_t \,|\, t \in \mathbb{Z}\} \subset \mathcal{L}$ tiene la propiedad $D(k)$. En este caso, $T_a\mathbf{V}$ tiene la propiedad y $\mathcal{D}(k+1)$.*

**Demostración.** Observe que $T_a V_t \in \operatorname{span}\{V_t, V_{t-1}\}$, lo cual implica que, para cualquier $r \in \mathbb{N}$, se tiene que $\operatorname{span}\{T_a V_t | -r \leq t \leq k+1+r\} \subset \operatorname{span}\{V_t | -(r+1) \leq t \leq k+(r+1)\}$. A continuación, note que, debido a que la sucesión $\mathbf{V}$ tiene la propiedad $\mathcal{D}(k)$, el espacio en el lado derecho de la anterior igualdad tiene dimensión menor o igual a $k+r+1$ y, por lo tanto,

$$\dim \operatorname{span}\{T_a V_t \,|\, -r \leq t \leq (k+1)+r\} \leq (k+1)+r.$$

Esta relación muestra que $T_a\mathbf{V}$ tiene la propiedad $\mathcal{D}(k+1)$, puesto que el número $r$ fue arbitrario. $\square$

Los siguientes dos Lemas relacionan la propeiedad $\mathcal{D}(k)$ con sucesiones de la forma $\mathbf{V}^\varphi$; vea la Definición 4.5.1.

**Lema 4.5.3** *Sea $\varphi(z)$ un polinomio de grado $p \geq 1$ y suponga que $\varphi(1) = 0$ ó $\varphi(-1) = 0$. En este caso $\mathbf{V}^\varphi$ tiene la propiedad $\mathcal{D}(p)$.*

**Demostración.** Primero suponga que $p = 1$. Cuando $\varphi(1) = 0$ el polinomio $\varphi(z)$ estạdo por $\varphi(z) = \varphi(0)(1-z)$ y, usando (4.3.13), se desprende que $\overrightarrow{\varphi} = -\overleftarrow{\varphi}$, lo cual impica que, para $n \geq 0$, se tiene $V^\varphi_{-n} = s^n(\overrightarrow{\varphi}) = -s^n(\overleftarrow{\varphi}) = -V^\varphi_{n+1}$. Si $\varphi(-1) = 0$ se desprende que $\varphi(z) = \varphi(0)(1+z)$, y entonces $\overrightarrow{\varphi} = \overleftarrow{\varphi}$, lo que conduce a $V^\varphi_{-n} = V^\varphi_{n+1}$ para todo $n \in \mathbb{N}$. Luego, en cualquier caso para cada entero $r \in \mathbb{N}$,

$$\operatorname{span}\{V^\varphi_t \mid -r \leq t \leq 1+r\} = \operatorname{span}\{V^\varphi_t \mid 1 \leq t \leq 1+r\}$$

y puesto que el espacio en el lado derecho tiene $r+1$ generadores, se sigue que $\dim\operatorname{span}\{V^\varphi_t \mid -r \leq t \leq 1+r\} \leq r+1$, *i.e.*, $\mathbf{V}^\varphi$ tiene la propiedad $\mathcal{D}(1)$; vea la Definición 4.5.1. La demostración se completará ahora por inducción. Suponga que el resultado ocurre para $p = k$, y sea $\varphi$ un polinomio de grado $k+1$ anulandose en 1 ó $-1$. En este caso es claro que es posible factorizar $\varphi(z)$ como $\varphi(z) = (1+az)\theta(z)$, donde $a \in \mathbb{C}$ y $\theta(z)$ tiene grado $k$ y satisface $\theta(1) = 0$ o $\theta(-1) = 0$. Por la hipótesis de inducción, $\mathbf{V}^\theta$ tiene la propiedad $\mathcal{D}(k)$ y, usando los Lemas 4.5.1 y 4.5.2, se sigue que $\mathbf{V}^\varphi = T_a\mathbf{V}^\theta$ tiene la propiedad $\mathcal{D}(k+1)$. $\square$

El siguiente Lema es la etapa final antes de la demostración del Teorema 4.5.1.

**Lema 4.5.4** *Sea $\varphi(z)$ un polinomio de grado $p \geq 2$ el cual satisface $\varphi(b) = \varphi(1/b) = 0$ para algún $b \in \mathbb{C} \setminus \{0, 1, -1\}$. Entonces, $\mathbf{V}^\varphi$ tiene la propiedad $\mathcal{D}(p)$.*

**Demostración.** El argumento es similar al usado para establecer el Lema 4.5.3. Primero, suponga que $p = 2$. En este caso

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

y usando (4.3.12) se desprende que $\overrightarrow{\varphi} = \overleftarrow{\varphi}$; por la Definición 4.5.1 (i), esto implica que $V^{\varphi}_{-n} = V^{\varphi}_{1+n}$ para todo $n \geq 0$, y entonces, para cada $r \in \mathbb{N}$,

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

y puesto que el espacio vectorial en el lado derecho tiene $r+2$ generadores se desprende que

$$\dim \operatorname{span}\{V^{\varphi}_t \,|\, -r \leq t \leq 2+r\} \leq r+2,$$

esto es, $\mathbf{V}^{\varphi}$ tiene la propiedad $\mathcal{D}(2)$. El resultado para $p$ arbitrario se obtiene por medio de un argumento de inducción similar al usado para demostrar el Lema 4.5.3. $\square$

Los lemas anteriores se usarán a continuación para demostrar el Teorema 4.5.1.

**Demostración del Teorema 4.5.1.** 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 $b \in \mathbb{C} \setminus \{0\}$

(i) Cuando $b = 1$ o $b = -1$, el Lema 4.5.3 implica que $\mathbf{V}^{\varphi}$ tiene la propiedad $\mathcal{D}(p)$, y entonces el Lema 4.5.4 permite obtener la misma conclusión cuando $b \neq 1, -1$. Entonces, por la Definición 4.5.2(i), se desprende que $\dim \operatorname{span}\{V^{\varphi}_{-1}, V^{\varphi}_0, \ldots, V^{\varphi}_p, V^{\varphi}_{p+1}\} \leq p+1$.

(ii) Note que

$$\operatorname{span}\{V^{\varphi}_r + aV^{\varphi}_{r-1} \,|\, r = 0, 1, \ldots, p+1\} \subset \operatorname{span}\{V^{\varphi}_t \,|\, -1 \leq t \leq p+1\}$$

y entonces $\dim \operatorname{span}\{V^{\varphi}_r + aV^{\varphi}_{r-1} \,|\, r = 0, 1, \ldots, p+1\} \leq p+1$, por la parte (i). Se sigue que los $p+2$ vectores $V^{\varphi}_r + aV^{\varphi}_{r-1}$, $r = 0, 1, 2, \ldots, (p+1)$ son linealmente dependientes en $\mathcal{L}$, lo cual implica la dependencia lineal de las filas de la matriz $M_{p+2}(V^{\varphi}_0 + aV^{\varphi}_{-1}, \cdots, V^{\varphi}_{p+1} + aV^{\varphi}_p)$; como una consecuencia de este hecho se tiene que

$$\operatorname{Det} M_{p+2}(V^{\varphi}_0 + aV^{\varphi}_{-1}, \cdots, V^{\varphi}_{p+1} + aV^{\varphi}_p) = 0$$

vea, por ejemplo, el capítulo 5 de Hoffman and Kunze (1971).

(iii) Defina $\psi(z) := (1+az)\varphi(z)$. En este caso $\psi(z)$ tiene grado $p+1$, y usando (4.5.17) con $p+1$ y $\psi$ en lugar de $p$ y $\varphi$, respectivamente, se desprende que

$$\begin{aligned} M(\psi) &= M_{p+2}(V^{\psi}_0, V^{\psi}_1, \cdots, V^{\psi}_{p+1}) \\ &= M_{p+2}(V^{\varphi}_0 + aV^{\varphi}_{-1}, \cdots, V^{\varphi}_{p+1} + aV^{\varphi}_p \end{aligned}$$

donde el Lema 4.5.1 se utilizó para establecer la segunda igualdad; a partir de este punto, la parte precedente implica que $\mathrm{Det}M(\psi) = 0$. $\square$

Esta sección concluye con un hecho que será útil en la demostración del Teorema 4.4.1.

**Lema 4.5.5** *Sea $\varphi(z)$ un polinomio de grado $p$ con $\varphi(0) = 1$. En este caso*

$$\mathrm{Det}M(\varphi) = DetM_{p+2}(V_0^{\varphi}, V_1^{\varphi}, \ldots, V_{p+1}^{\varphi}).$$

**Demostración.** Defina $L := M_{p+2}(V_0^{\varphi}, V_1^{\varphi}, \ldots, V_{p+1}^{\varphi})$ y observe los siguientes hechos (a) y (b):

(a) Combinando (4.3.11) con la Definición 4.5.1(i) se sigue que la submatriz obtenida eliminando la última fila y la última columna de $L$ es $M_{p+1}(V_0^{\varphi}, V_1^{\varphi}, \ldots, V_p^{\varphi})$.

A continuación, se evaluán las componentes en la última columna de $L$. Primero, recuerde la Definición 4.5.1 y note que $L_{0\,p+1} = V_0^{\varphi}(p+1) = s^0(\overrightarrow{\varphi})(p+1) = \overrightarrow{\varphi}(p+1) = 0$. Por otro lado, para $1 \leq n < p$, se tiene que $L_{n\,p+1} = \varphi n + p + 1 + \varphi_{n-p-1} = 0$, donde la convención 4.3.13) se utilizó 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 $L_{p+1\,p+1} = s(\overleftarrow{\varphi})(p+1) = \overleftarrow{\varphi}(p) = \varphi_0 = \varphi(0) = 1$.

En resumen:

(b) La última columna de $L$ consiste enteramente de ceros, excepto por el elemento en la última fila, el cual es 1.

Para concluir, expanda $\mathrm{Det}L$ por la última columna, y usando los anteriores hechos (a) y (b) se desprende que $\mathrm{Det}L = M_{p+1}(V_0^{\varphi}, V_1^{\varphi}, \ldots, V_p^{\varphi}) = \mathrm{Det}M(\varphi)$, donde la última igualdad es consecuencia de (4.5.17). $\square$

## 4.6. Demostración del Teorema 4.4.1

Los resultados preliminares establecidos en la sección anterior se usarán ahora para demostrar el principal resultado de este trabajo.

**Demostración del Teorema 4.4.1.** Como ya se mencionó, es suficiente establecer la parte (i). Sea $\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p$ un polinomio de grado $p$ y factorice $\varphi$ como

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

donde las raíces de $\varphi(z)$ son $-1/a_i$, $i = 1, 2, \cdots, p$. Con esta notación, (4.4.16) es equivalente a

$$\text{Det}M\left[\prod_{i=1}^{p}(1 + a_i z)\right] = \prod_{l \leq i < j \leq p}[1 - a_i a_j]\prod_{i=1}^{p}(1 - a_i^2), \tag{4.6.18}$$

igualdad que se verificó en la Sección 4.4 para $p = 1$ y $p = 2$. La demostración de (4.6.18) se completará ahora por inducción. Suponga que (4.6.18) es válido para $p = n \geq 2$, y sean $a_1, a_2, \cdots, a_{n+1}$ números complejos distintos de cero. Defina

$$\psi(z) := \prod_{i=1}^{n}(1 + a_i z) \tag{4.6.19}$$

y, para cada $c \in \mathbb{C}$, defina

$$F(c) := \text{Det}M_{n+2}(V_0^\psi + cV_{-1}^\psi, \cdots, V_{n+1}^\psi + cV_n^\psi). \tag{4.6.20}$$

Combinando los Lemas 4.5.1 y 4.5.17 se desprende que

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

$$F(a_{n+1}) = \text{Det}M\left[\prod_{i=1}^{n+1}(1 + a_i z)\right]. \tag{4.6.21}$$

Usando la multilinealidad de la función determinante, (4.6.20) implica que

(b) $F(c)$ es un polinomio de de grado $\leq n + 2$ en $c$; vea, por ejemplo, el capítulo 5 de Hoffmann and Kunze (1971):

Ahora, se determinarán las raíces del polinomio $F(c)$. Primero observe

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

Para verificar esta afirmación, defina $\psi^*(z) := (1+z)\prod_{i=2}^{p}(1+a_i z)$ y note que $\psi^*(-1) = 0$, y $(1+z)\psi(z) = (1+a_1 z)\psi^*(z)$; vea (4.6.19). Entonces, una aplicación de la anterior propiedad (a) implica que

$$F(1) = \mathrm{Det}M[(1+z)\psi(z)] = \mathrm{Det}M[(1+a_1 z)\psi^*(z)] = 0,$$

donde la última igualdad se desprende del Teorema 4.5.1(iii) con $\psi^*(z)$ y $-1$ en lugar de $\varphi(z)$ y $b$, respectivamente. De forma similar puede mostrarse que $F(-1) = 0$.

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

Para establecer esta afirmación, sea $k, i \in \{1, 2, \ldots, n\}$ un entero fijo con $k \neq i$ y defina

$$\tilde{\psi}(z) :\ = (1 + z/a_i)\prod^{(k)}(1 + a_j z), \tag{4.6.22}$$

donde $\prod^{(k)}$ indica el producto sobre todos los enteros $j$ entre 1 y $n$ que satisfacen $j \neq k$; note que

$$(1 + z/a_i)\psi(z) = (1 + a_k z)\tilde{\psi}(z). \tag{4.6.23}$$

A partir de (4.6.22) se obtiene que $\tilde{\psi}(-a_i) = 0$ y, puesto que $k \neq i$, el polinomio $\tilde{\psi}(z)$ contiene al factor $(1 + a_i z)$, y entonces $\tilde{\psi}(-1/a_i) = 0$. Por lo tanto, a partir de (a) y (4.6.23) se sigue que $F(1/a_i) = \mathrm{Det}M[(1 + z/a_i)\psi(z)] = \mathrm{Det}M[(1 + a_k z)\tilde{\psi}(z)]$, y una aplicación del Teorema 4.5.1(iii) con $\tilde{\psi}$ y $-a_i$ en lugar de $\varphi$ y $b$, respectivamente, lleva a la conclusión de que $F(1/a_i) = 0$.

Para continuar, suponga por el momento que $a_1, a_2, \ldots, a_n$ son números diferentes en $\mathbb{C}\backslash\{0, 1, -1\}$. En este caso, los puntos anteriores (c) y (d) muestran que el polinomio $F(c)$ tiene $n+2$ raíces diferentes, a saber, $1, -1$, y $1/a_i$ , $i = 1, 2, \cdots, n$. Combinando este hecho con (b) se desprende que $F(\cdot)$ tiene grado $n + 2$ y que puede factorizarse como

$$F(c) = F(0)(1 - c)(1 + c)\prod_{i=1}^{n}(1 - a_i c).$$

Conociendo $c = a_{n+1}$ y usando (a) se sigue que

$$DetM\left[\prod_{i=1}^{n+1}(1 + a_i z)\right] = F(0)(1 - a_i^2)\prod_{i=1}^{n}(1 - a_i a_{n+1}). \tag{4.6.24}$$

A continuación, usando (4.6.20) se ve que $F(0) = \mathrm{Det} M_{n+2}(V_0^{\psi}, V_1^{\psi}, \cdots, V_{n+1}^{\psi})$, y entonces $F(0) = M(\psi)$, por el Lema 4.5.5 aplicado al polinomio $\psi$, el cual tiene grado $n$, y entonces la hipótesis de inducción permite concluir que

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

Combinando esta igualdad con (4.6.24) se desprende que

$$\mathrm{Det} M\left[\prod_{i=1}^{n+1}(1 + a_i z)\right] = \prod_{i=1}^{n+1}(1 - a_i^2) \prod_{1 \le i < j \le n+1}(1 - a_i a_j), \tag{4.6.25}$$

relación que es (4.6.18) con $p = n + 1$. Aunque (4.6.25) ha sido establecida bajo el supuesto de que $a_1, \ldots, a_{n+1}$ son números diferentes en $\mathbb{C} \setminus \{0, 1, -1\}$, la igualdad ocurre para $a_i, \ldots, a_{n+1} \in \mathbb{C} \setminus \{0\}$ arbitrarios, puesto que ambos lados de (4.6.25) dependen continuamente de las $a_i$'s. En resúmen, suponiendo que la igualdad (4.6.18) es válida $p = n$, se ha demostrado que también ocurre para $p = n + 1$, concluyendo el argumento de inducción. $\square$

## 4.7. Observaciones Finales

Dado un polinomio $\varphi(z)$ con $\varphi(0) = 1$, se ha establecido una condición necesaria y suficiente para el sistema correspondiente de Yule-Waker tenga solución única, y se ha demostrado que tal condición se saisface cuando el polinomio $\varphi(z)$ es causal. Además de proporcionar una base firme pare el procedimeinto de dos etapas descrito en la Sección 4.2, hay otras partes en la teoría de las series de tiempo en donde es importante saber que el sistema de Yule-Walker tiene una solución única. Por ejemplo, considere el siguiente resultado: Dado $p$ >0 y una función de autocovarianza $\gamma(\cdot)$, con $\gamma(h) \to 0$ conforme $h \to \infty$, existe un proceso $\{Y_t\}$ autorregresivo de orden $p$ cuya función de autocovarianza $\gamma_Y(\cdot)$ coincide con $\gamma(\cdot)$ para $h = 0, 1, \cdots, p$. Una demostración de este resultado se encuentra en Brockwell and Davis (1987), pp. 232-233, y es oportuno mencionar que en un pasaje del argumento se utiliza que, para cierto polinomio causal, $\varphi(\cdot)$, tanto $\{\gamma(h)|0 \leq h \leq p\}$ como $\{\gamma_Y(h)\,|\,0 \leq h \leq p\}$ satisfacen (4.2.8), y entonces se concluye de inmediato que $\gamma(h) = \gamma_Y(h)$ para $0 \leq h \leq p$. Así, se está suponiendo en el mencionado argumento que el sistema de Yule-Walker de un polinomio causal tiene una solución única, un hecho que, por el Teorema 4.4.1, es realmente cierto.

# Bibliografía

[1] L. Alfhors (1980), Complex Variables, *McGraw-Hill*, New York.

[2] T. M. Apostol (1980), Mathematical Analysis, *Addison Wesley*, Reading, Massachusetts.

[3] A. A. Borovkov (1999), Mathematical Statistics, *Gordon and Breach*, New York

[4] P. J. Brockwell y R. A. Davis (1998), Time Series: Theory and Methods, *Springer-Verlag*, New York.

[5] E. Dudewica y S. Mishra (1998). Mathematical Statistics, *Wiley*, New York.

[6] W. A. Fuller (1998), Introduction to Statistical Time Series ,*Wiley*, New York.

[7] W. Fulks (1980), Cálculo Avanzado, *Limusa*, México, D. F.

[8] F. A. Graybill (2000), Theory and Application of the Linear Model, *Duxbury*, New York.

[9] F. A. Graybill (2001), Matrices with Applications in Statistics *Duxbury*, New York.

[10] D. A. Harville (2008), Matrix Algebra Form a Statistician's Perspective, *Springer-Verlaf*, New York.

[11] K. Hoffman y R. Kunze (1975), Linear Algebra, *Prentice-Hall*, New York.

[12] A. I. Khuri (2002), Advanced Calculus with Applications in Statistics, *Wiley*, New York.

[13] S. Lipschutz (1995), Linear Algebra, *McGraw-Hill*, New York.

[14] M. Loève (1984), Probability Theory, I, Springer-Verlag, New York.

[15] J. C. Chacón Hernández (2010), Un Criterio Integral de Causalidad para Procesos ARMA, Tesis de Maestría en Estadística Aplicada, Universidad Autónoma Agraria Antonio Narro, Saltillo Coah., MÉXICO.

[16] W. Rudin (1984), Real and Complex Analysis, *McGraw-Hill*, New York.

[17] H. L. Royden (2003), Real Analysis, *MacMillan*, London.

[18] J. Shao (2010), Mathematical Statistics, *Springer*, New York.

[19] R. H. Shumway y D. S. Stoffer (2006), Time Series Analysis and Its Applications With R Examples, *Springer-Verlag*, New York. Edition

[20] D. Wackerly, W. Mendenhall y R. L. Scheaffer (2009), Mathematical Statistics with Applications, *Prentice-Hall*, New York.