

00321



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

40

FACULTAD DE CIENCIAS

la Dirección General de Bibliotecas de la
para difundir en formato electrónico e impreso el
entendido de mi trabajo profesional.

COMITE: Gabriela Hernández Godínez

FECHA: 11 de agosto de 2003

APUNTES PARA UN CURSO SOBRE SERIES DE TIEMPO A NIVEL PROFESIONAL

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I A

P R E S E N T A :

GABRIELA HERNÁNDEZ GODÍNEZ

DIRECTOR DE TESIS:

M. en C. JOSÉ ANTONIO FLORES DÍAZ



FACULTAD DE CIENCIAS 2003 SECCIÓN ESCOLAR



FACULTAD DE CIENCIAS UNAM

TESIS CON FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:
Apuntes para un curso sobre series de tiempo a nivel profesional,

realizado por Gabriela Hernández Godínez

con número de cuenta 9018066 - 4 , quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de Tesis

Propietario M. en C. José Antonio Flores Díaz

Propietario M. en A.P. María del Pilar Alonso Reyes

Propietario Act. Jaime Vázquez Alanilla

Suplente Act. Marypaola Janett Maya López

Suplente Act. Gerardo Chávez Heredia

Consejo Departamental de Matemáticas

M. en C. José Antonio Flores Díaz

A Dios por darme la vida.

Con cariño a mis padres, que en las buenas y en las malas sé que cuento con ellos.

A mi hermanas Erika, Julissa y Jacqueline que siempre estaré con ellas

Y por supuesto, al hombre que amo y admiro,

Mi esposo Raúl

*Que suerte he tenido de nacer,
para tener acceso a la fortuna
de ser río y no laguna,
de ser lluvia en lugar de ver llover.*

Contenido

Introducción

1

1. Introducción a las series de tiempo	2
1.1 Descripción general.....	2
1.1.1 Utilidad de los procesos estocásticos en el análisis de las series de tiempo.....	5
1.1.2 Series de tiempo discretas.....	5
1.1.3 Características de los operadores y polinomios de retraso.....	6
1.1.4 Procesos estocásticos lineales.....	11
1.1.5 Estacionariedad.....	12
1.2 Ecuaciones en diferencia como herramienta en el análisis de series de tiempo.....	25
1.2.1 Relación de procesos deterministas con ecuaciones en diferencia.....	25
1.2.2 Ecuaciones en diferencia de primer orden.....	26
1.2.3 Ecuaciones en diferencia de segundo orden.....	27
1.2.4 Ecuaciones en diferencia de orden p	31
1.2.5 Procesos divergentes.....	36
2. Modelos para series de tiempo	42
2.1 Modelos autorregresivos (AR).....	42
2.1.1 Caso general de los modelos AR(1).....	43
2.1.2 Caso para los modelos AR(2).....	45
2.1.3 Caso general para los modelos AR(p).....	49
2.2 Modelos de promedios móviles (MA).....	55
2.2.1 El modelo MA de orden uno MA(1).....	56
2.2.2 Caso para los modelos MA(2).....	58
2.2.3 Caso general para los modelos MA(q).....	60
2.3 Los modelos ARMA.....	62
2.3.1 El modelo ARMA(1,1).....	63
2.3.2 El modelo ARMA(1,2).....	67
2.3.3 Caso ARMA(p,q).....	71
2.4 Generalización de los modelos ARMA, los modelos ARIMA.....	73
2.4.1 Extensión de modelos ARIMA debido a una tendencia determinista.....	75

2.4.2	Pronósticos con modelos ARIMA.....	75
2.5	Modelos estacionales o temporales.....	78
2.5.1	Modelo multiplicativo.....	81
3.	Ajuste de series de tiempo con modelos ARIMA.....	82
3.1	Identificación.....	83
3.1.1	Estabilización de la varianza.....	84
3.1.2	Estabilización del nivel.....	88
3.1.3	Función de Autocorrelación y Función de Autocorrelación Parcial.....	90
3.2	Estimación.....	95
3.2.1	Método de momentos.....	95
3.2.2	Verificación.....	99
	Conclusiones	112
	Apéndice	113
	Bibliografía	120

INTRODUCCIÓN

Uno de los problemas fundamentales en diversas áreas de la actividad humana es el conocimiento de un fenómeno para controlarlo y predecirlo. Esto es de suma importancia debido a la fuerte implicación moral o económica que se deriva de una mala decisión o de la tardía acción cuando algo no marcha bien, si por el contrario, podemos pronosticar mediante información conocida y un correcto modelo matemático las tendencias a futuro, y se puede actuar a tiempo cuando algo no marcha correctamente, se pueden prevenir muchas situaciones desfavorables que pudieran ocasionar cuantiosas pérdidas sobre todo económicas. Por ejemplo, el tener el conocimiento de la tendencia de las tasas de rendimiento en los próximos años, para así poder tomar la mejor decisión en el presente para actuar con medidas oportunas ante beneficios y adversidades.

Las ventas de un cierto producto, la inflación, así como otros indicadores (no sólo de carácter económico) son cantidades que se observan periódicamente, dando origen a una serie temporal, de la cual se espera que, bajo ciertos supuestos, se pueda explicar una situación futura a partir de la presente.

Para lograr esto, se han propuesto una diversidad de métodos que se basan en intuición y en aspectos matemáticos formales, conocidos como series de tiempo. Orientado a personas de nivel profesional el presente trabajo trata de explicar en 3 capítulos los conceptos básicos de series de tiempo al mismo tiempo que se desarrollan algunos de los resultados que se utilizan de manera tradicional, así como algunos ejemplos para su mejor comprensión.

CAPITULO 1

Introducción a las series de tiempo

1.1 Descripción general

El concepto de series de tiempo esta definido como "El registro metódico de la medición u observación numérica, efectuada a intervalos de tiempo fijos, de tales características o variables."¹

Existen dos aspectos en el estudio de las series de tiempo, el análisis y el modelo de la serie. El análisis de los datos consiste básicamente en resumir las propiedades de una serie y caracterizarla por sus rasgos sobresalientes, para ello en el presente trabajo se hará uso de la teoría matemática, utilizando los procesos estocásticos y ciertas herramientas provenientes de la estadística. La principal razón de modelar una serie de tiempo es facilitar el pronóstico de valores en un futuro. La característica que distingue a una serie de tiempo de un modelo econométrico, por ejemplo, es que la primera no pretende explicar el comportamiento de y_t en términos de otras variables sino por el contrario lo explica en función de sus propios valores, observados en periodos anteriores a t , de manera ponderada.

Por esta razón, en este primer capítulo se presentarán los conceptos básicos para entender la teoría de las series de tiempo. En estadística se utilizan principalmente dos tipos de análisis, estos son el descriptivo y el inferencial.

- 1) El descriptivo. Consiste esencialmente en resumir y describir en forma concisa, mediante gráficas o a través de unas cuantas medidas descriptivas, la información con la que se cuenta.
- 2) El inferencial. Tiene por objeto utilizar muestras representativas para realizar inferencias que sean válidas para toda la población de donde se obtuvo la muestra y dependen del tipo de análisis o modelo que se utilice.

Para ilustrar el análisis descriptivo, en la gráfica 1, se muestra el comportamiento de los datos del costo de captación mensual de los pasivos a plazo denominados en moneda nacional, (CCP). Se puede observar que los datos tienen una tendencia decreciente desde su inicio hasta la actualidad, aunque hay periodos en los cuales se sale de ese patrón, el más significativo esta en el intervalo de septiembre de 1998 a abril de 1999, dentro del cual se alcanza un máximo de 31.43% en el mes de octubre de 1998. Para el mes de mayo de 1999 nuevamente

¹ Victor M. Guerrero. Análisis Estadístico de Series de Tiempo Económicas. UAM Iztapalapa. 1991. Pag.1.

vuelve a su tendencia natural. También hay un periodo en el cual la media y la varianza podrían ser suficientes para describir a la parte de la serie que cubre el periodo mayo de 1997 a julio de 1998; Faltaría considerarse la explicación que dé el cálculo de la media y la varianza, así como de máximos y mínimos locales y globales.

Cuadro 1.
Costo de Captación de los Pasivos a Plazo denominados en Moneda Nacional (CCP)
 (Tasas mensuales)

MES	1996	1997	1998	1999	2000	2001	2002	2003
Enero		25.29	17.91	30.01	16.28	15.96	6.19	6.42
Febrero	37.22	22.09	17.95	28.64	16.07	15.82	6.76	7.16
Marzo	40.33	21.82	18.36	24.21	14.42	15.12	6.54	
Abril	36.61	21.91	18.65	20.31	13.12	14.33	5.46	
Mayo	30.54	19.59	17.72	18.95	13.29	12.65	5.71	
Junio	27.96	19.6	18.18	19.75	14.42	10.15	6.15	
Julio	30.01	18.97	18.86	19.09	13.84	8.97	6.37	
Agosto	28.58	18.21	20.07	19.18	13.98	8.04	5.81	
Septiembre	25.84	18.12	29.57	18.96	14.16	8.16	5.97	
Octubre	25.79	17.36	31.43	18.11	14.44	8.53	6.47	
Noviembre	29.01	18.63	29.45	17.10	15.56	7.22	6.38	
Diciembre	28.19	18.87	30.47	16.34	15.55	6.49	6.21	

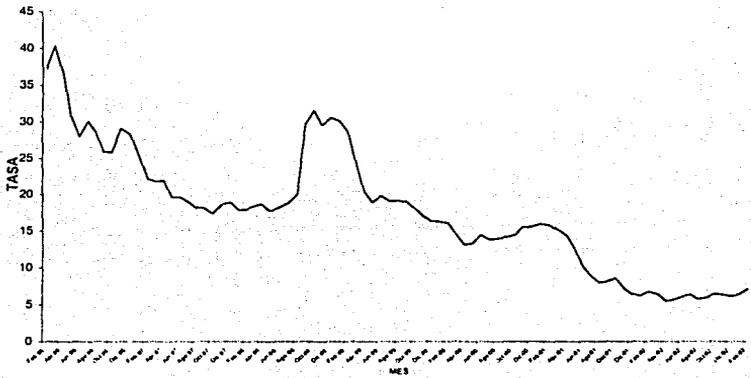
Cabe mencionar que dada una serie de datos su análisis puede ser realizado de maneras distintas, el método considerado como clásico, y con el cual se trabajará aquí, es el de *descomposición de la serie*², el cual contempla los siguientes componentes:

- a) Movimiento estacional; representa al efecto producido por algún fenómeno observado con cierta constancia o en un determinado periodo.

² Tesis titulada "Análisis del comportamiento de los índices de precios considerados como series cronológicas" de Rodríguez Araiza Albino.

- b) Cambios Cíclicos; parte del efecto estacional, algunas series de tiempo muestran variaciones en intervalos fijos de tiempo, debido a alguna otra causa fisica o bien son predecibles en alguna extensión.
- c) Cambios Cíclicos; parte del efecto estacional, algunas series de tiempo muestran variaciones en intervalos fijos de tiempo, debido a alguna otra causa fisica o bien son predecibles en alguna extensión.

GRÁFICA 1
CCP Moneda Nacional.



- d) Tendencia; se refiere al crecimiento o decrecimiento de los valores de la serie a través del tiempo.
- e) Irregularidad; se refiere a los movimientos imprevisibles de la serie debido a situaciones no controlables que afectan los valores que toma la serie y considerados como aleatorios.

Es muy importante tener bien identificada a cada una de las componentes de la serie, además se dice que los elementos de tendencia ciclo y estacionalidad, constituyen la parte determinista o semideterminista de la serie, pues se pueden

determinar mediante el comportamiento de los datos de la serie, mientras que la componente irregular viene a ser la parte no-determinista o estocástica, pues entra aquí la probabilidad de ocurrencia o no, del evento imprevisible.

1.1.1 Utilidad de los procesos estocásticos en el análisis de las series de tiempo.

Una serie no determinista se puede considerar como una realización de un proceso estocástico, el cual, se define como una familia de variables aleatorias asociadas a un conjunto índice de números reales, de forma tal que a cada elemento del conjunto le corresponde una y sólo una variable aleatoria, es decir, $\{Z(\gamma); \gamma \in T\}$, en donde T es el conjunto índice y $Z(\gamma)$ es la variable aleatoria correspondiente al elemento γ de T . Si T es un intervalo de números reales, cerrado o abierto, se dice que el proceso estocástico es continuo y si T es un conjunto finito o infinito pero numerable, el proceso estocástico es discreto, no obstante este hecho no dice nada acerca de las variables aleatorias involucradas, pues éstas a su vez pueden ser continuas o discretas.

1.1.2 Series de tiempo discretas.

Una serie de tiempo es entonces una sucesión de observaciones generadas por un proceso estocástico cuyo conjunto índice se toma con relación al tiempo, por lo que la inferencia a realizar será sobre las características del proceso estocástico generador de la serie observada. Por otro lado, así como existen procesos estocásticos continuos y discretos también existen series continuas y discretas. De aquí en adelante se trabajará con la teoría para series de tiempo discretas cuyas observaciones son hechas a intervalos equidistantes de tiempo es decir, si se toman en los momentos $\gamma_1, \gamma_2, \dots, \gamma_N$, el proceso estocástico respectivo se denotará por $\{Z(\gamma_1), Z(\gamma_2), \dots, Z(\gamma_N)\}$, no se hará distinción entre una variable aleatoria Z y su valor observado que se denotará de manera semejante, una vez establecido esto, cuando se tengan N valores sucesivos de una serie de tiempo se indicarán como $Z_1, Z_2, \dots, Z_t, \dots, Z_N$ a las observaciones hechas a intervalos equidistantes $\gamma_0+h, \gamma_0+2h, \dots, \gamma_0+th, \dots, \gamma_0+Nh$, donde γ_0 es el punto de origen y h es la longitud del intervalo que separa a dos observaciones inmediatas o seguidas. Para fines prácticos la serie de tiempo se podrá denotar simplemente como $\{Z_t\}$, con el supuesto implícito de que t toma los valores $1, 2, \dots, N$.

Con esta notación se facilita el manejo de las observaciones, pues no será necesario para cada una de ellas, indicar la fecha en que fue observada, así si se tiene una serie de tiempo diaria cuya primera observación fue hecha el 1° de enero de 2001, bastará con escribir Z_{366} para referirse a la observación del 1° de enero de 2002.

Una serie de tiempo es una realización de un proceso estocástico, y una ejecución del mismo proceso puede tener un comportamiento distinto al que se observó en la realidad, debido al elemento probabilístico en la serie de tiempo; mismo que conduce a tomar en cuenta a la función de densidad conjunta de todas las variables aleatorias que constituyen el proceso estocástico. En estadística el comportamiento de una variable aleatoria Z está definido por su función de densidad $f(Z)$, del mismo modo dos variables aleatorias Z_1 y Z_2 quedarán completamente definidas por su función de densidad conjunta $f(Z_1, Z_2)$. Generalizando N variables aleatorias y con ello un proceso estocástico, podrán describirse mediante la función de densidad conjunta $f(Z_1, Z_2, \dots, Z_N)$.

Como es sabido, en casi todo análisis estadístico van implícitas las hipótesis de que las variables aleatorias son independientes de tal forma que con el sólo conocimiento de las funciones de densidad individuales es posible obtener la conjunta. En el análisis de series de tiempo hay una excepción, pues aquí se supone que existe toda una estructura de correlación entre las observaciones, por lo que no es tan directa la obtención de la función de densidad conjunta y por lo que deberá emplearse algún otro camino para caracterizar a las variables aleatorias que intervienen en el análisis. En éste capítulo y el siguiente se presentan y justifican la utilización de algunos operadores y polinomios de uso frecuente en el análisis de series de tiempo.

1.1.3 Características de los operadores y polinomios de retraso.

El operador llamado **de retraso**, denotado por la letra B (del Inglés Backward) esta definido como:

$$BZ_t = Z_{t-1} \quad \forall t$$

y si se aplica sucesivamente el operador de retraso B se tiene que

$$B^2 Z_t = B(BZ_t) = BZ_{t-1} = Z_{t-2}$$

$$B^3 Z_t = B(B^2 Z_t) = BZ_{t-2} = Z_{t-3}$$

$$\vdots$$

$$B^k Z_t = B(B^{k-1} Z_t) = BZ_{t-k+1} = Z_{t-k}$$

por lo que se obtiene la expresión

$$B^k Z_t = Z_{t-k} \quad \text{para } k = 0, 1, 2, \dots \text{ y } \forall t.$$

es decir, al aplicar k -veces el operador B , se obtiene la variable retrasada k períodos y ya que $B^0 = I$ entonces $B^0 Z_t = Z_t$. Hay que tener en cuenta que el operador modifica a toda la sucesión de valores $\{Z_1, Z_2, \dots, Z_t, \dots, Z_N\}$ para transformarla en la nueva sucesión $\{Z_{1-k}, Z_{2-k}, \dots, Z_{t-k}, \dots, Z_{N-k}\}$. Es decir, si la serie originalmente contaba con N observaciones al aplicar el operador B^k , las observaciones $Z_{1-k}, Z_{2-k}, \dots, Z_0$ no se tendrán, quedando así una serie de $N-k$ observaciones.

Para dar un ejemplo de la aplicación de los operadores de retraso se tomarán los datos del Costo de Captación de los Pasivos en Moneda Nacional (CCP) que aparecen en el cuadro 1

Si se denota a dichos datos por la serie de tiempo $\{Z_t\}$, se tiene que

$Z_1 = 37.22$	$Z_4 = 30.54$	$Z_7 = 28.58$	$Z_{10} = 29.01$	$Z_{13} = 22.09$	$Z_{16} = 19.59$
$Z_2 = 40.33$	$Z_5 = 27.96$	$Z_8 = 25.84$	$Z_{11} = 28.19$	$Z_{14} = 21.82$	$Z_{17} = 19.60$
$Z_3 = 36.61$	$Z_6 = 30.01$	$Z_9 = 25.79$	$Z_{12} = 25.29$	$Z_{15} = 21.91$	$Z_{18} = 18.97$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Z_{17} = 28.64$	$Z_{20} = 18.95$	$Z_{23} = 19.18$	$Z_{26} = 17.10$	$Z_{29} = 16.07$	$Z_{32} = 13.29$
$Z_{18} = 24.21$	$Z_{21} = 19.75$	$Z_{24} = 18.96$	$Z_{27} = 16.34$	$Z_{30} = 14.42$	$Z_{33} = 14.42$
$Z_{19} = 20.31$	$Z_{22} = 19.09$	$Z_{25} = 18.11$	$Z_{28} = 16.28$	$Z_{31} = 13.12$	$Z_{34} = 13.84$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Z_{28} = 8.16$	$Z_{31} = 6.49$	$Z_{34} = 6.54$	$Z_{37} = 6.15$	$Z_{40} = 5.97$	$Z_{43} = 6.21$
$Z_{29} = 8.53$	$Z_{32} = 6.19$	$Z_{35} = 5.46$	$Z_{38} = 6.37$	$Z_{41} = 6.47$	$Z_{44} = 6.42$
$Z_{30} = 7.22$	$Z_{33} = 6.76$	$Z_{36} = 5.71$	$Z_{39} = 5.81$	$Z_{42} = 6.38$	$Z_{45} = 7.16$

si se construyen tres nuevas series $\{P_t\}$, $\{Q_t\}$ y $\{R_t\}$ definidas como $P_t = BZ_t$, $Q_t = B^2 Z_t$ y $R_t = B^{12} Z_t$, se obtendrían los siguientes valores, respectivamente.

$P_1 = -$	$P_2 = 36.61$	$P_3 = 30.01$	$P_4 = 25.79$	$P_5 = 25.29$	$P_6 = 21.91$
$P_7 = 37.22$	$P_8 = 30.54$	$P_9 = 28.58$	$P_{10} = 29.01$	$P_{11} = 22.09$	$P_{12} = 19.59$
$P_{13} = 40.33$	$P_{14} = 27.96$	$P_{15} = 25.84$	$P_{16} = 28.19$	$P_{17} = 21.82$	$P_{18} = 19.60$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$P_{19} = 30.01$	$P_{20} = 20.31$	$P_{21} = 19.09$	$P_{22} = 18.11$	$P_{23} = 16.28$	$P_{24} = 13.12$
$P_{25} = 28.64$	$P_{26} = 18.95$	$P_{27} = 19.18$	$P_{28} = 17.10$	$P_{29} = 16.07$	$P_{30} = 13.29$
$P_{31} = 24.21$	$P_{32} = 19.75$	$P_{33} = 18.96$	$P_{34} = 16.34$	$P_{35} = 14.42$	$P_{36} = 14.42$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$P_{37} = 8.04$	$P_{38} = 7.22$	$P_{39} = 6.76$	$P_{40} = 5.71$	$P_{41} = 5.81$	$P_{42} = 6.38$
$P_{43} = 8.16$	$P_{44} = 6.49$	$P_{45} = 6.54$	$P_{46} = 6.15$	$P_{47} = 5.97$	$P_{48} = 6.21$
$P_{49} = 8.53$	$P_{50} = 6.19$	$P_{51} = 5.46$	$P_{52} = 6.37$	$P_{53} = 6.47$	$P_{54} = 6.42$

$Q_1 = -$	$Q_2 = -$	$Q_3 = 37.22$	$Q_4 = 30.54$	$Q_5 = 28.58$	$Q_6 = 29.01$
$Q_7 = -$	$Q_8 = -$	$Q_9 = 40.33$	$Q_{10} = 27.96$	$Q_{11} = 25.84$	$Q_{12} = 28.19$
$Q_{13} = -$	$Q_{14} = -$	$Q_{15} = 36.61$	$Q_{16} = 30.01$	$Q_{17} = 25.49$	$Q_{18} = 25.29$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Q_{19} = 20.07$	$Q_{20} = 29.45$	$Q_{21} = 28.64$	$Q_{22} = 18.95$	$Q_{23} = 19.18$	$Q_{24} = 17.10$
$Q_{25} = 29.57$	$Q_{26} = 30.47$	$Q_{27} = 24.21$	$Q_{28} = 19.75$	$Q_{29} = 18.96$	$Q_{30} = 16.34$
$Q_{31} = 31.43$	$Q_{32} = 30.01$	$Q_{33} = 20.23$	$Q_{34} = 19.09$	$Q_{35} = 18.11$	$Q_{36} = 16.28$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$Q_{37} = 15.12$	$Q_{38} = 10.15$	$Q_{39} = 8.16$	$Q_{40} = 6.49$	$Q_{41} = 6.54$	$Q_{42} = 6.15$
$Q_{43} = 14.33$	$Q_{44} = 8.97$	$Q_{45} = 8.53$	$Q_{46} = 6.19$	$Q_{47} = 5.46$	$Q_{48} = 6.37$
$Q_{49} = 12.65$	$Q_{50} = 8.04$	$Q_{51} = 7.22$	$Q_{52} = 6.76$	$Q_{53} = 5.71$	$Q_{54} = 5.81$

$R_1 = -$	$R_2 = -$	$R_3 = -$	$R_4 = -$	$R_5 = 37.22$	$R_6 = 30.54$
$R_7 = -$	$R_8 = -$	$R_9 = -$	$R_{10} = -$	$R_{11} = 40.33$	$R_{12} = 27.96$
$R_{13} = -$	$R_{14} = -$	$R_{15} = -$	$R_{16} = -$	$R_{17} = 36.61$	$R_{18} = 30.01$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$R_{19} = 17.95$	$R_{20} = 17.72$	$R_{21} = 20.07$	$R_{22} = 29.45$	$R_{23} = 28.64$	$R_{24} = 18.95$
$R_{25} = 18.36$	$R_{26} = 18.18$	$R_{27} = 29.57$	$R_{28} = 30.47$	$R_{29} = 24.21$	$R_{30} = 19.75$
$R_{31} = 18.65$	$R_{32} = 18.86$	$R_{33} = 31.43$	$R_{34} = 30.01$	$R_{35} = 20.31$	$R_{36} = 19.09$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$R_{37} = 14.16$	$R_{38} = 15.55$	$R_{39} = 15.12$	$R_{40} = 10.15$	$R_{41} = 8.16$	$R_{42} = 6.49$
$R_{43} = 14.44$	$R_{44} = 15.96$	$R_{45} = 14.33$	$R_{46} = 8.97$	$R_{47} = 8.53$	$R_{48} = 6.19$
$R_{49} = 15.56$	$R_{50} = 15.82$	$R_{51} = 12.65$	$R_{52} = 8.04$	$R_{53} = 7.22$	$R_{54} = 6.76$

De donde se puede observar que tanto $\{P_t\}$, $\{Q_t\}$ y $\{R_t\}$ tiene los mismos datos que la serie original $\{Z_t\}$, excepto el último, los seis último y los doce último pero retrasados por uno, seis y doce periodos, respectivamente, por otro lado, nótese también que no hay valor definido para las observaciones $P_1, Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8, R_9, R_{10}, R_{11}, R_{12}$.

El siguiente operador de uso frecuente en el análisis de series de tiempo y que esta ligado con el de retraso B es el de **diferencia** ∇ . La característica principal de éste es que expresa las relaciones del tipo $Y_t = Z_t - Z_{t-1}$, es decir, el cambio cuantitativo de la observación entre un periodo y otro. Si se considera la expresión

$$\nabla Z_t = Y_t = Z_t - Z_{t-1} \quad \text{para toda } t$$

Se puede observar que este nuevo operador esta ligado con el de retraso B, visto anteriormente,

$$\text{ya que} \quad \nabla Z_t = Z_t - BZ_t = (1 - B)Z_t,$$

$$\text{entonces} \quad \nabla = (1 - B)$$

Generalizando mediante la aplicación sucesiva del operador diferencia ∇^k

$$\nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-j} \quad \text{para } k=0, 1, 2, \dots \quad \text{y } \forall t.$$

No hay necesidad de efectuar cada una de las aplicaciones para comprobar la validez de la forma general, basta con utilizar el Teorema del Binomio³ para desarrollar $\nabla^k = (1 - B)^k$, que no es más que un binomio elevado a la k-ésima potencia, es decir

$$\begin{aligned} (1 - B)^k &= (1 + (-B))^k = \sum_{j=0}^k \binom{k}{j} (1)^{k-j} (-B)^j \\ &= \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j B^j \end{aligned}$$

y al multiplicar ambos lados por Z_t , se obtiene la generalización de la aplicación sucesiva del operador ∇ .

Utilizando, nuevamente la serie del CCP para moneda nacional, para ejemplificar la utilización del operador diferencia, se obtendría la variación mensual del CCP $S_t = \nabla Z_t$, mostrada en el cuadro 2.

³ Este Teorema también se le conoce como el Teorema del Binomio de Newton.

Para fines del análisis de series de tiempo los operadores de retraso se utilizan en forma de polinomios de la siguiente manera; dado el proceso

$$Z_t - g_1 Z_{t-1} - g_2 Z_{t-2} - \dots - g_k Z_{t-k} = Z_t - \sum_{j=1}^k g_j Z_{t-j}$$

se puede expresar como $G(B)Z_t$, es decir, el polinomio asociado tiene la forma

$$G(B) = 1 - g_1 B - g_2 B^2 - \dots - g_k B^k = 1 - \sum_{j=1}^k g_j B^j$$

$$\begin{aligned} \text{pues } Z_t - g_1 Z_{t-1} - g_2 Z_{t-2} - \dots - g_k Z_{t-k} &= Z_t - g_1 B Z_t - g_2 B^2 Z_t - \dots - g_k B^k Z_t \\ &= (1 - g_1 B - g_2 B^2 - \dots - g_k B^k) Z_t \end{aligned}$$

donde los coeficientes g_j tienen la finalidad de ponderar la importancia de los retrasos con los cuales están asociados, de hecho para $k = 1, 2, 3, \dots$

Cuadro 2.
Costo de Captación de los Pasivos a Plazo denominados en Moneda Nacional (CCP)
(Variación Mensual en Puntos)

MES	1996	1997	1998	1999	2000	2001	2002
Enero		-2.90	-0.96	-0.46	-0.06	0.41	-0.30
Febrero		-3.20	0.04	-1.37	-0.21	-0.14	0.57
Marzo	3.11	-0.27	0.41	-4.43	-1.65	-0.70	-0.22
Abril	-3.72	0.09	0.29	-3.90	-1.30	-0.79	-1.08
Mayo	-6.07	-2.32	-0.93	-1.36	0.17	-1.68	0.25
Junio	-2.58	0.01	0.46	0.80	1.13	-2.50	0.44
Julio	2.05	-0.63	0.68	-0.66	-0.58	-1.18	0.22
Agosto	-1.43	-0.76	1.21	0.09	0.14	-0.93	-0.56
Septiembre	-2.74	-0.09	9.50	-0.22	0.18	0.12	0.16
Octubre	-0.05	-0.76	1.86	-0.85	0.28	0.37	0.50
Noviembre	3.22	1.27	-1.98	-1.01	1.12	-1.31	-0.09
Diciembre	-0.82	0.24	1.02	-0.76	-0.01	-0.73	

En la práctica, es común trabajar con polinomios de retraso racionales ya que se expresan como el cociente de dos polinomios de retraso y esto permite restringir los coeficientes de ponderación g_j . El hecho de restringir el valor de g se toma del siguiente argumento, supóngase que se requiere de trabajar con el polinomio de retraso $G(B) = 1 + gB + g^2 B^2 + g^3 B^3 + \dots$ con $|g| < 1$ sobre el proceso Z_t , entonces

$$G(B) Z_t = (1 + gB + g^2 B^2 + g^3 B^3 + \dots) Z_t = \sum_{j=0}^{\infty} g^j Z_{t-j}$$

suponiendo que $Z_t = \bar{Z}$, para toda t , se tiene que

$$G(B) Z_t = \sum_{j=0}^{\infty} g^j \bar{Z} = \bar{Z} \sum_{j=0}^{\infty} g^j$$

Expresión que no tendría sentido práctico a menos que se cumpla con la condición establecida $|g| < 1$.

Los polinomios de retraso utilizados para el análisis de series de tiempo, son una herramienta que muestra de manera clara y concisa el comportamiento de ciertos modelos que resultan ser de utilidad para representar fenómenos reales, dentro de ellos se encuentran los de *promedios móviles (MA)*, los *autorregresivos (AR)*; así como las combinaciones de estos *autorregresivos de promedios móviles (ARMA)* y por último junto con la aplicación del operador diferencia se tiene a los modelos *autorregresivos integrados y de promedios móviles (ARIMA)*.

1.1.4 Procesos estocásticos lineales.

En el desarrollo de los modelos se parte de la idea de que una serie de tiempo, cuyos valores consecutivos pueden ser altamente dependientes unos con otros, puede considerarse como generada a partir de una sucesión de variables aleatorias independientes $\{a_t\}$ conocida como ruido blanco⁴ que tiene la propiedad de tener media y varianza constantes, generalmente, media igual a cero y varianza σ^2 . Yule (1927) expresó al proceso $\{Z_t\}$ en función de $\{a_t\}$ mediante la relación lineal

$$\begin{aligned} Z_t &= \mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \\ &= \mu + \psi(B) a_t \end{aligned} \quad (1.1)$$

donde μ es un parámetro que determina el nivel (no necesariamente la media) del proceso y $\psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots$ es el polinomio de retraso que convierte al proceso $\{a_t\}$ en el proceso $\{Z_t\}$.

⁴ Para mayores detalles Andrew C. Harvey. Time Series Models. Second Edition. The MIT Press. 1994. Pag. 11.

1.1.5 Estacionariedad.

Cuando una serie de tiempo alcanza un punto de equilibrio o mejor dicho se estabiliza a partir de un cierto momento alrededor de una constante, se dice que es una serie estacionaria, será estrictamente estacionaria si se cumple que la distribución conjunta de $Z_{i_0}, Z_{i_1}, Z_{i_2}, \dots, Z_{i_n}$ es la misma que $Z_{i_0+k}, Z_{i_0+k+1}, Z_{i_0+k+2}, \dots, Z_{i_0+k+n}$ $\forall i_1, i = 0, 1, 2, \dots, \forall k \in \mathbb{N}$ y $\forall n$.

En el análisis de series de tiempo el concepto de estacionariedad de un proceso es de gran importancia, pues como se mencionó anteriormente, para poder caracterizar completamente a un proceso estocástico es necesario conocer la función de densidad conjunta de todas las variables aleatorias involucradas, lo cual no siempre ocurre. Por otro lado, ya que los dos primeros momentos de las variables aleatorias resumen en buena medida a su distribución, pues de ellos se derivan la media, la varianza y la covarianza de las variables involucradas en el proceso, serán entonces el objeto de estudio en lo que sigue a continuación.

Para un caso general, supóngase que la media del proceso Z_i es μ_i , utilizando la ecuación (1.1) se tiene

$$\begin{aligned}
 \mu_i &= E[Z_i] \quad \text{sustituyendo el valor de } Z_i, \\
 &= E[\mu + \psi(B)a_i] \\
 &= E[\mu + (1 - \sum_{l=1}^{\infty} \psi_l B^l) a_i] \\
 &= E[\mu + a_i - (\sum_{l=1}^{\infty} \psi_l B^l) a_i] \\
 &= E[\mu] + E[a_i] - (\sum_{l=1}^{\infty} \psi_l B^l) E[a_i] \quad \text{y como por hipótesis } E[a_i] = 0 \\
 &= \mu
 \end{aligned}$$

Por lo tanto,

$$E[Z_i] = \mu \quad (1.2)$$

Para el caso $E[a_i] \neq 0$ es necesaria la condición

$$\sum_{l=1}^{\infty} \psi_l B^l < \infty \quad (1.3)$$

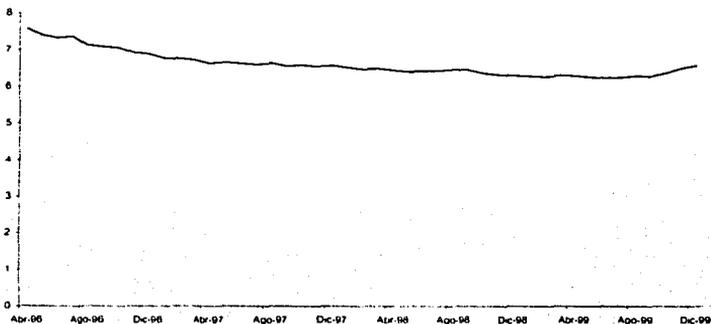
Pues de lo contrario no tendría sentido alguno.

Del resultado anterior se desprende que $E[Z_i] = \mu$, lo cual significa que la media del proceso no depende del tiempo, es decir, que aun cuando durante un cierto

período el proceso se aleje de la media éste siempre regresará a una vecindad de la misma.

Lo anterior se puede observar en la gráfica 2, en donde se utilizan los datos del CCP dólares para el período abril de 1996 a diciembre de 1999. La mayoría de los valores caen en una vecindad de ± 1 de la media de los datos que es de 6.6, aunque al principio los valores se encuentran muy alejados de ella con el tiempo se acercan bastante por un período de año y medio, luego en el mes de agosto de 1998 se alejan por un año pero la tendencia es regresar a la media para el mes de septiembre de 1999.

Gráfica 2.
Costo de Captación de los Pasivos a Plazo denominados en Dólares de los EE.UU.A. (CCP-dólares) de Abril de 1996 a Diciembre de 1999



La varianza del proceso (1.1) denotada por γ_0 , se obtiene por definición como

$$\begin{aligned} \gamma_0 &= E[(Z_t - \mu)^2], \text{ sustituyendo el valor de } Z_t - \mu \\ &= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)^2] \\ &= E[a_t^2 + \psi_1^2 a_{t-1}^2 + \psi_2^2 a_{t-2}^2 + \psi_3^2 a_{t-3}^2 + \dots] - 2 * E[a_t (\sum_{i=1}^{\infty} \psi_i a_{t-i})] + \\ &\quad 2 * E[\text{Productos cruzados } \psi_i a_{t-i} \text{ y } \psi_j a_{t-j}] \quad i, j = 1, 2, 3, \dots \end{aligned}$$

ahora por hipótesis $\{a_t\}$ es un proceso de ruido blanco, entonces

$$\begin{aligned} E[a_t (\psi_i a_{t-i})] &= \psi_i E[a_t * a_{t-i}] = \psi_i E[a_t] * E[a_{t-i}] = 0 \quad \text{y} \\ E[\psi_i a_{t-i} (\psi_j a_{t-j})] &= \psi_i \psi_j E[a_{t-i} * a_{t-j}] = \psi_i \psi_j E[a_{t-i}] * E[a_{t-j}] = 0 \quad i \neq j \end{aligned} \quad (1.4)$$

entonces la expresión de la varianza se reduce a

$$\begin{aligned} &= E[a_t^2] + \psi_1^2 E[a_{t-1}^2] + \psi_2^2 E[a_{t-2}^2] + \dots \\ &= \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 \end{aligned} \quad (1.5)$$

Luego entonces para que la varianza en cuestión exista, se tiene que $\sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2$ debe ser acotada, esto es $\sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 < +\infty$, lo cual está garantizado por (1.3).

Al considerar a una serie de tiempo como un proceso estocástico, se toman en cuenta a todas las variables aleatorias, $Z_1, Z_2, Z_3, \dots, Z_t, \dots, Z_{t+k}$, motivo por lo que resulta necesario estudiar la covarianza entre Z_t y Z_{t+k} , la cual se denota por $\gamma_k = \text{cov}(Z_t, Z_{t+k})$ y está dada como,

$$\begin{aligned} \gamma_k &= E[(Z_t - \mu)(Z_{t+k} - \mu)] \\ &= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)(a_{t+k} - \psi_1 a_{t+k-1} - \psi_2 a_{t+k-2} - \dots)] \\ &= E[-\psi_k a_t^2 + \psi_1 \psi_{k+1} a_{t-1}^2 + \psi_2 \psi_{k+2} a_{t-2}^2 + \dots] - E[\text{Productos cruzados } \psi_i a_{t-i} \text{ y } \psi_j a_{t-j}] \end{aligned}$$

Para simplificar esta expresión se utiliza el resultado de (1.4), para el producto de términos cruzados, el cual indica que este será cero, por lo cual sólo se trabajará con la primera parte.

$$\begin{aligned}
&= -\psi_k E[a_i^2] + \psi_1 \psi_{k+1} E[a_{i-1}^2] + \psi_2 \psi_{k+2} E[a_{i-2}^2] + \dots \\
&= \sigma_a^2 [-\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \dots] \\
&= \sigma_a^2 \left[\sum_{i=1}^{\infty} \psi_i \psi_{k+i} - \psi_k \right]
\end{aligned} \tag{1.6}$$

Un resultado que se puede adicionar y que se utilizará más adelante es el siguiente.

$$\begin{aligned}
\gamma_{-k} &= E[(Z_{i-k} - \mu)(Z_i - \mu)] \\
&= E[(a_{i-k} - \psi_1 a_{i-(k-1)} - \psi_2 a_{i-(k-2)} - \dots)(a_i - \psi_1 a_{i-1} - \psi_2 a_{i-2} - \dots)] \\
&= E[-\psi_k a_i^2 + \psi_1 \psi_{k+1} a_{i-1}^2 + \psi_2 \psi_{k+2} a_{i-2}^2 + \dots] - E[\text{Productos cruzados } \psi_1 a_{i-1} \text{ y } \psi_1 a_{i-2}] \\
&= -\psi_k E[a_i^2] + \psi_1 \psi_{k+1} E[a_{i-1}^2] + \psi_2 \psi_{k+2} E[a_{i-2}^2] + \dots \\
&= \sigma_a^2 [-\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \dots] \\
&= \sigma_a^2 \left[\sum_{i=1}^{\infty} \psi_i \psi_{k+i} - \psi_k \right] \\
&= \gamma_k
\end{aligned}$$

Nuevamente, para el cálculo de la covarianza, se vuelve a requerir la condición (1.3) para justificar la convergencia de $\sum_{i=0}^{\infty} \psi_i \psi_{k+i}$.

Se puede concluir que la covarianza es una función simétrica, es decir, toma el mismo valor tanto para valores positivos como negativos.

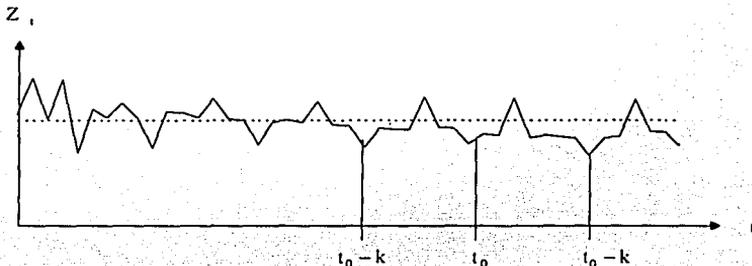
Cabe señalar que en las expresiones (1.2), (1.5) y (1.6) no aparece el factor t , lo cual significa que ni la media de la serie, ni la varianza, ni la covarianza dependen del tiempo.

También se puede concluir que dentro de la covarianza existe una dependencia en la separación (k) que hay entre las variables, lo cual implica una similitud en el comportamiento en términos generales sin importar el momento en el que se observe, es decir, si se graficara un cierto número de observaciones contiguas de una serie, la gráfica obtenida sería muy semejante al graficar el mismo número de observaciones contiguas, pero k periodos hacia delante o hacia atrás de los considerados inicialmente; en otros términos, el comportamiento de $(Z_{i_0-k}, Z_{i_0-k+1}, \dots, Z_{i_0-1})$ y $(Z_{i_0}, Z_{i_0+1}, \dots, Z_{i_0+k})$ será muy parecido para cualquier valor k que se considere, como puede observarse en la gráfica 3.

A los procesos que tienen la característica de que sus momentos de primero y segundo orden no dependen del tiempo se les llama **ESTACIONARIOS DE SEGUNDO ORDEN**. Por lo tanto, si (1.2), (1.5) y (1.6) se cumplen para una serie

de tiempo, bajo la condición de que $\sum_{i=0}^{\infty} |\psi_i|$ es convergente, se dice que la serie $\{Z_t\}$ es estacionaria de segundo orden.

Gráfica 3.
Ejemplo de una serie de tiempo estacionaria



Se dice que un proceso estocástico es estrictamente estacionario si dada la función de densidad para un conjunto arbitrario de variables $(Z_t, Z_{t+1}, Z_{t+2}, \dots, Z_{t+m})$, ésta cumple con

$$f(Z_t, Z_{t+1}, \dots, Z_{t+m}) = f(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+m}) \quad \forall t, m \text{ y } k.$$

de esta expresión, si los primeros momentos de f existen, entonces

$$E[Z_t] = E[Z_{t+m}] = \mu$$

luego;

$$E[(Z_t - \mu)^2] = E[(Z_{t+m} - \mu)^2] = \gamma_0$$

y además

$$E[(Z_t - \mu)(Z_{t+k} - \mu)] = E[(Z_{t+m} - \mu)(Z_{t+m+k} - \mu)] = \gamma_k \quad \forall t, m \text{ y } k$$

Por lo tanto, la estacionariedad estricta implica la de segundo orden pero no necesariamente lo inverso es cierto siempre, a menos que el proceso tenga como distribución conjunta a la normal multivariada.

Con este último resultado, y ya que en la práctica es común suponer que la distribución asociada con las series es la normal, basta entonces con conocer la

media y la función de autocovarianza $\{\gamma_k\}$ (que se definirá como la covarianza entre las observaciones de la misma variable), para caracterizar completamente a una serie estacionaria. Para evitar problemas de dispersión se trabajará con datos estandarizados, es decir, con las autocorrelaciones $\rho_0, \rho_1, \rho_2, \dots$ las cuales están dadas por

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{E(Z_t - \mu)^2} = \frac{\gamma_k}{\gamma_0}, \quad k = 0, \pm 1, \pm 2, \dots$$

y que dan origen a la **función de autocorrelación** $\{\rho_k\}$, la cual junto con μ y σ^2 sirven para denotar a un proceso estocástico estacionario. Para el caso de las series de tiempo debe recordarse que se cuenta con una realización finita del proceso, por lo cual es necesario obtener una estimación para dichos estadísticos. En dicha estimación se supone que el proceso posee ciertas propiedades **ergódicas**⁵ que permiten la equivalencia entre valores esperados y promedios muestrales obtenidos de una realización suficientemente larga del proceso; así la media del proceso se obtiene a través de la media muestral de la serie observada,

$$\bar{Z} = \frac{1}{N} \sum_{i=1}^N Z_i$$

el estimador para γ_k viene dado por

$$C_k = \frac{1}{N} \sum_{i=1}^{N-k} (Z_i - \bar{Z})(Z_{i+k} - \bar{Z}), \quad k = 0, 1, 2, 3, \dots$$

y por último el estimador para ρ_k , es

$$\alpha_k = \frac{C_k}{C_0} = \frac{\sum_{i=1}^{N-k} (Z_i - \bar{Z})(Z_{i+k} - \bar{Z})}{\sum_{i=1}^N (Z_i - \bar{Z})^2}, \quad k = 0, 1, 2, 3, \dots$$

Una característica típica de un proceso estacionario es que su función de autocorrelación decae rápidamente a cero, pues la dependencia de una observación con las subsiguientes tiende a desaparecer por las características del proceso y entre más distantes se encuentren menos se relacionan.

Para ver una aplicación de los conceptos anteriores se toman los 12 datos del CCP en moneda nacional para el año 2000 obteniéndose los siguientes resultados

⁵ La ergodicidad no será definida formalmente en este trabajo, pero lo que ésta básicamente implica es que aun cuando las observaciones estén suficientemente separadas deben ser no correlacionadas.

Media muestral

$$\overline{CCPMN} = \frac{1}{12} \sum_{i=1}^{12} CCPMN_i = 14.59$$

de la cual se desprende la varianza muestral

$$C_0 = \frac{1}{12} \sum_{i=1}^{12} (CCPMN_i - \overline{CCPMN})^2 = 0.9966$$

con ella se obtienen a su vez las autocovarianzas C_1, C_2, \dots que para este caso solo se calcularon

$$C_1 = \frac{1}{12} \sum_{i=1}^{11} (CCPMN_i - \overline{CCPMN})(CCPMN_{i+1} - \overline{CCPMN}) = 0.5283,$$

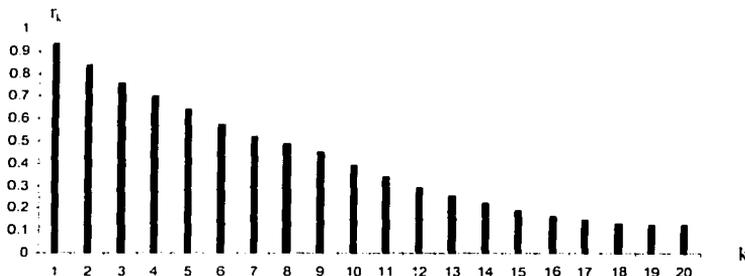
$$C_2 = \frac{1}{12} \sum_{i=1}^9 (CCPMN_i - \overline{CCPMN})(CCPMN_{i+2} - \overline{CCPMN}) = 0.3424$$

y así se obtienen

$$r_1 = \frac{C_1}{C_0} = 0.5301 \quad \text{y} \quad r_2 = \frac{C_2}{C_0} = 0.3436$$

Si se consideran todos los elementos de la serie en cuestión la gráfica de la función de autocorrelación muestral quedaría como sigue

Gráfica 4.
Función de autocorrelación muestral para la serie del CCP moneda nacional



donde se puede observar que las autocorrelaciones muestrales no consideran valores negativos de k y tienden hacia el valor cero de una manera lenta, característica de las series de tiempo no estacionarias.

Es importante mencionar que en la práctica lo más frecuente es encontrarse con series de tiempo no estacionarias, pero existen formas para transformar a las series que no lo son.

Se dice que la estacionariedad es homogénea cuando únicamente el nivel de la serie se ve afectado por la no-estacionariedad debido a la existencia de alguna tendencia polinomial, la cual es posible eliminar y por lo tanto volver a la serie estacionaria mediante la aplicación del operador ∇ un número apropiado de veces. Retomando del resultado (A.1) del apéndice se tiene que

$$\nabla^d Z_t = (d!) \delta_0 + \nabla^d a_t$$

expresión que define básicamente el comportamiento de $\nabla^d Z_t$, pues se puede observar que la nueva serie $\{\nabla^d Z_t\}$ tiene las mismas características de $\{\nabla^d a_t\}$ excepto por el nivel de ellas que para la primera es $(d!) \delta_0$ y para la segunda es cero, pero en ambas el nivel no depende de t . Así la conclusión es que la nueva serie obtenida es estacionaria en cuanto a su nivel.

En la gráfica 5 se pueden apreciar dos ejemplos de tendencia polinomial adaptiva de orden 1 y 2 respectivamente

Otro resultado que se observa de la demostración del apéndice (A.1) es que al aplicar el operador diferencia a una serie que ya es estacionaria se obtiene otra serie estacionaria por lo que se puede caer en el error de cometer una sobrediferenciación a una serie y con ello acarrear con problemas de identificación de algún modelo para representarla, incrementar la varianza de la serie y la pérdida innecesaria de observaciones, pues como se vio al aplicar el operador ∇ d -veces, se pierden d observaciones, por lo que es de importancia tener cuidado con este tipo de error.

La función de autocorrelación asociada con procesos homogéneos no-estacionarios no decae rápidamente a cero, esto es consecuencia de la tendencia polinomial la cual produce que los nuevos elementos dependan de observaciones distantes en el tiempo haciendo que las autocorrelaciones sean altas aún para retrasos muy separados. Por esta razón se dice que la función de autocorrelación para procesos homogéneos no-estacionarios mide lo marcado de la tendencia polinomial y la falta es eliminada mediante la aplicación del operador diferencia un número apropiado de veces, las suficientes para hacer las correlaciones muestrales prácticamente cero después de un número pequeño de retrasos (generalmente tres o cuatro).

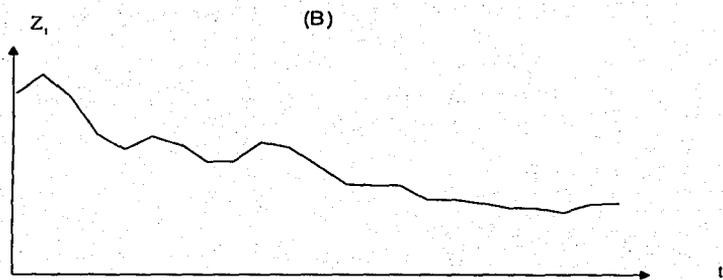
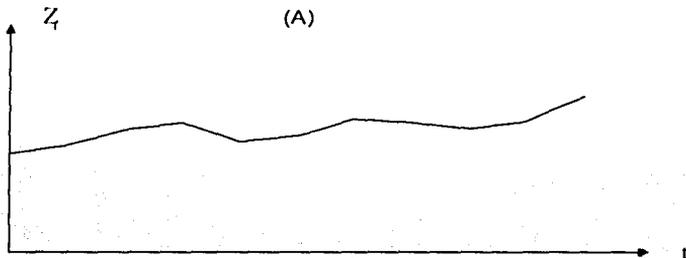
Para ver esto último de manera más clara, de la gráfica de CCP en moneda nacional se puede apreciar una tendencia lineal, es decir, según la teoría esta tendencia puede ser eliminada con una sola aplicación del operador ∇ y entonces la nueva serie con la cual se deberá trabajar es ∇CCPMN_t , y si se continúa aplicando el operador diferencia una o dos veces más a parte de la primera se puede caer en el error de perder información, pues al aplicar ∇^d se pierden d observaciones. Entonces si la serie original CCPMN_t cuenta con 85 observaciones, ∇CCPMN_t , $\nabla^2\text{CCPMN}_t$ y $\nabla^3\text{CCPMN}_t$ contarán con 84, 83 y 82 respectivamente.

Las gráficas de las series $\{\nabla\text{CCPMN}_t\}$, $\{\nabla^2\text{CCPMN}_t\}$ y $\{\nabla^3\text{CCPMN}_t\}$, así como sus correspondientes funciones de autocorrelación muestral, aparecen en las gráficas 5, 6 y 7, respectivamente. Se puede observar que en la serie $\{\nabla\text{CCPMN}_t\}$ originada con una sola aplicación del operador diferencia desaparece la tendencia lineal que presenta la serie original y las autocorrelaciones decaen rápidamente a cero, lo cual se considera como indicador de estacionariedad en la serie en cuanto al nivel se refiere. Aunque también la serie $\{\nabla^3\text{CCPMN}_t\}$ elimina la tendencia lineal, pero se puede observar que su varianza presenta un crecimiento y en cuanto a las autocorrelaciones éstas ya no se acercan a cero tan rápidamente, pues empiezan a dar brincos en una vecindad de este.

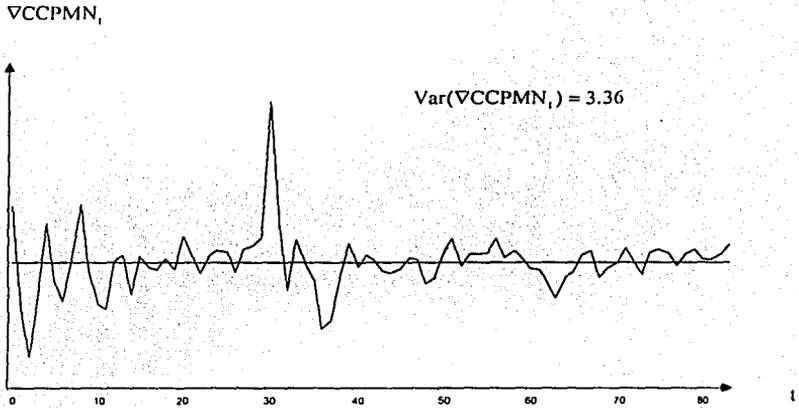
Por último, la serie $\{\nabla^3\text{CCPMN}_t\}$ es un claro ejemplo de sobrediferenciar a una serie, los valores empiezan a salirse de un intervalo acotado de su media, las autocorrelaciones no tienden a cero tan lentamente como para el caso de $\{\nabla\text{CCPMN}_t\}$ y la varianza es mucho mayor.

Así se ha presentado de manera breve los conceptos principales en el análisis de series de tiempo. Ahora se pasará a los detalles del análisis.

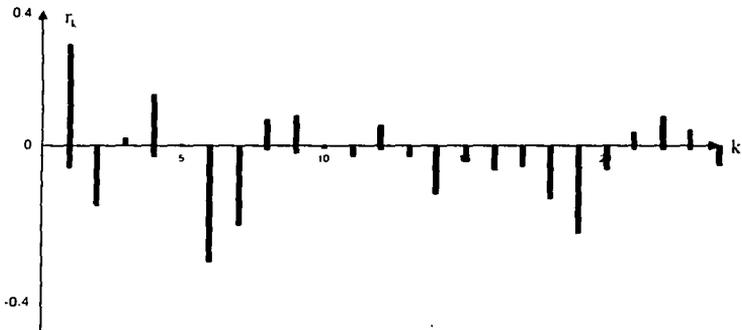
Gráfica 5.
Series de tiempo con tendencia polinomial adaptiva de (A) primer orden y
(B)segundo orden



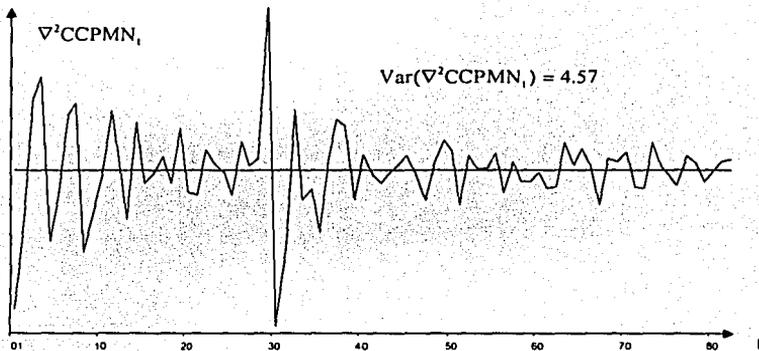
Gráfica 6.
A) Gráfica de la serie $\{\nabla\text{CCPMN}_t\}$



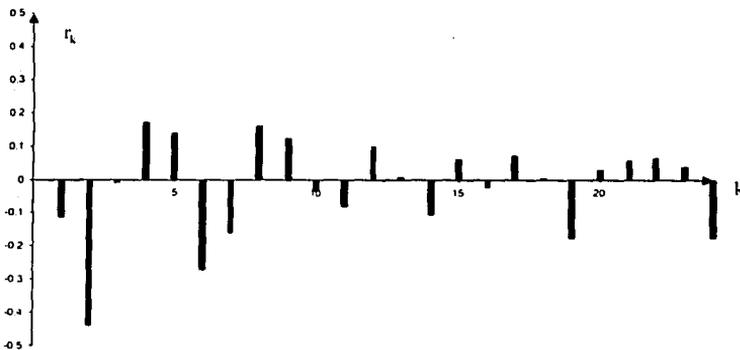
B) Función de Autocorrelación Muestral para $\{\nabla\text{CCPMN}_t\}$



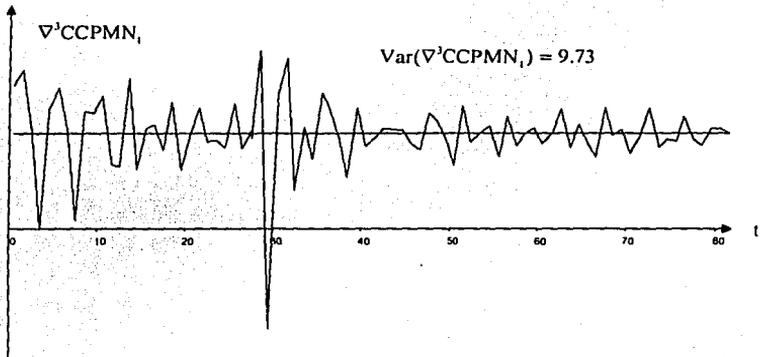
Gráfica 7.
A) Gráfica de la serie $\{\nabla^2\text{CCPMN}_t\}$



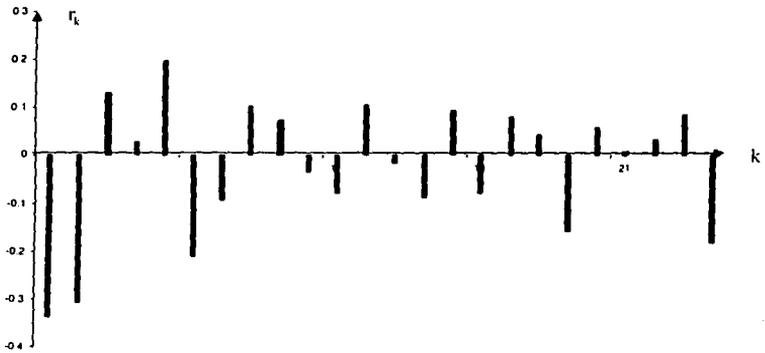
B) Función de Autocorrelación Muestral para $\{\nabla^2\text{CCPMN}_t\}$



Gráfica 8.
A) Gráfica de la serie $\{\nabla^3\text{CCPMN}_t\}$



B) Función de Autocorrelación Muestral para $\{\nabla^3\text{CCPMN}_t\}$



1.2 Ecuaciones en diferencia como herramienta en el análisis de series de tiempo.

En el apartado anterior se dió una introducción a lo que es el área de las series de tiempo y se explicó estadísticamente cómo se pueden generar series de tiempo a partir de procesos estocásticos muy generales. Ahora se verá que en ocasiones también la teoría correspondiente al fenómeno en estudio puede indicar cuál es el proceso generador de la serie y para ello se presenta a las ecuaciones en diferencia las cuales son de gran utilidad en el desarrollo de las ideas propuestas por Box y Jenkins (1970) para construir modelos de series de tiempo.

Es importante mencionar que existe una relación estrecha entre los procesos deterministas (cuyo comportamiento esta determinado por ecuaciones en diferencia) y las series de tiempo que admiten la representación autorregresiva, además el concepto de equilibrio eventual para los primeros esta relacionado con el de estacionariedad para los segundos.

1.2.1 Relación de procesos deterministas con ecuaciones en diferencia.

Las **ecuaciones en diferencia** son el equivalente discreto de las ecuaciones diferenciales que involucran variables en función del tiempo. Es decir, dada una variable $Z(t)$ y considerando que el tiempo es continuo, entonces el comportamiento de la variable queda determinado por sus derivadas

$$\frac{dZ}{dt}, \frac{d^2Z}{dt^2}, \dots, \frac{d^k Z}{dt^k}, \dots$$

en cambio si la misma variable es ahora observada en un tiempo discreto, entonces el comportamiento de $Z(t)$ esta dictado por sus diferencias

$$\nabla Z_t, \nabla^2 Z_t, \dots, \nabla^k Z_t, \dots$$

la manera correcta de denotar este comportamiento discreto sería $\nabla Z_t / \nabla t$, sin embargo, ya que t sólo toma valores enteros contiguos se sabe que $\nabla t = 1$; por lo tanto, ∇Z_t es el equivalente de dZ/dt cuando t toma los valores $\dots, -2, -1, 0, 1, 2, \dots$. La notación común para ecuaciones en diferencia es mediante el **operador incremento** definido por $\Delta Z_t = Z_{t+1} - Z_t$, donde se puede utilizar la relación $\nabla Z_t = \Delta_{t-1}$ que liga a los operadores incremento y diferencia, para conservar el concepto de variable retrasada utilizando simplemente ∇Z_t .

1.2.2 Ecuaciones en diferencia de primer orden.

La ecuación en diferencia más simple es la de orden uno, la cual se denota por

$$Z_t = a_0 + a_1 Z_{t-1} \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (1.7)$$

donde a_0 y $a_1 \neq 0$ son constantes. De manera simplificada

$$(1 - a_1) Z_t = a_0 \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (1.8)$$

El método más sencillo para resolver ecuaciones en diferencia es el llamado método iterativo, el cual parte de algún valor inicial Z_0 y por aplicación sucesiva (iteraciones) encuentra los valores para Z_1, Z_2, \dots , dando por resultado la posibilidad de inferir el comportamiento de Z_t en general. El método iterativo se emplea de la siguiente manera, considérese a la ecuación de primer orden (1.7) y el valor inicial Z_0 , entonces

$$Z_1 = a_0 + a_1 Z_0$$

$$Z_2 = a_0 + a_1 Z_1 = a_0 + a_1(a_0 + a_1 Z_0) = a_0(1 + a_1) + a_1^2 Z_0$$

$$Z_3 = a_0 + a_1 Z_2 = a_0 + a_1(a_0(1 + a_1) + a_1^2 Z_0) = a_0(1 + a_1 + a_1^2) + a_1^3 Z_0$$

en general

$$Z_t = a_0 \sum_{j=0}^{t-1} a_1^j + a_1^t Z_0, \quad \text{para } t \geq 1$$

si $a_1 \neq 1$, se tiene

$$\sum_{j=0}^{t-1} a_1^j = \frac{1 - a_1^t}{1 - a_1}$$

y así, la solución general para este tipo de ecuaciones en diferencia⁶, esta dada por

$$Z_t = a_1^t Z_0 + \left[\frac{1 - a_1^t}{1 - a_1} \right] a_0 \quad \text{con } a_1 \neq 1, \quad Z_0 = \text{Condición inicial} \quad (1.9)$$

Se hará la limitación al caso $|a_1| < 1$, debido a que se requieren procesos estables. Por otro lado, si usamos la ecuación (1.9) en su forma

⁶ Para saber más acerca de la solución de las ecuaciones en diferencia ver "Takehito Takahashi, Ecuaciones en Diferencia con Aplicaciones, Grupo Editorial Iberoamerica, México 1990"

$$Z_t = a_1 Z_{t-1} + \frac{a_0}{1-a_1} - \left[\frac{a_0}{1-a_1} \right] a_1^t \quad \text{para } t = \dots, -2, -1, 0, 1, 2, \dots \quad (1.10)$$

Ahora, nótese que debido al supuesto de que $|a_1| < 1$, $\lim_{t \rightarrow \infty} a_1^t = 0$ y por lo tanto que

$$\lim_{t \rightarrow \infty} Z_t = \frac{a_0}{1-a_1}$$

lo que significa que la serie tiende a estabilizarse en el punto $\frac{a_0}{1-a_1}$ conforme el tiempo crece; si $|a_1| > 1$ entonces Z_t nunca se estabilizará, es decir, el proceso no convergerá. Para el caso $|a_1| = 1$ y si se observa (1.9), el proceso será divergente.

Si se define una nueva variable m , como

$$m = Z_0 - \frac{a_0}{1-a_1}$$

y se reduce la expresión (1.10) en términos de esta nueva

$$Z_t = \frac{a_0}{1-a_1} + m a_1^t \quad \text{con } a_1 \neq 1$$

se obtiene una solución general equivalente a (1.9) para las ecuaciones en diferencia de primer orden. Esta simplificación será de utilidad para los resultados generales de las soluciones de ecuaciones en diferencia de orden mayor, y facilitará la obtención de restricciones de los valores relacionados con el proceso que se este modelando.

La ecuación en diferencia (1.7) se clasifica como lineal, ya que ningún término de Z aparece elevado a alguna potencia distinta de uno; de primer orden, porque interviene a lo más una diferencia y por lo tanto a lo más un sólo retraso para Z_t , en comparación a una ecuación en diferencia de orden $p > 1$, en donde pueden intervenir hasta p retrasos, es decir, $Z_t, Z_{t-1}, \dots, Z_{t-p}$. Este último caso se estudiará más adelante.

1.2.3 Ecuaciones en diferencia de segundo orden.

Este tipo de ecuaciones en diferencia de segundo orden, tienen la forma

$$(1 - a_1 B - a_2 B^2)Z_t = a_0 \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (1.11)$$

donde para ser de segundo orden necesariamente $a_2 \neq 0$.

Para la solución, se necesitan de dos condiciones iniciales para poder obtener la general, por lo que bajo el mismo argumento que en el caso anterior, para la ecuación de primer orden se introducirán dos nuevas variables, que las llevan implícitas y cuyo desarrollo ya no se elaborará aquí, pues no es el propósito del presente trabajo el encontrar las soluciones para cada caso, sino utilizar los resultados. Por lo que dicha solución tendrá la forma

$$Z_t = \frac{a_0}{1 - a_1 - a_2} + r_1 g_1^t + r_2 g_2^t \quad a_2 \neq 0 \quad (1.12)$$

donde r_1 y r_2 son constantes que se pueden determinar mediante dos condiciones iniciales o colaterales. Los valores de g_1 y g_2 se relacionan con los coeficientes a_1 y a_2 de (1.11) mediante las relaciones

$$a_1 = g_1 + g_2 \quad \text{y} \quad a_2 = -g_1 g_2 \quad (1.13)$$

y que surgen de la siguiente factorización

$$\begin{aligned} (1 - a_1 B - a_2 B^2) &= (1 - g_1 B)(1 - g_2 B) \\ &= 1 - g_1 B - g_2 B + g_1 g_2 B^2 \\ &= 1 - (g_1 + g_2)B + g_1 g_2 B^2 \end{aligned} \quad (1.14)$$

igualando coeficientes, se obtienen las ecuaciones de (1.13)

Ahora se verificará que (1.12) es la solución de (1.11), multiplicando ambos lados de ella por la nueva factorización de $(1 - a_1 B - a_2 B^2)$, entonces

$$(1 - g_1 B)(1 - g_2 B)Z_t = a_0 + (1 - g_2 B)(1 - g_1 B)r_1 g_1^t + (1 - g_1 B)(1 - g_2 B)r_2 g_2^t$$

donde se sabe que

$$(1 - g_2 B)(1 - g_1 B)r_1 g_1^t = (1 - g_1 B)(1 - g_2 B)r_2 g_2^t = 0$$

y por la factorización (1.14) se obtiene la ecuación en diferencia original. Por lo tanto, (1.12) es la solución general para una ecuación en diferencia de segundo orden.

En cuanto a los valores de g_1 y g_2 , éstos se calculan para valores dados de a_1 y a_2 , de la siguiente manera. Dada la ecuación

$$(1 - g_1 x)(1 - g_2 x) = 0$$

que tiene como raíces a $x_1 = g_1^{-1}$ y $x_2 = g_2^{-1}$ y por la relación que existe entre el polinomio $1 - a_1 x - a_2 x^2$ y la factorización (1.14), las raíces g_1^{-1} y g_2^{-1} se encuentran al resolver la ecuación característica del proceso:

$$1 - a_1 x - a_2 x^2 = 0$$

para dos valores de x . Así, g_1 y g_2 se obtienen como el recíproco de las raíces dadas por

$$x = \frac{-a_1 \pm \sqrt{a_1^2 + 4a_2}}{2a_2} \quad (1.15)$$

la cual puede resolverse una vez que se conocen a_1 y a_2 . Existen tres casos en cuanto al valor de las raíces, éstos son:

Caso 1. Si $a_1^2 + 4a_2 > 0$, entonces (1.15) proporciona dos valores reales y diferentes, $x_1 = g_1^{-1} \neq x_2 = g_2^{-1}$, con la condición $|x_1^{-1}| < 1$ y $|x_2^{-1}| < 1$. Si se calcula el límite cuando t tiende a infinito a la solución general, se tendría que

$$\lim_{t \rightarrow \infty} Z_t = \lim_{t \rightarrow \infty} \left\{ \frac{a_0}{1 - a_1 - a_2} + r_1 g_1^t + r_2 g_2^t \right\} = \frac{a_0}{1 - a_1 - a_2}$$

Si por el contrario $|g_1| > 1$ y $|g_2| > 1$, entonces se tendrá que el término $r_1 g_1^t + r_2 g_2^t$ tenderá a crecer rápidamente y no existirá convergencia, lo mismo pasa cuando se tiene el caso $|g_1| > 1$ y $|g_2| < 1$, pues $r_1 g_1^t$ tenderá a crecer mientras que $r_2 g_2^t$ tenderá a cero, con lo cual Z_t no converge y lo mismo ocurre para el caso contrario.

Cuando $|g_1| = 1$ ó $|g_2| = 1$ se estudiará de manera general más adelante.

Caso 2. Si $a_1^2 + 4a_2 < 0$, entonces las dos raíces de (1.15) serán los complejos conjugados

$$g_1^{-1} = u + iv \quad \text{y} \quad g_2^{-1} = u - iv$$

éstos pueden escribirse en coordenadas polares para ver si Z_t converge o no, es decir,

$$\begin{aligned} g_1^{-1} &= s e^{i\theta} = s[\cos\theta + i\text{sen}\theta] \\ g_2^{-1} &= s e^{-i\theta} = s[\cos\theta - i\text{sen}\theta] \end{aligned} \quad (1.16)$$

en donde

$$s = \sqrt{u^2 + v^2} = \sqrt{g_1^{-1} \cdot g_2^{-1}}$$

y θ es el ángulo, en radianes, que cumple con que $\cos\theta = \frac{u}{s}$ y $\operatorname{sen}\theta = \frac{v}{s}$ para $\theta \in [0, 2\pi]$. A partir de (1.15) se obtienen los valores de g_1 y g_2

$$g_1 = r[\cos\theta - i\operatorname{sen}\theta] \quad \text{y} \quad g_2 = r[\cos\theta + i\operatorname{sen}\theta]$$

con $r = s^{-1} = \sqrt{g_1 \cdot g_2}$; así la parte que determina la convergencia o no de Z_1 , al punto de equilibrio $a_0/1 - a_1 - a_2$, está dada por

$$\begin{aligned} r_1 g_1^t + r_2 g_2^t &= r_1 r^t [\cos(\theta t) - i\operatorname{sen}(\theta t)] + r_2 r^t [\cos(\theta t) + i\operatorname{sen}(\theta t)] \\ &= r^t [(r_1 + r_2) \cos(\theta t) + i(r_2 - r_1) \operatorname{sen}(\theta t)] \end{aligned}$$

la cual sigue un patrón de fluctuaciones cíclicas al llevar las funciones de seno y coseno que se encuentran acotadas, pero que tenderán a intensificarse o a disminuir dependiendo del factor r^t . Para $r < 1$ el patrón cíclico se verá obligado a desaparecer y habrá convergencia; si $r > 1$ se observarán oscilaciones explosivas y no habrá convergencia para Z_1 (sólo que se dieran las condiciones iniciales $Z_0 = Z_1 = a_0 / [(1 - g_1)(1 - g_2)]$, en cuyo caso se pueden determinar los valores de r_1 y r_2 y el proceso ya estaría en equilibrio); por último para el caso $r = 1$ el patrón oscilatorio no cambiará y el punto de equilibrio es periódico.

Caso 3. Si $a_1^2 + 4a_2 = 0$, la ecuación característica tiene dos raíces iguales $g_1^{-1} = g_2^{-1} = -a_1/2a_2$, en cuyo caso si se considera $g = g_1 = g_2$, la nueva ecuación en estudio será

$$(1 - gB)^2 Z_1 = a_0 \tag{1.17}$$

cuya solución general está dada por

$$\begin{aligned} Z_1 &= (1 - gB)^{-2} a_0 + r_1 g^t + r_2 t g^t \\ &= \sum_{j=0}^{\infty} (1 + j) g^j a_0 + r_1 g^t + r_2 t g^t \end{aligned} \tag{1.18}$$

se puede observar que ésta es la solución general multiplicando ambos lados de ella por $(1 - gB)^2$, con lo cual se obtiene (1.17). Para una solución en particular se establecen los valores para r_1 y r_2 , mediante dos condiciones iniciales. Para que

el término $\sum_{j=0}^n (1+j)g^j$ que aparece en (1.18) sea finito se requiere que $|g| < 1$ y con ello el proceso Z_t convergerá al punto de equilibrio, para este caso, es

$$\frac{a_0}{(1-g)^2}$$

En conclusión se puede observar que la condición para que exista la convergencia en los tres casos anteriores es que los módulos de g_1 y g_2 sean menores que la unidad y r_1 y r_2 se determinan mediante las condiciones iniciales Z_0 y Z_1 .

Entonces la convergencia de un proceso está en términos de las raíces g_1 y g_2 , pero también pueden estar expresadas en función de los parámetros originales a_1 y a_2 que determinan a la ecuación en diferencia. La manera es la siguiente, dadas las condiciones $|g_1| < 1$ y $|g_2| < 1$ se puede considerar que lo anterior implica los siguientes casos:

$$\begin{array}{lll} \text{i) } g_1 < 1, & \text{ii) } -g_1 < 1, & \\ \text{iii) } g_2 < 1, & \text{iv) } -g_2 < 1 & \text{v) } |g_1 g_2| < 1 \end{array} \quad (1.19)$$

a su vez, de i) y iii) se puede obtener que $g_1(1-g_2) < (1-g_2)$, de igual manera de ii) y iv) puede derivarse el resultado $-g_1(1+g_2) < (1+g_2)$, desarrollando y despejando las restricciones se llega a que

$$g_1 + g_2 - g_1 g_2 < 1, \quad -g_1 - g_2 - g_1 g_2 < 1 \quad \text{y} \quad |g_1 g_2| < 1$$

que en términos de los parámetros originales y por la factorización (1.14) se obtienen las nuevas condiciones para la estabilidad o convergencia del proceso:

$$a_1 + a_2 < 1, \quad -a_1 + a_2 < 1 \quad \text{y} \quad |a_2| < 1 \quad (1.20)$$

Estas últimas son más fáciles de verificar que las condiciones (1.19), si es que no hay necesidad de calcular los valores g_1 y g_2 , lo cual ocurre en el caso de que no se requiera encontrar la solución para la ecuación en diferencia y simplemente se desee indicar si el proceso alcanzará o no su equilibrio eventualmente.

1.2.4 Ecuaciones en diferencia de orden p.

Ahora se verá a grandes rasgos la metodología para resolver ecuaciones en diferencia de orden $p \geq 2$. Si se considera la ecuación general de la forma

$$(1 - a_1 B - a_2 B^2 - \dots - a_p B^p) Z_t = a_0 \quad t = \dots, -2, -1, 0, 1, 2, \dots$$

$$a_p \neq 0 \quad (1.21)$$

y utilizando el Teorema Fundamental del Álgebra⁷, el polinomio de retraso involucrado queda como sigue

$$(1 - g_1 B)(1 - g_2 B) \dots (1 - g_p B) = G(B)$$

en donde las raíces de la ecuación característica $G(x)$ son $x = g_1^{-1}, g_2^{-1}, \dots, g_p^{-1}$; si se supone que todas las raíces son distintas, entonces la solución general de (1.21), haciendo uso del polinomio $G(B)$, es

$$Z_t = \frac{a_0}{(1 - g_1)(1 - g_2) \dots (1 - g_p)} + r_1 g_1^t + r_2 g_2^t + \dots + r_p g_p^t \quad (1.22)$$

donde los coeficientes r_1, r_2, \dots, r_p se pueden determinar con base a p condiciones iniciales. Si se tiene una raíz real m veces repetida con $1 < m \leq p$, es decir,

$$g_1 = g_2 = \dots = g_m = g$$

se tendría la solución general dada por

$$Z_t = \frac{a_0}{(1 - g)^m (1 - g_{m+1}) \dots (1 - g_p)} + r_1 g^t + r_2 t g^t + \dots + r_m t^{m-1} g^t + r_{m+1} g_{m+1}^t + \dots + r_p g_p^t \quad (1.23)$$

y bajo el supuesto de que $|g| < 1$ y tomando a $(1 - g)^{-1}$ como una función f con argumento g , se tiene el siguiente procedimiento

$$f(g) = \frac{1}{1 - g} = \sum_{j=0}^{\infty} g^j$$

cuya primera derivada de f con respecto a g , en términos de sumas, esta dada por

$$f'(g) = \frac{1}{(1 - g)^2} = \sum_{j=1}^{\infty} j g^{j-1} = \sum_{j=0}^{\infty} (1 + j) g^j$$

tomando una segunda derivada de f , se tiene la expresión

⁷ Sean a_0, a_1, \dots, a_{n-1} números complejos cualesquiera. Entonces existe un número complejo Z tal que $Z^n + a_{n-1} Z^{n-1} + a_{n-2} Z^{n-2} + \dots + a_0 = 0$. Michael Spivak, "Cálculo infinitesimal", editorial Reverté, México 1981.

$$f''(g) = \frac{2}{(1-g)^3} = \sum_{j=2}^{\infty} j(j-1)g^{j-2}$$

de donde se puede deducir que

$$\frac{1}{(1-g)^3} = \sum_{j=0}^{\infty} \frac{(2+j)(1+j)}{2} g^j$$

generalizando al considerar el cociente $1/(1-g)^p$ para cualquier $p \geq 2$ se tiene que la expresión quedaría como

$$\frac{1}{(1-g)^p} = \sum_{j=0}^{\infty} \frac{(p-1+j)(p-2+j) \cdots (2+j)(1+j)}{(p-1)!} g^j$$

Por lo tanto, dada la condición $|g| < 1$, la solución general de (1.23) queda expresada como

$$Z_t = a_0 \sum_{j=0}^{\infty} \frac{(p-1+j)(p-2+j) \cdots (2+j)(1+j)}{(p-1)!} g^j + g^t \sum_{i=1}^p r_i t^{i-1}$$

En el caso de las ecuaciones en diferencia de primero y segundo orden se observará que, el que un proceso (descrito por una ecuación en diferencia) alcance su punto de equilibrio, depende de que el recíproco del módulo de cada una de las raíces de la ecuación característica, sea menor que la unidad. Así pues para el caso general (1.20), se tiene que para que dicha condición se cumpla es equivalente a que se verifique el siguiente resultado al cual se le conoce con el nombre de **Teorema de Schur**⁸:

Los módulos de las raíces de la ecuación

$$g^p - a_1 g^{p-1} - a_2 g^{p-2} - \cdots - a_{p-1} g - a_p = 0 \quad (1.24)$$

serán todos menores que la unidad, **si y sólo si** los p determinantes que se muestran a continuación son positivos.

⁸ Para mayor detalle del teorema revisar R. Barbofía, P. Sanz. Álgebra lineal y teoría de matrices. Prentice may, 1998.

$$D_1 = \begin{vmatrix} -1 & | & a_p \\ \hline a_p & | & -1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} -1 & 0 & | & a_p & a_{p-1} \\ a_{p-1} & -1 & | & 0 & a_p \\ \hline a_p & 0 & | & -1 & a_{p-1} \\ a_{p-1} & a_p & | & 0 & -1 \end{vmatrix},$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 & | & a_p & a_{p-1} & a_{p-2} \\ a_{p-2} & -1 & 0 & | & 0 & a_p & a_{p-1} \\ a_{p-1} & a_{p-2} & -1 & | & 0 & 0 & a_p \\ \hline a_p & 0 & 0 & | & -1 & a_{p-2} & a_{p-1} \\ a_{p-1} & a_p & 0 & | & 0 & -1 & a_{p-2} \\ a_{p-2} & a_{p-1} & a_p & | & 0 & 0 & -1 \end{vmatrix} \dots$$

$$D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & | & a_p & a_{p-1} & \dots & a_1 \\ a_1 & -1 & \dots & 0 & | & 0 & a_p & \dots & a_2 \\ \dots & \dots & \dots & \dots & | & \dots & \dots & \dots & \dots \\ a_{p-1} & a_{p-2} & \dots & -1 & | & 0 & 0 & \dots & a_p \\ \hline a_p & 0 & \dots & 0 & | & -1 & a_1 & \dots & a_{p-1} \\ a_{p-1} & a_p & \dots & 0 & | & 0 & -1 & \dots & a_{p-2} \\ \dots & \dots & \dots & \dots & | & \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_p & | & 0 & 0 & \dots & -1 \end{vmatrix}$$

(1.25)

Utilizando este resultado para resolver la ecuación característica del proceso representado por (1.20), la cual es

$$1 - a_1 x - a_2 x^2 - \dots - a_{p-1} x^{p-1} - a_p x^p = 0$$

y multiplicándola por x^{-p} para tener la forma de (1.24), con $x^{-1} = g$ para que exista relación entre el Teorema de Schur y el criterio para determinar la convergencia del proceso, entonces se tiene la expresión

$$(x^{-1})^p - a_1 (x^{-1})^{p-1} - a_2 (x^{-1})^{p-2} - \dots - a_{p-1} x^{-1} - a_p = 0$$

en donde lo único que falta para verificar si una cierta ecuación en diferencia de orden p representa a un proceso convergente, es calcular los p determinantes especificados en (1.25), y si **todos** son positivos podrá concluirse que el proceso tenderá a equilibrarse eventualmente; y con **uno solo** de los determinantes no sea positivo bastará para indicar que el proceso no se estabilizará jamás.

Aplicando el Teorema de Schur para el caso de una ecuación en diferencia de orden dos, es decir, para la ecuación $(1 - a_1 B - a_2 B^2)Z_t = a_0$, se tiene que los determinantes correspondientes son sólo dos, a saber

$$D_1 = \begin{vmatrix} -1 & a_2 \\ a_1 & -1 \end{vmatrix} = 1 - a_2^2 \quad y$$

$$D_2 = \begin{vmatrix} -1 & 0 & a_2 & a_1 \\ a_1 & -1 & 0 & a_2 \\ a_2 & 0 & -1 & a_1 \\ a_1 & a_2 & 0 & -1 \end{vmatrix}$$

$$\begin{aligned} &= -1 \begin{vmatrix} -1 & 0 & a_2 \\ 0 & -1 & a_1 \\ a_2 & 0 & -1 \end{vmatrix} - a_1 \begin{vmatrix} 0 & a_2 & a_1 \\ 0 & -1 & a_1 \\ a_2 & 0 & -1 \end{vmatrix} + a_2 \begin{vmatrix} 0 & a_2 & a_1 \\ -1 & 0 & a_2 \\ a_2 & 0 & -1 \end{vmatrix} - a_1 \begin{vmatrix} 0 & a_2 & a_1 \\ -1 & 0 & a_2 \\ 0 & -1 & a_1 \end{vmatrix} \\ &= -1[-1(1 + a_2(a_2))] - a_1[a_2(a_2 a_1 + a_1)] + a_2[-a_2 + a_2(a_2)^2] - a_1[a_2 a_1 + a_1] \\ &= -1[-1 + a_2^2] - a_1[a_2^2 a_1 + a_1 a_2] + a_2[a_2^3 - a_2] - a_1[a_2 a_1 + a_1] \\ &= 1 - a_2^2 - a_1^2 a_2^2 - a_1^2 a_2 + a_2^3 - a_2^2 - a_1^2 a_2 - a_1^2 \\ &= 1 - 2a_2^2 - a_1^2 a_2^2 - 2a_1^2 a_2 - a_1^2 + a_2^3 \quad \text{factorizando,} \\ &= (1 + a_2)^2 [(1 - a_2)^2 - a_1^2] \end{aligned}$$

de la condición $D_1 > 0$ se puede ver que implica

$1 - a_2^2 > 0$ luego entonces,
 $a_2^2 < 1$ sacando raíz cuadrada a ambos términos
 $\sqrt{a_2^2} < 1$ donde el lado izquierdo corresponde al valor absoluto, entonces
 $|a_2| < 1$

Por otro lado, la condición $D_2 > 0$ equivale a especificar que $(1+a_2)^2[(1-a_2)^2 - a_1^2] > 0$ o bien que $(1+a_2)^2[(1-a_2)^2 - a_1^2] < 0$. El primer caso implica que $(1+a_2)^2 > 0$ y que $[(1-a_2)^2 - a_1^2] > 0$. Considerando la primera desigualdad se tiene que $(1+a_2)^2 > 0 \Leftrightarrow 0 < (1+a_2) < 0 \Leftrightarrow -1 < a_2 < -1 \Leftrightarrow a_2 \neq -1$, pero de la condición $D_1 > 0$ se obtuvo que $|a_2| < 1$ por lo que $-1 < a_2 < 1$ y al considerar las dos situaciones conjuntamente se observa que la condición que prevalece o satisface a ambas es $-1 < a_2 < 1$.

De la segunda desigualdad se observa que $[(1-a_2)^2 - a_1^2] > 0 \Leftrightarrow (1-a_2)^2 > a_1^2 \Leftrightarrow 1-a_2 > a_1$ ó bien $1-a_2 > -a_1$. En el primer caso se tiene que dicha condición implica que $1 > a_1 + a_2$ y la condición segunda sugiere que $1 > a_2 - a_1$ o alternativamente que $a_1 - a_2 > -1$.

Para el segundo caso se tiene que $(1+a_2)^2 < 0$ y que $[(1-a_2)^2 - a_1^2] < 0$, donde la primera desigualdad es el mismo caso de arriba, pues $(1+a_2)^2 < 0 \Leftrightarrow 0 < (1+a_2) < 0 \Leftrightarrow -1 < a_2 < -1 \Leftrightarrow a_2 \neq -1$ y por el mismo argumento del caso anterior $-1 < a_2 < 1$.

El caso $[(1-a_2)^2 - a_1^2] < 0$ no se puede dar por las propiedades del valor absoluto.

En consecuencia se puede observar que pedir que $D_1 > 0$ y $D_2 > 0$, es equivalente a considerar que $|a_2| < 1$ y que $a_1 + a_2 < 1$ y por último que $a_2 - a_1 < 1$. Condiciones todas estas estipuladas en (1.20).

1.2.5 Procesos divergentes.

Aquí se verá el caso especial de procesos divergentes en el que intervienen raíces unitarias de la ecuación característica, asunto que quedó pendiente para el caso 1 en que $a_1^2 + 4a_2 > 0$. Este tipo de procesos puede ser representado por una ecuación en diferencia de orden $p+d$ dada por

$$\begin{aligned}
A(B)Z_t &= a_0, && \text{con} \\
A(B) &= G(B)(1-B)^d && \text{donde,} \\
A(B) &= (1-a_1B - a_2B^2 - \dots - a_{p+d}B^{p+d}) && \text{y} \\
G(B) &= (1-g_1B)(1-g_2B) \dots (1-g_pB) && (1.26)
\end{aligned}$$

y debido a que el comportamiento de un proceso representado por una ecuación en diferencia, depende del tipo de raíces de su ecuación característica entonces esta última se puede denotar como

$$(1-g_1x)(1-g_2x) \dots (1-g_px)(1-x)^d = 0$$

donde se observa que existen d raíces unitarias, y p raíces no unitarias $g_1^{-1}, g_2^{-1}, \dots, g_p^{-1}$, las cuales se pueden obtener como solución de dos ecuaciones características a saber

$$(1-g_1x)(1-g_2x) \dots (1-g_px) = 0 \quad \text{y} \quad (1-x)^d = 0$$

y que además podrían ser asociadas con los procesos

$$G(B)Z_t = a_0 \quad \text{y} \quad \nabla^d Z_t = 0$$

respectivamente. El primero de ellos ya ha sido estudiado en cuanto a su convergencia al punto de equilibrio y se vio que la estabilidad eventual del proceso depende de la condición $|g_i| < 1$ para $i=1, 2, \dots, p$. En cuanto al proceso $\nabla^d Z_t = 0$ es divergente, ya que las raíces de su ecuación característica son todas unitarias, pero es interesante conocer el comportamiento que sigue dicho proceso. Se tomarán los casos más sencillos y de ahí al general. Considérese la ecuación

$$\nabla Z_t = 0, \quad t = 1, 2, \dots \quad (1.27)$$

de la cual se tiene que $\nabla Z_t = Z_t - Z_{t-1} = 0 \Leftrightarrow Z_t = Z_{t-1}$

luego, obsérvese lo siguiente

$$\begin{aligned}
\nabla Z_1 &= Z_1 - Z_0 = 0 \Rightarrow Z_1 = Z_0 \\
\nabla Z_2 &= Z_2 - Z_1 = 0 \Rightarrow Z_2 = Z_1 \\
\nabla Z_3 &= Z_3 - Z_2 = 0 \Rightarrow Z_3 = Z_2 \\
\nabla Z_4 &= Z_4 - Z_3 = 0 \Rightarrow Z_4 = Z_3 \\
&\vdots \\
\nabla Z_t &= Z_t - Z_{t-1} = 0 \Rightarrow Z_t = Z_{t-1} \\
\Rightarrow Z_t &= Z_{t-1} = \dots = Z_1 = Z_0
\end{aligned}$$

en general

$$Z_t = Z_{t-1} \quad \text{para } i = 1, 2, \dots, t$$

que es lo mismo que

$$Z_t = Z_0, \quad t = 1, 2, \dots$$

Por lo tanto, el proceso queda completamente determinado al conocer la condición inicial. Ahora considérese a la ecuación

$$\nabla^2 Z_t = 0, \quad t = 2, 3, 4, \dots \quad (1.28)$$

donde desarrollando el lado derecho de la igualdad se tiene que

$$\nabla(\nabla Z_t) = \nabla(Z_t - Z_{t-1}) = \nabla Z_t - \nabla Z_{t-1} = Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2}$$

igualando a cero se tiene que

$$\begin{aligned}
Z_t - 2Z_{t-1} + Z_{t-2} &= 0 \\
Z_t &= 2Z_{t-1} - Z_{t-2}
\end{aligned}$$

y resolviendo de manera iterativa se obtiene que

$$\begin{aligned}
Z_t &= 2Z_{t-1} - Z_{t-2} \\
&= 2(2Z_{t-2} - Z_{t-3}) - Z_{t-2} \\
&= 4Z_{t-2} - 2Z_{t-3} - Z_{t-2} \\
&= 3Z_{t-2} - 2Z_{t-3} \\
&= 3(2Z_{t-3} - Z_{t-4}) - 2Z_{t-3} \\
&= 6Z_{t-3} - 3Z_{t-4} - 2Z_{t-3} \\
&= 4Z_{t-3} - 3Z_{t-4} \\
&= 4(2Z_{t-4} - Z_{t-5}) - 3Z_{t-4} \\
&= 8Z_{t-4} - 4Z_{t-5} - 3Z_{t-4} \\
&= 5Z_{t-4} - 4Z_{t-5} \\
&\vdots \\
&= nZ_{t-(n-1)} - (n-1)Z_{t-n} \quad n \geq 2
\end{aligned}$$

Ahora bien $t - (n - 1) = 0$ ó $t - n = 0$ lo anterior implica $t = n$ ó $t = n - 1$ y se elige como máximo valor posible de n en consecuencia

$$\begin{aligned} Z_t &= tZ_{t-(t-1)} - (t-1)Z_0 \\ &= tZ_t - tZ_0 + Z_0 \end{aligned}$$

Por último, considérese ahora el caso $\nabla^3 Z_t = 0$, $t = 3, 4, 5, \dots$,

$$\begin{aligned} \nabla^2(\nabla Z_t) &= \nabla^2(Z_t - Z_{t-1}) = \nabla(\nabla Z_t - \nabla Z_{t-1}) = \nabla(Z_t - Z_{t-1} - Z_{t-1} + Z_{t-2}) = \nabla Z_t - 2\nabla Z_{t-1} + \nabla Z_{t-2} \\ &= Z_t - Z_{t-1} - 2(Z_{t-1} - Z_{t-2}) + Z_{t-2} - Z_{t-3} = Z_t - 3Z_{t-1} + 3Z_{t-2} - Z_{t-3} \end{aligned}$$

$$\begin{aligned} \nabla^3 Z_t &= Z_t - 3Z_{t-1} + 3Z_{t-2} - Z_{t-3} = 0 \\ \Rightarrow Z_t &= 3Z_{t-1} - 3Z_{t-2} + Z_{t-3} \end{aligned} \tag{1.29}$$

aplicando el método iterativo, se tiene que

$$\begin{aligned} Z_t &= 3(3Z_{t-2} - 3Z_{t-3} + Z_{t-4}) - 3Z_{t-2} + Z_{t-3} \\ &= 9Z_{t-2} - 9Z_{t-3} + 3Z_{t-4} - 3Z_{t-2} + Z_{t-3} \\ &= 6Z_{t-2} - 8Z_{t-3} + 3Z_{t-4} \\ &= 6(3Z_{t-3} - 3Z_{t-4} + Z_{t-5}) - 8Z_{t-3} + 3Z_{t-4} \\ &= 18Z_{t-3} - 18Z_{t-4} + 6Z_{t-5} - 8Z_{t-3} + 3Z_{t-4} \\ &= 10Z_{t-3} - 15Z_{t-4} + 6Z_{t-5} \\ &= 10(3Z_{t-4} - 3Z_{t-5} + Z_{t-6}) - 15Z_{t-4} + 6Z_{t-5} \\ &= 30Z_{t-4} - 30Z_{t-5} + 10Z_{t-6} - 15Z_{t-4} + 6Z_{t-5} \\ &= 15Z_{t-4} - 24Z_{t-5} + 10Z_{t-6} \\ &\vdots \\ &= \frac{n^2 - n}{2} Z_{t-(n-2)} - (n-2)nZ_{t-(n-1)} + \frac{(n-2)(n-1)}{2} Z_{t-n} \\ &= \frac{n(n-1)}{2} Z_{t-(n-2)} - (n-2)nZ_{t-(n-1)} + \frac{(n-2)(n-1)}{2} Z_{t-n} \end{aligned}$$

Ahora,

$$\begin{array}{lll} t - (n - 2) = 0 & t - (n - 1) = 0 & t - n = 0 \\ t - n + 2 = 0 & t - n + 1 = 0 & n = t \\ n = t + 2 & n = t + 1 & \end{array}$$

Si n toma como valor más grande a t entonces

$$t - (n-2) = t - t + 2 = 2; \quad t - t + 1 = 1; \quad t - t = 0$$

$$\begin{aligned} Z_1 &= \frac{t(t-1)}{2} Z_2 - (t-2)tZ_1 + \frac{(t-2)(t-1)}{2} Z_0 \\ &= \frac{t^2-t}{2} Z_2 - (t^2-2t)Z_1 + \left(\frac{t^2-3t+2}{2} \right) Z_0 \\ &= \frac{t^2}{2} Z_2 - \frac{t}{2} Z_2 - t^2 Z_1 + 2tZ_1 + \frac{t^2}{2} Z_0 - \frac{3t}{2} Z_0 + Z_0 \\ &= t^2 \left[\frac{Z_2}{2} - Z_1 + \frac{Z_0}{2} \right] + t \left[-\frac{Z_2}{2} + 2Z_1 - \frac{3Z_0}{2} \right] + Z_0 \\ &= \frac{1}{2} [Z_2 - 2Z_1 + Z_0] t^2 + \frac{1}{2} [-Z_2 + 4Z_1 - 3Z_0] t + Z_0 \end{aligned}$$

Retomando estos tres casos se puede decir que para la ecuación en diferencia de primer orden (1.27) se obtiene como solución general una constante, es decir, un polinomio de grado cero, necesitando una condición inicial para la solución particular; mientras que para la ecuación en diferencia de segundo orden (1.28), la solución general es un polinomio de grado uno y requiere de dos condiciones iniciales para la solución particular, y por último para el caso de la ecuación en diferencia de tercer orden (1.29), cuya solución general es un polinomio de grado dos, se necesitan de tres condiciones iniciales para la solución general. En consecuencia para el caso general de la ecuación en diferencia de orden d se tendrá como solución general, a un polinomio de grado $d-1$ y requerirá de d condiciones iniciales para la solución específica.

Del resultado anterior, se puede afirmar que el proceso representado por (1.26), con $|g_i| < 1$ para $i=1, 2, \dots, p$, es divergente debido a que existe una tendencia polinomial, la cual al ser eliminada, el proceso que resulta es convergente. De manera que si se considera una nueva variable

$$W_i = \nabla^d Z_i$$

el proceso queda entonces definido por

$$G(B)W_i = a_n$$

el cual tenderá a estabilizarse eventualmente.

Cabe hacer notar que teniendo la expresión

$$A(B)Z_i = G(B)\nabla^d Z_i = G(B)W_i = a_n$$

el proceso en términos de la variable Z es divergente (por las d raíces unitarias), pero en términos de la variable W sí puede ser convergente, siempre y cuando se cumpla la condición $|g_i| < 1$ para $i=1, 2, \dots, p$.

CAPITULO 2

Modelos para series de tiempo

En la práctica los fenómenos reales son ciertamente más complicados de representar por medio de las ecuaciones en diferencia lineales mostradas en el capítulo anterior, a pesar de que ellas son de gran utilidad en la práctica, tienen la desventaja de ser completamente deterministas, motivo por el cual se hace necesario introducir un componente aleatorio a las ecuaciones en diferencia, que permita mayor flexibilidad a la representación del fenómeno. Este componente aleatorio da origen a los modelos autorregresivos.

En el transcurso del capítulo aparecen otro tipo de necesidades que generan a los modelos de promedios móviles y posteriormente, las combinaciones de los ya mencionados, dan lugar a modelos más generales conocidos como los modelos ARMA y los modelos ARIMA.

2.1 Modelos autorregresivos (AR).

En el capítulo anterior se vieron las ecuaciones del tipo

$$A(B)Z_t = \text{constante}$$

donde $A(B)$ representa a un polinomio de retraso. Una generalización de este tipo de ecuaciones tiene lugar al introducir una variable aleatoria al lado derecho de la ecuación, obteniendo así la expresión

$$A(B)Z_t = \text{constante} + a_t \quad (2.1)$$

aquí se supone que $\{a_t\}$ es un proceso referido como ruido blanco. Las ecuaciones en diferencia del tipo (2.1) permiten representar a los *procesos autorregresivos*, ya mencionados, para los cuales se tienen las siguientes características

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)Z_t = \text{constante} + a_t \quad (2.2)$$

donde la constante considerada que es del tipo $(1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu$, para el caso en que el proceso Z_t tenga un nivel medio constante, dado por $E[Z_t] = \mu \quad \forall t$, provocando que la ecuación (2.2) se transforme en la expresión

$$\psi(B)Z_t = a_t, \quad \text{con } Z_t = Z_{t-1} - \mu \quad (2.3)$$

el término de autorregresivo (AR) que se le da al proceso representado por (2.3), se refiere al hecho de que también puede expresarse como

$$Z_t = (1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t \quad (2.4)$$

que como se puede ver corresponde a una ecuación de regresión lineal, pero con la diferencia de que aquí el valor de la variable dependiente Z_t no está en función de los valores de un cierto conjunto de variables independientes, sino de sus propios valores observados en periodos anteriores a t y ponderados según los coeficientes autorregresivos $\phi_1, \phi_2, \dots, \phi_p$.

La idea de convergencia para una ecuación en diferencia en donde interviene algún elemento aleatorio no es aplicable debido a las fluctuaciones aleatorias que siempre estarán presentes, aún alrededor del punto de equilibrio. Entonces para denotar a la convergencia para este tipo de ecuaciones se utiliza el concepto de equilibrio estocástico, mejor conocido como estacionariedad. Por lo tanto, mientras que para un proceso determinista se habla de equilibrio, para uno estocástico se hablará de estacionariedad. Así, un proceso AR podrá ser estacionario o no-estacionario, dependiendo (como ya se vio anteriormente) de los valores que tomen las raíces de la ecuación característica

$$\phi(x) = 0 \quad (2.5)$$

la cual establece el comportamiento del proceso autorregresivo. Las raíces se obtienen escribiendo a $\phi(x)$ como el caso general de una ecuación en diferencia

$$\phi(B) = (1 - g_1 B)(1 - g_2 B) \dots (1 - g_p B)$$

de donde se sabe que el proceso AR definido por $\phi(x)$ será estacionario siempre y cuando se cumpla la condición

$$|g_i| < 1 \quad \text{para } i=1,2,\dots,p$$

o bien sí y sólo si las raíces de (2.5), que son $g_1^{-1}, g_2^{-1}, \dots, g_p^{-1}$, se encuentran fuera del círculo unitario, en el plano complejo.

2.1.1 Caso general de los modelos AR(1).

El modelo autorregresivo de orden uno AR(1), es el caso más simple y esta representado por

$$Z_i - \phi Z_{i-1} = a_i \quad (2.6)$$

expresión también conocida como serie de Markov⁹. Para que dicha serie sea estacionaria se requiere que la raíz de la ecuación

$$1 - \phi x = 0$$

se encuentre fuera del círculo unitario en el plano complejo; es decir, que se cumpla la condición $|\phi| < 1$.

Este proceso autorregresivo de orden uno también puede ser representado en términos de la serie de errores aleatorios, de la siguiente manera

$$Z_i - \phi Z_{i-1} = (1 - \phi B)Z_i \quad \text{entonces}$$

$$(1 - \phi B)Z_i = a_i$$

despejando a Z_i o lo que es lo mismo multiplicando ambos lados por $(1 - \phi B)^{-1}$ se tiene que

$$\begin{aligned} Z_i &= (1 - \phi B)^{-1} a_i \\ &= (1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \dots) a_i \\ &= a_i + \phi a_{i-1} + \phi^2 a_{i-2} + \dots \end{aligned}$$

y si se cumple que $|\phi| < 1$, se tienen los siguientes resultados

$$E[Z] = E[a_i] + \phi E[a_{i-1}] + \phi^2 E[a_{i-2}] + \dots = 0$$

$$\begin{aligned} \gamma_0 = \text{var}[Z] &= \text{var}(a_i) + \phi^2 \text{var}(a_{i-1}) + \phi^4 \text{var}(a_{i-2}) + \dots \\ &\text{y como por hipótesis } a_i \text{ es ruido blanco,} \\ &= \sigma_a^2 + \phi^2 \sigma_a^2 + \phi^4 \sigma_a^2 + \phi^6 \sigma_a^2 + \dots \\ &= \sigma_a^2 (1 + \phi^2 + \phi^4 + \dots) \end{aligned}$$

Se puede observar que tanto la media como la varianza de la serie son constantes y además sustituyendo la suma $(1 + \phi^2 + \phi^4 + \dots)$ por su expresión equivalente, se tiene que

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

Las autocovarianzas se obtienen a partir de la expresión general (1.6), haciendo la sustitución $\psi_i = -\phi^i$, $i = 1, 2, \dots$, dando por resultado la expresión

⁹P.S. Joel Hoel, S.C. Port, C.J. Stone. Introduction to stocastics processes. Houghton-Miffling.

$$\begin{aligned}
\gamma_k &= \sigma_a^2 \left[\sum_{i=1}^{\infty} \phi^i \phi^{k+i} + \phi^k \right] \\
&= \sigma_a^2 \left[\sum_{i=1}^{\infty} \phi^k \phi^{2i} + \phi^k \right] \\
&= \sigma_a^2 \phi^k \left[\sum_{i=1}^{\infty} \phi^{2i} + 1 \right] \\
&= \sigma_a^2 \phi^k \left[\sum_{i=0}^{\infty} \phi^{2i} \right] \quad k = 1, 2, \dots
\end{aligned}$$

y ya que por hipótesis $|\phi| < 1$, se tiene que

$$\begin{aligned}
\gamma_k &= \sigma_a^2 \phi^k / (1 - \phi^2) \\
&= \phi (\sigma_a^2 \phi^{k-1} / (1 - \phi^2)) \\
&= \phi \gamma_{k-1}
\end{aligned}$$

Además, por la propiedad de que $\gamma_k = \gamma_{-k}$, se obtiene la fórmula general

$$\gamma_k = \sigma_a^2 \phi^{|k|} / (1 - \phi^2), \quad k = 0, \pm 1, \pm 2, \dots \quad (2.7)$$

por lo que las autocorrelaciones tienen la forma

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sigma_a^2 \phi^{|k|} / (1 - \phi^2)}{\sigma_a^2 / (1 - \phi^2)} = \phi^{|k|}, \quad k = 0, \pm 1, \pm 2, \dots \quad (2.8)$$

Expresión que indica que conforme $k > 0$ crece, la función de autocorrelación (que se abreviará con las siglas **FAC**), tiende a cero con decaimiento del tipo exponencial cuando $0 < \phi < 1$ y con signos alternados cuando $-1 < \phi < 0$.

Si se utiliza (2.8) con $k = 0$ para cambiar la condición de estacionariedad del proceso AR(1), $|\phi| < 1$, y dejarla en términos de las autocorrelaciones se obtiene que

$$|\rho| < 1.$$

2.1.2 Caso para los modelos AR(2).

Ahora se trabajará con el modelo autorregresivo de segundo orden AR(2), el cual está definido por

$$(1 - \phi_1 B - \phi_2 B^2)Z_t = a_t \quad (2.9)$$

expresión introducida por Yule (1927). Para que (2.9) sea estacionario se requiere que las raíces de

$$1 - \phi_1 x - \phi_2 x^2 = 0 \quad (2.10)$$

se encuentren fuera del círculo unitario.

Para encontrar las condiciones de estacionariedad se hace uso del Teorema de Schur mencionado en el capítulo 2, esto debido a la equivalencia entre equilibrio de un proceso determinista y estacionariedad de uno estocástico. Así, para que un proceso AR(2) sea estacionario se requiere que

$$D_1 = \begin{vmatrix} -1 & \phi_2 \\ \phi_2 & -1 \end{vmatrix} > 0 \quad \text{y} \quad D_2 = \begin{vmatrix} -1 & 0 & \phi_2 & \phi_1 \\ \phi_1 & -1 & 0 & \phi_2 \\ \phi_2 & 0 & -1 & \phi_1 \\ \phi_1 & \phi_2 & 0 & -1 \end{vmatrix} > 0$$

y por el resultado (1.25), las condiciones son equivalentes a requerir que

$$\phi_2 + \phi_1 < 1, \quad \phi_2 - \phi_1 < 1 \quad \text{y} \quad |\phi_2| < 1 \quad (2.11)$$

si el proceso cumple con éstas condiciones, es decir, si resulta ser estacionario, lo único que falta para que quede descrito completamente es obtener su función de autocovarianza, la cual queda de la siguiente manera:

si $E[Z_t] = \mu \quad \forall t$, entonces $E[Z_t] = E[Z_t - \mu] = 0$ y calculando la autocovarianza entre Z_t y Z_{t+k} se tiene lo siguiente.

De (2.9) se sabe que $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t$, en consecuencia

$$\begin{aligned} E[Z_t Z_{t+k}] &= E[Z_t (\phi_1 Z_{t+k-1} + \phi_2 Z_{t+k-2} + a_{t+k})] \\ &= E[\phi_1 Z_t Z_{t+k-1} + \phi_2 Z_t Z_{t+k-2} + Z_t a_{t+k}] \\ &= \phi_1 E[Z_t Z_{t+k-1}] + \phi_2 E[Z_t Z_{t+k-2}] + E[Z_t a_{t+k}] \end{aligned}$$

Expresión que se divide en los casos:

Para $k = 0$

$$\begin{aligned} E[Z_t Z_t] &= \phi_1 E[Z_t Z_{t-1}] + \phi_2 E[Z_t Z_{t-2}] + E[Z_t a_t] \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + E[Z_t a_t] \end{aligned}$$

donde $\forall t \quad E[Z_t a_t] = \phi_1 E[Z_{t-1} a_t] + \phi_2 E[Z_{t-2} a_t] + E[a_t^2]$ y como Z_{t-1} depende de los errores aleatorios hasta el tiempo $t-1$ y no depende de a_t ; lo mismo que ocurre

con Z_{t-2} en el tiempo $t-2$, entonces se tiene el siguiente resultado $\forall i \geq 1$ $E[Z_{t-i}, a_t] = E[Z_{t-i}] \cdot E[a_t] = 0$, quedando sólo el término $E[a_t^2] = \sigma_a^2$ y por lo tanto $\forall t$ $E[Z_t, a_t] = \sigma_a^2$. Así, para el caso $k=0$ se tiene que

$$E[Z_t, Z_t] = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2$$

Para $k > 0$

$$\begin{aligned} E[Z_t, Z_{t+k}] &= \phi_1 E[Z_t, Z_{t+k-1}] + \phi_2 E[Z_t, Z_{t+k-2}] + E[Z_t, a_{t+k}] \\ &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + E[Z_t, a_{t+k}] \end{aligned}$$

y debido a que $k > 0$ y por el resultado del caso anterior se tiene que $E[Z_t, a_{t+k}] = 0$, por lo tanto, para $k > 0$

$$E[Z_t, Z_{t+k}] = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}$$

Juntando los resultados anteriores, para obtener la función de autocovarianza en general, se tiene

$$\gamma_k = \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2 & \text{si } k = 0 \\ \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} & \text{si } k > 0 \end{cases} \quad (2.12)$$

Para obtener los valores de las autocorrelaciones ρ_1 y ρ_2 se hará uso de las ecuaciones de Yule-Walker, las cuales se definen como sigue: si $\gamma_0 < \infty$ entonces se pueden dividir a las autocovarianzas γ_1 y γ_2 , definidas como en (2.12), entre γ_0 obteniendo así las ecuaciones

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \quad (2.13)$$

Las cuales, como se puede ver, permiten obtener los valores de ρ_1 y ρ_2 en términos de los parámetros autorregresivos ϕ_1 y ϕ_2 .

La forma general de la FAC, quedaría entonces como sigue:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k \geq 3$$

Expresión que es de gran utilidad por su comportamiento autorregresivo similar al de (2.9), pero con la ventaja de que aquí no aparece ningún término aleatorio. Por lo tanto, las raíces de la ecuación característica (2.10) sirven para indicar el comportamiento dinámico de la función de autocorrelación; y utilizando los casos 1 y 3 de la expresión (1.15), se tiene que: si $\phi_1^2 + 4\phi_2 \geq 0$, entonces las raíces son

reales y además las autocorrelaciones decaerán exponencialmente a cero; para el caso $\phi_1^2 + 4\phi_2 < 0$ las raíces serán complejas, por lo que la FAC correspondiente seguirá un comportamiento senoidal convergente a cero.

Resolviendo las ecuaciones (2.13) en términos de los coeficientes autorregresivos, se tiene

$$\begin{aligned}\phi_1 &= \rho_1(1 - \rho_2)/(1 - \rho_1^2) \\ \phi_2 &= (\rho_2 - \rho_1^2)/(1 - \rho_1^2)\end{aligned}\quad (2.14)$$

Resultado con el cual se obtienen estimaciones para los parámetros ϕ_1 y ϕ_2 al sustituir los valores de las autocorrelaciones ρ_1 y ρ_2 por sus valores estimados.

De esta manera con (2.14) se pueden expresar las condiciones de estacionariedad (2.11), pero ahora en términos de las autocorrelaciones como sigue.

$$\begin{aligned}1) \phi_2 + \phi_1 < 1: & \quad (\rho_2 - \rho_1^2)/(1 - \rho_1^2) + \rho_1(1 - \rho_2)/(1 - \rho_1^2) < 1 \\ \text{desarrollando} & \quad (\rho_2 - \rho_1^2 + \rho_1 - \rho_1\rho_2)/(1 - \rho_1^2) < 1 \\ \text{despejando} & \quad \rho_2 - \rho_1^2 + \rho_1 - \rho_1\rho_2 < 1 - \rho_1^2\end{aligned}$$

$$\begin{aligned}2) \phi_2 - \phi_1 < 1: & \quad (\rho_2 - \rho_1^2)/(1 - \rho_1^2) - \rho_1(1 - \rho_2)/(1 - \rho_1^2) < 1 \\ \text{desarrollando} & \quad (\rho_2 - \rho_1^2 - \rho_1 + \rho_1\rho_2)/(1 - \rho_1^2) < 1 \\ \text{despejando} & \quad \rho_2 - \rho_1^2 - \rho_1 + \rho_1\rho_2 < 1 - \rho_1^2\end{aligned}$$

$$\begin{aligned}3) |\phi_2| < 1: & \quad |\rho_2 - \rho_1^2|/(1 - \rho_1^2) < 1 \\ & \quad |\rho_2 - \rho_1^2| < |1 - \rho_1^2| \\ \text{despejando} & \quad |\rho_2 - \rho_1^2| < |1 - \rho_1^2|\end{aligned}$$

De las expresiones 1) a 3) se puede obtener lo siguiente.

$$\text{De 3) } |\phi_2| < 1: \quad |\rho_2 - \rho_1^2| < |1 - \rho_1^2|$$

se tienen los casos

$$\text{a) } \rho_2 - \rho_1^2 < 1 - \rho_1^2 \quad \text{despejando,}$$

$$\rho_2 < 1$$

$$\begin{aligned}
 \text{b) } & -\rho_2 + \rho_1^2 < 1 - \rho_1^2 \\
 & -\rho_2 + 2\rho_1^2 < 1 \\
 & 2\rho_1^2 < 1 + \rho_2 \\
 & \rho_1^2 < (1 + \rho_2) / 2
 \end{aligned}$$

$$\begin{aligned}
 \text{De 1) } \phi_2 + \phi_1 < 1: & \rho_2 - \rho_1^2 + \rho_1 - \rho_1\rho_2 < 1 - \rho_1^2 \\
 \text{reduciendo} & \rho_2 + \rho_1 - \rho_1\rho_2 < 1 \\
 & \rho_2 + \rho_1(1 - \rho_2) < 1 \\
 & \rho_1(1 - \rho_2) < 1 - \rho_2 \\
 & \rho_1 < 1
 \end{aligned}$$

$$\begin{aligned}
 \text{De 2) } \phi_2 - \phi_1 < 1: & \rho_2 - \rho_1^2 - \rho_1 + \rho_1\rho_2 < 1 - \rho_1^2 \\
 \text{reduciendo} & \rho_2 - \rho_1 + \rho_1\rho_2 < 1 \\
 & \rho_2 - \rho_1(1 - \rho_2) < 1 \\
 & -\rho_1(1 - \rho_2) < 1 - \rho_2 \\
 & -\rho_1 < 1 \\
 & \rho_1 > -1
 \end{aligned}$$

En resumen, se obtienen las condiciones de estacionariedad del proceso autorregresivo de segundo orden, en términos de sus autocorrelaciones

$$\begin{aligned}
 \rho_2 & < 1 \\
 \rho_1^2 & < (1 + \rho_2) / 2 \\
 -1 & < \rho_1 < 1
 \end{aligned} \tag{2.15}$$

2.1.3 Caso general para los modelos AR(p).

Ahora se considerara el caso general de un proceso autorregresivo de orden p AR(p), el cual se describe mediante la ecuación (2.4) y en su forma equivalente como

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad \text{con } \tilde{Z}_t = Z_t - \mu \tag{2.16}$$

y como se ha visto, un proceso AR(p) será estacionario si y sólo si las raíces de la ecuación característica correspondiente

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0$$

se encuentran fuera del círculo unitario. El Teorema de Schur puede ser utilizado para encontrar las condiciones de estacionariedad en términos de los parámetros $\phi_1, \phi_2, \dots, \phi_p$, siempre y cuando los p determinantes siguientes sean positivos

$$D_1 = \begin{vmatrix} -1 & \phi_p \\ \phi_p & -1 \end{vmatrix}$$

$$D_2 = \begin{vmatrix} -1 & 0 & \phi_p & \phi_{p-1} \\ \phi_{p-1} & -1 & 0 & \phi_p \\ \phi_p & 0 & -1 & \phi_{p-1} \\ \phi_{p-1} & \phi_p & 0 & -1 \end{vmatrix}$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 & \phi_p & \phi_{p-1} & \phi_{p-2} \\ \phi_{p-2} & -1 & 0 & 0 & \phi_p & \phi_{p-1} \\ \phi_{p-1} & \phi_{p-2} & -1 & 0 & 0 & \phi_p \\ \phi_p & 0 & 0 & -1 & \phi_{p-2} & \phi_{p-1} \\ \phi_{p-1} & \phi_p & 0 & 0 & -1 & \phi_{p-2} \\ \phi_{p-2} & \phi_{p-1} & \phi_p & 0 & 0 & -1 \end{vmatrix} \dots$$

$$D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & \phi_p & \phi_{p-1} & \dots & \phi_1 \\ \phi_1 & -1 & \dots & 0 & 0 & \phi_p & \dots & \phi_2 \\ \dots & \dots \\ \phi_{p-1} & \phi_{p-2} & \dots & -1 & 0 & 0 & \dots & \phi_p \\ \phi_p & 0 & \dots & 0 & -1 & \phi_1 & \dots & \phi_{p-1} \\ \phi_{p-1} & \phi_p & \dots & 0 & 0 & -1 & \dots & \phi_{p-2} \\ \dots & \dots \\ \phi_1 & \phi_2 & \dots & \phi_p & 0 & 0 & \dots & -1 \end{vmatrix}$$

Otra manera de verificar el supuesto de estacionariedad del proceso es utilizando las ecuaciones de Yule-Walker correspondientes al proceso AR(p), es decir,

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\
&\dots \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \dots + \phi_p
\end{aligned}
\tag{2.17}$$

de las cuales se pueden obtener los valores de las primeras p autocorrelaciones en función de los parámetros autorregresivos $\phi_1, \phi_2, \dots, \phi_p$; y las posteriores se obtienen de la relación

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \dots + \phi_p\rho_{k-p}, \quad k \geq p+1 \tag{2.18}$$

Esta última ecuación se puede expresar como una ecuación en diferencia de la forma

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)\rho_k = 0, \quad k = p+1, p+2, \dots$$

con (2.17) como condiciones iniciales. Se puede observar (para los procesos AR(p)), que de ser $\{Z_t\}$ un proceso estacionario, la varianza de cualquier combinación lineal de N variables aleatorias, sean estas $Z_1, Z_{1-1}, \dots, Z_{1-N+1}$, por ejemplo

$$L_t = c_1 Z_t + c_2 Z_{t-1} + \dots + c_N Z_{t-N+1}$$

se calcula de la siguiente manera

$$\begin{aligned}
\text{Var}(L_t) &= E\left[(L_t - E[L_t])^2 \right] \\
&= E\left[(L_t)^2 \right] \quad \text{pues } E[L_t] = 0 \\
&= E\left[\left(\sum_{j=1}^N c_j Z_{t-j+1} \right)^2 \right] \\
&= E\left[\sum_{j=1}^N c_j Z_{t-(j-1)} \sum_{i=1}^N c_i Z_{t-(i-1)} \right] \\
&= E\left[\sum_{j=1}^N \sum_{i=1}^N c_j c_i Z_{t-(j-1)} Z_{t-(i-1)} \right]
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^N \sum_{j=1}^N c_i c_j E[Z_{i-(j-i)} Z_{i-(i-i)}] \\
&= \sum_{i=1}^N \sum_{j=1}^N c_i c_j \gamma_{|i-(j-i)-i+(i-i)|} \\
&= \sum_{i=1}^N \sum_{j=1}^N c_i c_j \gamma_{|-j+i+i-i|} \\
&= \sum_{i=1}^N \sum_{j=1}^N c_i c_j \gamma_{|i-j|}
\end{aligned}$$

pero como $i-j$ puede ser >0 , <0 ó 0 y como además $\gamma_k = \gamma_{-k}$, lo anterior se simplifica con la expresión

$$= \sum_{i=1}^N \sum_{j=1}^N c_i c_j \gamma_{|i-j|}$$

Que tiene la característica de ser siempre positiva (por propiedad de varianza), a excepción de que los coeficientes c_1, c_2, \dots, c_N sean todos cero. Si se expresa a dicha varianza en términos matriciales se tiene que

$$\text{Var}(L_T) = C^* T_N C > 0$$

donde $C^* = (c_1, c_2, \dots, c_N)$ y T_N es la matriz simétrica

$$T_N = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{N-2} \\ & & & \cdots & \\ \gamma_{N-1} & \gamma_{N-2} & \gamma_{N-3} & \cdots & \gamma_0 \end{bmatrix}$$

la cual por consiguiente debe ser **positiva definida**¹⁰. Por lo tanto, la matriz de autocorrelaciones A_N , que está dada por

$$A_N = \gamma_0^{-1} T_N = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{N-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{N-2} \\ & & & \cdots & \\ \rho_{N-1} & \rho_{N-2} & \rho_{N-3} & \cdots & 1 \end{bmatrix}$$

¹⁰ "Sean A una matriz de $m \times m$ y X un vector de $m \times 1$ de modo que si $X^* A X > 0 \forall X \neq 0$ entonces A es positiva definida."

debe ser también positiva definida y por consecuencia el determinante de A_N , así como todos sus menores principales, deben ser positivos (resultado de álgebra lineal). Es decir, para que el proceso AR(p) sea estacionario se debe cumplir que los determinantes

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix},$$

$$\begin{vmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{vmatrix}, \dots, \begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_p \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_p & \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{vmatrix}$$

sean todos positivos.

Entonces si el proceso AR(p) resulta ser estacionario, de acuerdo con lo visto en las secciones 1.2 y 1.3, puede ser representado como una suma ponderada de choques aleatorios, es decir, se puede expresar como sigue

$$Z_t = a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \quad (2.19)$$

donde $\sum_{i=1}^{\infty} |\psi_i| = \text{cte} < \infty$ y los coeficientes ψ_i , $i = 1, 2, \dots$ se obtienen del hecho de que un proceso AR(p) estacionario se puede expresar como

$$\phi(B)Z_t = a_t \quad \text{y} \quad Z_t = \psi(B)a_t$$

y entonces

$$\psi(B) = 1/\phi(B) \quad \text{o bien} \quad 1 = \phi(B)\psi(B) \quad (2.20)$$

Ahora, como se sabe $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{p-1} B^{p-1} - \phi_p B^p)$, de manera que las ψ_i 's se pueden determinar de manera autorregresiva en función de los coeficientes $\phi_1, \phi_2, \dots, \phi_{p-1}, \phi_p$, es decir,

$$\begin{aligned}
I &= (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \dots - \phi_{p-1} B^{p-1} - \phi_p B^p)(1 - \psi_1 B - \psi_2 B^2 - \psi_3 B^3 - \dots) \\
&= (1 - \psi_1 B - \psi_2 B^2 - \psi_3 B^3 - \dots) - \phi_1 (B - \psi_1 B^2 - \psi_2 B^3 - \dots) - \\
&\quad \phi_2 (B^2 - \psi_1 B^3 - \psi_2 B^4 - \psi_3 B^5 - \dots) - \phi_3 (B^3 - \psi_1 B^4 - \psi_2 B^5 - \psi_3 B^6 - \dots) - \dots - \\
&\quad \phi_{p-1} (B^{p-1} - \psi_1 B^p - \psi_2 B^{p+1} - \psi_3 B^{p+2} - \dots) - \phi_p (B^p - \psi_1 B^{p+1} - \psi_2 B^{p+2} - \psi_3 B^{p+3} - \dots) \\
&= 1 - (\psi_1 + \phi_1)B - (\psi_2 - \phi_1 \psi_1 + \phi_2)B^2 - (\psi_3 - \phi_1 \psi_2 - \phi_2 \psi_1 + \phi_3)B^3 - \dots - \\
&\quad (\psi_{p-1} - \phi_1 \psi_{p-2} - \phi_2 \psi_{p-3} - \dots + \phi_{p-1})B^{p-1} - (\psi_p - \phi_1 \psi_{p-1} - \phi_2 \psi_{p-2} - \dots - \phi_{p-1} \psi_1 + \phi_p)B^p - \dots
\end{aligned}$$

para que (2.20) se cumpla, todos los coeficientes que aparecen multiplicando a B^i deben ser cero $\forall i \geq 1$, con lo cual resulta que

$$\begin{aligned}
\psi_1 &= -\phi_1 \\
\psi_2 &= \phi_1 \psi_1 - \phi_2 \\
\psi_3 &= \phi_1 \psi_2 + \phi_2 \psi_1 - \phi_3 \\
&\dots \\
\psi_{p-1} &= \phi_1 \psi_{p-2} + \phi_2 \psi_{p-3} + \dots - \phi_{p-1} \\
\psi_p &= \phi_1 \psi_{p-1} + \phi_2 \psi_{p-2} + \dots + \phi_{p-1} \psi_1 - \phi_p
\end{aligned}$$

generalizando,

$$\begin{aligned}
\psi_i &= \phi_1 \psi_{i-1} + \phi_2 \psi_{i-2} + \dots + \phi_{i-1} \psi_1 - \phi_i, \quad \text{para } i \geq 3 \\
\text{con } \phi_i &= 0 \quad \text{para } i > p
\end{aligned}$$

Como se puede observar el mismo proceso AR(p) se puede representar a través de p parámetros autorregresivos o bien por medio de un número infinito de coeficientes ψ asociados ambos con el proceso $\{a_t\}$. Por lo que resulta preferible desarrollar el primer caso por cuestiones de realizar siempre el mínimo esfuerzo, por supuesto si la explicación que se obtiene del fenómeno es la misma. Pero no siempre es cierto esto, pues tómesese como ejemplo en que p fuese infinito y el proceso estacionario estuviera dado por

$$(1 - \phi B - \phi^2 B^2 - \dots)Z_t = a_t, \quad \text{con } |\phi| < 1 \quad (2.21)$$

que puede escribirse también como

$$\begin{aligned}
Z_t &= (1 - \phi B - \phi^2 B^2 - \dots)^{-1} a_t, \quad \text{con } |\phi| < 1 \\
&= (1 + \phi B)^{-1} a_t, \\
&= (1 - \psi_1 B - \psi_2^2 B^2 - \dots) a_t, \quad \text{por (2.20)}
\end{aligned} \quad (2.22)$$

con $\psi_1 = -\phi$ y $\psi_i = 0$ para $i \geq 2$. Como se ve el ahorro en parámetros en este caso es infinito, por lo que la representación (2.22) es preferible a la representación (2.21), aún cuando ambas son equivalentes.

Con el ejemplo anterior se puede ver que existen procesos que pueden representarse mejor en términos de choques aleatorios, que en términos de autorregresiones, a este tipo se le denomina de promedios móviles y se le denota por las siglas MA.

Por último, cabe mencionar que para los modelos autorregresivos existe solamente una FAC asociada con el proceso AR en estudio, debido a que esta en función de los parámetros y autocorrelaciones propias del proceso en estudio.

2.2 Modelos de promedios móviles (MA).

La idea básica de estos modelos¹¹ consiste en representar a un proceso estocástico $\{Z_t\}$, cuyos valores pueden ser dependientes unos de otros, como una suma finita ponderada de choques aleatorios independientes $\{a_t\}$, es decir

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (2.23)$$

donde $\{Z_t\}$ representa a las desviaciones del proceso $\{Z_t\}$ con respecto a su nivel medio μ y $\theta_1, \theta_2, \dots, \theta_q$ son las ponderaciones (parámetros de promedios móviles) asociados con los choques aleatorios en los periodos $t-1, t-2, \dots, t-q$, respectivamente.

Aunque el término de promedios móviles da la idea de que el modelo se obtiene como un promedio de los choques aleatorios que intervienen, esto no es así puesto que los parámetros no tienen que ser necesariamente positivos ni su suma debe ser la unidad, como es característica. Como se puede ver (2.23) tiene la forma general de (1.1) y la suma $\sum_{i=1}^q |\theta_i|$, por ser sobre un número finito de sumandos, debe tener un resultado finito (a menos que algún θ_i fuese infinito, pero eso implicaría que Z_t es también infinito). Por lo tanto, **todo MA¹² es estacionario**. Entonces para los modelos MA: dado un proceso que se encuentre en equilibrio, las fluctuaciones alrededor del punto de equilibrio, $\{Z_t\}$, son causadas por choques asociados con eventos inesperados; los cuales no necesariamente se asimilan de manera instantánea, sino que pueden seguir

¹¹ Estos modelos fueron introducidos por Yule (1926) y Slutsky (1927).

¹² El término MA proviene de las siglas en Inglés Moving Average

causando efectos aun después de transcurrir un cierto número de períodos y además la intensidad del choque se refleja en el valor de su ponderación θ .

2.2.1 El modelo MA de orden uno MA(1).

El proceso de promedios móviles de orden uno, MA(1) es el más simple y se expresa como sigue

$$Z_t = (1 - \theta B)a_t, \quad (2.24)$$

del cual se puede obtener de manera inmediata que

$$\begin{aligned} E[Z_t] &= E[(a_t - \theta a_{t-1})] = 0 \quad \text{y} \\ \gamma_0 &= \text{Var}[Z_t] = \text{Var}[(a_t - \theta a_{t-1})] = \sigma_a^2(1 + \theta^2) \end{aligned} \quad (2.25)$$

luego, las autocovarianzas están dadas por la función

$$\begin{aligned} \gamma_k &= E[(Z_t, Z_{t-k})] \\ &= E[(a_t - \theta a_{t-1})(a_{t-k} - \theta a_{t-k-1})] \\ &= E[a_t a_{t-k} - \theta a_t a_{t-k-1} - \theta a_{t-1} a_{t-k} + \theta^2 a_{t-1} a_{t-k-1}] \\ &= E[a_t a_{t-k}] - \theta E[a_t a_{t-k-1}] - \theta E[a_{t-1} a_{t-k}] + \theta^2 E[a_{t-1} a_{t-k-1}] \\ &= \begin{cases} -\theta \sigma_a^2 & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases} \end{aligned} \quad (2.26)$$

por lo que la FAC está dada por

$$\rho_k = \gamma_k / \gamma_0 = \begin{cases} \frac{-\theta}{1 + \theta^2} & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases} \quad (2.27)$$

Se puede concluir, del hecho de que las autocorrelaciones para retrasos mayores que un período sean cero, que el proceso MA(1) no tiene memoria para más allá de lo ocurrido en un periodo anterior. Pero, aunque la primera autocorrelación es distinta de cero, no puede ser muy grande, puesto que este hecho indicaría que existiría una fuerte dependencia de la observación actual con la anterior y así sucesivamente, y entonces sería más adecuado pensar en un modelo

autorregresivo para esa situación; por otro lado, se puede demostrar que dada la ecuación (2.27) y debido a que $|0| < 1$, se obtiene que

$$|\rho_1| \leq 0.5 \quad (2.28)$$

Demostración:

Por hipótesis $|0| < 1$, entonces se tienen

$$0^2 < 1 \quad (1)$$

$$-1 < 0 < 1 \quad (2)$$

$$-1 < -0 < 1 \quad (3)$$

al sumar 1 en ambos lados de (1) e invertir se tiene que $1 + 0^2 < 2$, y en consecuencia

$$\frac{1}{1+0^2} > \frac{1}{2} \quad (4);$$

al multiplicar (4) por $-0 > -1$, de (3) se consigue que $\frac{-0}{1+0^2} > -\frac{1}{2}$;

por otro lado, si se multiplica a (4) por $-1 < 0$ de (2)

$$\frac{0}{1+0^2} > -\frac{1}{2} \quad (5);$$

que al multiplicar por -1 ,

$$\text{da } -\frac{0}{1+0^2} < \frac{1}{2}$$

entonces al unir ambos resultados, se tiene que

$$-\frac{1}{2} < \frac{-0}{1+0^2} < \frac{1}{2}$$

por lo tanto,

$$\left| \frac{-0}{1+0^2} \right| < \frac{1}{2}$$

Resultado que satisface la conclusión. Así, se puede afirmar que aún cuando una cierta FAC muestre un sólo pico, correspondiente al retraso 1, ésta representará a un proceso MA(1) sólo si se satisface la restricción (2.28).

Regresando al final de la sección 2.1.3, donde se vio que los procesos autorregresivos estacionarios también pueden ser representados por modelos de promedios móviles, como es el caso del proceso AR(∞) descrito por (2.21) y cuya representación equivalente es el proceso MA(1), siempre y cuando $|\phi| < 1$. Esto implica entonces que el proceso descrito por (2.24) se puede escribir de manera autorregresiva siempre y cuando $|0| < 1$.

Generalizando, cuando un proceso puede expresarse mediante un modelo AR, se dice que dicho proceso es *invertible*, es decir que puede ser representado mediante

$$\pi(B)Z_t = a_t,$$

donde

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

es un polinomio de retraso que tiene la característica de que la suma

$$\pi(x) = 1 - \sum_{i=1}^{\infty} \pi_i x^i$$

converge dentro o sobre el círculo unitario, y que puede interpretarse como una restricción sobre los coeficientes π_1, π_2, \dots (que están asociados con las variables Z_{t-1}, Z_{t-2}, \dots), de manera tal que mientras mayor sea el retraso de la variable, menor será el valor de la π correspondiente, es decir, que es menor la influencia de dicha variable retrasada sobre las más recientes.

Puede observarse que en los párrafos anteriores aparece la idea de DUALIDAD, entre estacionariedad para un proceso AR e invertibilidad para un proceso MA. En adición, las condiciones para invertibilidad pueden ser obtenidas de manera similar a las de estacionariedad, como se puede apreciar para el caso (2.24) donde se necesita que la suma

$$\pi(x) = 0^{-1}(x) = 1 + \sum_{i=1}^{\infty} \theta^i x^i$$

converja dentro o sobre el círculo unitario, lo cual ocurre siempre y cuando $|\theta| < 1$. Por lo tanto, la condición de invertibilidad para un proceso MA se puede expresar también en términos del polinomio $\theta(B)$ siempre y cuando las raíces de la ecuación

$$\theta(x) = 0$$

se encuentren fuera del círculo unitario. El concepto de invertible es de gran importancia, pues **todo proceso que lo es, está determinado de manera única por su FAC**, lo cual no ocurre con procesos que no lo son.

2.2.2 Caso para los modelos MA(2).

El proceso de promedios móviles de orden dos MA(2) se expresa como sigue

$$Z_t = (1 - \theta_1 B - \theta_2 B^2)a_t \quad (2.29)$$

cuya media y varianza son

$$E(Z_t) = 0 \quad \text{y} \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_a^2 \quad (2.30)$$

y función de autocovarianza dada por

$$\begin{aligned} \gamma_k &= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-k} - \theta_1 a_{t-k-1} - \theta_2 a_{t-k-2})] \\ &= E[a_t a_{t-k} - \theta_1 a_t a_{t-k-1} - \theta_2 a_t a_{t-k-2} - \theta_1 a_{t-1} a_{t-k} + \theta_1^2 a_{t-1} a_{t-k-1} + \theta_1 \theta_2 a_{t-1} a_{t-k-2} - \\ &\quad \theta_2 a_{t-2} a_{t-k} + \theta_2 \theta_1 a_{t-2} a_{t-k-1} + \theta_2^2 a_{t-2} a_{t-k-2}] \\ &= E[a_t a_{t-k} - \theta_1 (a_t a_{t-k-1} + a_{t-1} a_{t-k}) - \theta_2 (a_t a_{t-k-2} + a_{t-2} a_{t-k}) + \\ &\quad \theta_1 \theta_2 (a_{t-1} a_{t-k-2} + a_{t-2} a_{t-k-1}) + \theta_1^2 a_{t-1} a_{t-k-1} + \theta_2^2 a_{t-2} a_{t-k-2}] \end{aligned} \quad (2.31)$$

$$= \begin{cases} (-\theta_1 + \theta_1 \theta_2)\sigma_a^2 & \text{si } k = 1 \\ (-\theta_2)\sigma_a^2 & \text{si } k = 2 \\ 0 & \text{si } k \geq 3 \end{cases}$$

entonces la correspondiente FAC esta dada por

$$\rho_k = \begin{cases} \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{si } k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{si } k = 2 \\ 0 & \text{si } k \geq 3 \end{cases} \quad (2.32)$$

Por lo anterior, se puede concluir que el proceso es estacionario; y para que sea también invertible se necesita que las raíces de la ecuación

$$1 - \theta_1 x - \theta_2 x^2 = 0$$

se encuentren fuera del círculo unitario, para lo cual se utiliza el Teorema de Schur y las relaciones (2.10) y (2.11), con lo que se obtiene que las condiciones de invertibilidad son

$$|\theta_2| < 1, \quad \theta_2 + \theta_1 < 1 \quad \text{y} \quad \theta_2 - \theta_1 < 1 \quad (2.33)$$

Ahora, así como se restringió el valor de la primera autocorrelación para el caso MA(1), de igual manera se hará para las de un proceso MA(2); mediante

$$\rho_1^2 \leq 0.5 \quad \text{y} \quad |\rho_2| \leq 0.5 \quad (2.34)$$

2.2.3 Caso general para los modelos MA(q).

La representación general de un proceso de promedios móviles de orden $q \geq 1$ es la siguiente

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (2.35)$$

con $Z_t = Z_t - \mu$, con μ el nivel del proceso, $\theta_1, \theta_2, \dots, \theta_q$ los parámetros de promedios móviles y $\{a_t\}$ es ruido blanco, bajo la hipótesis de que tiene media cero y varianza constante σ_a^2 . Como se vio todo proceso MA es estacionario, y por lo tanto un MA(q) no es la excepción, pues ni la media, ni la varianza, ni las covarianzas dependen del tiempo:

$$E[Z_t] = E[Z_t - \mu] = 0$$

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

$$\begin{aligned} \gamma_k &= E[(a_t - \theta_1 a_{t-1} - \dots - \theta_{q-k} a_{t-q+k} - \dots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \dots - \\ &\quad \theta_{q-k} a_{t-k-(q-k)} - \dots - \theta_q a_{t-k-q})] \\ &= E[a_t a_{t-k} + \theta_1^2 a_{t-1} a_{t-1-k} - \theta_k (a_{t-1} a_{t-1-k} + a_{t-2} a_{t-2-k}) + \theta_1 \theta_{k+1} (a_{t-1} a_{t-1-k-1} + a_{t-1} a_{t-2-k-1}) + \dots + \\ &\quad \theta_{q-k} \theta_q (a_{t-q} a_{t-q-k} + a_{t-q+k} a_{t-k-q}) + \dots + \theta_{q-1} \theta_q (a_{t-q+1} a_{t-k-q} + \theta_q a_{t-q} \theta_{t-k-q-1})] \\ &= \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_a^2, & \text{si } k = 1, \dots, q \\ 0 & \text{si } k \geq q+1 \end{cases} \end{aligned} \quad (2.36)$$

definiendo a $\theta_0 = 0$. Por lo que la FAC esta definida como

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & \text{si } k = 1, \dots, q \\ 0 & \text{si } k \geq q+1 \end{cases} \quad (2.37)$$

con lo cual se puede ver que el proceso MA(q) tiene una memoria limitada a q periodos. Las expresiones de (2.36) permiten apreciar que no es posible obtener un conjunto de ecuaciones, similares a las de Yule-Walker, que muestren a los parámetros de promedios móviles en términos de las autocovarianzas, pues dichas ecuaciones serían no-lineales y por lo tanto, no tendrían una solución única, a menos que se establecieran algunas restricciones sobre los valores de los parámetros; y precisamente estas últimas, son las condiciones de invertibilidad, las cuales van a permitir asociar un sólo proceso MA a una FAC.

Para obtener las condiciones de invertibilidad del proceso (2.35) se utiliza la dualidad con las restricciones de estacionariedad para un proceso AR(q), entonces haciendo uso del Teorema de Schur, se tiene que para que un MA(q) sea invertible, se necesita que los q determinantes

$$D_1 = \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} -1 & 0 & 0 & \theta_{q-1} \\ 0 & -1 & 0 & \theta_q \\ 0 & 0 & -1 & \theta_{q-1} \\ 0 & 0 & 0 & -1 \end{vmatrix},$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 & \theta_q & \theta_{q-1} & \theta_{q-2} \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & \theta_q \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & \theta_q & 0 & 0 & -1 & \theta_{q-2} \\ 0 & \theta_{q-2} & \theta_{q-1} & \theta_q & 0 & -1 \end{vmatrix} \dots$$

$$D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & \theta_q & \theta_{q-1} & \dots & \theta_1 \\ 0 & -1 & \dots & 0 & 0 & \theta_q & \dots & \theta_2 \\ \dots & \dots \\ \theta_{q-1} & \theta_{q-2} & \dots & -1 & 0 & 0 & \dots & \theta_q \\ \theta_q & 0 & \dots & 0 & -1 & \theta_1 & \dots & \theta_{q-1} \\ \theta_{q-1} & \theta_q & \dots & 0 & 0 & -1 & \dots & \theta_{q-2} \\ \dots & \dots \\ \theta_1 & \theta_2 & \dots & \theta_p & 0 & 0 & \dots & -1 \end{vmatrix} \quad (2.38)$$

sean todos positivos, y bastará con que alguno no lo sea para concluir que el proceso no es invertible. Luego, si resulta serlo entonces se podrá expresar también de la forma

$$\pi(B)Z_t = 1 - \pi_1 Z_{t-1} - \pi_2 Z_{t-2} - \dots \quad (2.39)$$

con $\sum_{j=1}^{\infty} |\pi_j| < \infty$, y los coeficientes π_1, π_2, \dots se pueden obtener de la relación

$$\pi(B)\theta(B) = 1$$

entonces se tendría que

$$\begin{aligned} 1 &= (1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \pi_4 B^4 - \dots)(1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4 - \dots - \theta_q B^q) \\ &= (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4 - \dots - \theta_q B^q) - \\ &\quad \pi_1 (B - \theta_1 B^2 - \theta_2 B^3 - \theta_3 B^4 - \dots - \theta_q B^{q+1}) - \\ &\quad \pi_2 (B^2 - \theta_1 B^3 - \theta_2 B^4 - \theta_3 B^5 - \dots - \theta_q B^{q+2}) - \\ &\quad \pi_3 (B^3 - \theta_1 B^4 - \theta_2 B^5 - \theta_3 B^6 - \dots - \theta_q B^{q+3}) - \\ &\quad \pi_4 (B^4 - \theta_1 B^5 - \theta_2 B^6 - \theta_3 B^7 - \dots - \theta_q B^{q+4}) - \dots \\ &= 1 - (\theta_1 + \pi_1)B - (\theta_2 - \pi_1\theta_1 + \pi_2)B^2 - (\theta_3 - \pi_1\theta_2 - \pi_2\theta_1 + \pi_3)B^3 - \\ &\quad (\theta_4 - \pi_1\theta_3 - \pi_2\theta_2 - \pi_3\theta_1 + \pi_4)B^4 - \dots \end{aligned} \quad (2.40)$$

de manera que

$$\begin{aligned} \pi_1 &= -\theta_1 \\ \pi_2 &= \pi_1\theta_1 - \theta_2 \\ \pi_3 &= \pi_2\theta_1 + \pi_1\theta_2 - \theta_3 \\ \pi_4 &= \pi_3\theta_1 + \pi_2\theta_2 + \pi_1\theta_3 - \theta_4 \\ &\dots \\ \pi_j &= \pi_{j-1}\theta_1 + \pi_{j-2}\theta_2 + \dots + \pi_1\theta_{j-1} - \theta_j, \quad j > q \\ \text{y } \theta_k &= 0 \text{ para } j > q. \end{aligned} \quad (2.41)$$

2.3 Los modelos ARMA.

La generalización de los modelos AR Y MA es la combinación de estos y se les conoce como modelos *autorregresivos y de promedios móviles* (ARMA). El proceso ARMA(p,q) esta representado por

$$\phi(B)Z_t = \theta(B)a_t \quad (2.42)$$

en donde $\phi(B)$ y $\theta(B)$ son los polinomios de retraso de orden p y q respectivamente, $\{a_t\}$ es un proceso de ruido blanco y por último Z_t es la serie de desviaciones de la variable Z_t respecto a su nivel μ .

Esta idea de juntar a los procesos surge por el hecho de que las series de tiempo que se observan en la práctica, muchas veces presentan características tanto de procesos AR como MA; por lo que es necesario una combinación de ellos para obtener un modelo más completo de representación.

2.3.1 El modelo ARMA(1,1).

Aunque es el caso más sencillo de los procesos ARMA, proporciona representaciones adecuadas para muchas series de fenómenos reales. El modelo ARMA(1,1) está definido por

$$(1 - \phi B)Z_t = (1 - \theta B)a_t \quad (2.43)$$

y aunque posee características tanto autorregresivas como de promedios móviles no tiene por qué ser invertible ni estacionario, pero las condiciones para estos dos aspectos se derivan precisamente de los procesos AR(1) y MA(1); en otras palabras, para que (2.43) sea invertible y estacionario se necesita que tanto la raíz de $1 - \theta x = 0$ como de $1 - \phi x = 0$ se encuentren fuera del círculo unitario.

Y si el proceso resulta ser estacionario e invertible, entonces las expresiones

$$\begin{aligned} Z_t &= \psi(B)a_t = a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \quad y \\ \pi(B)Z_t &= Z_t - \pi_1 Z_{t-1} - \pi_2 Z_{t-2} - \dots = a_t \end{aligned}$$

son tales que las sumas $\sum_{i=1}^{\infty} |\psi_i|$ y $\sum_{i=1}^{\infty} |\pi_i|$ son convergentes. La comprobación se puede realizar al escribir (2.43) como

$$\begin{aligned} Z_t &= \left[\frac{1 - \theta B}{1 - \phi B} \right] a_t \\ &= \left[(1 - \theta B)(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) \right] a_t \\ &= \left[(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) - \theta B(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) \right] a_t \\ &= \left[1 - (\theta - \phi)B - \phi(\theta - \phi)B^2 - \phi^2(\theta - \phi)B^3 - \phi^3(\theta - \phi)B^4 - \dots \right] a_t \end{aligned}$$

lo cual tiene significado siempre y cuando se cumpla la condición de estacionariedad $|\phi| < 1$ y entonces las ponderaciones ψ_i están dadas por

$$\psi_i = \phi^{i-1}(\theta - \phi) \quad i = 1, 2, \dots$$

expresión de la cual se observa que, conforme el índice i crece, la ponderación ψ_i , tiende a cero y la suma infinita de ponderaciones $\sum_{i=1}^{\infty} \psi_i$ será convergente. Lo mismo ocurre con la representación de (2.43) en la forma

$$\begin{aligned} a_i &= \left[\frac{1 - \phi B}{1 - \theta B} \right] Z_i \\ &= \left[(1 - \phi B)(1 + \theta B + (\theta B)^2 + (\theta B)^3 + \dots) \right] Z_i \\ &= \left[(1 + \theta B + (\theta B)^2 + (\theta B)^3 + \dots) - \phi B(1 + \theta B + (\theta B)^2 + (\theta B)^3 + \dots) \right] Z_i \\ &= \left[1 - (\phi - \theta)B - 0(\phi - \theta)B^2 - \theta^2(\phi - \theta)B^3 - \theta^3(\phi - \theta)B^4 - \dots \right] Z_i \end{aligned}$$

la cual es válida si se cumple con la restricción $|\theta| < 1$, que es la condición de invertibilidad, por otro lado, los coeficientes se definen como

$$\pi_j = \theta^{j-1}(\phi - \theta), \quad j = 1, 2, \dots$$

y con respecto a la suma se tiene que

$$\sum_{j=1}^{\infty} \pi_j = (\phi - \theta) \sum_{j=1}^{\infty} \theta^{j-1} = \frac{\phi - \theta}{1 - \theta} < \infty$$

Para el cálculo de la varianza y autocovarianzas se utiliza a la expresión (2.43) en su forma equivalente

$$Z_i = \phi Z_{i-1} + a_i - \theta a_{i-1} \quad (2.44)$$

y ya que la media de Z_i es cero, se pasa al cálculo de la varianza haciendo uso de (2.44)

$$\begin{aligned} \gamma_0 &= \text{Var}(Z_i) \\ &= E[Z_i Z_i] \\ &= E[\phi Z_{i-1} Z_{i-1} + Z_i a_i - \theta Z_{i-1} a_{i-1}] \\ &= \phi E(Z_{i-1} Z_{i-1}) + E(Z_i a_i) - \theta E(Z_{i-1} a_{i-1}) \end{aligned}$$

en donde

$E(Z_t, Z_{t-1}) = \gamma_1$ es la autocovarianza entre Z_t y Z_{t-1}

$$E(Z_t, a_t) = \sigma_a^2$$

$$E(Z_t, a_{t-1}) = \phi E(Z_{t-1}, a_{t-1}) + E(a_t, a_{t-1}) - \theta E(a_{t-1}^2)$$

por el argumento de que $E(Z_t, a_t) = 0$, ya que

$$\begin{aligned} E(Z_t, a_t) &= E[\phi Z_{t-1} a_t + a_t^2 - \theta a_{t-1}] \\ &= \phi E(Z_{t-1} a_t) + E(a_t^2) - \theta E(a_{t-1}) \end{aligned}$$

en donde el término $E(Z_{t-1} a_t)$ depende de los errores aleatorios

hasta el tiempo $t-1$ y no depende de a_t , por lo que $E(Z_{t-1} a_t) = 0$

ya que $E(a_t) = 0$, en cuanto al término $E(a_{t-1}) = 0$. Así, el único

término que queda es $E(a_t^2) = \sigma_a^2$. Por lo tanto,

$$= \phi \sigma_a^2 - \theta \sigma_a^2$$

por lo que la varianza para este modelo se convierte en

$$\begin{aligned} \gamma_0 &= \phi \gamma_1 + \sigma_a^2 - \theta(\phi \sigma_a^2 - \theta \sigma_a^2) \\ &= \gamma_1 + \sigma_a^2 - \theta(\phi - \theta) \sigma_a^2 \\ &= \phi \gamma_1 + (1 - \theta(\phi - \theta)) \sigma_a^2 \end{aligned} \tag{2.45}$$

En cuanto a las autocovarianzas se tiene que

$$\begin{aligned} \gamma_k &= E[Z_t Z_{t-k}] \\ &= E[\phi Z_{t-1} Z_{t-k} + a_t Z_{t-k} - \theta a_{t-1} Z_{t-k}] \\ &= \phi E(Z_{t-1} Z_{t-k}) + E(a_t Z_{t-k}) - \theta E(a_{t-1} Z_{t-k}) \end{aligned}$$

Ahora, se observan dos casos

Cuando $k = 1$

$$\gamma_1 = \phi E(Z_{t-1} Z_{t-1}) + E(a_t Z_{t-1}) - \theta E(a_{t-1} Z_{t-1})$$

en donde se sabe que

$$E(Z_{t-1} Z_{t-1}) = \text{var}(Z_{t-1}) = \gamma_0$$

$$E(a_t Z_{t-1}) = 0 \quad \text{y}$$

$$E(a_{t-1} Z_{t-1}) = \sigma_a^2 \quad \text{así, al sustituir valores}$$

$$\gamma_1 = \phi \gamma_0 - \theta \sigma_a^2$$

Cuando $k \geq 2$

$$\gamma_k = \phi E(Z_{t-1} Z_{t-k}) + E(a_t Z_{t-k}) - 0E(a_{t-1} Z_{t-k})$$

aquí,

$$E(Z_{t-1} Z_{t-k}) = \gamma_{k-1}, \quad \text{resultado visto anteriormente}$$

$$E(a_t Z_{t-k}) = E(a_{t-1} Z_{t-k}) = 0 \quad \text{luego entonces,}$$

$$\gamma_k = \phi \gamma_{k-1}$$

Uniendo ambos resultados,

$$\gamma_k = \begin{cases} \phi \gamma_0 - 0 \sigma_a^2 & \text{si } k = 1 \\ \phi \gamma_{k-1} & \text{si } k \geq 2 \end{cases} \quad (2.46)$$

en donde se puede observar que para más de un retraso las autocovarianzas siguen la misma tendencia que un proceso AR(1), debido a que la parte de promedios móviles es de orden uno y por lo tanto sólo la primera autocovarianza muestra a los parámetros del tipo MA. También se deduce que al conocer γ_0 y γ_1 , las autocovarianzas restantes se obtienen de manera inmediata. Y para obtener γ_0 y γ_1 se resuelve el siguiente sistema de ecuaciones, obtenido de la expresión para la varianza y para el caso $k = 1$ de (2.46), respectivamente,

$$\begin{aligned} \gamma_0 &= \phi \gamma_1 + (1 - 0(\phi - 0)) \sigma_a^2 \\ \gamma_1 &= \phi \gamma_0 - 0 \sigma_a^2 \end{aligned}$$

cuya solución es

$$\begin{aligned} \gamma_0 &= \frac{(1 - 2\phi 0 + 0^2) \sigma_a^2}{1 - \phi^2} \\ \gamma_1 &= \frac{(1 - \phi 0)(\phi - 0) \sigma_a^2}{1 - \phi^2} \end{aligned} \quad (2.47)$$

Si se unen las expresiones (2.46) y (2.47), se obtiene la función de autovarianzas en general

$$\gamma_k = \frac{\phi^{k-1} (1 - \phi 0)(\phi - 0) \sigma_a^2}{1 - \phi^2}, \quad k = 1, 2, \dots$$

y que da origen a la FAC

$$\rho_k = \frac{\phi^{k-1} (1 - \phi 0)(\phi - 0)}{1 - 2\phi 0 + 0^2}, \quad k = 1, 2, \dots$$

la cual debido al factor ϕ^{k-1} y a la condición de estacionariedad $|\phi| < 1$, se puede afirmar que ρ_k tiene un decaimiento exponencial a cero después de ρ_1 . Por otra parte, el parámetro ϕ determina si la convergencia a cero de ρ_k será con o sin cambios alternados de signo, mientras que el parámetro θ interviene sólo para fijar el valor de ρ_1 .

2.3.2 El modelo ARMA(1,2).

Ahora se analizará el caso ARMA(1,2), el cual esta representado por la expresión

$$(1 - \phi B)Z_t = (1 - \theta_1 B - \theta_2 B^2)a_t \quad (2.48)$$

en donde se tiene que $\phi(B)$ y $\theta(B)$ son de orden 1 y 2, respectivamente. Para que la expresión anterior sea invertible se requiere que la raíz de $1 - \theta_1 x - \theta_2 x^2 = 0$ se encuentre fuera del círculo unitario, y para ser estacionario se necesita lo mismo para la raíz de $1 - \phi x = 0$.

Para probar que cumple con las condiciones de ser invertible y estacionario, se utilizan las expresiones

$$Z_t = \psi(B)a_t \quad \text{y} \\ \pi(B) = a_t$$

las cuales deben cumplir con que las sumas de $\sum_{j=1}^{\infty} |\psi_j|$ y $\sum_{j=1}^{\infty} |\pi_j|$ sean convergentes.

Para demostrar lo anterior, se despeja a Z_t de (3.3.8) y se trabaja con la expresión siguiente

$$\begin{aligned} Z_t &= (1 - \phi B)^{-1} * (1 - \theta_1 B - \theta_2 B^2) a_t \\ &= (1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) * (1 - \theta_1 B - \theta_2 B^2) a_t \\ &= [(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) - \theta_1 B(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots) - \\ &\quad \theta_2 B^2(1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots)] a_t \\ &= [1 - (\theta_1 - \phi)B - (\theta_2 + \phi\theta_1 - \phi^2)B^2 - \phi(\theta_2 + \phi\theta_1 - \phi^2)B^3 - \\ &\quad \phi^2(\theta_2 + \phi\theta_1 - \phi^2)B^4 - \phi^3(\theta_2 + \phi\theta_1 - \phi^2)B^5 - \dots] a_t \\ &= [1 - (\theta_1 - \phi)B - (\theta_2 + \phi(\theta_1 - \phi))B^2 - \phi(\theta_2 + \phi(\theta_1 - \phi))B^3 - \\ &\quad \phi^2(\theta_2 + \phi(\theta_1 - \phi))B^4 - \phi^3(\theta_2 + \phi(\theta_1 - \phi))B^5 - \dots] a_t \end{aligned}$$

expresión que es convergente, siempre y cuando se cumpla que $|\phi| < 1$ y las ψ_i , que se representarán como

$$\psi_i = \begin{cases} 1 & \text{si } i = 0 \\ (0_i - \phi) & \text{si } i = 1 \\ \phi^{i-2}(0_i + \phi(0_i - \phi)) & \text{si } i \geq 2 \end{cases}$$

y que conforme i crece, ψ_i tiende a cero y si se realiza la suma de todos los elementos, esta resulta ser finita, por lo que $\sum_{i=1}^{\infty} \psi_i$ es convergente.

Por otro lado, de (2.48) en la forma

$$a_i = \left(\frac{1 - \phi B}{1 - 0_1 B - 0_2 B^2} \right) Z_i \\ = (1 + 0_1 B + 0_2 B^2 + 0_1^2 B^2 + 20_1 0_2 B^3 + 0_1^3 B^4 + 0_1^2 0_2 B^3 + 30_1 0_2^2 B^4 + \dots)(1 - \phi B)Z_i$$

Para el cálculo de las medidas de dispersión se utilizará la expresión del proceso definido en (2.48), en su forma equivalente

$$Z_i - \phi Z_{i-1} = a_i - 0_1 a_{i-1} - 0_2 a_{i-2} \\ Z_i = \phi Z_{i-1} + a_i - 0_1 a_{i-1} - 0_2 a_{i-2}$$

y ya que $E[Z_i] = 0$, se pasa al cálculo de la varianza y autocovarianzas. Entonces

$$\gamma_0 = \text{Var}(Z_i) \\ = E[Z_i Z_i] \\ = E[\phi Z_i Z_{i-1} + Z_i a_i - 0 Z_i a_{i-1}] \\ = \phi E(Z_i Z_{i-1}) + E(Z_i a_i) - 0 E(Z_i a_{i-1})$$

Donde calculando por separado cada término, se tiene que

$E(Z_i Z_{i-1}) = \gamma_1$ es la autocovarianza entre Z_i y Z_{i-1}

$E(Z_i a_i) = \sigma_a^2$

$$\begin{aligned}
E(Z_{i-1} a_{i-1}) &= E[\phi Z_{i-1} a_{i-1} + a_{i-1} - 0_1 a_{i-1} a_{i-1} - 0_2 a_{i-1} a_{i-2}] \\
&= \phi E(Z_{i-1} a_{i-1}) + E(a_{i-1}) - 0_1 E(a_{i-1}^2) - 0_2 E(a_{i-1} a_{i-2}), \text{ ahora} \\
E(Z_{i-1} a_{i-1}) &= \sigma_a^2, \\
E(a_{i-1} a_{i-2}) &= E(a_{i-1} a_{i-2}) = 0, \text{ pues } a_{i-2}, a_{i-1} \text{ y } a_i \text{ son independientes} \\
E(a_{i-1}^2) &= \sigma_a^2, \text{ por ser } a_{i-1} \text{ ruido blanco, entonces} \\
&= \phi \sigma_a^2 + 0 - 0_1 * \sigma_a^2 - 0_2 * 0 \\
&= (\phi - 0_1) \sigma_a^2
\end{aligned}$$

$$\begin{aligned}
E(Z_i a_{i-2}) &= E[\phi Z_{i-1} a_{i-2} + a_{i-2} - 0_1 a_{i-1} a_{i-2} - 0_2 a_{i-2} a_{i-2}] \\
&= \phi E(Z_{i-1} a_{i-2}) + E(a_{i-2}) - 0_1 E(a_{i-1} a_{i-2}) - 0_2 E(a_{i-2}^2), \text{ ahora} \\
E(Z_{i-1} a_{i-2}) &= (\phi - 0_1) \sigma_a^2 \text{ del resultado anterior} \\
E(a_{i-1} a_{i-2}) &= E(a_{i-1} a_{i-2}) = 0, \text{ pues } a_{i-2}, a_{i-1} \text{ y } a_i \text{ son independientes} \\
E(a_{i-2}^2) &= \sigma_a^2, \text{ por ser } a_{i-2} \text{ ruido blanco, entonces} \\
&= \phi \sigma_a^2 (\phi - 0_1) + 0 - 0_1 * 0 - 0_2 * \sigma_a^2 \\
&= (\phi(\phi - 0_1) - 0_2) \sigma_a^2
\end{aligned}$$

Uniendo los resultados anteriores.

$$\begin{aligned}
\text{Var}[Z_i] &= \phi \gamma_1 + \sigma_a^2 - 0_1 (\phi - 0_1) \sigma_a^2 - 0_2 (\phi(\phi - 0_1) - 0_1) \sigma_a^2 \\
&= \phi \gamma_1 + [1 - 0_1 (\phi - 0_1) - 0_2 (\phi(\phi - 0_1) - 0_1)] \sigma_a^2 \\
&= \phi \gamma_1 + [1 - 0_1 \phi + 0_1^2 - 0_2 \phi(\phi - 0_1) + 0_2^2] \sigma_a^2
\end{aligned}$$

después de simplificar y reordenar términos, se tiene que por lo tanto

$$\gamma_0 = \text{Var}[Z_i] = \phi \gamma_1 + [1 + 0_1^2 + 0_2^2 + 0_2 \phi(0_1 - \phi) - 0_1 \phi] \sigma_a^2 \quad (2.49)$$

En cuanto al cálculo de las autocovarianzas se tiene lo siguiente

$$\begin{aligned}
\gamma_k &= \phi E(Z_i Z_{i-k}) + E(a_i Z_{i-k}) - 0_1 E(a_{i-1} Z_{i-k}) - 0_2 E(a_{i-2} Z_{i-k}) \\
&= \begin{cases} \phi \gamma_0 + (0_1(0_1 - \phi) - 0_1) \sigma_a^2 & \text{si } k = 1 \\ \phi \gamma_1 - 0_2 \sigma_a^2 & \text{si } k = 2 \\ \phi \gamma_{k-1} & \text{si } k > 2 \end{cases} \quad (2.50)
\end{aligned}$$

Se puede concluir que para más de dos retrasos las autocovarianzas tienen la misma tendencia que un proceso AR(1), como consecuencia de que la parte de promedios móviles es de orden dos, por lo que únicamente la primera y la segunda autocovarianzas muestran los parámetros del tipo MA.

Al conocer las expresiones para γ_0 y γ_1 , las autocovarianzas restantes se obtienen resolviendo las ecuaciones

$$(1) \quad \gamma_0 = \phi\gamma_1 + [1 + \theta_1^2 + \theta_2^2 + 0_2\phi(0_1 - \phi) - 0_1\phi]\sigma_a^2 \quad (2.51)$$

$$(2) \quad \gamma_1 = \phi\gamma_0 + [0_2(0_1 - \phi) - 0_1]\sigma_a^2$$

para expresar a las autocovarianzas en términos de los parámetros conocidos del modelo, se resuelve el siguiente sistema de ecuaciones

$$(1') \quad \gamma_0 - \phi\gamma_1 = [1 + \theta_1^2 + \theta_2^2 + 0_2\phi(0_1 - \phi) - 0_1\phi]\sigma_a^2$$

$$(2') \quad \phi\gamma_0 - \gamma_1 = [0_2(0_1 - \phi) - 0_1]\sigma_a^2$$

multiplicando a (2') por $-\phi$ y haciendo la suma de (1') + $(-\phi * (2'))$

$$-\phi^2\gamma_0 + \phi\gamma_1 = -\phi[0_2(0_1 - \phi) - 0_1]\sigma_a^2$$

$$(1 - \phi^2)\gamma_0 = [1 + \theta_1^2 + \theta_2^2 + 0_2\phi(0_1 - \phi) - 0_1\phi]\sigma_a^2 + \phi[0_2(0_1 - \phi) - 0_1]\sigma_a^2$$

$$= [1 + \theta_1^2 + \theta_2^2 + 0_2\phi(0_1 - \phi) - 0_1\phi + \phi 0_2(0_1 - \phi) - \phi 0_1]\sigma_a^2$$

$$= [1 + \theta_1^2 + \theta_2^2 + 2\theta_2\phi(0_1 - \phi) - 2\theta_1\phi]\sigma_a^2$$

Por lo tanto,

$$\gamma_0 = \frac{[1 + \theta_1^2 + \theta_2^2 + 2\theta_2\phi(0_1 - \phi) - 2\theta_1\phi]\sigma_a^2}{1 - \phi^2} \quad (2.52)$$

Ahora sustituyendo el valor obtenido de γ_0 en la ecuación (2) de (2.51) se obtiene el valor de γ_1 en términos de parámetros conocidos

$$\begin{aligned} \gamma_1 &= \phi \left[\frac{[1 + \theta_1^2 + \theta_2^2 + 2\theta_2\phi(0_1 - \phi) - 2\theta_1\phi]\sigma_a^2}{1 - \phi^2} \right] + [0_2(0_1 - \phi) - 0_1]\sigma_a^2 \\ &= \phi \left[\frac{[1 + \theta_1^2 + \theta_2^2 + 2\theta_2\phi(0_1 - \phi) - 2\theta_1\phi]\sigma_a^2 + (1 - \phi^2)[0_2(0_1 - \phi) - 0_1]\sigma_a^2}{1 - \phi^2} \right] \\ &= \left[\frac{\phi[1 + \theta_1^2 + \theta_2^2] + 2\theta_2\phi^2(0_1 - \phi) - 2\theta_1\phi^2 + 0_2(0_1 - \phi) - 0_1 - 0_2\phi^2(0_1 - \phi) + \theta_1\phi^2}{1 - \phi^2} \right] \sigma_a^2 \\ &= \left[\frac{\phi[1 + \theta_1^2 + \theta_2^2] + 0_2\phi^2(0_1 - \phi) - 0_1\phi^2 + 0_2(0_1 - \phi) - 0_1}{1 - \phi^2} \right] \sigma_a^2 \\ &= \left[\frac{\phi[1 + \theta_1^2 + \theta_2^2 + 0_2\phi(0_1 - \phi) - 0_1\phi] + 0_2(0_1 - \phi) - 0_1}{1 - \phi^2} \right] \sigma_a^2 \end{aligned}$$

(2.53)

En general,

$$Y_k = \begin{cases} \phi^{k-1} \sigma_a^2 \left[\phi \left[\frac{1 - 2\phi(\theta_1 + (\phi - \theta_1)\theta_2) + \theta_1^2 + \theta_2^2}{1 - \phi^2} \right] - \theta_1 - \theta_2(\phi - \theta_1) \right] \\ \phi^{k-1} \sigma_a^2 \left[\phi \left[\frac{1 - 2\phi(\theta_1 + (\phi - \theta_1)\theta_2) + \theta_1^2 + \theta_2^2}{1 - \phi^2} \right] - \theta_1 - \theta_2 - \theta_2(\phi - \theta_1) \right] \end{cases}$$

2.3.3 Caso ARMA(p,q).

El caso general de un proceso ARMA(p,q) se representa como sigue

$$\phi(B)Z_t = \theta(B)a_t,$$

en donde los polinomios $\phi(B)$ y $\theta(B)$ son de orden p y q respectivamente, o sea se tiene que

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)Z_t &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)a_t \\ Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} &= a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \end{aligned} \quad (2.54)$$

y para ser estacionario es necesario que las raíces de $\phi(x) = 0$ se encuentren fuera del círculo unitario, y se cumpla la misma condición con las raíces de la ecuación $\theta(x) = 0$ para que también sea invertible; si ambas condiciones se cumplen, entonces el proceso admite las representaciones AR y MA,

$$Z_t = \frac{\theta(B)}{\phi(B)} a_t = \psi(B)a_t, \quad \frac{\phi(B)}{\theta(B)} Z_t = \pi(B)Z_t = a_t$$

con $\sum_{i=1}^{\infty} |\psi_i| < \infty$ y $\sum_{j=1}^{\infty} |\pi_j| < \infty$, y los coeficientes ψ_i , $i = 1, 2, \dots$ y π_j , $j = 1, 2, \dots$, se pueden obtener igualando los coeficientes de potencias del operador B .

En cuanto a la media, varianza y autocovarianzas del proceso se utiliza la expresión (2.54) para el cálculo. Así, si el proceso es estacionario, la media es cero; y para $k \geq 0$, se tiene

$$\begin{aligned} Y_k &= E(Z_t Z_{t-k}) \\ &= \phi_1 E(Z_{t-1} Z_{t-k}) + \phi_2 E(Z_{t-2} Z_{t-k}) + \dots + \phi_p E(Z_{t-p} Z_{t-k}) \\ &\quad + E(a_t Z_{t-k}) - \theta_1 E(a_{t-1} Z_{t-k}) - \dots - \theta_q E(a_{t-q} Z_{t-k}) \end{aligned} \quad (2.55)$$

en donde, como Z_{t-k} está afectada por los choques aleatorios $a_{t-k}, a_{t-k-1}, \dots$, pero no de $a_{t-k+1}, a_{t-k+2}, \dots$, es independiente de éstos últimos, entonces

$$E(a_{t-i}, Z_{t-k}) = 0 \quad \text{si } k > i$$

Así de la ecuación (2.55) se desprende el siguiente resultado

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad \text{para } k > q \quad (2.56)$$

y para el caso $k \leq q$, γ_k involucrará, además a los parámetros $\theta_k, \theta_{k+1}, \dots, \theta_q$.

En particular, para la varianza se tendría

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2 - \theta_1 E(a_{t-1}, Z_t) - \dots - \theta_q E(a_{t-q}, Z_t)$$

que involucra a $\gamma_1, \gamma_2, \dots, \gamma_p$, por lo que tiene que ser resuelta simultáneamente con las ecuaciones para estas p autocovarianzas.

Un resultado que vale la pena observar, para hacer uso de él, es la igualdad para cada caso siguiente

$$E(a_{t-i}, Z_t) = E(a_{t-i-1}, Z_{t-1}) = E(a_{t-i-2}, Z_{t-2}) = \dots, \quad \text{para } i = 0, \pm 1, \pm 2, \dots$$

además de que

$$E(a_t, Z_t) = \sigma_a^2$$

y

$$E(a_{t-i}, Z_t) = \phi_1 E(a_{t-i-1}, Z_{t-1}) + \phi_2 E(a_{t-i-2}, Z_{t-2}) + \dots + \phi_p E(a_{t-i-p}, Z_{t-p}) - \theta_i \sigma_a^2,$$

con $i = 1, 2, \dots, \max(p, q)$

donde $\theta_{q+1} = \theta_{q+2} = \dots = \theta_q = 0$ si $q < p$ y $\phi_{p+1} = \phi_{p+2} = \dots = \phi_q = 0$ si $p < q$.

Por último en cuanto a las autocorrelaciones para procesos ARMA(p,q) se tienen los casos:

Para retrasos mayores a q , es decir para $k > q$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > q$$

Para retrasos menores a q , la autocorrelación involucrará a los parámetros $\theta_k, \theta_{k+1}, \dots, \theta_q$. Y la FAC estará dada por la ecuación en diferencia

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \dots - \phi_{p-1} B^{p-1} - \phi_p B^p) \rho_k = 0, \quad k > p \quad (2.57)$$

sujeta a las p condiciones iniciales dictadas por $\rho_1, \rho_2, \dots, \rho_p$; si $p \leq q$, habrá q autocorrelaciones iniciales $\rho_1, \rho_2, \dots, \rho_p$; que no sigan el comportamiento general de la FAC, pero las autocorrelaciones $\rho_{q+1}, \rho_{q+2}, \rho_{q+3}, \dots$ se comportarán de acuerdo con la ecuación en diferencia (2.57) con la condición $k > q$ y las p condiciones iniciales $\rho_{q-p+1}, \rho_{q-p+2}, \dots, \rho_q$.

Hasta aquí se ha trabajado con modelos estacionarios, pero lo más frecuente es encontrarse con series que no son estacionarias, ya sea porque se presenta algún tipo de tendencia o porque su varianza no es constante o porque están influenciadas por algún factor de tipo determinista como lo es la estacionalidad. Si el problema es una tendencia en el comportamiento de la serie y si dicha tendencia es de carácter polinomial, entonces puede ser eliminada mediante la aplicación adecuada del operador diferencia, lo cual da origen a los modelos ARIMA que se verán a continuación. Si además la no-estacionariedad se debe también a que la varianza no es constante, en este caso, un argumento derivado del trabajo de Bartlett (1947), lleva a determinar una transformación potencia del tipo

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases} \quad (2.58)$$

(Ver Apéndice (A.2)). Esta transformación llega a ser útil para estabilizar a la varianza de la serie, antes de cancelar la posible tendencia polinomial adaptiva.

2.4 Generalización de los modelos ARMA, los modelos ARIMA.

Los *modelos autorregresivos e integrados de promedios móviles* (ARIMA) pueden considerarse como una generalización de modelos ARMA, ya estudiados. Como se vio al final de los capítulos I y II la aplicación del operador diferencia ∇^d para eliminar la tendencia polinomial de orden d , presente en la serie que se analiza, permite tener una gran flexibilidad a los modelos ARMA. Cuando una serie es no estacionaria se puede deber a que la media es función de un polinomio de grado d en t , entonces se pueden hacer d diferencias de la serie para convertirla en estacionaria. Si se tiene al proceso original $\{Z_t\}$, con la característica de no estacionariedad causada por una tendencia polinomial, es posible construir un proceso estacionario $\{W_t\}$ como sigue

$$W_t = \nabla^d Z_t \quad \text{para toda } t \quad (2.59)$$



serie para la cual ya es posible obtener un modelo ARMA: $\Phi(B)W_t = \Theta(B)a_t$, que es equivalente a considerar el modelo ARIMA

$$\Phi(B)\nabla^d Z_t = \Theta(B)a_t, \quad d \geq 1 \quad (2.60)$$

Donde $\Phi(B)$ es un polinomio de retraso de grado p , $\Theta(B)$ es un polinomio de promedios móviles de orden q y a_t es un proceso de ruido blanco. A este modelo se le conoce como ARIMA (p,d,q) . Se puede observar que estos modelos son una generalización de los ARMA (p,q) , por lo cual su función de autocorrelación y autocorrelación parcial son las correspondientes a este último modelo.

En el cuadro siguiente se muestran cuáles son las principales características de algunos procesos ARIMA más comunes de encontrar (no se demostrará el proceso para llegar a ellos).

Cuadro 3.
Características de algunos procesos ARIMA

MODELO	FUNCION DE AUTOCORRELACIÓN	FUNCION DE AUTOCORRELACIÓN PARCIAL
(1,d,0)	Decaimiento exponencial	$\phi_{kk} = 0 \quad \forall k > 1$
(2,d,0)	Decaimiento exponencial	$\phi_{kk} = 0 \quad \forall k > 2$
(p,d,0)	Decaimiento exponencial	$\phi_{kk} = 0 \quad \forall k > p$
(0,d,1)	$\rho_k = 0 \quad \forall k > 1$	Decaimiento exponencial
(0,d,2)	$\rho_k = 0 \quad \forall k > 2$	Decaimiento exponencial
(0,d,p)	$\rho_k = 0 \quad \forall k > q$	Decaimiento exponencial
(p,d,q)	Decaimiento exponencial después de $q-p$ retrasos	Decaimiento exponencial después de $q-p$ retrasos

2.4.1 Extensión de modelos ARIMA debido a una tendencia determinista.

Por último, se verá el caso en el que la media del proceso estacionario $\{W_t\}$ es distinta de cero (y no como hasta ahora que se ha considerado igual a cero), es decir, una media constante μ_w del proceso, caso con el cual los modelos ARIMA adquieren una mayor generalidad. Si se considera a

$$\phi(B)(W_t - \mu_w) = \theta(B)a_t$$

entonces la expresión general de un modelo ARIMA quedaría

$$\phi(B)W_t = \theta_0 + \theta(B)a_t \quad (2.61)$$

en donde $\theta_0 = \mu_w(1 - \phi_1 - \dots - \phi_p)$. La idea básica de introducir el parámetro θ_0 es que representa para $\{Z_t\}$ una tendencia polinomial determinista, que no depende de los choques aleatorios $\{a_t\}$ y por lo tanto es predecible; motivo por el cual se considera $\theta_0 = 0$, a menos que el comportamiento del fenómeno en estudio o los mismos datos indiquen lo contrario.

2.4.2 Pronósticos con modelos ARIMA

Sea Z_t una serie estacionaria que admite una representación del tipo ARMA(p,q), es decir

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

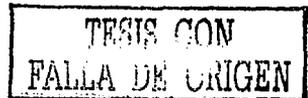
ya que por hipótesis el modelo es estacionario, se puede representar como

$$Z_t = \psi(B)a_t$$

donde

$$\begin{aligned} \psi(B) &= \frac{\theta(B)}{\phi(B)} \\ &= \sum_{j=0}^{\infty} \psi_j B^j \quad \text{con } \psi_0 = 1 \end{aligned}$$

sustituyendo esta última expresión



$$Z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

donde para $t = n + f$

$$Z_{n+f} = \sum_{j=0}^{\infty} \psi_j a_{n+f-j} \quad (2.62)$$

supóngase que en el tiempo $t = n$ se tienen las observaciones $Z_n, Z_{n-1}, Z_{n-2}, \dots$ y se quiere saber el valor de Z_{n+f} a partir de la combinación lineal de las observaciones mencionadas. Entonces

$$Z_n(f) = c_f a_n + c_{f+1} a_{n-1} + c_{f+2} a_{n-2} + \dots \quad (2.63)$$

es un pronóstico para Z_{n+f} .

Dado que existe un gran número de posibles combinaciones de valores $c_f, c_{f+1}, c_{f+2}, \dots$, el criterio que se ocupa para decidir qué valores seleccionar es el error cuadrático medio mínimo. Para la expresión (2.63) el error está dado por

$$E[Z_{n+f} - Z_n(f)]^2$$

Sustituyendo el valor de las expresiones, para desarrollar términos se tiene

$$\begin{aligned} E[Z_{n+f} - Z_n(f)]^2 &= E \left[\sum_{j=0}^{f-1} \psi_j^2 a_{n+f-j}^2 + \sum_{j=f}^{\infty} (\psi_j - c_j)^2 a_{n+f-j}^2 + 2 \sum_{j=0, \dots, j}^{\infty} \psi_j (\psi_f - c_f) a_n a_j \right] \\ &= \sum_{j=0}^{f-1} \psi_j^2 E(a_{n+f-j}^2) + \sum_{j=f}^{\infty} (\psi_j - c_j)^2 E(a_{n+f-j}^2) \\ &= \sigma_a^2 \left[\sum_{j=0}^{f-1} \psi_j^2 + \sum_{j=f}^{\infty} (\psi_j - c_j)^2 \right] \end{aligned}$$

el mínimo de esta expresión se obtiene haciendo $\psi_j = c_j$. Por lo tanto,

$$Z_n(f) = \psi_f a_n + \psi_{f+1} a_{n-1} + \psi_{f+2} a_{n-2} + \dots$$

ya que las observaciones están hechas hasta $t = n$, se tiene

$$E(a_{n+j} / Z_n, Z_{n-1}, Z_{n-2}, \dots) = \begin{cases} 0 & \text{para } j > 0 \\ a_{n+j} & \text{para } j \leq 0 \end{cases}$$

Por otro lado, utilizando la expresión (2.62),

$$E(Z_{n,r}/Z_n, Z_{n-1}, Z_{n-2}, \dots) = E(a_{n,r} + \psi_r a_{n,r-1} + \psi_2 a_{n,r-2} + \dots + \psi_r a_n + \psi_{r+1} a_{n-1} + \dots) \\ = \psi_r a_n + \psi_{r+1} a_{n-1} + \dots$$

De aquí se concluye que el pronóstico que proporciona el mínimo error cuadrático medio para $Z_{n,r}$ está dado por la esperanza condicional.

El error del pronóstico está dado por

$$c_n(r) = Z_{n,r} - \hat{Z}_n(r) \\ = \sum_{j=0}^{\infty} \psi_j a_{n,r-j} - \sum_{j=r}^{\infty} \psi_j a_{n,r-j} \\ = \sum_{j=0}^{r-1} \psi_j a_{n,r-j}$$

Calculando el valor esperado se tiene

$$E[c_n(r)] = E\left[\sum_{j=0}^{r-1} \psi_j a_{n,r-j}\right] \\ = \sum_{j=0}^{r-1} \psi_j E(a_{n,r-j}) \\ = 0 \tag{2.64}$$

Con este último resultado se puede apreciar que los pronósticos son insesgados dado que

$$E(Z_{n,r} - \hat{Z}_n(r)) = 0$$

y además implica que

$$E(Z_{n,r}) = E(\hat{Z}_n(r))$$

La varianza está dada por

$$\text{var}[c_n(r)] = E\left[\sum_{j=0}^{r-1} \psi_j a_{n,r-j}\right]^2 \\ = \sum_{j=0}^{r-1} \psi_j^2 E(a_{n,r-j})^2 + 2 \sum_{j=0}^{r-1} \psi_j \psi_j E(a_j)$$

$$= \sigma_e^2 \sum_{j=0}^{f-1} \psi_j^2 \quad (2.65)$$

Expresión de la cual se puede observar que la varianza del pronóstico va en aumento en la medida que f crece.

Para $f = 1$

$$\begin{aligned} e_n(1) &= Z_{n+1} - Z_n(1) \\ &= a_{n+1} \end{aligned}$$

lo cual muestra que los errores del pronóstico un periodo hacia adelante no están correlacionados.

Una vez dados los pronósticos, es conveniente dar un intervalo que represente la confiabilidad de los mismos. Para ello, bajo el supuesto de que $a_n \sim N(0, \sigma_a^2) \forall n$ y dadas las ecuaciones (2.64) y (2.65), se tiene

$$e_n(f) / \{Z_n, Z_{n-1}, Z_{n-2}, \dots\} \sim N(0, \text{Var}[e(f)])$$

Esta última expresión equivale a escribir

$$[Z_{n+f} - Z_n(f)] / \{Z_n, Z_{n-1}, Z_{n-2}, \dots\} \sim N(0, \text{Var}[e(f)])$$

Por lo cual, un intervalo de confianza al $100(1-\alpha)\%$ para Z_{n+f} , condicionado por el conocimiento de las observaciones $Z_n, Z_{n-1}, Z_{n-2}, \dots$ es

$$Z_n(f) \pm z_{\alpha/2} \left(\sum_{j=0}^{f-1} \psi_j^2 \right)^{1/2} \sigma_a$$

2.5 Modelos estacionales o temporales

Para poder manejar series que presenten *temporalidad E* (lo que significa que existe un patrón en el comportamiento de la serie cada E observaciones) se define el *operador diferencia estacional* como

$$\begin{aligned} \nabla_E Z &= (1 - B^E)Z \\ &= Z_t - Z_{t-E} \end{aligned}$$

Así se tiene que

$$\begin{aligned}\nabla_E^k Z_t &= (1 - B^E)^k Z_t \\ &= \left[\sum_{j=0}^k \frac{k!}{j!(k-j)!} (-B^E)^j \right] Z_t\end{aligned}$$

También se define al *polinomio de retraso estacional* de orden k con coeficientes constantes como

$$\begin{aligned}G(B^E) &= 1 - g_1 B^E - g_2 B^{2E} - \dots - g_k B^{kE} \\ &= 1 - \sum_{j=1}^k g_j B^{jE}\end{aligned}$$

Con ∇_E^k y $G(B^E)^k$ se pretende quitar la estacionalidad de la serie; logrando esto, se puede intentar ajustar un modelo ARIMA $(P, D, Q)_E$ de la forma

$$\Phi(B^E) \nabla_E^D(Z_t) = \Theta(B^E) a_t,$$

donde

- $\Phi(B^E)$ representa a un polinomio autorregresivo estacional de orden P .
- $\Theta(B^E)$ representa a un polinomio de promedios móviles estacional de orden Q .
- $\{a_t\}$ es un proceso de ruido blanco.

La forma de la función de autocorrelación del modelo ARIMA $(P, D, Q)_E$ está muy relacionada con la forma del respectivo modelo ARIMA (p, d, q) .

Para tener más clara esta relación, considere los modelos ARIMA $(0,0,1)$ y ARIMA $(0,0,1)_E$, es decir

$$\begin{aligned}Z_t &= a_t - 0, a_{t-1} \\ W_t &= a_t - \Theta_1 a_{t-E}\end{aligned} \tag{2.66}$$

Las esperanzas de estos modelos son

$$\begin{aligned}E\{Z_t\} &= 0 \\ E\{W_t\} &= 0\end{aligned}$$

al emplear la función de autocorrelación de los modelos MA(q) y dado que $q = 1$ se tiene que $\rho_k = 0$ para todo $k > 1$, y para $k = 1$

$$\begin{aligned}\rho_k &= \frac{\sum_{i=0}^1 \theta_i \theta_{i+k}}{\sum_{i=0}^1 \theta_i^2} \\ &= \frac{-\theta_1}{1 + \theta_1^2}\end{aligned}$$

Si ahora se trabaja con (2.66)

$$\begin{aligned}\text{Var}(W_t) &= E(a_t - \Theta_1 a_{t-E})^2 \\ &= E[a_t^2 - 2\Theta_1 a_t a_{t-E} + \Theta_1^2 (a_{t-E})^2] \\ &= E(a_t)^2 - 2\Theta_1 E(a_t a_{t-E}) + \Theta_1^2 E(a_{t-E})^2\end{aligned}$$

por ser a_t y a_{t-E} independientes, $E(a_t a_{t-E}) = 0$ y por lo tanto $\text{Var}(W_t) = \sigma_a^2(1 + \Theta_1^2)$.

Las autocovarianzas para W_t están dadas por

$$\begin{aligned}\gamma_k &= E[(a_t - \Theta_1 a_{t-E})(a_{t-k} - \Theta_1 a_{t-E-k})] \\ &= E[a_t a_{t-k} - \Theta_1 a_t a_{t-E-k} - \Theta_1 a_{t-E} a_{t-k} + \Theta_1^2 a_{t-E} a_{t-E-k}]\end{aligned}$$

Cuando $k = E$,

$$\begin{aligned}\gamma_k &= -\Theta_1 E(a_{t-E})^2 \\ &= -\Theta_1 \sigma_a^2\end{aligned}$$

Si $k \neq E$, $\gamma_k = 0$. De aquí que la función de autocorrelación este definida como

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{-\Theta_1}{1 + \Theta_1^2} & \text{si } k = E \\ 0 & \text{si } k \neq E \end{cases}$$

De este ejemplo se puede apreciar que las autocorrelaciones de los modelos estacionales son E- veces más lentas en su decaimiento.

2.5.1 Modelo multiplicativo

El *Modelo Multiplicativo Estacional* permite estudiar series que presentan un comportamiento temporal y otro no temporal. Este modelo tiene la forma

$$\Phi(B^E) \nabla_E^D(Z_t) = \Theta(B^E) \alpha_t, \quad (2.67)$$

donde $\{\alpha_t\}$ ya no se considera un ruido blanco, sino un ARIMA(p,d,q), es decir

$$\phi(B) \nabla^d \alpha_t = \theta(B) a_t, \quad (2.68)$$

con $\{a_t\}$ un ruido blanco.

Por (2.67) y (2.68) se obtiene el modelo

$$\phi(B) \Phi(B^E) \nabla^d \nabla_E^D(Z_t) = \theta(B) \Theta(B^E) a_t, \quad (2.69)$$

conocido como ARIMA(p,d,q) × (P,D,Q)_E.

Debido a que tratar de obtener las autocorrelaciones de este modelo a través de las definiciones resulta muy laborioso, normalmente lo que se hace es emplear la función generadora de autocovarianza y recordar que la función generadora de autocovarianzas para un modelo del tipo (2.69) es el producto de las funciones generadoras de autocovarianzas de sus respectivas componentes, Box y Jenkins (1970). Es decir, si $\gamma(B)$ es la función generadora de autocovarianza de (2.68) y $\gamma(B^E)$ es la función generadora de autocovarianza de (2.67), la función generadora de autocovarianza de (2.69) es

$$\gamma(B) \gamma(B^E)$$

Guerrero (1991) desarrolla la forma de las autocorrelaciones para los procesos del tipo multiplicativo más usados en la práctica.

La flexibilidad del modelo multiplicativo radica en que toma en cuenta las relaciones entre observaciones consecutivas dentro del periodo estacional y las relaciones de las observaciones Z_t y Z_{t-1} .

CAPITULO 3

Ajuste de series de tiempo con modelos ARIMA

Teniendo en mente toda la base teórica de series de tiempo es necesario indicar cuales son los pasos a seguir para poder asignar a una serie de datos, el modelo estocástico que mejor la represente.

Box y Jenkins han sugerido cuatro pasos:

1. *Identificación.* Consiste en encontrar los valores p , d , q que especifican al modelo ARIMA.
2. *Estimación.* Se refiere a la estimación de los coeficientes de los polinomios autorregresivos y de promedios móviles que ocupe el modelo.
3. *Verificación.* Hay que revisar que se cumplan los supuestos del modelo.
4. *Uso del Modelo.* Puede ser pronosticar, simular, etc. el fenómeno.

Antes de comenzar un análisis exhaustivo de las características de una serie de tiempo, es conveniente graficarla para tener una visión general de su comportamiento.

La gráfica de una serie de tiempo, puede proporcionar una idea de la presencia de estacionariedad, temporalidad, discontinuidades y "outliers" que presenta.

Para ir ejemplificando cada uno de los pasos de este procedimiento de ajuste se tomará la serie del Costo de Captación de los Pasivos a Plazo denominados en dólares de los EE.UU.A., de aquí en adelante CCP dólares para trabajar con ella y así mostrar el uso de esta herramienta para modelar series.

Dicha serie fue tomada de la Central de Información (CIF) del Banco de México. Las observaciones van de abril de 1996 a febrero de 2003, los datos aparecen en el cuadro 4 y son representados en la gráfica 9.

El Banco de México inicio el cálculo mensual del costo de captación por concepto de tasa de interés de los pasivos a plazo en dólares de los EE.UU.A., incluidos los préstamos de bancos del extranjero, a cargo de instituciones de banca múltiple, excepto aquéllos que se deriven de obligaciones subordinadas susceptibles de convertirse en títulos representativos del capital social de instituciones de crédito, del otorgamiento de avales, de la celebración de operaciones entre instituciones de crédito, así como de los financiamientos recibidos de los Export-Import Banks, de la Commodity Credit Corporation y de otros organismos similares.

Dicho costo de captación a plazo de pasivos denominados en dólares de los EE.UU.A. (CCP-Dólares), podrá ser utilizado como referencia para determinar la tasa de interés de créditos denominados en dólares de los EE.UU.A.¹³

3.1 Identificación

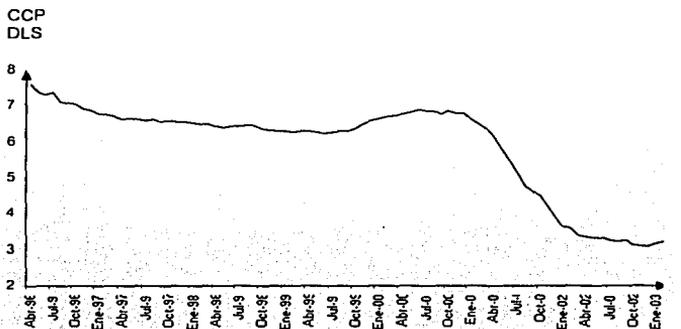
Uno de los de los supuestos de los modelos ARIMA es que las series que se estudian son estacionarias. En la práctica este supuesto no necesariamente se cumple y por ello se hace necesario aplicar transformaciones a los datos que nos permitan tener una serie estacionaria. Las transformaciones que aquí se sugieren están encaminadas a estabilizar la varianza y la media.

Cuadro 4.
Costo de Captación de los Pasivos a Plazo denominados en Dólares de los EE.UU.A (CCP Dólares)
(Tasas Mensuales)

MES	1996	1997	1998	1999	2000	2001	2002	2003
Enero	-	6.77	6.51	6.28	6.60	6.61	3.66	3.17
Febrero	-	6.75	6.47	6.26	6.68	6.49	3.58	3.21
Marzo	-	6.72	6.48	6.30	6.71	6.34	3.37	
Abril	7.59	6.63	6.43	6.28	6.75	6.17	3.34	
Mayo	7.39	6.65	6.40	6.25	6.80	5.82	3.33	
Junio	7.32	6.62	6.41	6.23	6.88	5.50	3.32	
Julio	7.36	6.59	6.42	6.24	6.86	5.16	3.26	
Agosto	7.12	6.63	6.46	6.28	6.83	4.76	3.21	
Septiembre	7.08	6.55	6.44	6.27	6.76	4.61	3.24	
Octubre	7.05	6.58	6.35	6.36	6.84	4.48	3.12	
Noviembre	6.92	6.54	6.31	6.48	6.78	4.17	3.09	
Diciembre	6.88	6.56	6.29	6.57	6.77	3.88	3.10	

¹³ Publicado en el Diario Oficial de la Federación el día 3 de mayo de 1996.

Gráfica 9.
Costo de Captación de los Pasivos a Plazo denominados en Dólares de los EE.UU.A. (CCP-dólares) de Abril de 1998 a Febrero de 2003.



En una primera inspección a la gráfica, ésta presenta una tendencia decreciente a través del tiempo, aunque en los meses de noviembre de 1999 a enero de 2001 se presenta una etapa creciente quizás originado por circunstancias externas, posteriormente vuelve a su mismo comportamiento. Aunque en las dos últimas observaciones se puede observar una tendencia a crecer nuevamente.

3.1.1 Estabilización de la varianza

El elegir una transformación que establezca la varianza tiene su origen en la hipótesis, planteada por Bartlett (1974)¹⁴.

Para encontrar la función T que transforme la escala de la serie se considerara la transformación dada por

¹⁴ "...si la variación tiende a cambiar con el nivel de las mediciones, la varianza solamente será estabilizada con un cambio apropiado de escala".

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases}$$

el desarrollo de cómo llegar a esta transformación la puede consultar el lector en el apéndice del presente trabajo.

Pero determinar cuál es el valor de λ adecuado resulta muy complicado, por lo que se sugiere mejor utilizar una potencia de λ que satisfaga la relación

$$\sigma_t^2 / \mu_t^{1-\lambda} = c \quad \text{para } t = 1, 2, \dots, N, \quad c = \text{constante} \quad (3.1)$$

donde σ_t y μ_t representan a la desviación estándar y a la media de la variable Z_t , y N es el número de observaciones que se tienen para la serie $\{Z_t\}$.

Ahora como para cada t se tiene una observación de Z_t , no es posible obtener una estimación de σ_t para cada t , por lo que se utilizará el siguiente procedimiento para estabilizar a la varianza: dividase a las N observaciones de la serie en H grupos que contengan $R = (N - n) / H$ observaciones contiguas cada uno, no incluyendo a n observaciones ($0 \leq n < R$) del principio o del final de la serie; la única restricción que se pide es que exista homogeneidad entre los grupos y que todos ellos contengan el mismo número de observaciones, para calcular estimaciones de la media y de la desviación estándar dentro de cada grupo, de manera que se tendrán H parejas de valores comparables $\{S_h, Z_h\}$, $h = 1, \dots, H$, con los cuales se puede construir el cuadro siguiente:

En donde, si $Z_{h,r}$ es la r -ésima observación del grupo H , se tiene que

$$Z_h = \sum_{r=1}^R Z_{h,r} / R, \quad S_h = \sqrt{\sum_{r=1}^R (Z_{h,r} - Z_h)^2 / (R - 1)}$$

y el coeficiente de variación se calcula como la desviación estándar $DE(\lambda)$, entre la media $M(\lambda)$, es decir

$$CV(\lambda) = DE(\lambda) / M(\lambda), \quad (3.2)$$

$$M(\lambda) = \sum_{h=1}^H (S_h / Z_h^{1-\lambda}) / H \quad \text{y}$$

$$DE(\lambda) = \sqrt{\sum_{h=1}^H [(S_h / Z_h^{1-\lambda}) - M(\lambda)]^2 / (H - 1)}$$

ahora ya que se busca satisfacer la condición (3.1) con los valores muestrales, deberá elegirse la potencia que proporcione el mínimo coeficiente de variación, pues ese mismo valor satisfará de manera aproximada la relación

$$\frac{S_h}{Z_h^{1-\lambda}} = c, \quad \text{para } h = 1, \dots, H \quad (3.3)$$

Cuadro 4.
Cálculos para Seleccionar la Potencia λ de la Transformación Estabilizadora de Varianza

Grupo	Potencia (λ)				
	-1	-0.5	0	0.5	1
1	S_1 / Z_1^2	$S_1 / Z_1^{1.5}$	S_1 / Z_1	$S_1 / Z_1^{0.5}$	S_1
2	S_2 / Z_2^2	$S_2 / Z_2^{1.5}$	S_2 / Z_2	$S_2 / Z_2^{0.5}$	S_2
...					
h	S_h / Z_h^2	$S_h / Z_h^{1.5}$	S_h / Z_h	$S_h / Z_h^{0.5}$	S_h
...					
H	S_H / Z_H^2	$S_H / Z_H^{1.5}$	S_H / Z_H	$S_H / Z_H^{0.5}$	S_H
Coefficiente de variación	CV(-1)	CV(-0.5)	CV(0)	CV(0.5)	CV(1)

En el ejemplo de este trabajo, la varianza se aprecia constante, sin embargo, se realizará el análisis de varianza aquí mencionado. Para ello se dividieron los datos en cinco grupos con 16 observaciones cada uno de ellos, como se puede apreciar en el cuadro 5, dejando fuera únicamente a dos elementos del total de datos. La media, varianza y desviación estándar sobre cada grupo aparecen en el cuadro 6, los valores de λ escogidos, así como los resultados se presentan en el cuadro 7. De éste último se concluye que no se requiere una función que transforme la varianza, pues el coeficiente de variación más pequeño lo da el valor de $\lambda = 1$, lo cual implica que la serie se queda igual.

Cuadro 5.
Nueva agrupación de la serie el CCP dólares

MES	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
1	7.59	6.63	6.29	6.75	4.76
2	7.39	6.55	6.28	6.80	4.61
3	7.32	6.58	6.26	6.88	4.48
4	7.36	6.54	6.30	6.86	4.17
5	7.12	6.56	6.28	6.83	3.88
6	7.08	6.51	6.25	6.76	3.66
7	7.05	6.47	6.23	6.84	3.58
8	6.92	6.48	6.24	6.78	3.37
9	6.88	6.43	6.28	6.77	3.34
10	6.77	6.40	6.27	6.61	3.33
11	6.75	6.41	6.36	6.49	3.32
12	6.72	6.42	6.48	6.34	3.26
13	6.63	6.46	6.57	6.17	3.21
14	6.65	6.44	6.60	5.82	3.24
15	6.62	6.35	6.68	5.50	3.12
16	6.59	6.31	6.71	5.16	3.09

Cuadro 6.
Estadísticas principales de cada grupo

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
Media	6.97	6.47	6.38	6.46	3.65
Varianza	0.101	0.007	0.028	0.283	0.311
σ	0.32	0.09	0.17	0.53	0.56

Cuadro 7.
Determinación de λ para la serie CCP dólares

Grupos	λ				
	-1	-0.5	0	0.5	1
1	0.0066	0.0173	0.0457	0.1207	0.3186
2	0.0021	0.0053	0.0134	0.0340	0.0866
3	0.0041	0.0104	0.0263	0.0665	0.1681
4	0.0128	0.0324	0.0824	0.2094	0.5322
5	0.0419	0.0800	0.1528	0.2919	0.5577
Media	0.013474	0.029079	0.064123	0.144513	0.332635
Varianza	0.0003	0.0009	0.0031	0.0112	0.0446
σ	0.0164	0.0302	0.0560	0.1058	0.2111
C.V.	1.214718	1.039505	0.873043	0.732206	0.634719

3.1.2 Estabilización del nivel

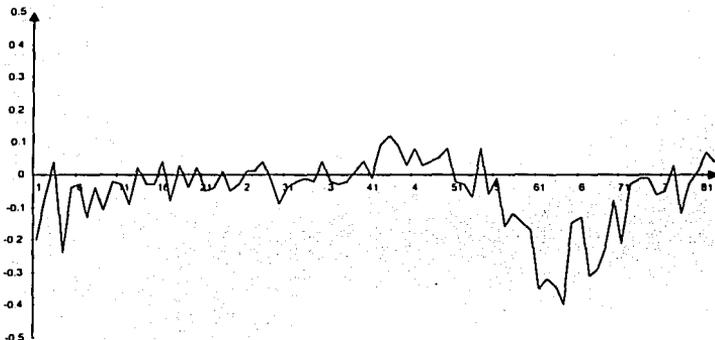
Una vez que se ha logrado conseguir una estabilización en la varianza de la serie, es necesario saber si ésta es estacionaria o no. Para ver si una serie es estacionaria se usa el autocorrelograma y para transformar una serie no estacionaria a estacionaria se emplea el operador diferencia. Se sugiere que el grado de diferenciación se puede estimar calculando la desviación estándar muestral de las series diferenciadas y tomando como valor d de aquélla que induzca el mínimo de estas desviaciones. En la práctica no se llevan a cabo más de 3 diferencias a una serie.

Para el caso de la serie del CCP dólares, para conocer el grado de diferenciación, se ocuparon las gráficas de la primera, segunda y tercera diferencias no estacionales gráficas 10, 11 y 12, respectivamente. Además, se calcularon las principales medidas sobre cada serie para identificar la diferencia adecuada (cuadro 8), según el método de Kendall y Stuart.

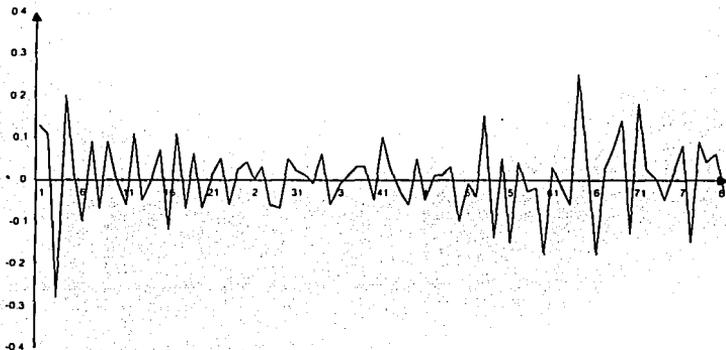
Cuadro 8.
Principales medidas para las diferencias del CCP dólares

	Serie	Primera	Segunda	Tercera
	Original	Diferencia	Diferencia	Diferencia
Media	5.8834	-0.0534	0.0030	-0.0020
Varianza	1.7861	0.0118	0.0079	0.0226
σ	1.3365	0.1087	0.0887	0.1505

Gráfica 10.
Primera diferencia del CCP dólares



Gráfica 11.
Segunda diferencia del CCP dólares



TESIS CON
FALLA DE ORIGEN

Gráfica 12
Tercera diferencia del CCP dólares



Se puede apreciar que al aplicar una diferencia (gráfica 10), persiste una ligera tendencia en los últimos términos, la cual es eliminada en la segunda y tercera diferencias (gráficas 11 y 12, respectivamente).

Debido a que se busca llevar la tendencia de la serie a una tendencia estacional y no es recomendable sobrediferenciar la serie, se tomará la segunda diferencia como la óptima, según el análisis visual.

Esta opción se reafirma si se toma en consideración el método de Kendall y Stuart (1968), pues si se observa el cuadro 8, el aplicar una diferencia no estacional de orden 2 nos da la mínima varianza que una diferencia de orden 1 o 3.

Una vez que se logró volver estacionaria a la serie, el paso siguiente es encontrar el modelo que mejor se le ajuste para ello se hará uso de la Función de Autocorrelación y Función de Autocorrelación Parcial.

3.1.3 Función de Autocorrelación y Función de Autocorrelación Parcial

Si se trabaja con W_t como la serie original y que denota a $\nabla^d T(Z_t)$ y además tiene media y varianza ya estabilizada, el siguiente paso sería



1. Calcular las primeras k autocorrelaciones muestrales y autocorrelaciones parciales muestrales de W_t . Las primeras 24 autocorrelaciones explican en buena medida el comportamiento de la función en general.
2. Graficar las autocorrelaciones y autocorrelaciones parciales
3. Asociar el comportamiento de los valores muestrales a las funciones de autocorrelación teórica y autocorrelación parcial teórica de un proceso ARMA (p, q)

Ahora, para continuar hay que establecer cuáles autocorrelaciones son significativamente diferentes de cero.

Anderson (1942) demostró que bajo la hipótesis de que las autocorrelaciones teóricas de ρ_k sean cero, las estimaciones r_k , divididas entre su desviación estándar, se distribuyen aproximadamente una función Normal. Por tal motivo, los valores $\pm 1.96 \times$ (desviación estándar) constituyen los límites del intervalo que con un 95% de confianza, aseguran que las autocorrelaciones son cero.

Bartlett (1946) encontró que la varianza de las autocorrelaciones muestrales de un proceso generado por un ruido blanco con distribución Normal está dada por

$$\text{Var}(r_k) \cong \frac{1}{N} \sum_{v=-\infty}^{\infty} (\rho_v^2 + \rho_{v+k} \rho_{v-k} - 4\rho_k \rho_v \rho_{v-k} + 2\rho_k^2 \rho_v^2) \quad (3.4)$$

Para un modelo del tipo MA (q) , en el que $\rho_k = 0$ para toda k mayor a q , se tiene que para $v \geq q+1$ la expresión anterior se modifica a

$$\text{Var}(r_k) \cong \frac{1}{N} \sum_{v=-q}^q (\rho_v^2 + \rho_{v+k} \rho_{v-k} - 4\rho_k \rho_v \rho_{v-k} + 2\rho_k^2 \rho_v^2)$$

$$\text{Var}(r_k) \cong \frac{1}{N} \sum_{v=-q}^q \rho_v^2$$

Utilizando los resultados $\rho_0 = 1$ y $\rho_k = \rho_{-k}$ la forma final de la varianza es

$$\text{Var}(r_k) \cong \frac{1}{N} \left(1 + 2 \sum_{v=1}^q \rho_v^2 \right) \quad k \geq q$$

En la práctica se sustituyen las autocorrelaciones teóricas por las muestrales, es decir,

$$\sigma(r_k) \cong \frac{1}{N^2} (1 + 2(r_1^2 + r_2^2 + \dots + r_q^2))^{1/2} \quad k \geq q$$

Por otro lado, Quenouille (1949) demostró que suponiendo que el proceso es autorregresivo de orden p , las autocorrelaciones parciales estimadas de orden mayor o igual a $p+1$ se distribuyen independientes con varianza

$$\text{Var}(\hat{\phi}_{kk}) \cong \frac{1}{N} \quad k \geq p+1$$

Por lo tanto, su error estándar está dado por

$$\hat{\phi}_{kk} \cong \frac{1}{\sqrt{N}} \quad k \geq p$$

De lo anterior se desprende que

1. Las autocorrelaciones serán significativamente diferentes de cero si

$$|r_k| > 1.96 \sqrt{\frac{1}{N} \left(1 + 2 \sum_{v=1}^q r_v^2 \right)}$$

2. Las autocorrelaciones parciales serán significativamente diferentes de cero si

$$|\hat{\phi}_{kk}| > 1.96 \left(\frac{1}{N} \right)^{1/2}$$

Para el ejercicio de la serie del CCP dólares, ya que se logró volver estacionaria a la serie, lo siguiente es encontrar el modelo que mejor se le ajuste. Para ello, se calcularon la función de autocorrelación y la función de autocorrelación parcial de la serie dos veces diferenciada (Cuadros 9 y 10, respectivamente). Del cuadro 9 se puede ver que $r_k = 0$ para $k > 1$ y puesto que ninguna autocorrelación muestral con retraso $k \geq 2$ es cero se propone un modelo IMA(2,1) e IMA(1,1). Por otro lado, de las autocorrelaciones parciales se observa que $\phi_{ii} = 0$ para $i > 2$, esto ya que aquellas cuyo valor absoluto sea menor 0.217777 son consideradas cero.

De las gráficas 13 y 14 se puede observar que la función de autocorrelación y la función de autocorrelación parcial de la serie presentan un tipo de estacionalidad cada seis meses aproximadamente, por lo cual habrá de considerarse un parámetro autorregresivo estacional Q con frecuencia de aparición semestral. De la gráfica 13 sólo la primera autocorrelación es significativamente distinta de cero y en la gráfica 14 las dos primeras autocorrelaciones parciales cumplen lo mismo.

Una vez analizada la información relacionada con las funciones de autocorrelación y autocorrelación parcial se proponen 3 modelos para representar a la serie del CCP dólares:

- Modelo 1. ARIMA(2,2,1)x(0,1,1)₆.
- Modelo 2. ARIMA(2,1,1)x(0,1,1)₆ y
- Modelo 3. ARIMA(1,1,1)x(0,1,0)₆

Cuadro 9.
Autocorrelaciones estimadas para la serie del CCP dólares diferenciada

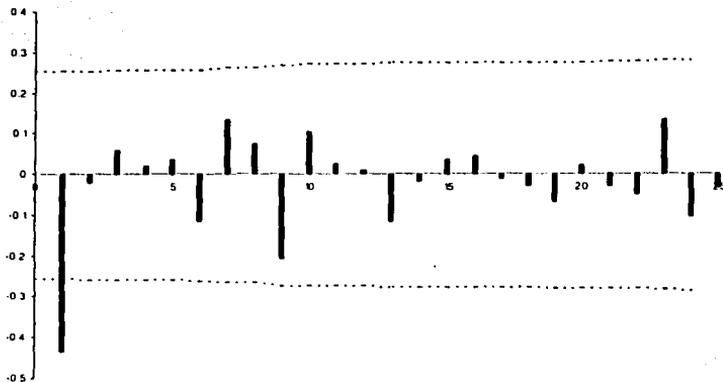
Retraso	Autocorr.	σ	Retraso	Autocorr.	σ
1	-0.43499	0.25569	13	-0.11589	0.27547
2	-0.02117	0.25577	14	-0.01883	0.27553
3	0.05949	0.25642	15	0.03623	0.27575
4	0.02050	0.25650	16	0.04433	0.27609
5	0.03650	0.25675	17	-0.01067	0.27611
6	-0.11595	0.25922	18	-0.02969	0.27626
7	0.13280	0.26243	19	-0.06831	0.27706
8	0.07503	0.26344	20	0.02069	0.27714
9	-0.20764	0.27109	21	-0.02790	0.27727
10	0.10503	0.27302	22	-0.04902	0.27768
11	0.02535	0.27313	23	0.13439	0.28075
12	0.00999	0.27315	24	-0.10333	0.28255

Cuadro 10.
Autocorrelaciones parciales para la serie del CCP dólares diferenciada

Retraso	Autocorr.	Retraso	Autocorr.
1	-0.43499	13	-0.09327
2	-0.25949	14	-0.15188
3	-0.08594	15	-0.16041
4	0.01405	16	0.02739
5	0.09409	17	0.11513
6	-0.05887	18	0.01998
7	0.06975	19	-0.12832
8	0.19159	20	-0.08701
9	-0.07420	21	-0.04503
10	-0.02528	22	-0.12442
11	0.01556	23	0.04970
12	0.04120	24	-0.02822

TESIS CON
 FALLA DE ORIGEN

Gráfica 13.



Gráfica 14.



TESIS CON
FALLA DE ORIGEN

3.2 Estimación

Una vez determinado el grado de p y q se requiere conocer los valores de los parámetros de los polinomios autorregresivo y de promedios móviles para que el modelo (3.5) quede completo.

$$\phi(B)\nabla^d T(Z_t) = \mu + 0(B)a_t \quad (3.5)$$

A continuación se explicarán tres maneras diferentes de estimar dichos parámetros, así como sus ventajas y desventajas.

Para el análisis de la serie del CCP dólares, en las etapas de estimación, verificación y pronóstico de la serie se utilizará el paquete Statistica, esto debido a la complejidad y cantidad de cálculos que se tendrían que llevar a cabo para analizar cada uno de los procesos propuestos para modelar a nuestra serie en cuestión. Primero se presenta la teoría para mostrar un poco los fundamentos sobre los cuales están contruidos los resultados obtenidos del paquete estadístico, posteriormente se muestran dichos resultados.

3.2.1 Método de momentos

Consiste en sustituir los momentos muestrales (media, varianza y función de autocorrelación) por los teóricos.

Cuando el modelo es del tipo

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t$$

entonces, la media $\mu = E(Z_t)$ puede ser estimada a través de \tilde{Z}_t , y al momento de sustituir las autocorrelaciones teóricas por las muestrales en las ecuaciones de Yule Walker se obtiene

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned}$$

De este sistema se puede despejar a $\phi_1, \phi_2, \dots, \phi_p$ y conseguir los estimadores $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p$.

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_2 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}$$

Utilizando la varianza de un AR(p) y del hecho de que $\rho_k = \frac{\gamma_k}{\gamma_0}$ se sigue que

$$\sigma_a^2 = \gamma_0 (1 - \hat{\phi}_1 \rho_1 - \hat{\phi}_2 \rho_2 - \dots - \hat{\phi}_p \rho_p)$$

Ahora, esto no resulta muy sencillo de aplicar cuando se tiene un proceso que tiene la forma de un MA o un ARMA. Considérese el proceso

$$Z = a_t - a_{t-1}$$

de lo visto en la teoría de los modelos de promedios móviles se sabe que

$$\rho_1 = \frac{-0}{1+0^2}$$

Cuando se sustituye ρ_1 por ρ_1 , resulta la ecuación

$$0^2 + \frac{0}{\rho_1} + 1 = 0$$

Resolviendo para 0

$$0 = \frac{-1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1}$$

Cuando $\rho_1 = \pm 0.5$, la solución es $\hat{\theta}_1 = \pm 1$ la cual conduce a un proceso no invertible; si $|\rho_1| > 0.5$ la solución no es real por lo tanto $|\rho_1| < 0.5$ da los valores adecuados para ρ_1 , sólo falta analizar cual de las dos raíces conlleva a un proceso invertible.

Si para $q=1$ el análisis resultó laborioso, para casos de $q > 1$ el análisis se complica demasiado y lo que comúnmente se hace es recurrir al método de Máxima Verosimilitud.

Máxima Verosimilitud

Máxima verosimilitud condicional

Si se considera al proceso (3.5) en su forma

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

o lo que es equivalente

$$a_t = \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} - \mu + W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p}, \quad (3.6)$$

donde $W_t = \nabla^d Z_t$ y $\{a_t\}$ es ruido blanco con distribución $N(0, \sigma_a^2)$.

Dado que $\bar{\phi} = (\phi_1, \dots, \phi_p)$, μ , $\bar{\theta} = (\theta_1, \dots, \theta_q)$ y σ_a^2 son parámetros desconocidos la función de verosimilitud de ellos, sería

$$L.(\bar{\phi}, \mu, \bar{\theta}, \sigma_a^2 / a) = (2\pi\sigma_a^2)^{-\frac{N}{2}} \exp\left[-\frac{1}{2\sigma_a^2} \sum_{i=1}^N a_i^2\right] \quad (3.7)$$

Para el caso $t = 1$

$$a_1 = \theta_1 a_0 + \dots + \theta_q a_{1-q} - \mu + W_1 - \phi_1 W_0 - \dots - \phi_p W_{1-p}$$

y se necesitan dos condiciones iniciales $W_0 = (W_{-p}, \dots, W_{-1}, W_0)$ y $a_0 = (a_{1-q}, \dots, a_{-1}, a_0)$ para poder evaluar a (3.7) para este caso.

Si se define

$$S.(\bar{\phi}, \mu, \bar{\theta}) = \sum_{i=1}^N a_i^2(\bar{\phi}, \mu, \bar{\theta} / W_0, a_0, W), \quad (3.8)$$

y se obtiene el algoritmo de la verosimilitud, entonces

$$\ln L.(\bar{\phi}, \mu, \bar{\theta}, \sigma_a^2) = -\frac{N}{2} \ln 2\pi\sigma_a^2 - \frac{S.(\bar{\phi}, \mu, \bar{\theta})}{2\sigma_a^2}. \quad (3.9)$$

Se nota que $S.(\bar{\phi}, \mu, \bar{\theta})$ es la suma de cuadrados condicionada a W_0, a_0 y W , por esto a los estimadores máximo verosímiles $\hat{\phi}, \hat{\mu}, \hat{\theta}$ se les conoce como los estimadores de máxima verosimilitud condicionales. Debido a que $L.(\bar{\phi}, \mu, \bar{\theta}, \sigma_a^2 / a)$

involucra a $\bar{\phi}, \mu, \bar{\theta}$ sólo a través de $S(\bar{\phi}, \mu, \bar{\theta})$, $\hat{\phi}, \hat{\mu}$ y $\hat{\theta}$ son los estimadores de mínimos cuadrados condicionales.

Para σ_a^2 el estimador de máxima verosimilitud está dado por

$$\sigma_a^2 = \frac{S(\hat{\phi}, \hat{\mu}, \hat{\theta})}{N}$$

Debido a que existen pocas formas de determinar los valores de a_t y Z_t , para $t \leq 0$, normalmente se hace uso de los supuestos de estacionariedad de $\{Z_t\}$ y de la distribución como proceso de ruido blanco $N(0, \sigma_a^2)$ de $\{a_t\}$ para sustituir por ceros a a_t y por Z_t a Z_t .

Máxima verosimilitud incondicional

Box y Jenkins (1979) sugirieron la función de máxima verosimilitud incondicional con el fin de obtener una mejor aproximación en la estimación, la cual aplicando logaritmo tiene la forma

$$\ln L(\bar{\phi}, \mu, \bar{\theta}, \sigma_a^2) = -\frac{N}{2} \ln 2\pi\sigma_a^2 - \frac{S(\bar{\phi}, \mu, \bar{\theta})}{2\sigma_a^2} \quad (3.10)$$

donde,

$$S(\bar{\phi}, \mu, \bar{\theta}) = \sum_{i=-\infty}^N [E(a_t | (\bar{\phi}, \mu, \bar{\theta}, W))]^2 \quad (3.11)$$

Estas dos expresiones no están condicionadas, es decir, dan por hecho que se tiene conocimiento de los valores a_{q-1}, \dots, a_0 y W_{p-1}, \dots, W_0 . Box y Jenkins parten del hecho de que es posible pronosticar valores futuros del proceso y consecuentemente también debe ser posible pronosticar valores anteriores del mismo.

Si en el modelo

$$(1 - \phi_1 B - \dots - \phi_p B^p) W_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (3.12)$$

se reemplaza el operador de retraso B por el adelanto F , donde $F^j W_t = W_{t+j}$, entonces W_t puede escribirse como

$$(1 - \phi_1 F - \dots - \phi_p F^p) W_t = (1 - \theta_1 F - \dots - \theta_q F^q) \epsilon_t \quad (3.13)$$

donde ϵ_t es un proceso de ruido blanco.

Esta forma permite encontrar los valores $W_{t-p}, \dots, W_{t-1}, W_t$ a través de las observaciones conocidas W_1, \dots, W_N .

En la práctica (3.11) se aproxima a través de

$$S(\bar{\phi}, \mu, \bar{\theta}) = \sum_{t=-M}^N [E(a_t | (\bar{\phi}, \mu, \bar{\theta}, W))]^2 \quad (3.14)$$

Donde el valor de M es entero y tal que

$$|E(W_t | \bar{\phi}, \mu, \bar{\theta}, W) - E(W_{t-1} | \bar{\phi}, \mu, \bar{\theta}, W)| < \epsilon \quad \forall t \leq -(M+1).$$

es decir, el número de pronósticos anteriores a considerar se detiene cuando $E(a_t | (\bar{\phi}, \mu, \bar{\theta}, Z))$ es "insignificante".

Hasta ahora, ya se ha utilizado la verosimilitud condicional y la incondicional, se ha superado el problema de los valores iniciales sin considerar la existencia de otro inconveniente. Para obtener los estimadores máximo verosímiles se debe derivar la función de verosimilitud, encontrar los puntos críticos y evaluar la derivada en estos puntos para demostrar que existe efectivamente un máximo. Para obtener el máximo de (3.7) o de (3.10), se requiere minimizar ya sea (3.8) o (3.11); debido a que las parciales de estas últimas con respecto a alguno de los parámetros resulta ser no lineal, sólo se puede encontrar aproximaciones de $\bar{\phi}, \mu, \bar{\theta}$, asignando valores a estos y graficando la función para encontrar el máximo.

3.3 Verificación

En esta etapa el objetivo principal es revisar que no existan violaciones a los supuestos del modelo.

Para verificar los supuestos se emplearán los residuos. Un residuo a_t , se define como la diferencia entre W_t y el pronóstico \hat{W}_t . Es decir, $a_t = W_t - \hat{W}_t$ es el residuo al tiempo t. Cuando el tamaño de la muestra es grande, los errores aleatorios $\{a_t\}$ y los residuos $\{a_t\}$, son casi iguales, es por ello que lo que se verifica es que las características de $\{a_t\}$ las cumpla a_t .

Con base a la teoría expuesta en el capítulo anterior, los errores aleatorios $\{a_t\}$, deberían ser una realización de un proceso de ruido blanco. Por ello, la verificación del modelo consistirá principalmente en comprobar si esto se cumple.

Adicionalmente, la teoría de series de tiempo para efectuar pronósticos supone que los errores $\{a_t\}$ tienen distribución normal para toda t . Para verificar este supuesto, a través de los residuos, existen tres caminos:

1. Graficar el Histograma de los residuos y por inspección revisar que exista simetría en los mismos.
2. Graficar los residuos contra el tiempo y verificar que el 95% de las observaciones se encuentran en el intervalo $(-1.96\sigma_a, 1.96\sigma_a)$.
3. Elaborar la gráfica de papel normal, que tiene su origen al graficar los cuantiles observados contra los cuantiles de la distribución en hipótesis. Si a_1, a_2, \dots, a_N , son los N residuos (ordenados por el tiempo) y $a_{(1)}, a_{(2)}, \dots, a_{(N)}$ son los residuos ordenados de menor a

mayor, se dirá que $a_{(i)}$ es el $\frac{i-0.5}{N}$ cuantil observado. Por otro lado, si $F(i)$ es la distribución acumulada "supuesta", el cuantil p de $F(i)$ es el valor a para el cual se ha acumulado p de probabilidad.

Si los residuos provienen de una distribución normal, la gráfica de probabilidad normal en una escala adecuada, es una línea recta, y que debido a fluctuaciones aleatorias puede tener observaciones que se acumulen alrededor de la línea.

Ya que se puede dar por hecho que la distribución de los residuos es normal, es conveniente revisar que cumpla con las características de un ruido blanco, es decir, que su media sea cero y su varianza constante. Para probar la hipótesis $H_0: \mu = 0$ vs $H_1: \mu \neq 0$, se puede utilizar el hecho de que

$$\left(m(\hat{a}) - 1.96\sqrt{\frac{\sigma_a^2}{N}}, m(\hat{a}) + 1.96\sqrt{\frac{\sigma_a^2}{N}} \right)$$

donde $N' = N - d - p$, constituye un intervalo de confianza al 95% de la media de una normal.

Se rechazará la hipótesis nula si el cero no está dentro de este intervalo o bien, rechaza la hipótesis nula si $\left| \frac{\sqrt{N'} m(\hat{a})}{\hat{\sigma}_a} \right| \geq 1.96$, donde

$$m(\hat{a}) = \sum_{t=d+p+1}^N \frac{a_t}{N'} \quad \text{y}$$

$$\hat{\sigma}_a = \sqrt{\frac{\sum_{t=d+p+1}^N [a_t - m(\hat{a})]^2}{N'}}$$

Por el contrario, si $\left| \frac{\sqrt{N'} m(\hat{a})}{\hat{\sigma}_a} \right| < 1.96$ se dirá que la media de los residuos es cero.

Si al concluir la prueba se obtiene que la media de los residuos no es cero, puede ser debido a dos causas:

1. Existe una parte semideterminista en la serie que podría corregirse con una diferencia más (o agregando un parámetro autorregresivo)
2. El problema es determinista y se corrige añadiendo una constante μ al modelo que se calcula conjuntamente con los otros parámetros y cuyo valor inicial es $m(\hat{a})$

Una forma de revisar que la varianza de los residuos es constante consiste en graficar los residuos contra el tiempo y observar si la varianza lleva un patrón creciente, decreciente o totalmente aleatorio.

Otro camino, es utilizar la teoría desarrollada por Barlett (1946) de la sección (3.1.1) de este capítulo para estabilizar la varianza. Si el resultado del análisis es que los residuos no requieren transformación, se podrá decir que la varianza es constante.

Cuando se describió el proceso de ruido blanco se dió la forma que sigue su función de autocorrelación, es por ello que hay que verificar que las autocorrelaciones de los residuos para $k > 1$ sean cero. Para ello hay que proceder como sigue:

Calcular la Función de Autocorrelación de los residuos, que en caso de tener media cero está dada por

$$r_k(a) = \frac{\sum_{t=c}^{N-k} a_t a_{t+k}}{\sum_{t=c}^N a_t^2}$$

donde $c = d + p + 1$.

Anteriormente se habla obtenido que

$$\text{Var}(r_k) \cong \frac{1}{N} \left\{ 1 + 2 \sum_{v=1}^q \rho_v^2 \right\} \quad k \geq q$$

Dado que para un proceso de ruido blanco $\rho_k = 0$ para $k > 1$ las autocorrelaciones serán significativamente diferentes de cero si $|r_k(a)| \geq \frac{1.96}{\sqrt{N-d-p}}$.

Box y Ljung (1978) encontraron que una prueba para verificar que los errores son ruido blanco, puede hacerse a través del estadístico Q' dado por

$$Q' = (N-d-p)(N-d-p+2) \sum_{k=1}^K \frac{r_k^2(a)}{(N-d-p-k)}$$

Si $K > 20$, $Q' \sim X_{K-p-q}^2$. Por lo tanto, se dirá que el modelo es inadecuado si $Q' > X_{K-p-q}^2$.

Cuando las autocorrelaciones no sean las de un ruido blanco, hay que revisar las FAC de los residuos para tratar de identificar algún patrón conocido e incorporar los parámetros que el patrón determine al modelo original.

Otra forma de tener más elementos que faciliten la elección de un modelo, es revisar que sus parámetros no estén de más, es decir, que todos sean necesarios para explicarlo. Si $\hat{\theta}$ es el parámetro a verificar, el intervalo de confianza es

$$\left[\hat{\theta} - 1.96 \sqrt{\widehat{\text{Var}}(\hat{\theta})}, \hat{\theta} + 1.96 \sqrt{\widehat{\text{Var}}(\hat{\theta})} \right]$$

Si el valor creído se encuentra dentro de este intervalo es posible que el parámetro respectivo sea redundante. Puede suceder, por conocimiento del fenómeno en estudio, que aún cuando el valor de un parámetro sea estadísticamente cero, no se elimine por considerarse necesario.

Una vez presentadas las bases sobre las cuales están construidas las herramientas para llevar a cabo un análisis rápido y confiable, se pasa a la etapa de selección del proceso que represente mejor a la serie en cuestión. Se mostrará la estimación de los parámetros de cada proceso las faltas que tenga cada uno de ellos descartando a dos de ellos y dejar al que mejor represente a la serie del CCP dólares.

**Cuadro 11.
Estimación de los parámetros del modelo 1.**

Modelo 1				Intervalo de confianza al 95%	
Parámetros	Estimación	Desv. Est.	p	Izquierdo	Derecho
p(1)	-0.448599	0.262525	0.091859	-0.972059	0.074861
p(2)	-0.264772	0.170941	0.125849	-0.605618	0.076074
q(1)	0.224558	0.265890	0.401199	-0.305612	0.754728
Q(1)	0.048776	0.166976	0.771051	-0.284164	0.381716

**Cuadro 12.
Estimación de los parámetros del modelo 2.**

Modelo 2				Intervalo de confianza al 95%	
Parámetros	Estimación	Desv. Est.	p	Izquierdo	Derecho
p(1)	0.680759	0.213482	0.002115	0.255191	1.106327
p(2)	0.140937	0.154084	0.363415	-0.166224	0.448098
q(1)	0.419798	0.190041	0.030358	0.040959	0.798637
Q(1)	0.057695	0.167482	0.731484	-0.276173	0.391564

**Cuadro 13.
Estimación de los parámetros del modelo 3.**

Modelo 3				Intervalo de confianza al 95%	
Parámetros	Estimación	Desv. Est.	p	Izquierdo	Derecho
p(1)	0.853394	0.086985	0.000000	0.680072	1.026716
q(1)	0.521459	0.122191	0.000058	0.277987	0.764931

En los cuadros 11 a 13 aparecen los parámetros estimados para cada uno de los tres modelos propuestos, respectivamente, así como los intervalos de confianza, desviación estándar y probabilidad de error de cada uno de ellos. Lo que se busca es un modelo que tenga parámetros altamente significativos con desviación estándar y probabilidad mínimos.

Así de los cuadros 11, 12 y 13 se puede observar lo siguiente. Para el modelo 1 (cuadro 11), los cuatro parámetros son muy pequeños y su intervalo de confianza contiene al cero, tienen desviación estándar y probabilidad de error mayor que la de los parámetros de los modelos 2 y 3, por lo tanto, se descarta al modelo 1 y se prefieren a los modelos 2 y 3. Para el modelo 2 (cuadro 12) los parámetros para $p(2)$ y $Q(1)$ siguen siendo muy pequeños y los únicos que figuran dentro del modelo son $p(1)$ y $q(1)$. El modelo 3 (cuadro 13) que considera los parámetros significativos del modelo 2, uno autorregresivo y uno de promedios móviles, respectivamente, son significativamente distintos de cero, tienen desviación mucho menor que los parámetros de los dos modelos anteriores y la probabilidad de error es casi cero. Por lo tanto, se considera al modelo 3 como el óptimo que proporciona menores errores estadísticamente.

Lo anterior se refuerza con las pruebas de normalidad mostradas en las gráficas 15 a 20, de las cuales se puede observar que la correspondiente al modelo 3 es la que mejor normaliza los residuos, pues el histograma es el que mejor simetría tiene a pesar de que en ninguno de los tres casos existen observaciones fuera del intervalo $(-1.96\sigma_s^2, 1.96\sigma_s^2)$. La gráfica de papel correspondiente al modelo 3 es también la que más se asemeja a una línea recta.

De las funciones de autocorrelación y autocorrelación parcial de los residuos para los tres modelos (gráficas 21 a 26), se observa que los valores de ellas pueden considerarse estadísticamente cero, esto es, se pueden considerar los de un ruido blanco. Así que no existe algún inconveniente en seguir considerando al modelo 3 como el que mejor explique a la serie del CCP.

Así para este ejemplo en particular se obtiene el modelo que mejor representa a la serie del CCP dólares es

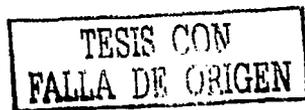
Modelo 3. ARIMA(1,1,1)x(0,1,0)₆, es decir,

$$W_t = a_t - 0a_{t-1} + \phi W_{t-1}$$

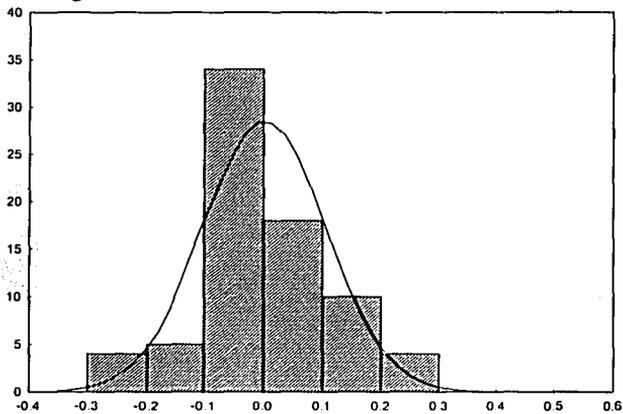
Con $W_t = \nabla^d \nabla Z_t$, y $Z_t =$ Serie original del CCP dólares.

Sustituyendo los valores de los parámetros estimados,

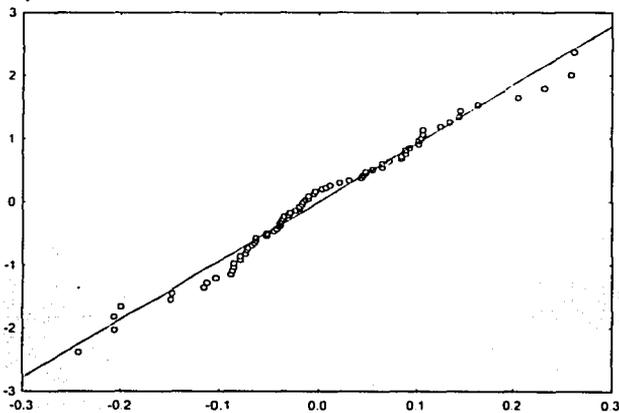
$$W_t = a_t - (0.52146)a_{t-1} + (0.85339)W_{t-1}$$



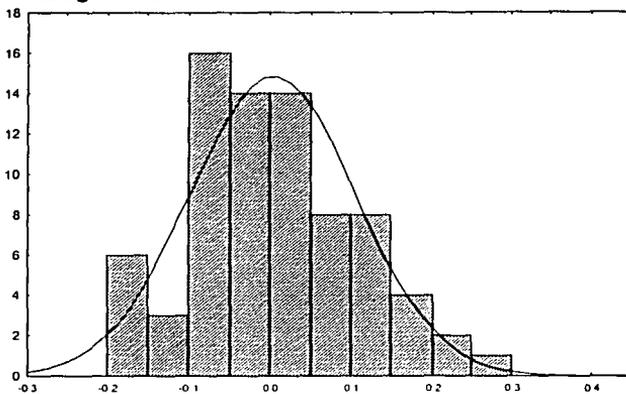
Gráfica 15.
Histograma de los residuos del modelo 1 del CCP dólares



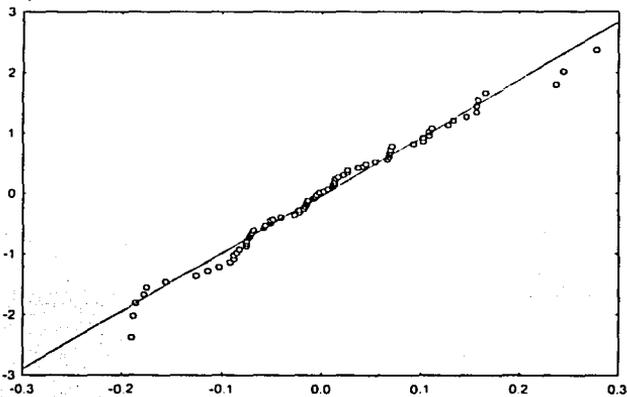
Gráfica 16.
Gráfica de probabilidad normal de los residuos del modelo 1 del CCP dólares



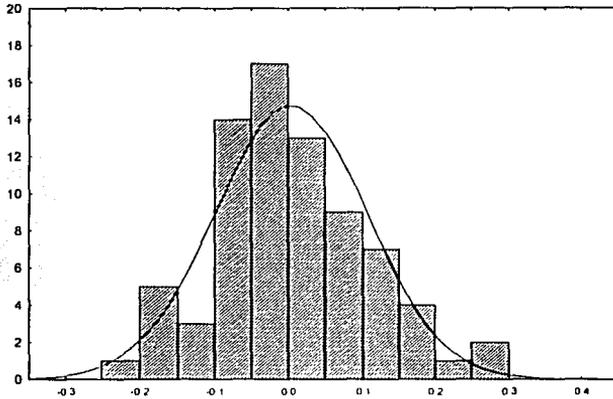
Gráfica 17.
Histograma de los residuos del modelo 2 del CCP dólares



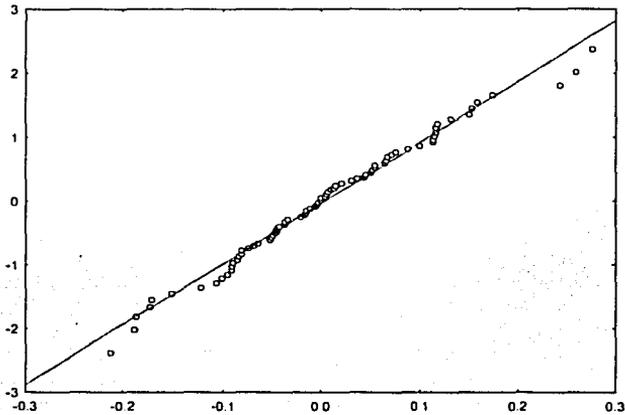
Gráfica 18.
Gráfica de probabilidad normal de los residuos del modelo 2 del CCP dólares



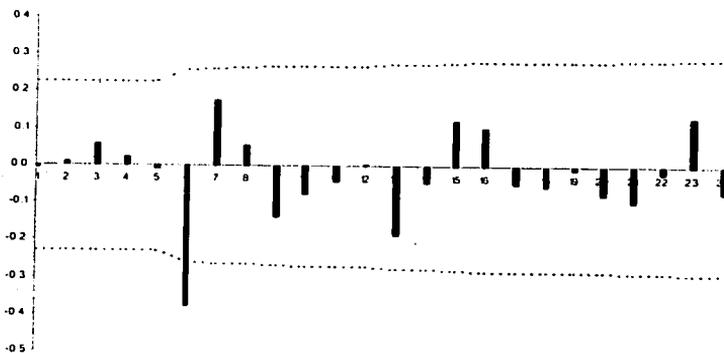
Gráfica 19.
Histograma de los residuos del modelo 3 del CCP dólares



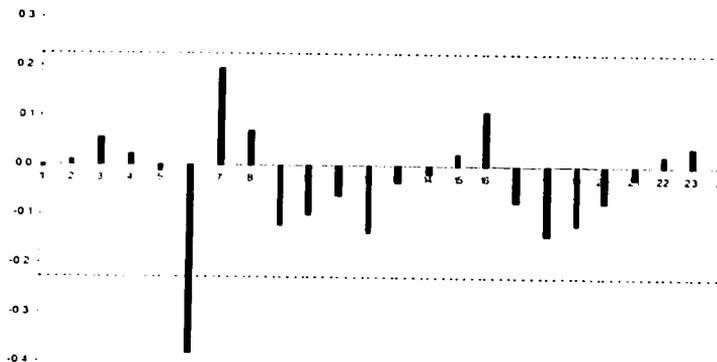
Gráfica 20.
Gráfica de probabilidad normal de los residuos del modelo 1 del CCP dólares



Gráfica 21.
Autocorrelaciones muestrales de los residuos del modelo 1 del CCP dólares

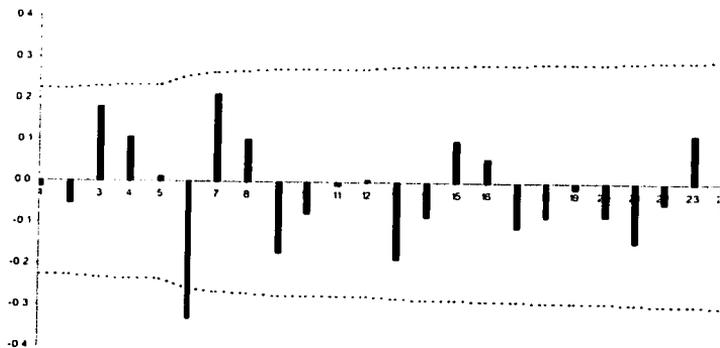


Gráfica 22.
Autocorrelaciones parciales muestrales de los residuos del modelo 1 del CCP dólares

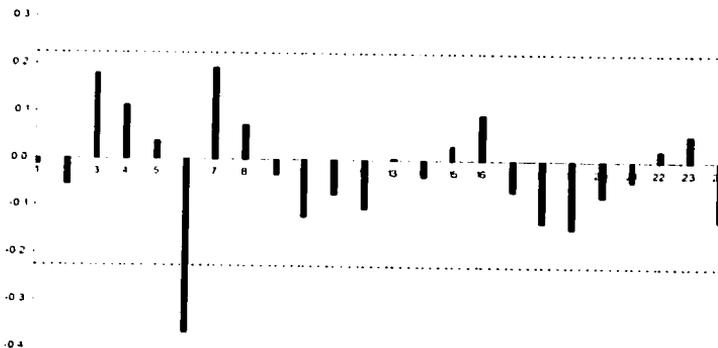


TESIS CON
 FALLA DE CIRCUN

Gráfica 23.
Autocorrelaciones muestrales de los residuos del modelo 2 del CCP dólares

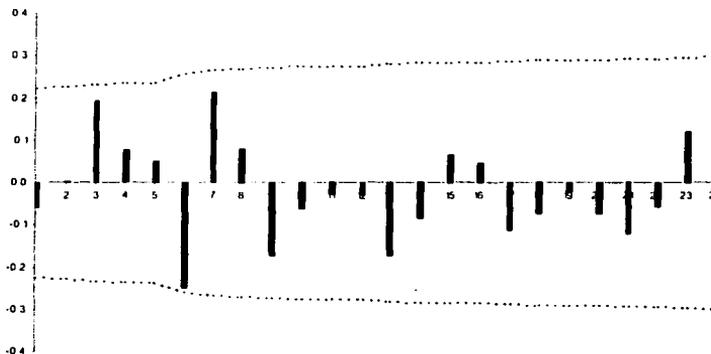


Gráfica 24.
Autocorrelaciones parciales muestrales de los residuos del modelo 2 del CCP dólares

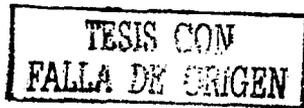
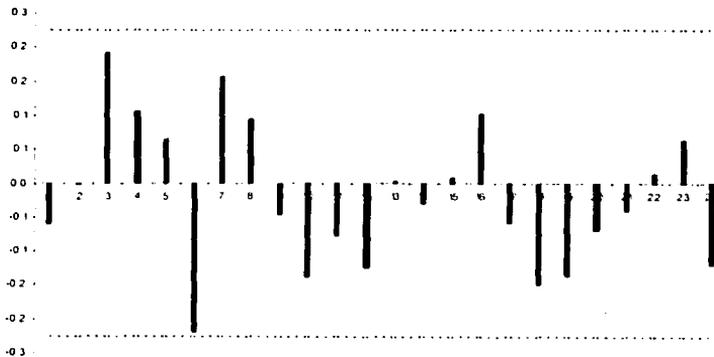


TESIS CON
 FALLA DE ORIGEN

Gráfica 25.
Autocorrelaciones muestrales de los residuos del modelo 3 del CCP dólares



Gráfica 26
Autocorrelaciones parciales muestrales de los residuos del modelo 3 del CCP dólares



En el cuadro 16 se pueden observar los valores pronosticados por el modelo sugerido.

Cuadro 16.
Pronósticos del modelo 3

Observación	Valores originales	Pronósticos	Residuales
80	3.09	3.14	-0.05
81	3.10	3.16	-0.06
82	3.17	3.13	0.04
83	3.21	3.10	0.11
84	3.12	3.15	-0.03
85	2.91	3.04	-0.13
86		3.08	
87		3.11	
88		3.08	
89		3.06	

Conclusiones

Los pronósticos obtenidos por medio de series de tiempo son instrumentos fundamentales para la toma de decisiones, pero para llegar a ellos es indispensable conocer la teoría de las series de tiempo, pues se puede caer en el error de obtener modelos mal fundamentados que a la larga originan malas decisiones. También se debe tener en mente que para modelar series de tiempo se requiere de experiencia y conocimiento de lo que se está modelando pues a pesar de que la teoría parece abarcar lo suficiente, siempre existen pequeñas aberraciones como por ejemplo el que no se pueda estabilizar la varianza de los datos, que los residuales no sean ruido blanco, por mencionar algunos problemas y que si el lector se adentra en este trabajo de modelar series encontrará muchos más.

La serie del ejemplo aquí planteado dice a grandes rasgos en que nivel se encuentra el fondeo en dólares de los bancos en el país, es decir, cuanto le está costando al sistema de banca múltiple obtener sus recursos en dólares. Esto da una idea para pronosticar en tiempos venideros el comportamiento de los recursos de los bancos en dólares, pues si las tasas en dólares se mantienen bajas y como son tasas referenciadas a las que se utilizan en países hasta cierto punto con estabilidad económica los bancos preferirán obtener sus recursos en dólares en lugar de pesos mexicanos si las tasas para moneda nacional tienden a subir o viceversa.

Según el análisis realizado, un modelo adecuado para el CCP dólares es el $ARIMA(1,1,1) \times (0,1,0)_6$. Las características de este modelo son

- La periodicidad es semestral
- La tendencia es de orden 1
- Se ocupan un componente de promedios móviles y uno autorregresivo
- Los residuos son ruido blanco con distribución normal.

Por último, cabe mencionar que no en todos los casos se encuentra un modelo ARIMA para modelar una serie, pues existe una gran variedad de alteraciones que no permiten la completa representación como lo puede ser el no poder estabilizar la varianza o bien el que los residuos no sean una realización de ruido blanco.

APÉNDICE

A.1. La d -ésima diferencia de una serie que sigue una tendencia polinomial de orden d , es estacionaria.

La demostración de este resultado consta de dos partes; en la primera se supone que dado $\{a_t\}$ un proceso de ruido blanco entonces, la serie $\{\nabla^d a_t\}$, con $d \geq 0$ es estacionaria; en la segunda se supone que la serie $\{Z_t\}$ que sigue una tendencia polinomial de orden $d \geq 0$ del tipo

$$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_d t^d + a_t$$

en donde $\{a_t\}$ un proceso de ruido blanco y $\beta_0, \beta_1, \dots, \beta_d$ son constantes, y se demostrara la igualdad $\nabla^d Z_t = d! \beta_d + \nabla^d a_t$. Así, al unir estos dos resultados se concluirá que $\{\nabla^d Z_t\}$ es una serie estacionaria.

Parte 1:

Si $\{a_t\}$ es ruido blanco cuya varianza esta dada por $\text{Var}(a_t) = \sigma_a^2$, y se define a la variable $Y_t = \nabla^d a_t$, entonces se tiene lo siguiente

$$Y_t = \nabla^d a_t = \sum_{j=0}^d \binom{d}{j} (-1)^j a_{t-j}$$

con

$$E(Y_t) = \sum_{j=0}^d \binom{d}{j} (-1)^j E(a_{t-j}) = 0$$

$$\text{pues } \sum_{j=0}^d \binom{d}{j} (-1)^j = (1-1)^d$$

por otro lado,

$$\begin{aligned} \text{Cov}(Y_t, Y_{t+k}) &= E \left\{ \left[\sum_{j=0}^d \binom{d}{j} (-1)^j a_{t-j} \right] \left[\sum_{i=0}^d \binom{d}{i} (-1)^i a_{t+k-i} \right] \right\} \\ &= E \left[\sum_{j=0}^d \sum_{i=0}^d \binom{d}{j} \binom{d}{i} (-1)^{j+i} a_{t-j} a_{t+k-i} \right] \end{aligned}$$

de aquí se observan dos casos:

cuando $j \neq i - k$

$$\sum_{j=0}^d \sum_{i=0}^d \binom{d}{j} \binom{d}{i} (-1)^{j+i} = 0, \quad \text{pues } \sum_{j=0}^d \binom{d}{j} (-1)^j \cdot (1-1)^d = 0$$

y cuando $j = i - k$

$$\sum_{i=k}^d \binom{d}{i-k} \binom{d}{i} (-1)^{2i-k} = \binom{2d}{d-k} (-1)^k$$

Por lo tanto, la covarianza esta dada por la expresión

$$\text{Cov}(Y_i, Y_{i+k}) = \begin{cases} \sum_{j=0}^d \sum_{i=0}^d \binom{d}{j} \binom{d}{i} (-1)^{j+i} E(a_{i-j} a_{i+k-i}) + \\ j \neq i - k \\ \sum_{i=k}^d \binom{d}{i-k} \binom{d}{i} (-1)^{2i-k} E(a_{i+k-i}^2) & \text{si } k \leq d \\ \sum_{j=0}^d \sum_{i=0}^d \binom{d}{j} \binom{d}{i} (-1)^{j+i} E(a_{i-j} a_{i+k-i}) + 0 & \text{si } k > d \\ j \neq i - k \end{cases}$$

$$= \begin{cases} \binom{2d}{d-k} (-1)^k \sigma_a^2 & \text{si } k \leq d \\ 0 & \text{si } k > d \end{cases}$$

con $k=0$ se obtiene la varianza, dada por $\text{Var}(Y_i) = \binom{2d}{d} \sigma_a^2$; así, se puede apreciar que ni la media, ni la varianza, ni la covarianza de $Y_i = \nabla^d a_i$ dependen del tiempo.

Parte 2:

Primero se demostrara que al tomar una primera diferencia se reduce en un grado la tendencia polinomial, si se usa la expresión $\gamma_i^{(0)} = \beta_i$, $i=0,1,\dots,d$ en donde, el supraíndice sirve para denotar el grado de la diferenciación del proceso, entonces originalmente se tendría

$$Z_i = \sum_{l=0}^d \gamma_i^{(0)} t^l + a_i$$

luego, al aplicar una diferencia se obtiene

$$\begin{aligned} \nabla Z_i &= \sum_{l=0}^d \gamma_i^{(0)} t^l + a_i - \sum_{l=0}^d \gamma_i^{(0)} (t-1) - a_{i-1} \\ &= \sum_{l=0}^d \gamma_i^{(0)} \left[t^l - \sum_{j=0}^l \binom{l}{j} (-1)^j t^{l-j} \right] + \nabla a_i \\ &= \sum_{l=1}^d \gamma_i^{(0)} \left[- \sum_{j=1}^l \binom{l-1}{j-1} (-1)^j t^{l-j} \right] + \nabla a_i, \quad \text{utilizando que } \binom{l}{j} = \frac{l}{j} \binom{l-1}{j-1} \end{aligned}$$

haciendo el cambio de variable $l = j-1$, se tiene que

$$= \sum_{l=1}^d \gamma_i^{(0)} \left[\sum_{r=0}^{l-1} \frac{l}{l+1} \binom{l-1}{r} (-1)^l t^{l-1-r} \right] + \nabla a_i$$

sustituyendo por $r = i-1$

$$\begin{aligned} &= \sum_{r=0}^{d-1} \sum_{l=0}^r \gamma_i^{(0)} \frac{r+1}{l+1} \binom{r}{l} (-1)^l t^{r-1-l} + \nabla a_i \\ &= \sum_{l=0}^{d-1} \gamma_i^{(1)} t^l + \nabla a_i \end{aligned}$$

De las últimas dos expresiones se puede observar que $\gamma_0^{(1)}, \dots, \gamma_{d-1}^{(1)}$ dependen de $\gamma_1^{(0)}, \dots, \gamma_d^{(0)}$ y que para obtener los nuevos coeficientes se igualan uno a uno con las potencias de t , es decir, $\gamma_0^{(1)}$ es el coeficiente de t^0 y se obtiene cuando $l=r$, o sea

$$\gamma_0^{(1)} = \gamma_1^{(0)} - \gamma_2^{(0)} + \gamma_3^{(0)} - \dots + (-1)^{d-1} \gamma_d^{(0)}$$

y así sucesivamente para los demás coeficientes de $\gamma_i^{(1)}$, hasta llegar al último coeficiente $\gamma_{d-1}^{(1)}$ que corresponde a t^{d-1} y cuando $l = r - d + 1$, es decir,

$$\gamma_{d-1}^{(1)} = d\gamma_d^{(0)}$$

Al aplicar una vez más el argumento anterior, se tiene

$$\begin{aligned} \nabla^2 Z_t &= \sum_{i=0}^{d-1} \gamma_i^{(1)} t^i + \nabla a_t - \sum_{i=0}^{d-1} \gamma_i^{(1)} (t-1)^i - \nabla a_{t-1} \\ &= \sum_{i=0}^{d-2} \gamma_i^{(2)} t^i + \nabla^2 a_t \end{aligned}$$

y si se aplica $d-1$ veces

$$\nabla^{d-1} Z_t = \sum_{i=0}^1 \gamma_i^{(d-1)} t^i + \nabla^{d-1} a_t$$

por lo que al diferenciar una vez más se tiene

$$\begin{aligned} \nabla^d Z_t &= \sum_{i=0}^1 \gamma_i^{(d-1)} t^i + \nabla^{d-1} a_t - \sum_{i=0}^1 \gamma_i^{(d-1)} (t-1)^i - \nabla^{d-1} a_{t-1} \\ &= \gamma_0^{(d)} + \nabla^d a_t \end{aligned}$$

y para obtener los valores de los nuevos coeficientes se igualan término a término las potencias de t , y da por resultado

$$\begin{aligned} \gamma_0^{(d)} &= \gamma_1^{(d-1)} \\ &= 2\gamma_2^{(d-2)} \\ &= 2 \cdot 3 \cdot \gamma_3^{(d-3)} \\ &\dots \\ &= 2 \cdot 3 \cdot \dots \cdot (d-1) \gamma_{d-1}^{(1)} \\ &= d! \gamma_d^{(0)} \end{aligned}$$

Por lo tanto, $\{\nabla^d Z_t\}$ es estacionaria.

A.2 Desarrollo del procedimiento para elegir una transformación potencia que estabilice a la varianza de una serie de tiempo.

Si se parte de la idea planteada por Bartlett (1947) "... si la variación tiende a cambiar con el nivel medio de las mediciones, la varianza solamente será estabilizada con un cambio apropiado de escala", se tiene el siguiente procedimiento.

Suponiendo que la varianza σ_i^2 de la variable aleatoria Z_i puede expresarse en función de su media μ_i , es decir,

$$\sigma_i^2 = f(\mu_i)$$

luego, si se considera a $T(\cdot)$ una función cuya primera derivada existe, entonces se puede utilizar la aproximación lineal a $T(Z_i)$, dada por

$$T(Z_i) = T(\mu_i) + \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) (Z_i - \mu_i)$$

que tiene como aproximación lineal a la esperanza y varianza dadas por

$$\begin{aligned} E(T(Z_i)) &= E \left[T(\mu_i) + \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) (Z_i - \mu_i) \right] \\ &= T(\mu_i) + \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) [E(Z_i) - \mu_i] \end{aligned}$$

$\text{Var}(T(Z_i)) = E \left\{ [T(Z_i) - E(T(Z_i))]^2 \right\}$ sustituyendo valores,

$$\begin{aligned} &= E \left\{ \left[T(\mu_i) + \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) (Z_i - \mu_i) - T(\mu_i) - \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) (E(Z_i) - \mu_i) \right]^2 \right\} \\ &= E \left\{ \left[\left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right) (Z_i - E(Z_i)) \right]^2 \right\} \\ &= \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right)^2 E \left[(Z_i - E(Z_i))^2 \right] \\ &= \left(\frac{dT}{dZ_i} \Big|_{Z_i = \mu_i} \right)^2 \sigma_i^2 \end{aligned}$$

y ya que se tiene el objetivo de que $T(Z_1)$ tenga una varianza constante, por decir a^2 , entonces se sigue que

$$\frac{dT}{dZ_1} \Big|_{Z_1 = \mu_1} = \frac{a}{\sqrt{f(\mu_1)}}$$

Luego, integrando de ambos lados,

$$T(\mu_1) = \int \frac{a}{\sqrt{f(\mu_1)}} d\mu_1 \quad (\text{A.2.1})$$

Como se puede observar es necesario conocer la función f para poder utilizar (A.2.1) y así determinar la transformación T que establezca la varianza. Si se restringe el tipo de transformaciones a la familia de transformaciones potencia, la cual es utilizada debido a los buenos resultados que se observan en la práctica, es factible entonces aplicar esta restricción a (A.2.1). Ahora si la variable Z_1 es positiva y si es posible suponer que σ_1^2 es proporcional a $\mu_1^{2(1-\lambda)}$ para alguna λ ; es decir,

$$f(\mu_1) \approx \mu_1^{2(1-\lambda)}$$

entonces se tendría para (A.2.1) que

si $\lambda \neq 0$

$$\int \frac{1}{\sqrt{\mu_1^{2(1-\lambda)}}} d\mu_1 = \int \mu_1^{\lambda-1} d\mu_1 = c\mu_1^\lambda$$

si $\lambda = 0$

$$\int \frac{1}{\mu_1} d\mu_1 = c \ln(\mu_1)$$

Por lo tanto,

$$T(\mu_1) \approx \begin{cases} \mu_1^\lambda & \text{si } \lambda \neq 0 \\ \ln(\mu_1) & \text{si } \lambda = 0 \end{cases}$$

Resultado que sugiere que la función T , que vuelve aproximadamente constante a la varianza de Z_1 , es la transformación potencia definida por

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \ln(Z_t) & \text{si } \lambda = 0 \end{cases}$$

válida si $Z_t > 0$ para toda t .

BIBLIOGRAFÍA

- Análisis estadístico de series de tiempo económicas.
Victor M. Guerrero
Universidad Autónoma Metropolitana, 1991.
- Time Series Models.
Andrew C. Harvey
The MIT Press, 1994.
Second Edition.
- Ecuaciones en Diferencia con Aplicaciones.
Takehito Takahashi
Grupo Editorial Iberoamerica, 1990.
- Cálculo infinitesimal.
Michael Spivak
Editorial Reverté, 1981.
- Álgebra lineal y teoría de matrices.
R. Barbolla, P. Sanz
Prentice Hall, 1998.
- Introduction to the Theory of Statistics.
M. Mood y A. Graybill
Mc. Graw Hill, 1992.
- Álgebra Lineal.
Kenneth Hoffman, Ray Kunze
Prentice Hall Hispanoamericana, 1994.
- Introduction to Stochastic Processes.
P. S. Hoel, S.C. Port y C.J. Stone
Houghton-Miffling.

- **Análisis del comportamiento de los índices de precios considerados como series cronológicas**
Tesis de Albino Rodríguez Araiza
UNAM, 1980
- **Aplicación del análisis de intervención de Box y Tiao a la serie de tiempo del INPC en México.**
Tesis de María Odette Barrón Rivero
UNAM, 1986
- **Pronósticos para series de tiempo con modelos ARIMA.**
Tesis Claudia Tapia Rangel
UNAM, 1995.