

13  
2ej



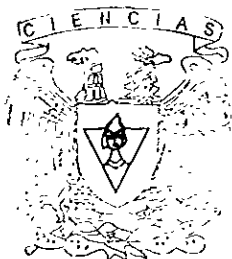
# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

## FACULTAD DE CIENCIAS

Series de tiempo ARIMA.  
Aplicación a dos series  
económicas relevantes.

T E S I S  
Que para obtener el titulo de  
A C T U A R I O  
p r e s e n t a

LETICIA GARCIA VERGARA



FACULTAD DE CIENCIAS  
UNAM

Director de Tesis:  
Dr. Joaquín Curiel Cañedo

1999  
FACULTAD DE CIENCIAS  
SECCION DE REGISTRO

256



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.



Mi primer agradecimiento es para mi director de tesis, el Dr Joaquín Curiel y para los profesores Alejandro Hazas, Sergio Hugo Delgado, Juan León Montañez y Hugo Villaseñor, quienes ocuparon parte de su tiempo en revisar el presente texto.

Agradezco a mis padres y hermanos por el apoyo que me brindaron durante la elaboración de este trabajo.

Mi mayor agradecimiento a Víctor Hugo, por su apoyo incondicional, sus observaciones y por todo lo demás.

# Contenido

Prólogo.....	I
1 Series de tiempo: Introducción .....	1
1.1 Definición y utilidad .....	1
1.2 Una perspectiva histórica de las series temporales ..	4
1.3 Bases del enfoque de Box y Jenkins .....	8
2 Modelos para series de tiempo.....	15
2.1 Modelos autorregresivos .....	15
2.2 Modelos de promedios móviles .....	20
2.3 Modelos autorregresivos y de promedios móviles .....	22
2.4 Modelos autorregresivos e integrados de promedios móviles.....	23
3 Construyendo el modelo.....	26
3.1 Identificar el modelo.....	26
3.2 Estimación.....	35
3.3 Verificación: ¿Es adecuado el modelo seleccionado? .....	47
4 Aplicación del modelo en pronósticos .....	57
4.1 Pronósticos óptimos para series estacionarias.....	57
4.2 Pronósticos óptimos para series transformadas.....	63
4.3 Intervalos de confianza para los pronósticos.....	66
4.4 Para actualizar los pronósticos .....	67
4.5 Capacidad de pronóstico del modelo .....	68
5 Un caso especial: Series de tiempo estacionales.....	71
5.1 Operadores y modelos estacionales .....	71

5.2	Modelo multiplicativo estacional.....	72
5.3	Identificación de procesos estacionales ....	76
5.4	Construcción de modelos estacionales.....	78
6	Serie de tiempo: Ingreso federal .....	80
6.1	Identificación.....	82
6.2	Estimación.....	88
6.3	Verificación.....	88
6.4	Pronósticos .....	102
7	Serie de tiempo: Exportaciones petroleras .....	110
7.1	Identificación.....	111
7.2	Estimación.....	116
7.3	Verificación.....	122
7.4	Pronósticos .....	137
	Conclusiones .....	152
	Tabla A.....	155
	Tabla B.....	160
	Tabla C.....	165
	Bibliografía.....	167

## Prólogo

La apertura de las fronteras comerciales de México en 1994 ha traído consigo la necesidad de que las empresas, tanto públicas como privadas, optimicen sus procesos productivos y administrativos con la finalidad de ser más competitivos a nivel internacional

La planeación es una de las actividades de mayor importancia para lograr esa optimización y dentro de ésta uno de los pasos más importantes es la predicción de valores para las variables que influyen directamente en el desarrollo de la corporación. De la precisión de las predicciones logradas en esta etapa dependerá en gran medida el éxito que pueda obtenerse en el seguimiento de los planes.

Así, con la idea sobre la necesidad de predicción de valores para variables (sobre todo de carácter económico) decidí elaborar mi tesis acerca de un método de predicción. Encontré que algunos autores señalaban el fundamental papel que la teoría de análisis de series de tiempo bajo el enfoque estocástico ha venido desempeñando en tiempos recientes y del importante desarrollo que se continúa haciendo para mejorarla, esto fue lo que me hizo decidirme a enfocar mi trabajo sobre esta metodología de la que al principio tenía sólo algunas ideas. Hoy puedo decir que he ampliado un poco mi conocimiento sobre el tema y este trabajo es el resultado de mi investigación.

Otro punto importante que debo mencionar en esta sección introductoria, es que la teoría de series de tiempo se basa en un supuesto de dependencia entre observaciones que impide la aplicación directa a las series de tiempo del análisis de regresión. Esta indicación es muy significativa ya que es bastante lógico pensar que

en la vida real mucha de la información que se recopila y que se requiere analizar presenta dependencia en sus observaciones. El tratamiento a esta dependencia es uno de los puntos que no se revisa en cursos obligatorios de la carrera de Actuaría y por eso también fue interesante para mí conocer las bases de la teoría de series de tiempo.

Existen dos enfoques básicos para el análisis de series de tiempo: el clásico y el estocástico. Como señalé líneas arriba, me centraré en el enfoque estocástico para el cual se han desarrollado diversas metodologías respecto a las series según el tipo al que correspondan; así existe una teoría para las multivariadas y otra para las univariadas (más adelante se especifican las características de éstas). La teoría que presento en este trabajo versa sobre series univariadas bajo una representación lineal y se desarrolla de manera elemental, sin dar una amplia demostración de todos los resultados ya que para ello cito las obras a las que se puede consultar para ampliar la información.

El enfoque estocástico para análisis de series de tiempo presentado inicialmente por Box y Jenkins propone la construcción de modelos para la serie en estudio, de tal forma que los datos mismos sean quienes sugieran el modelo más adecuado para su representación. Las etapas de las que consta el análisis de series de tiempo son cuatro: identificación del modelo, estimación de los parámetros, verificación del modelo y uso del mismo.

La organización con que cuenta este escrito es la siguiente: en el primer capítulo, presento varios resultados y supuestos que son indispensables para el desarrollo de la teoría de series de tiempo según el enfoque de Box y Jenkins; en el segundo, realizo la presentación de los diferentes modelos con los que se cuenta

para representar a la serie de tiempo que se esté analizando, en el tercero, explico brevemente los pasos que deben seguirse para especificar el modelo que mejor describa a la serie, en el cuarto presento la metodología para realizar pronósticos mediante el modelo construido, en el quinto menciono la teoría correspondiente a las series temporales que presentan un comportamiento estacional y, finalmente, en el sexto y séptimo aplico la teoría mostrada en los cinco capítulos anteriores para modelar el comportamiento del ingreso federal y las exportaciones petroleras respectivamente.

La selección de las dos series señaladas se debió principalmente a la trascendencia que han tenido y tienen para el desarrollo económico sustentado de nuestro México. Todos sabemos de la trascendencia que tiene en nuestro país el ingreso federal ya que esos recursos le permiten a las entidades gubernamentales atender las necesidades de sus habitantes. Por otro lado, la importancia de las exportaciones petroleras radica en el hecho de que éstas han representado desde hace mucho tiempo una fuente de ingresos muy importante para nuestro país, tanto así que en los últimos años los recursos obtenidos por la exportación de petróleo han representado aproximadamente 40% del ingreso federal.



# 1 Series de tiempo: Introducción

En este capítulo versaré sobre algunos temas que servirán de base para conocer lo que son las series de tiempo, en qué campos del conocimiento es posible encontrarlas y cómo ha ido evolucionando su análisis pues, como se comprenderá, para cada tipo de serie existe una metodología que permite describirla.

En la primera sección presentaré una definición del término *serie de tiempo* y un esbozo de los campos en los que estas series pueden ser de gran utilidad; la segunda estará formada básicamente por referencias a autores que han contribuido al desarrollo de la teoría analítica de las series de tiempo, pero sin hacer una revisión de las metodologías propuestas por ellos. Finalmente, en la última sección mencionaré algunos resultados útiles y supuestos esenciales para el análisis de series de tiempo según el enfoque moderno

## 1.1 Definición y utilidad

Lo primero que se debe conocer para adentrarse en el estudio de las series de tiempo es precisamente su definición, así que daré a continuación una definición basada en la lectura de las referencias bibliográficas:

*Se dice que una serie de tiempo es una secuencia de mediciones u observaciones numéricas ordenadas que se realizan metódicamente sobre una variable de interés particular. Dicha ordenación se realiza a través del tiempo (aunque el análisis también puede desarrollarse en el dominio de la frecuencia) y particularmente en términos de intervalos de igual longitud, aunque esto no es indispensable.*

A partir de esta definición, se sigue que la ordenación que el tiempo induce en los datos es esencial para el análisis de la serie, ya que justamente esa evolución en el tiempo es la que determinará el modelo que la describe. Se considera entonces que una serie de tiempo es una función bidimensional en donde la variable dependiente es la magnitud que se quiere analizar y la independiente es el tiempo.

En la práctica se pueden encontrar distintos tipos de series de tiempo: estacionarias, no estacionarias, estacionales y no estacionales, de las que más adelante presentaré sus características. También existen series multivariadas y univariadas, así como lineales y no lineales. Las series multivariadas son aquellas cuyas observaciones dependen o son el resultado de combinar los valores tomados por otras series temporales; las univariadas son las que sus valores no dependen de otras series si no sólo del tiempo. En cuanto a las series lineales, puedo decir que son aquellas en cuya representación no se requieren variables dependientes del tiempo elevadas a una potencia mayor a 1, en tanto que las series no lineales son las que requieren variables dependientes del tiempo elevadas a una potencia mayor a 1. Cuando las series de tiempo pueden ser registradas continuamente en el tiempo se las llama *series de tiempo continuas* y a las que sólo toman valores en intervalos de tiempo específicos se las llama *series de tiempo discretas*.

Las series que analizaré en este trabajo son del tipo continuo, con observaciones mensuales (intervalos de tiempo "iguales", aunque se sabe que no todos los meses tienen el mismo número de días) y la teoría considerada será la aplicable a series univariadas lineales.

Existen muchos objetivos que hacen importante el estudio de las también llamadas series temporales, cronológicas o históricas; a continuación señalo los principales:

- a) El entendimiento y descripción del mecanismo generador de la serie.
- b) El pronóstico óptimo de valores futuros de la serie.
- c) El control adecuado de un sistema a fin de minimizar los errores en variables dependientes de la variable estudiada.

El uso más común de una serie de tiempo es la predicción de valores, tanto futuros como pasados; entre los cuales los primeros son los más requeridos, ya que las decisiones para planeación en casi cualquier campo se basan en lo que se espera ocurrirá en el futuro. En este sentido, las series de tiempo son de gran importancia en el marco económico debido a que mucha información económica sobre las naciones, las empresas y los individuos que es registrada numéricamente debe ser analizada para poder llevar a cabo la planeación del crecimiento y desarrollo de esas entidades.

La aplicación de la metodología para el análisis de las series de tiempo se da en muy diversas áreas, los siguientes son algunos ejemplos:

- a) Planeación de las finanzas y de la economía
- b) Planeación de la producción
- c) Control de los inventarios
- d) Control de la producción
- e) Estudios psico-sociológicos

Las series temporales discretas son generadas en la práctica mediante dos formas básicas: por muestreo de una serie de tiempo continua o por acumulación de una serie de tiempo, ya sea continua o discreta, sobre un periodo de tiempo dado.

A continuación presento sucintamente el desarrollo de la teoría para análisis de series históricas.

## **1.2 Una perspectiva histórica de las series temporales**

Aunque es posible descubrir gráficas de series de tiempo anteriores al siglo XI, el primer diagrama reconocido como tal parece ser atribuido a Playfair y data de 1821 (en A Letter on our Agricultural Distress, Londres).

Se sabe también que durante el siglo XIX se desarrollaron diferentes métodos para el suavizamiento de datos, una de tales técnicas fue el uso de los promedios móviles; sin embargo, este recurso fue aplicado a las *tablas de vida*<sup>1</sup> y no a las series de tiempo como tales. El trabajo que Buys-Ballot desarrolló analizando series temporales en el área meteorológica para establecer comportamientos regulares en las mismas permitió que Schuster desarrollara en 1898 el *periodograma* que es una función basada en las series de Fourier, formadas por la suma de una cantidad ilimitada de los llamados armónicos puros o simples de periodo  $c$  definidos como la suma de funciones del tipo  $a \cdot \text{Sen}(b \cdot t/c)$ , donde  $a$  y  $b$  son parámetros que permiten variar la amplitud y frecuencia de la función  $\text{Sen}(t)$  y  $c$  es el periodo estacional del movimiento descrito por la serie en estudio, que permite simular la componente estacional, lo que permite además descubrir periodicidades de las series.

---

<sup>1</sup> Compendio de probabilidades de vida por edad y sexo.

Con este resultado se inició una etapa de investigaciones sobre periodicidades ocultas y en la década de los años cuarenta surgieron los primeros trabajos sobre análisis espectral moderno que se basan en la posibilidad de representar una serie cronológica mediante una sucesión de puntos que muestran oscilaciones dadas por los armónicos y, por medio de éstos, convertir la serie del dominio del tiempo al de las frecuencias; el propósito de este tipo de análisis es determinar en qué medida la variación de las observaciones en el tiempo se debe a ciertas frecuencias.

En 1927, el desarrollo de las series autorregresivas formulado por Yule representó un importante adelanto en el modelado de las series temporales. Ésto, combinado con el Teorema Ergódico de Birkhoff y Khinchin y la demostración de Wold (A Study in The Analysis of Stationary Time Series, Suecia) de que la parte no determinista de cualquier serie estacionaria puede representarse como un promedio móvil, preparó el terreno para el florecimiento del análisis aplicando procesos autorregresivos de promedios móviles que serán descritos más adelante.

El análisis de las series cronológicas lineales se ha enriquecido ampliamente en los últimos años con la publicación del libro de Box y Jenkins (1970) y con la clarificación estructural de Akaike (1974). Por su parte, Ozaki y Tong desarrollaron en 1975 un conjunto de pruebas para modelos autorregresivos lineales no estacionarios mediante la linealización de sus parámetros.

En el campo de las de series de tiempo no lineales, se han desarrollado diferentes modelos entre los que destacan:

- a) El modelo bilineal de Priestley (1978) y Rao (1977),
- b) El modelo polinomial AR de T. Ozaki y Oda (1977),
- c) El modelo del umbral de Tong (1978), y

d) El modelo exponencial de Ozaki.

En cuanto al análisis de las series de tiempo basado sobre el dominio de la frecuencia, el trabajo de Wiener (Acta Math y The Extrapolation, Interpretation & Smoothing of Time Series, Nueva York) tiene gran importancia por las bases que propone para este tipo de análisis. Otra contribución significativa fue el trabajo de Abbe (1863) sobre la distribución del coeficiente de correlación serial. T. N. Thiele anticipó muchas de las ideas de los filtros de Kalman que actualmente son esenciales en este tipo de análisis.

Las técnicas más usuales en los últimos tiempos para el análisis de las series de tiempo univariadas lineales son. la clásica (o de enfoque determinista) y la moderna (o de enfoque estocástico), ésta última será la que abordaré más ampliamente en este trabajo.

El análisis clásico considera que toda serie de tiempo está formada por cuatro componentes. la tendencia **T** (movimiento general a largo plazo de la serie), la estacionalidad **E** (oscilaciones en los valores o comportamiento de la serie que se presentan con un periodo igual o inferior al año, que se reproducen de manera reconocible en los diferentes años; estas oscilaciones se deben frecuentemente a condiciones climatológicas o a las consecuencias de éstas en la actividad económica), la ciclicidad **C** (oscilaciones que se producen con un periodo mayor al año —y que en el caso de series económicas se deben a épocas de prosperidad y depresión en la actividad económica—) y una componente aleatoria **R** (movimientos que no muestran un carácter periódico reconocible y que se consideran originados por fenómenos singulares que afectan a la variable en estudio de forma casual y no

permanente). Cuando una serie presenta la componente estacional se dice que se trata de una serie de tiempo estacional y se la llama no estacional cuando no la incluye.

Los diferentes esquemas de combinación de las variables mencionadas con base en este enfoque son:

- Esquema aditivo:  $T + C + E + R$  (1.2.1)

- Esquema multiplicativo:  $T * C * E * R$  (1.2.2)

- Esquema mixto:  $T * C * E + R$  (1.2.3)

Para este tipo de análisis es necesario el supuesto de independencia de la componente aleatoria; esto se cumple en las representaciones (1.2.1) y (1.2.3), pero no en (1.2.2) donde la serie presenta un esquema multiplicativo; en ese caso, se toma el logaritmo de la serie para eliminar la dependencia, por lo que si  $Z_t$  es el valor de la serie en el tiempo  $t$ :

$$Z_t = T_t * C_t * E_t * R_t$$

se tendrá:

$$\text{Log}(Z_t) = \text{Log}(T_t * C_t * E_t * R_t)$$

$$\text{Log}(Z_t) = \text{Log}(T_t) + \text{Log}(C_t) + \text{Log}(E_t) + \text{Log}(R_t) \quad (1.2.4)$$

en (1.2.4) la componente aleatoria ya está representada en forma independiente.

Los pasos a seguir para el análisis clásico de una serie cronológica son: análisis de la tendencia, eliminación de las variaciones estacionales e identificación de las variaciones cíclicas.

En lo referente a la metodología de Box y Jenkins, también se consideran las cuatro componentes ya mencionadas, pero describe a una serie cronológica como un

proceso estocástico caracterizándolo a partir de los primeros momentos de las variables que lo integran.

### **1.3 Bases del enfoque de Box y Jenkins**

Para introducir el análisis de las series cronológicas según el enfoque de Box y Jenkins empezaré con la descripción de un proceso estocástico según la siguiente definición

*Un proceso estocástico es una familia de variables aleatorias asociadas a un conjunto índice de números reales, de tal forma que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria.<sup>1</sup>*

Con base en esta definición se tiene que una serie de tiempo es un proceso estocástico cuyo conjunto índice se toma con relación al tiempo. En este contexto, es importante hacer notar que dado que una serie de tiempo es resultado de una realización del proceso estocástico que la genera, bien pudo haberse registrado una realización diferente del mismo; esto refleja su carácter azaroso o probabilístico. En adelante, aún bajo el riesgo de causar confusión, denotaré por  $Z_t$  tanto a la variable aleatoria observada en el tiempo  $t$  como al valor observado de la misma.

Para describir el comportamiento del proceso estocástico es necesario conocer la función de densidad conjunta de las variables que lo integran. No es fácil determinar esta función, dado que el supuesto principal en materia de series de tiempo es que existe dependencia entre las observaciones registradas, por lo que deben considerarse los estimadores de los momentos de primero y segundo orden del proceso para resumirlo en forma más o menos adecuada, es decir, se emplean a

---

<sup>1</sup> Guerrero, Víctor M., Análisis estadístico de series de tiempo económicas, p.5.



$E(Z_t) = \mu_t$ ,  $\text{Var}(Z_t) = E[(Z_t - \mu_t)^2]$  y  $\text{Cov}(Z_t, Z_{t+k}) = E[(Z_t - \mu_t)(Z_{t+k} - \mu_{t+k})]$  —esta última en la práctica se sustituye por la autocorrelación<sup>1</sup> para eliminar la influencia de las unidades de medida—).

La caracterización que menciono en el párrafo anterior sólo es posible si el proceso puede ser considerado *ergódico*. Se dice que un proceso estocástico es ergódico si bajo ciertas condiciones es posible estimar consistentemente sus características a partir de una realización del mismo. Para que un proceso sea ergódico es necesario que  $E(Z_t) < \infty$ ,  $\lim_{n \rightarrow \infty} \sum_{k=1}^n \rho_k = 0$  ( $\rho_k$  es la autocorrelación entre las variables  $Z_t$  y  $Z_{t+k}$ ) y que sea un proceso estacionario (al menos de segundo orden), entre otras condiciones.

La *estacionariedad* se presenta cuando a que existen funciones del proceso estocástico que no varían con el tiempo, es decir, que llegan a un nivel de equilibrio cuyo valor no depende del tiempo. Existen dos tipos de estacionariedad: la *débil* o de *segundo orden*, en la que los momentos de primer y segundo orden no dependen del tiempo, y la *estricta*, que implica que la función de densidad para un conjunto arbitrario de variables que están incluidas en el proceso no varía como consecuencia de desplazamientos en el tiempo, es decir,  $f(Z_t, Z_{t+1}, \dots, Z_{t+m}) = f(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+m})$ . En los procesos estocásticos con función de densidad normal multivariada la estacionariedad débil implica la estacionariedad estricta; así pues, como lo más común es suponer que la distribución correspondiente a las series de tiempo es la normal, cuando en este trabajo haga referencia a la estacionariedad deberá entenderse que

---

<sup>1</sup> Los términos autocorrelación y autocovarianza se refieren al hecho de que tales medidas se toman sobre la misma variable pero con retraso.

me refiero a la estacionariedad de segundo orden (por lo que sólo será necesario conocer la media y la función de autocovarianza de las variables del proceso).

El concepto de *estacionariedad* es de la mayor importancia ya que la teoría de Box y Jenkins parte del supuesto de que la serie a analizar es estacionaria. Es cierto que en la práctica la mayoría de las series temporales no son estacionarias, pero afortunadamente se han desarrollado métodos para volver estacionarias a las series que en un principio no lo son; esos métodos serán presentados en el capítulo tres pero adelantaré que para estacionarizar la serie habrá que aplicar el operador diferencia a la serie.

Los dos operadores que a continuación describo son un elemento importante en el modelado de Box y Jenkins:

Operador	Símbolo	Significado
<i>De retraso</i>	B	$B^k Z_t = Z_{t-k}$
<i>Diferencia</i>	$\nabla$	$\nabla^k Z_t = (1-B)^k Z_t = \sum_{j=0}^k \frac{k!(-1)^j Z_{t-j}}{j!(k-j)!}$

utilizando estos operadores se definen los llamados *polinomios de retraso* que tienen la siguiente representación

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

Mediante el uso de estos polinomios es como se representan los modelos de las series en el enfoque estocástico.

El penúltimo tema introductorio a los modelos de Box y Jenkins para series de tiempo se refiere a las *ecuaciones lineales en diferencia* para procesos deterministas, que son el *equivalente discreto* de las ecuaciones diferenciales, es decir, miden la variación entre los valores de la variable respecto del tiempo transcurrido entre las

observaciones. Considerando que en el presente caso el tiempo relacionado con tales observaciones es discreto y varía en una unidad en cada registro, entonces las ecuaciones en diferencia se reducen a medir las variaciones entre los valores de la variable.

Una ecuación en diferencia de orden  $p$ , cuyo formato es:

$$Z_t - a_1 Z_{t-1} - a_2 Z_{t-2} - a_3 Z_{t-3} \dots - a_p Z_{t-p} = a_0 \quad (1.3.1)$$

$$\text{con } t=0, \pm 1, \pm 2, \dots, \pm p; \quad a_i = \text{constante, } j = 0, 1, \dots, p; \quad a_p \neq 0$$

Puede expresarse, utilizando el operador de retraso, como.

$$(1 - a_1 B - a_2 B^2 - a_3 B^3 \dots - a_p B^p) Z_t = a_0 \quad (1.3.2)$$

Por el *Teorema Fundamental del Álgebra* este polinomio de retraso puede expresarse como:

$$(1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_p B) = G(B) \quad (1.3.3)$$

De este modo:

$$(1 - a_1 B - a_2 B^2 - a_3 B^3 \dots - a_p B^p) = (1-g_1 B)(1-g_2 B) (1-g_3 B) \dots (1-g_p B) \quad (1.3.4)$$

Las raíces de la ecuación característica  $G(x)=0$  relativa a este polinomio serán  $x=g_1^{-1}, g_2^{-1}, g_3^{-1}, \dots, g_p^{-1}$  y permitirán conocer la solución general de la ecuación (1.3.1); dicha solución estará dada por

$$Z_t = a_0 / [(1-g_1)(1-g_2) \dots (1-g_p)] + s_1 g_1^t + s_2 g_2^t + s_3 g_3^t + \dots + s_p g_p^t \quad (1.3.5)$$

Cuyos valores  $s_i$   $i=1,2,\dots,p$  son determinados con base en  $p$  condiciones iniciales conocidas que pueden ser los valores  $Z_1$  a  $Z_p$ , el nivel medio del proceso o alguna

cota superior o inferior. Para confirmar que (1.3.5) es solución general de (1.3.1) sólo es necesario multiplicar (1.3.5) por (1.3.3) y considerar la equivalencia expresada en (1.3.4), esto es, para  $i \leq p$  tenemos:

$$\begin{aligned}
 s_i g_i^t G(B) &= s_i g_i^t (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_i B) \dots (1-g_p B) \\
 &= s_i g_i^t (1-g_1 B) (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_{i-1} B) (1-g_{i+1} B) \dots (1-g_p B) \\
 &= ((s_i g_i^t)(1) - s_i g_i B g_i^t) (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_{i-1} B) (1-g_{i+1} B) \dots (1-g_p B) \\
 &= (s_i g_i^t - s_i g_i g_i^{t-1}) (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_{i-1} B) (1-g_{i+1} B) \dots (1-g_p B) \\
 &= (s_i g_i^t - s_i g_i^t) (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_{i-1} B) (1-g_{i+1} B) \dots (1-g_p B) \\
 &= 0
 \end{aligned}$$

Con lo que el producto de (1.3.5) por (1.3.3) se reduce a

$$\begin{aligned}
 (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_p B) Z_t &= \\
 &= a_0 / [(1-g_1 B)(1-g_2 B) \dots (1-g_p B)] * (1-g_1 B)(1-g_2 B)(1-g_3 B) \dots (1-g_p B) \\
 \Rightarrow (1 - a_1 B - a_2 B^2 - a_3 B^3 - \dots - a_p B^p) Z_t &= a_0
 \end{aligned}$$

En este punto debo señalar que la convergencia del proceso expresado como (1.3.5) a un nivel de equilibrio dependerá del módulo de los valores  $g_i$ ,  $i=1,2,\dots,p$ : el proceso será convergente si los  $p$  valores cumplen con  $|g_i| < 1$ , es decir, que  $g_i^{-1}$  se encuentre fuera del círculo unitario en el plano complejo. Estas condiciones surgen del hecho de que al representar al proceso determinista como:

$$Z_t = (1-g_1 B)^{-1} (1-g_2 B)^{-1} \dots (1-g_p B)^{-1} a_0$$

Este sólo tiene sentido en la práctica si:

$$(1-g_i B)^{-1} = \sum_{n=0}^{\infty} (g_i B)^n = \sum_{n=0}^{\infty} g_i^n = \frac{1}{(1-g_i)} \quad \text{si } |g_i| < 1, \quad i=1,2,\dots,p \quad (1.3.6)$$

Además esto permite que:

$$\lim_{t \rightarrow \infty} s_i g_i^t = 0 \quad \text{si } |g_i| < 1, \quad i=1,2,\dots,p \quad (1.3.7)$$

Las condiciones dadas en (1.3.6) y (1.3.7) no se cumplirían si  $|g_i| \geq 1$ .

El último tema base para el análisis de series temporales bajo el enfoque moderno se refiere a la función de autocorrelación, **FAC**, y la función de autocorrelación parcial, **FACP**. La importancia de estas funciones es que permiten describir algunas características del proceso generador de la serie por medio de las cuales es posible identificar el modelo correspondiente a ésta. Las características de dichas funciones son:

*Función de autocorrelación:* permite evaluar la correlación entre dos observaciones distintas de una misma variable en dos puntos temporales diferentes, es decir, separadas por k periodos:

$$\rho_k = \frac{\text{Cov}(Z_t, Z_{t+k})}{[\text{Var}(Z_t) \cdot \text{Var}(Z_{t+k})]^{1/2}} \quad (1.3.8)$$

*Función de autocorrelación parcial:* mide la correlación entre dos observaciones distintas de una misma variable en dos puntos temporales diferentes, mientras las correlaciones de los periodos intermedios son controladas o mantenidas constantes, es decir, la **FACP** es una autocorrelación condicionada en la que el valor de esta función entre dos observaciones está mediatizado por los valores de las correlaciones para las observaciones intermedias. Como consecuencia, la **FACP** no puede obtenerse por medio del uso de una sola ecuación, como ocurre con la **FAC**, pues los términos y la forma de la ecuación irán cambiando a medida que los

retardos se incrementen, pues serán más las observaciones intermedias implicadas.

El cálculo de esta función se realiza mediante la regla de Cramer:

$$\begin{aligned}
 \phi_{11} &= \rho_1 \\
 \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \\
 \phi_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}} \quad (1.3.9) \\
 \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}
 \end{aligned}$$

Una ventaja de las funciones mencionadas es que, al estar referidas a las correlaciones y ser éstas adimensionales, no se ven influidas por cambios extremos en los valores de la serie.

Ahora, puedo presentar los diferentes modelos base que propone el enfoque estocástico para caracterizar a una serie de tiempo.

## 2 Modelos para series de tiempo

En este capítulo haré una breve introducción a los modelos para series de tiempo univariadas, lineales, discretas, sobre intervalos de tiempo iguales y según el enfoque estocástico utilizados para analizar las series cronológicas que se presentan en la práctica. Estos modelos se comportan a lo largo del tiempo de acuerdo con las leyes de probabilidad y suponen que el proceso a describir es ergódico.

Cabe señalar que estos prototipos sólo son aplicables a series que no presentan estacionalidad, ya que los procesos estacionales requieren de un tratamiento diferente que detallaré en el capítulo cinco.

Del apartado 2.1 al 2.3 trataré modelos para series estacionarias, mientras que en el punto 2.4 presentaré el modelo para series no estacionarias en la media. Para todos ellos elaboraré una breve revisión de su estructura y características.

### 2.1 Modelos autorregresivos

Estos modelos tienen una representación similar a la de las ecuaciones en diferencia mostradas en (1.3.1), pero consideran una variable aleatoria  $a_t$  en vez de una constante, es decir, el modelo se convierte en:

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \dots - \phi_p Z_{t-p} = c + a_t \quad (2.1.1)$$

$$t=0, \pm 1, \pm 2, \pm 3, \dots; \phi_i, c = \text{constante}, j=1, 2, \dots, p; \phi_p \neq 0; a_t \text{ variable aleatoria}$$

Se llaman modelos autorregresivos debido a que el valor actual de la variable se estima con base en los valores que ésta misma tomó en tiempos anteriores a  $t$ , valores que se ponderan con los coeficientes autorregresivos  $\phi_i, i=1, 2, \dots, p$ . La

necesidad de introducir a la variable aleatoria se debe a que una serie cronológica no es un proceso determinista, si no resultado de un patrón influido por movimientos aleatorios

Si se cumple que  $Z_t$  tenga media constante  $E(Z_t) = \mu$  y se hace el cambio de variable  $Z_t = Z_t - \mu$  entonces el modelo en (2.1.1) puede representarse también como:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.1.2)$$

En el cual es más notorio el origen del nombre de estos procesos. A (2.1.2) se le llama modelo autorregresivo de orden  $p$  y se le simboliza como **AR(p)**. Otra forma de representar a (2.1.1) es:

$$\phi(B)Z_t = a_t \quad (2.1.3)$$

Considerando en general y por simplicidad que  $\{a_t\}$  es un proceso de ruido blanco<sup>1</sup> con media cero, varianza  $\sigma_a^2$  y distribución normal que mide las variaciones o errores aleatorios del proceso.

En el caso de este tipo de procesos, es imposible decir que se alcance un punto de equilibrio debido a la parte aleatoria que involucra; sin embargo, puede hacerse referencia a su estacionaridad. Recordando lo revisado en el capítulo anterior sobre ecuaciones en diferencia, se tiene que por analogía la estacionaridad del proceso dependerá de los módulos de las raíces de la ecuación característica  $\phi(x) = 0$ .

---

<sup>1</sup> Un proceso de ruido blanco es una sucesión de choques aleatorios (realizaciones independientes) de una variable aleatoria con media constante (generalmente se le considera igual a cero) y varianza  $\sigma_a^2$ .



Retomando que  $\phi(B)$  puede escribirse como  $(1-g_1B)(1-g_2B)(1-g_3B)\dots(1-g_pB) = \phi(B)$  se llega a que (2.1.2) será un proceso estacionario si  $|g_i| < 1$  para  $i=1, 2, \dots, p$ . Esto es equivalente a pedir que los valores  $g_1^{-1}, g_2^{-1}, g_3^{-1}, \dots, g_p^{-1}$  se encuentren todos fuera del círculo unitario (en el plano complejo). Otro resultado que permite determinar la estacionaridad de la serie es el *Teorema de Schur*<sup>1</sup>, que hace posible deducir que el proceso es estacionario si los siguientes  $p$  determinantes son **todos** positivos:

$$\begin{array}{l}
 D_1 = \begin{vmatrix} -1 & \phi_p \\ \phi_p & -1 \end{vmatrix} \\
 D_2 = \begin{vmatrix} -1 & 0 & \phi_p & \phi_{p-1} \\ \phi_1 & -1 & 0 & \phi_p \\ \phi_p & 0 & -1 & \phi_1 \\ \phi_{p-1} & \phi_p & 0 & -1 \end{vmatrix} \\
 D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & \phi_p & \phi_{p-1} & \dots & \phi_1 \\ \phi_1 & -1 & \dots & 0 & 0 & \phi_p & \dots & \phi_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \phi_{p-1} & \phi_{p-2} & \dots & -1 & 0 & 0 & \dots & \phi_p \\ \phi_p & 0 & \dots & 0 & -1 & \phi_1 & \dots & \phi_{p-1} \\ \phi_{p-1} & \phi_p & \dots & 0 & 0 & -1 & \dots & \phi_{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \phi_1 & \phi_2 & \dots & \phi_p & 0 & 0 & \dots & -1 \end{vmatrix} \quad (2.1.4)
 \end{array}$$

Si (2.1.3) resulta ser estacionario entonces  $\phi(B)Z_t = a_t$  puede expresarse como una suma infinita en términos de los errores aleatorios  $Z_t = \phi(B)^{-1}a_t = \Psi(B)a_t = a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots$  con  $\sum_{i=1}^{\infty} \Psi_i < \infty$ .

A esta propiedad se le conoce con el nombre de *invertibilidad*. Adicionalmente, puedo mencionar que es mejor representar el modelo con un número finito de

<sup>1</sup> Los módulos de las raíces de la ecuación  $d^p - \phi_1 d^{p-1} - \phi_2 d^{p-2} - \dots - \phi_{p-1} d - \phi_p = 0$  serán todos menores que la unidad, si y sólo si los  $p$  determinantes mostrados en (2.1.4) son todos positivos.

parámetros que con una infinidad de los mismos; a esta medida de elegir el modelo con el menor número de parámetros se le llama *Principio de Parsimonia*.

Considerando que el proceso  $AR(p)$  es estacionario, la media no depende del tiempo, ya que  $E(Z_t)=0$  y los valores de la varianza y autocovarianza, basados en la representación invertida de  $Z_t$ , serán:

$$\begin{aligned} \gamma_0 &= \text{Var}(Z_t) = E[(Z_t - E(Z_t))^2] = E(Z_t^2) \\ \gamma_0 &= E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots)^2] \\ \gamma_0 &= E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) a_t] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_1 a_{t-1}] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_2 a_{t-2}] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_3 a_{t-3}] + \dots \\ \gamma_0 &= E(a_t^2) + \Psi_1^2 E(a_{t-1}^2) + \Psi_2^2 E(a_{t-2}^2) + \Psi_3^2 E(a_{t-3}^2) + \dots + \sum_{j=1}^{\infty} \Psi_j \Psi_j E[a_{t-j} a_{t+k-j}] \\ \gamma_0 &= \sigma_a^2 \sum_{j=1}^{\infty} \Psi_j^2 \quad \text{con } \Psi_0=1 \end{aligned} \tag{2.1.5}$$

Puede advertirse que tampoco  $\gamma_0 = \text{Var}(Z_t)$  depende del tiempo, sino sólo de la convergencia de la suma infinita y, finalmente:

$$\begin{aligned} \gamma_k &= \text{Cov}(Z_t, Z_{t+k}) = E[(Z_t - E(Z_t))(Z_{t+k} - E(Z_{t+k}))] = E(Z_t Z_{t+k}) \\ \gamma_k &= E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots)(a_{t+k} + \Psi_1 a_{t+k-1} + \Psi_2 a_{t+k-2} + \dots + \Psi_{k-1} a_{t+k-(k-1)} + \Psi_k a_{t+k-k} \\ &\quad + \Psi_{k+1} a_{t+k-(k+1)} + \dots)] \\ \gamma_k &= E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) a_{t+k}] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_1 a_{t+k-1}] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_2 a_{t+k-2}] + \dots \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_{k-1} a_{t+k-(k-1)}] \\ &\quad + E[(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \Psi_3 a_{t-3} + \dots) \Psi_k a_{t+k-k}] \end{aligned}$$

$$+ E[(\alpha_t + \Psi_1 \alpha_{t-1} + \Psi_2 \alpha_{t-2} + \Psi_3 \alpha_{t-3} + \dots) \Psi_{k+1} \alpha_{t+k-(k+1)}] + \dots$$

$$\gamma_k = E(\Psi_k \alpha_t^2) + E(\Psi_1 \Psi_{k+1} \alpha_{t-1}^2) + E(\Psi_2 \Psi_{k+2} \alpha_{t-2}^2) + \dots + \sum_{j=k-1}^{\infty} \Psi_1 \Psi_j E[\alpha_{t-j} \alpha_{t+k-j}]$$

$$\gamma_k = \sigma_\alpha^2 [\Psi_k + \sum_{j=1}^{\infty} \Psi_j \Psi_{k+j}] \quad \text{con } \Psi_0 = 1 \quad (2.1.6)$$

Así, en (2.1.6),  $\gamma_k = \text{Cov}(\mathbf{Z}_t, \mathbf{Z}_{t+k})$  tampoco depende del tiempo, sino sólo de la separación de las variables y de la convergencia de la suma infinita.

Las primeras  $p$  autocorrelaciones que forman la **FAC** pueden obtenerse en términos de los parámetros  $\phi$  y las ecuaciones de Yule-Walker siendo éstas:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\dots \dots \dots \dots \dots \dots \dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (2.1.7)$$

Las demás autocorrelaciones se obtienen de la relación

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad \text{para } k \geq p+1 \quad (2.1.8)$$

La **FAC** converge a cero exponencialmente o bien con un comportamiento sinusoidal dependiendo de las raíces de  $\phi(B) = 0$  (el comportamiento sinusoidal aparece cuando alguna de las raíces es compleja). Todo proceso **AR(p)** tiene asociada una única **FAC** que se encuentra determinada por las condiciones (2.1.7) y (2.1.8), cuyas correlaciones obviamente no dependen del tiempo.

En cuanto a la **FACP**, se tiene que sólo las primeras  $p$  autocorrelaciones parciales son distintas de cero debido a que el modelo depende de  $p$  retrasos en  $Z_t$ ,

(la autocorrelación se ve influida sólo por  $p$  variables consecutivas). Para ver esto se tiene que para la autocorrelación parcial  $\phi_{kk}$  para  $k > p$ , señalada en (1.3.9), la última columna en la matriz del numerador puede ser escrita como combinación lineal de las columnas previas de la misma matriz por lo que para  $k > p$  la **FACP** desaparece.

## 2.2 Modelos de promedios móviles

Estos modelos conciben a la serie de tiempo  $\{Z_t\}$ , formada por observaciones dependientes, como si se hubiera generado por una combinación lineal finita de choques aleatorios independientes  $\{a_t\}$ . Si se toman las desviaciones de  $\{Z_t\}$  respecto de su media  $\mu$  es factible generar la serie  $\{Z_t\}$ , donde  $Z_t = Z_t - \mu$ , y formular el modelo de promedios móviles de orden  $q$  como a continuación se describe.

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \dots - \theta_q a_{t-q} \quad (2.2.1)$$

Equivalente a.

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q) a_t \quad (2.2.2)$$

Y a:

$$Z_t = \theta(B) a_t \quad (2.2.3)$$

Donde  $\theta_i$ ,  $i = 1, 2, \dots, q$  son las ponderaciones asociadas a los choques en los periodos  $t-1, t-2, \dots, t-q$ . Por las representaciones anteriores se nota que al ser  $Z_t$  una combinación lineal finita de variables aleatorias estacionarias —recuérdese que se supone que  $\{a_t\}$  es un proceso de ruido blanco— entonces todo proceso de promedios móviles es estacionario.

El proceso en (2.2.1) se simboliza como **MA**( $q$ ) y se le llama proceso de promedios móviles de orden  $q$ . Probablemente la denominación de promedios móviles proviene de la idea de promediar o ponderar los choques aleatorios, aunque

en este caso no se trata necesariamente de un promedio pues no es indispensable que la suma de las ponderaciones sea uno, ni que éstas sean positivas.

Las restricciones para invertibilidad de este modelo se obtienen por dualidad de las condiciones de estacionaridad del modelo autorregresivo mostradas en (2.1.4); es decir, (2.2.3) será un proceso invertible si los  $q$  determinantes que se presentan a continuación son **todos** positivos:

$$\begin{aligned}
 D_1 &= \begin{vmatrix} -1 & \theta_q \\ \theta_q & -1 \end{vmatrix} \\
 D_2 &= \begin{vmatrix} -1 & 0 & \theta_q & \theta_{q-1} \\ \theta_1 & -1 & 0 & \theta_q \\ \theta_q & 0 & -1 & \theta_1 \\ \theta_{q-1} & \theta_q & 0 & -1 \end{vmatrix} \\
 D_q &= \begin{vmatrix} -1 & 0 & \dots & 0 & \theta_q & \theta_{q-1} & \dots & \theta_1 \\ \theta_1 & -1 & \dots & 0 & 0 & \theta_q & \dots & \theta_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_{q-1} & \theta_{q-2} & \dots & -1 & 0 & 0 & \dots & \theta_q \\ \theta_q & 0 & \dots & 0 & -1 & \theta_1 & \dots & \theta_{q-1} \\ \theta_{q-1} & \theta_q & \dots & 0 & 0 & -1 & \dots & \theta_{q-2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_1 & \theta_2 & \dots & \theta_q & 0 & 0 & \dots & -1 \end{vmatrix} \quad (2.2.4)
 \end{aligned}$$

Si el proceso  $MA(q)$  resulta ser invertible podrá expresarse como:

$$Z_t = a_t + \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \pi_3 Z_{t-3} + \dots$$

Con  $\sum_{i=1}^{\infty} \pi_i < \infty$ .

Como mostré, dado que el proceso es estacionario ni la media, ni la varianza, ni la covarianza dependen del tiempo, pues es inmediato que  $E(Z_t) = 0$  y como consecuencia de (2.1.5) y (2.1.6)

$$\gamma_0 = \text{Var}(Z_t) = \sigma_e^2 \sum_{j=1}^q \theta_j^2 \quad \text{con } \theta_0=1 \quad (2.2.5)$$

$$\gamma_k = \text{Cov}(Z_t, Z_{t+k}) = \sigma_e^2 [-\theta_k + \sum_{j=1}^q \theta_j \theta_{k+j}] \quad (2.2.6)$$

Con base en esto se obtiene la función de autocorrelación

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \dots + \theta_q\theta_{q-k}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{si } k=1, 2, \dots, q \\ 0 & \text{si } k \geq q+1 \end{cases} \quad (2.2.7)$$

En la **FAC** y **FACP** de estos modelos ocurre lo contrario que en los **AR(p)**, es decir, sólo las primeras  $q$  autocorrelaciones son distintas de cero, pues sólo existe relación de la variable con los primeras  $q$  choques aleatorios, en tanto que la **FACP** converge a cero exponencialmente o bien con un comportamiento sinusoidal dependiendo de las raíces de  $\theta(B) = 0$  (el comportamiento sinusoidal aparece cuando alguna de las raíces es compleja).

### 2.3 Modelos autorregresivos y de promedios móviles

Los esquemas denotados por **ARMA(p,q)** surgen al combinar el modelo **AR(p)** con el **MA(q)** para poder cumplir mejor con el principio de parsimonia, es decir, utilizar el menor número de parámetros para estimar en forma eficiente la serie en estudio. El modelo **ARMA(p,q)** se representa así.

$$\phi(B)Z_t = \theta(B)a_t \quad (2.3.1)$$

Equivalente a:

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \phi_3 Z_{t-3} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \dots - \theta_q a_{t-q} \quad (2.3.2)$$

Con  $\phi(B)$  y  $\theta(B)$  polinomios de retraso de órdenes  $p$  y  $q$  respectivamente,  $\{a_t\}$  un proceso de ruido blanco y  $\{Z_t\}$  es la serie de desviaciones de  $\{Z_t\}$  respecto de su media  $\mu$ . Para que el proceso sea estacionario se requiere que las raíces de  $\phi(x)=0$  estén fuera del círculo unitario y para que sea invertible es necesario que las de

$\theta(x)=0$  lo estén también, por lo que deberá cumplirse que **todos** los determinantes mostrados en (2.1.4) y (2.2.4) sean positivos.

Como consecuencia de la estacionaridad y de la invertibilidad que presentan los procesos **AR**( $p$ ) y **MA**( $q$ ) que se combinan, se tiene que la esperanza, varianza, covarianza y autocorrelaciones del **ARMA**( $p,q$ ) no dependen del tiempo.

Para derivar la **FAC** del proceso basta con despejar a  $Z_t$  de (2.3.2), multiplicar el resultado por  $Z_{t+k}$  y tomar la esperanza de cada uno de los términos con lo que se obtiene:

$$\begin{aligned} \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \phi_3 \gamma_{k-3} + \dots + \phi_p \gamma_{k-p} \quad k \geq q+1 \Rightarrow \\ \rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + \dots + \phi_p \rho_{k-p} \quad k \geq q+1 \end{aligned} \quad (2.3.3)$$

Por lo que sigue la misma ecuación en diferencia que la **FAC** del proceso **AR**( $p$ ) expresada en (2.1.8), sólo que en (2.3.3) esta función inicia su convergencia a cero después del retraso  $k = q$ , en tanto que las primeras  $q$  autocorrelaciones dependen tanto de los parámetros autorregresivos como de promedios móviles, lo que implica que tienen un comportamiento irregular. Por su parte, la **FACP** será una sucesión infinita convergente a cero.

## **2.4 Modelos autorregresivos e integrados de promedios móviles**

Este tipo de representación es una extensión de los modelos **ARMA** en la que se considera una serie temporal que no es estacionaria en la media. La idea del modelo **ARIMA** es la siguiente: si la serie original  $\{Z_t\}$  tiene una tendencia polinomial no determinista (u homogénea, como también se la llama) que provoca que la serie no

sea estacionaria entonces es factible la construcción de un segundo proceso  $\{W_t\}$  que sí sea estacionario.

Como señalé en la sección 1.3, la forma para construir  $\{W_t\}$  sería aplicar a  $\{Z_t\}$  el operador diferencia  $d$  veces, suponiendo que la tendencia polinomial de  $\{Z_t\}$  es de orden  $d$ . Se define entonces a  $W_t$  como:

$$W_t = \nabla^d Z_t \quad \forall t \quad (2.4.1)$$

A esta serie se le podría hacer corresponder un proceso **ARMA**( $p,q$ ) de la forma  $\phi(B)W_t = \theta(B)a_t$  que sería equivalente a considerar el modelo **ARIMA**( $p,d,q$ ) con la representación:

$$\phi(B)\nabla^d Z_t = \theta(B)a_t, \quad d \geq 1 \quad (2.4.2)$$

Con  $\phi(B)$  y  $\theta(B)$  polinomios de retraso de órdenes  $p$  y  $q$  respectivamente,  $\{a_t\}$  un proceso de ruido blanco y  $\{Z_t\}$  es la serie en estudio. Para que este modelo sea estacionario e invertible se requiere que las raíces de  $\phi(x)=0$  y  $\theta(x)=0$  estén fuera del círculo unitario.

Nótese que he supuesto implícitamente que el nivel del proceso  $\{W_t\}$  es cero, pero de no ser así debería considerar la existencia de una media constante  $\mu_w$  en la serie con lo que tendría que considerar a un modelo **ARIMA** más general con la forma

$$\phi(B)(W_t - \mu_w) = \theta(B)a_t \quad (2.4.3)$$

O bien

$$\phi(B)W_t = \theta_0 + \theta(B)a_t \quad (2.4.4)$$



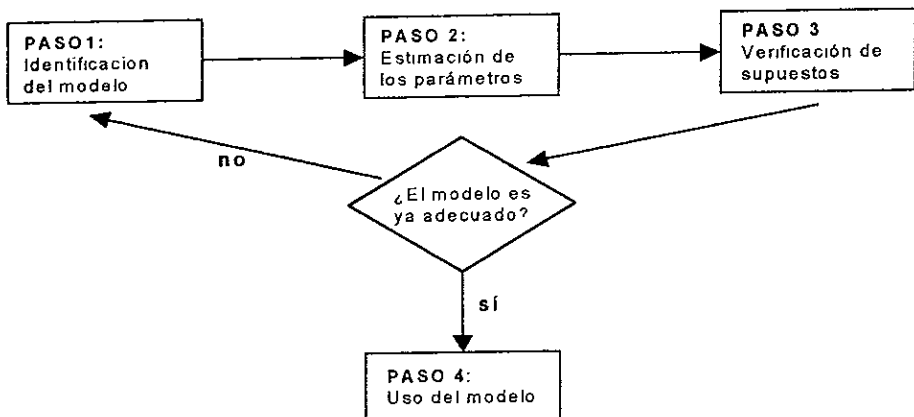
Con  $\theta_0 = \mu_w(1 + \phi_1 + \phi_2 + \phi_3 + \dots + \phi_p)$  Donde  $\theta_0$  es el resultado de una tendencia polinomial determinista (no dependiente de los choques aleatorios  $\{a_t\}$ ) no adaptiva y que por lo tanto es perfectamente predecible. Como esta condición es difícilmente identificable, casi siempre se supone que  $\theta_0 = 0$ .

Finalmente, en ocasiones se hace referencia a modelos tipo **ARI** o del tipo **IMA** lo que significa que el proceso **ARIMA** está formado sólo por un proceso autorregresivo diferenciado o por uno de promedio móviles diferenciado, respectivamente.

### 3 Construyendo el modelo

Según la teoría de Box y Jenkins, el procedimiento a seguir en la formulación de un modelo **ARIMA** consta de una estrategia iterativa de identificación, estimación y verificación.

En el siguiente esquema presento las etapas a seguir en la creación del modelo.



A continuación detallo la metodología que seguiré en los pasos 1 a 3 para construir el modelo.

#### 3.1 Identificar el modelo

En esta etapa se determina la transformación de la serie temporal necesaria para estacionarizar la varianza y el número  $d$  de veces que deberá diferenciarse la serie transformada para obtener una serie estacionaria; posteriormente, con base en los comportamientos de la **FAC** y la **FACP**, se establecen los órdenes de los polinomios autorregresivos y de promedios móviles  $(p,q)$  para el proceso **ARMA** relacionado con la serie estacionaria.

### 3.1.1 Estabilización de la varianza

Como advertí en la sección 1.3, es indispensable la estacionaridad en la media y en la varianza del proceso en estudio para poder aplicar la metodología de Box y Jenkins. La no estacionaridad debe resolverse primero para la varianza y posteriormente para la media.

Para empezar a identificar el modelo lo primero que se debe hacer es determinar si es posible considerar que el proceso tiene varianza constante. Para lograr este objetivo, dado que se cuenta sólo con una sola observación y no con un conjunto de observaciones para cada tiempo, se forman grupos homogéneos de observaciones contiguas, es decir, si se tienen  $N$  observaciones se dividen éstas en  $H$  grupos con  $R$  observaciones (i.e.  $R=[N/H]$ ) y si sobran  $n=N-R*H$  observaciones éstas deberán eliminarse del principio o del final de la serie para realizar estas pruebas —en la práctica se recomienda formar grupos de entre 8 y 15 observaciones—.

A continuación se determinan las medias y desviaciones estándares de cada grupo  $H$ : Sea  $Z_{h,r}$  la observación  $r$ -ésima del grupo  $h$ -ésimo

$$Z_h = \sum_{r=1}^R Z_{h,r}/R \quad h=1,2,\dots,H \quad (3.1.1.1)$$

$$S_h^2 = \sum_{r=1}^R (Z_{h,r} - Z_h)^2 / (R-1) \quad h=1,2,\dots,H \quad (3.1.1.2)$$

Si resulta que la varianza por grupo no es "estable" se determina el valor de  $\lambda$  (parámetro de transformación) para el cual la serie transformada cumpla con esa condición. El valor de  $\lambda$  seleccionado será el que minimice el coeficiente de variación de  $\lambda$  ( $CV(\lambda)$ ) con.

$$CV(\lambda) = S(\lambda)/M(\lambda) \quad (3.1.1.3)$$

$$M(\lambda) = \sum_{h=1}^H (S_h/Z_h^{1-\lambda})/H \quad (3.1.1.4)$$

$$S^2(\lambda) = \sum_{h=1}^H [(S_h/Z_h^{1-\lambda}) - M(\lambda)]^2/(H-1) \quad (3.1.1.5)$$

La selección de  $\lambda$  se realiza de esta forma porque se supone que ese valor satisfará aproximadamente la relación  $S_h/Z_h^{1-\lambda} = \text{constante}$ .

Para estacionarizar la varianza los métodos más comunes son el de Bartlett (1947) quien sugiere una transformación del tipo:

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \text{Log}(Z_t) & \text{si } \lambda = 0 \end{cases} \quad (3.1.1.6)$$

Y el de Box-Cox (1964) que transforma a la serie como:

$$T(Z_t) = \begin{cases} (Z_t^\lambda - 1)/\lambda & \text{si } \lambda \neq 0 \\ \text{Log}(Z_t) & \text{si } \lambda = 0 \end{cases} \quad (3.1.1.7)$$

Aunque esta transformación es básicamente muy parecida a la anterior, tiende a normalizar los datos que se obtienen. Otra versión de esta transformación es la siguiente:

$$T(Z_t) = \frac{(Z_t + F_2)^{F_1 - 1}}{F_1 C^{F_1 - 1}} \quad \text{si } F_1 \neq 0 \text{ (} Z_t > -F_2 \text{)}$$

$$T(Z_t) = \text{CLn}(Z_t + F_2) \quad \text{si } F_1 = 0 \text{ (} F_2 < 0 \text{)}$$

Donde  $C$  es la media geométrica simple de  $Z_1 + F_2$ ,  $F_1$  es el equivalente al parámetro  $\lambda$ ,  $F_2$  se elige de tal forma que  $Z_1 + F_2$  sea siempre positiva (por lo tanto  $F_2$  será cero si los datos con los que se realiza el análisis son todos positivos).

Finalmente, cabe hacer notar que no necesariamente son sólo estas las transformaciones que pueden utilizarse pero sí son las más frecuentes.

### 3.1.2 Estabilización del nivel

Para estacionarizar la media habrá que diferenciar la serie un número  $d$  de veces adecuado (generalmente de 1 a 3 veces) aplicando el operador  $\nabla$ ; la nueva serie obtenida será estacionaria y con  $d$  observaciones menos que la original. Esto supone que existe un polinomio de grado  $d$  que se ajusta aproximadamente al comportamiento de la serie. Para decidir el número  $d$  de veces que se debe aplicar el operador diferencia es aconsejable seguir las técnicas que a continuación describo:

Una de las formas para determinar  $d$  es a través de las gráficas de la **FAC** muestral de algunas diferencias de la serie analizada ( $\{T(Z_i)\}$ ,  $\{\nabla T(Z_i)\}$ ,  $\{\nabla^2 T(Z_i)\}$ ,  $\{\nabla^3 T(Z_i)\}$ , ...,  $\{\nabla^n T(Z_i)\}$ ) ya que un decaimiento rápido de la **FAC** hacia cero indica que un proceso es estacionario; entonces  $d$  será el orden del operador diferencia cuya **FAC** corresponda a un proceso estacionario. En la práctica es frecuente el uso de a lo más tres diferencias y no es recomendable la sobrediferenciación pues esto puede ampliar la varianza.

Otra opción es el método de *diferencia de la variable* que se basa en la idea de que al tomar diferencias sucesivas de una serie no estacionaria su varianza va disminuyendo hasta que la serie es estacionaria y comienza a crecer con la

sobrediferenciación. Basado en este supuesto Anderson (1976) sugiere seleccionar a  $d$  como:

$$S(d) = \min \{S(j), j = 0, 1, 2, 3, \dots, n\}$$

Donde

$$S^2(j) = \frac{\sum_{t=j+1}^N [\nabla^j T(Z_t) - \sum_{t=j+1}^N \nabla^j T(Z_t)/(N-j)]^2}{N-j-1}$$

Dado que este es, ante todo, un método empírico es mejor considerarlo como complemento del análisis de la **FAC** al decidir el valor más adecuado de  $d$ .

### 3.1.3 Análisis de la función de autocorrelación y de autocorrelación parcial

Obviamente, no se conocen los valores reales de la **FAC** ni de la **FACP** por lo que, si se cuenta con  $N$  observaciones, se procede a analizar los valores de la **FAC** muestral definida por

$$r_k = \frac{\sum_{t=1}^{N-k} (Z_t - M(Z_t))(Z_{t+k} - M(Z_{t+k}))}{\sum_{t=1}^N (Z_t - M(Z_t))^2} \quad (3.1.3.1)$$

Donde  $M(Z_t)$  es la media de las observaciones y  $M(Z_t) = M(Z_{t+k})$ . Asimismo, se revisan los valores de la **FACP** muestral que se obtienen al sustituir en (1.3.9) a  $\rho$  por su estimación  $r_t$ .

En este punto se debe asociar la **FAC** a un proceso **ARIMA**. Las aproximaciones básicas para las varianzas y autocovarianzas de las autocorrelaciones proporcionadas por Bartlett son

$$\text{Var}(r_k) \approx \sum_{j=-\infty}^{\infty} (\rho_j^2 + \rho_{j+k}\rho_{j-k} - 4\rho_k\rho_j\rho_{j-k} + 2\rho_k^2\rho_j^2)/N$$

$$\text{Cov}(r_k, r_{k+s}) \approx [\sum_{j=-\infty}^{\infty} \rho_j\rho_{j+s}]/N$$

a) Si el proceso es generado por ruido blanco con distribución normal, la aproximación para la varianza será:

$$\text{Var}(r_k) \approx 1/N$$

para decidir si realmente todas las correlaciones son cero se realiza un comparativo entre los valores  $r_k$  y sus desviaciones estándar; así  $r_k$  será considerada diferente de cero si

$$|r_k| > 2[1/N]^{1/2} \quad \forall k$$

b) Si el proceso es originado por un **MA(q)** las autocorrelaciones para atrasos mayores a  $q$  son cero y entonces la aproximación sería.

$$\text{Var}(r_k) \approx (1 + 2\sum_{j=1}^q \rho_j^2)/N \quad \text{para } k > q$$

para decidir si las correlaciones son cero a partir de cierto retraso  $q$  se realiza un comparativo entre los valores  $r_k$  y sus desviaciones estándar, bajo este criterio  $r_k$  es diferente de cero si

$$|r_k| > 2[(1 + 2\sum_{j=1}^q \rho_j^2)/N]^{1/2} \quad \text{para } k > q$$

esta medida se basa en el hecho de que para una muestra de más de cien observaciones la distribución de  $r_k$  se aproxima a una distribución normal con media cero.

Como no se cuenta con los valores de  $\rho_1$  utilizo su estimador  $r_1$  obteniendo aproximaciones válidas para muestras grandes de las fórmulas presentadas.

Por otro lado, el orden de un proceso **AR**( $\rho$ ) no es fácilmente identificable a partir de la **FAC** por lo que se hace uso de la **FACP** estimada; como esta función está influida por errores muestrales, para decidir si el valor  $\phi_{11}$  es cero se realiza una prueba a su estimador  $\hat{\phi}_{11}$  para cuantificar la variación muestral de éste. Según Quenouille (1949) si el proceso es **AR**( $\rho$ ) las autocorrelaciones parciales muestrales se distribuyen de manera independiente con media  $E(\hat{\phi}_{11}) = \phi_{11}$  y con

$$\text{Var}(\hat{\phi}_{11}) \approx 1/N \text{ para } i > \rho$$

Con base en esto se determina que  $\phi_{11}$  es distinto de cero si  $\hat{\phi}_{11}$  se encuentra fuera del intervalo definido por

$$\pm 1.96[\text{Var}(\hat{\phi}_{11})]^{1/2} \approx \pm 2[\text{Var}(\hat{\phi}_{11})]^{1/2} \approx \pm 2/(N)^{1/2}$$

entonces, si se cumple la condición anterior para  $i > \rho$  se tendrá asociado un proceso **AR**( $\rho$ ).

Finalmente, para identificar un modelo **ARMA**( $p, q$ ) se cuenta con diferentes referencias. La primera técnica de identificación es revisar que el proceso tenga asociada una **FACP** que no desaparecerá después de un número finito de retrasos, en tanto que la **FAC** tendrá un comportamiento irregular de las primeras  $q$



autocorrelaciones y después convergerá a cero. La segunda técnica consiste en revisar que si  $q > p$ , la **FAC** tendrá un comportamiento irregular de las primeras  $q - p$  autocorrelaciones y después convergerá a cero, en tanto que la **FACP** no desaparecerá después de un número finito de retrasos; o bien, si  $q < p$ , la **FACP** tendrá un comportamiento irregular de las primeras  $p - q$  autocorrelaciones y después convergerá a cero, en tanto que la **FAC** no desaparecerá después de un número finito de retrasos.

La última técnica de identificación de un modelo **ARMA**( $p,q$ ) se basa en que, según Beguín, Gouriéroux y Monfort (1980), si  $Z_t$  es un proceso **ARMA**( $p,q$ ) con la representación  $\phi(B)Z_t = \theta(B)a_t$  donde  $a_t$  es un proceso de ruido blanco y  $\phi(B)$  y  $\theta(B)$  son polinomios de retrasos de órdenes  $p$  y  $q$  respectivamente cuyas raíces tienen módulo superior a uno y no tienen raíces comunes, se puede establecer el siguiente resultado:

Si  $Z_t$  es un proceso estacionario se tendrá que sigue un proceso **ARMA**( $p,q$ ) si y sólo si se cumple que

$$\begin{aligned} \Delta(i,j) &= 0 \quad \forall i \geq q+1 \text{ y } \forall j \geq p+1 \\ \Delta(i,p) &\neq 0 \quad \forall i \geq q \\ \Delta(q,j) &\neq 0 \quad \forall j \geq p \end{aligned}$$

donde  $\Delta(i,j)$  es el siguiente determinante:

$$\Delta(i,j) = \begin{vmatrix} \rho_i & \rho_{i-1} & \rho_{i-2} & \dots & \rho_{i-j} & \rho_{i-j+1} \\ \rho_{i+1} & \rho_i & \rho_{i-1} & \dots & \rho_{i-j+1} & \rho_{i-j+2} \\ \rho_{i+2} & \rho_{i+1} & \rho_i & \dots & \rho_{i-j+2} & \rho_{i-j+3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{i+j-2} & \rho_{i+j-3} & \rho_{i+j-4} & \dots & \rho_i & \rho_{i-1} \\ \rho_{i+j-1} & \rho_{i+j-2} & \rho_{i+j-3} & \dots & \rho_{i+1} & \rho_i \end{vmatrix}$$

Para facilitar la identificación del modelo, se construye una tabla con base en los valores  $\Delta(i,j)$  calculados para diferentes valores de  $i$  y  $j$ :

ORDEN MA	ORDEN AR						
	j						
i	1	2	3	.....	$p$	$p+1$	.....
1	$\Delta(1,1)$	$\Delta(1,2)$	$\Delta(1,3)$	.....	$\Delta(1,p)$	$\Delta(1,p+1)$	.....
2	$\Delta(2,1)$	$\Delta(2,2)$	$\Delta(2,3)$	.....	$\Delta(2,p)$	$\Delta(2,p+1)$	.....
3	$\Delta(3,1)$	$\Delta(3,2)$	$\Delta(3,3)$	.....	$\Delta(3,p)$	$\Delta(3,p+1)$	.....
.....	.....	.....	.....	.....	.....	.....	.....
$q$	$\Delta(q,1)$	$\Delta(q,2)$	$\Delta(q,3)$	.....	$\Delta(q,p)$	$\Delta(q,p+1)$	.....
$q+1$	$\Delta(q+1,1)$	$\Delta(q+1,2)$	$\Delta(q+1,3)$	.....	$\Delta(q+1,p)$	0	0
.....	.....	.....	.....	.....	.....	0	0

Para detectar los órdenes  $p$  y  $q$  se buscará una ruptura entre las líneas  $i = q$  e  $i=q+1$  y entre las columnas  $j = p$  y  $j = p+1$

En resumen, si el proceso presenta su **FAC** y su **FACP** iguales a cero para todos sus retrasos entonces se tratará de ruido blanco, en tanto que un proceso **AR**( $p$ ) tiene sólo las primeras  $p$  autocorrelaciones parciales distintas de cero mientras que su **FAC** converge a cero y, por otra parte, un **MA**( $q$ ) tendrá todas sus primeras  $p$  autocorrelaciones distintas de cero aunque su **FACP** muestre convergencia a cero y, finalmente, un proceso **ARMA**( $p,q$ ) tendrá comportamientos irregulares en las dos funciones, **FACP** y **FAC**.

### 3.2 Estimación

Una vez determinados los grados de los polinomios  $AR(p)$  y  $MA(q)$  y el grado de diferenciación  $d$ , el siguiente paso es determinar los valores de los parámetros  $\phi_1, \phi_2, \phi_3, \dots, \phi_p, \theta_1, \theta_2, \theta_3, \dots, \theta_q$  utilizando un método estadístico apropiado.

Los métodos más usuales para la estimación de parámetros son: el de los momentos, el de los mínimos cuadrados y el de la máxima verosimilitud. El *método de los momentos* es el procedimiento más simple para la estimación de los parámetros  $\phi_i$  de los procesos  $AR(p)$ ; pero puede volverse complejo al intentar la estimación de los parámetros  $\theta_j$  de un proceso  $MA(q)$  ya que la condición de invertibilidad limita los valores que permiten una solución aceptable. Por otra parte, el *método de los mínimos cuadrados no lineales* no presenta problemas especiales para la estimación de los parámetros  $\phi_i$  y  $\theta_j$  pero una de sus principales limitantes es que no utiliza toda la información disponible para una estimación más exacta.

Finalmente, con base en el supuesto de que la función de distribución de  $a_t$  es la normal, es posible estimar los parámetros usando *el método de la máxima verosimilitud*. Si se condiciona la función de verosimilitud a los valores iniciales se obtiene una función de verosimilitud condicionada, y si no se condiciona la función se obtiene una estimación de la verosimilitud exacta.

#### 3.2.1 Función de verosimilitud

La principal razón de la importancia de la función de verosimilitud es el *principio de verosimilitud* que señala que (siendo correcto el modelo asumido) todo lo que los datos tienen que decir acerca los parámetros se encuentra expresado en la función de verosimilitud. Las bases para la generación de esta función son las siguientes:

Dada una muestra de  $N$  observaciones de la variable aleatoria  $\mathbf{Z}$  se le asocia una variable aleatoria  $N$ -dimensional cuya función de distribución de probabilidad conocida  $f(\mathbf{Z}|\xi)$  depende de algunos parámetros desconocidos  $\xi$ . En este caso, este conjunto de parámetros  $\xi$  podría estar formado por los  $p+q+1$  parámetros  $(\phi, \theta, \sigma_a)$  del modelo **ARIMA**( $p, d, q$ ). La función de verosimilitud  $\mathcal{L}(\xi|\mathbf{Z})$  permite conocer, a través de los valores observados por las variables aleatorias, la estimación de los parámetros que originaron dichos valores; en este sentido, la función de verosimilitud tiene la misma estructura que  $f(\mathbf{Z}|\xi)$  sólo que en el caso de la verosimilitud los valores conocidos como fijos son las observaciones  $\mathbf{Z}$  y los desconocidos son los del conjunto de parámetros  $\xi$ . Los valores de los parámetros que maximizan el valor de la función de verosimilitud son llamados estimadores máximo verosímiles.

### 3.2.1.1 Función de verosimilitud condicional

El problema general consiste en ajustar un modelo **ARMA**, cuyos parámetros  $\theta$  y  $\phi$  aseguren la estacionaridad e invertibilidad del modelo, a la serie  $W_t$   $t = 1, 2, \dots, n$  con  $n = N-d$  donde  $\{W_t\}$  está definida como en (2.4.1) y es un proceso estacionario,  $d$  es el grado de diferenciación aplicado a la serie original y  $N$  es el número de observaciones de la serie original. Posteriormente, se define la serie  $\{w_t\}$  en la forma  $w_t = W_t - \mu_w$  donde  $\mu_w$  es la media de  $\{W_t\}$  y se representa a la serie como:

$$w_t = w_t - \phi_1 w_{t-1} - \phi_2 w_{t-2} - \dots - \phi_p w_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (3.2.1.1.1)$$

cabe señalar que generalmente se considera  $\mu_w = 0$  por lo que si esto se cumple se tendrá que  $w_t = W_t$ . Es obvio que no se puede realizar una sustitución inmediata de los valores  $w_t$  en (3.2.1.1.1) debido a que se desconocen los valores anteriores a la

primera observación. De cualquier forma, supóngase que se cuenta con los  $p$  valores  $\mathbf{W}_*$  de la variable  $W$  anteriores a la primera observación y los  $q$  valores  $a_*$  de la variable  $a$  entonces, a partir de una elección inicial de parámetros  $(\theta, \phi)$ , es posible calcular sucesivamente a través de (3.2.1.1.1) un conjunto de valores condicionales  $a_t(\theta, \phi | \mathbf{W}_*, a_*, W)$  para los tiempos  $t = 1, 2, \dots, n$  (donde  $W$  es el conjunto de observaciones conocidas de la variable  $W$ ).

La construcción de la función de verosimilitud se hace con base en el supuesto de que la distribución de  $a$  es normal y como los choques son independientes se tiene que:

$$\mathbf{f}(a_1, a_2, \dots, a_n) = (2\pi)^{-n/2} (\sigma_a^2)^{-n/2} \exp\left\{-\sum_{t=1}^n a_t^2 / (2\sigma_a^2)\right\}$$

la función de verosimilitud condicional dado un conjunto de datos  $W$  será:

$$\mathcal{L}(\theta, \phi, \sigma_a^2 | \mathbf{W}_*, a_*, W) = (2\pi)^{-n/2} (\sigma_a^2)^{-n/2} \exp\{-\mathbf{S}_*(\theta, \phi) / (2\sigma_a^2)\}$$

donde

$$\mathbf{S}_*(\theta, \phi) = \sum_{t=1}^n a_t^2(\theta, \phi | \mathbf{W}_*, a_*, W) = \sum_{t=1}^n (\gamma_{t-1} \phi_1 \gamma_{t-1} - \phi_2 \gamma_{t-2} \dots - \phi_p \gamma_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q})^2$$

y por otro lado el logaritmo de la función de verosimilitud es:

$$\mathbf{Ln}_*[\mathcal{L}(\theta, \phi, \sigma_a^2 | \mathbf{W}_*, a_*, W)] = -(n/2) \mathbf{Ln}(2\pi) - (n/2) \mathbf{Ln}(\sigma_a^2) - \mathbf{S}_*(\theta, \phi) / (2\sigma_a^2) \quad (3.2.1.1.2)$$

El  $*$ , en  $\mathbf{S}_*$  y  $\mathbf{Ln}_*$  indica la dependencia que existe entre estas funciones y los valores iniciales  $\mathbf{W}_*$  y  $a_*$ . En (3.2.1.1.2) se nota que la única parte que contiene a los parámetros  $\theta$  y  $\phi$  es la suma de cuadrados de los errores por lo que esta estimación

por máxima verosimilitud es similar a la estimación por mínimos cuadrados y, adicionalmente, para cualquier  $\sigma_e$  dada,  $\mathbf{L}\mathbf{n}_t$  es función lineal de  $\mathbf{S}_t$ .

Por otro lado, la determinación de los valores iniciales de  $\mathbf{W}_t$  y  $\mathbf{a}_t$  se realiza de tal manera que se aproxime a una verosimilitud incondicional. Para la selección de estos valores existen dos métodos:

- a) Igualar  $\mathbf{W}_t$  y  $\mathbf{a}_t$  con sus esperanzas incondicionales, es decir,  $\mathbf{W}_t = 0$  (si  $E(\mathbf{W})=0$ ) y  $\mathbf{a}_t = 0$ ; sin embargo, esta aproximación no es muy buena cuando alguna raíz de  $\phi(B) = 0$  se encuentra cerca de los límites del círculo unitario
- b) Calcular los valores de  $\mathbf{a}_t$  a partir de  $\mathbf{a}_{p+1}, \mathbf{a}_{p+2}, \dots, \mathbf{a}_n$ , igualando con cero los valores de  $\mathbf{a}_t$  anteriores a  $\mathbf{a}_{p+1}$  y utilizando los valores obtenidos para  $\mathbf{W}$ . Usando, este método sólo es posible sumar  $n-p$  valores de  $\mathbf{a}_t$ , sin embargo, la pérdida de  $p$  valores no es muy significativa para valores grandes de  $n$ .

### 3.2.1.2 Función de verosimilitud incondicional

Como se muestra en la obra de Box y Jenkins<sup>1</sup> la función de verosimilitud incondicional o de verosimilitud exacta está dada por.

$$\mathcal{L}(\theta, \phi, \sigma_e^2 | W) = (2\pi\sigma_e^2)^{-n/2} f(\theta, \phi)^{1/2} \exp\{-\mathbf{S}(\theta, \phi)/(2\sigma_e^2)\}$$

cuyo logaritmo es:

$$\mathbf{L}\mathbf{n}[\mathcal{L}(\theta, \phi, \sigma_e^2 | W)] = -(n/2)[\mathbf{L}\mathbf{n}(2\pi) + \mathbf{L}\mathbf{n}(\sigma_e^2)] + (1/2)\mathbf{L}\mathbf{n}f(\theta, \phi) - \mathbf{S}_t(\theta, \phi)/(2\sigma_e^2) \quad (3.2.1.2.1)$$

donde  $\mathbf{S}(\theta, \phi) = \sum_{t=-\infty}^n (\mathbf{a}_t | \theta, \phi, W)^2$  con  $(\mathbf{a}_t | \theta, \phi, W) = E(\mathbf{a}_t | \theta, \phi, W)$  y  $f(\theta, \phi)$  es una función de  $\theta$  y  $\phi$ . Para series grandes  $\mathbf{S}(\theta, \phi)$  tiene mayor peso que  $f(\theta, \phi)$  y por tanto el perímetro

<sup>1</sup>Box & Jenkins, Time Series Analysis: Forecasting and Control, pp. 269-283.

de valores parametrales  $(\theta, \phi)$  correspondiente a  $\mathbf{S}$  está muy cercano al correspondiente a la función de verosimilitud y al logaritmo de la misma, es decir, la estimación por mínimos cuadrados de (3.2.1.2.1) puede proveer de valores muy próximos a los estimadores máximo verosímiles.

Para determinar la suma incondicional de cuadrados los valores de  $a_t$  se calculan recursivamente tomando esperanzas condicionales de (3.2.1.1.1). Para iniciar este cálculo se requieren los valores  $W_j, j=0, 1, 2, \dots$ ; con la finalidad de calcular estos valores se realizan los siguientes pasos:

- a) Expresar el modelo como  $\phi(F)W_t = \theta(F)a_t$  es decir,  $a_t = W_t - \phi_1 W_{t-1} - \phi_2 W_{t-2} - \dots - \phi_p W_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$ .
- b) Comenzar a calcular  $a_t$  a partir de  $t = n-p, n-p-1, n-p-2, \dots$  (a partir del final de la serie descontando las últimas  $p$  observaciones), haciendo  $a_{n-p+i} = 0$  para  $i=1, 2, \dots, q$  y considerando los valores conocidos de  $W_t$ .
- c) Hacer  $a_i = 0$  para  $i = 0, 1, 2, \dots$  y con base en estos valores, en los valores calculados en el punto anterior y en los valores conocidos de  $W_t$  obtener los valores de  $W_i$  para  $i = 0, 1, 2, \dots$ .
- d) Expresar el modelo como  $\phi(B)W_t = \theta(B)a_t$ , es decir,  $a_t = W_t - \phi_1 W_{t-1} - \phi_2 W_{t-2} - \dots - \phi_p W_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$  y con base en los datos obtenidos en los pasos anteriores se realiza ahora el cálculo de  $a_t$  [es cierto que la recursión para el cálculo de  $W_t$  puede ser infinita pero, dado el supuesto de estacionaridad, se considera que existe un entero  $Q$  "razonable" tal que  $W_t \approx 0$  para  $t \leq -Q$  por lo que se hace  $a_j = 0$  para  $j \leq -Q$ ; este valor  $Q$  puede ser detectado durante el cálculo de  $W_t$ ].

- e) La suma de los cuadrados de los errores se realiza a través de las  $a_t \neq 0$  o bien se utiliza  $a_t$  para  $t \geq 1$ .

### 3.2.2 Estimación por el método de mínimos cuadrados: Algoritmos

En esta sección trataré de mostrar algunos métodos para estimar los parámetros  $\theta$  y  $\phi$  que minimicen el valor de  $\mathbf{S}(\theta, \phi)$ ; estos métodos son los que se refieren a la estimación no lineal pues al elevar al cuadrado la expresión (3.2.1.1.1) se generan ecuaciones no lineales que involucran a los parámetros

Como aproximación al estimador máximo verosímil se utiliza la suma  $\mathbf{S}(\theta, \phi)$  definida anteriormente y que, para fines prácticos, puede reemplazarse por.

$$\mathbf{S}(\theta, \phi) = \sum_{t=1-Q}^n (a_t | \theta, \phi, W)^2$$

Sea  $\beta = (\theta, \phi)$  el conjunto de los  $k = p + q$  parámetros de un modelo **ARMA**( $p, q$ ), como lo que se requiere es minimizar  $\mathbf{S}(\theta, \phi) = \mathbf{S}(\beta)$ , se inicia por aproximar cada término  $(a_t | \beta, W)$  de la suma por su expansión en serie de Taylor alrededor de un conjunto de valores parametrales  $\beta_0 = (\beta_{1,0}, \beta_{2,0}, \dots, \beta_{k,0})$ , así:

$$a_t \approx a_{t,0} - \sum_{i=1}^k (\beta_i - \beta_{i,0}) \chi_{i,t} \quad (3.2.2.1)$$

donde

$$\chi_{i,t} \approx \left. \frac{\partial a_t}{\partial \beta_i} \right|_{\beta = \beta_0}$$

y los términos de orden superior a uno han sido despreciados al suponer que tanto las potencias de  $(\beta_i - \beta_{i,0})^2$ ,  $(\beta_i - \beta_{i,0})^3, \dots$  como las derivadas parciales de esos órdenes serán pequeñas en valor absoluto.



Sean  $\mathbf{a}$  y  $\mathbf{a}_0$  los vectores columna de dimensión  $(n+Q) \times 1$  que incluyen a las variables  $\alpha_t$  y  $\alpha_{t,0}$ ,  $\mathbf{X}$  la matriz  $(n+Q) \times k$  formada por las derivadas  $\chi_{it}$ ; basándose en estas equivalencias y en (3.2.2.1) se tiene que:

$$\mathbf{a}_0 \approx \mathbf{X}(\beta - \beta_0) + \mathbf{a}$$

de tal forma que  $\mathbf{S}(\theta, \phi) = \mathbf{S}(\beta) = \mathbf{a}'\mathbf{a}$  es la función que debe minimizarse.

Para calcular esta función lo que hace falta es poder calcular las derivadas  $\chi_{it}$  para lo cual se propone el siguiente método numérico que consiste en evaluar inicialmente  $(\alpha_t | W, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0})$  para  $t = 1-Q, \dots, n$ , después se calcula  $(\alpha_t | W, \beta_{1,0} + \delta_1, \beta_{2,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0})$ , luego  $(\alpha_t | W, \beta_{1,0}, \beta_{2,0} + \delta_2, \dots, \beta_{i,0}, \dots, \beta_{k,0})$  y así sucesivamente hasta  $(\alpha_t | W, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0})$ . Con estos valores se obtienen las aproximaciones deseadas a través de

$$\chi_{it} \approx \frac{(\alpha_t | W, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0}) - (\alpha_t | W, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0})}{\delta_i}$$

Una vez que se ha encontrado la solución que minimiza la función, su matriz de covarianza puede aproximarse por  $(\mathbf{X}'\mathbf{X})^{-1}\sigma_c^2$

### 3.2.2.1 Algoritmo de Marquardt

La función a aproximar es  $\mathbf{S}(\beta)$  cuya expansión por series de Taylor en forma matricial será:

$$\mathbf{S}(\beta) \approx \mathbf{S}(\beta_0) + \mathbf{d}\beta + \frac{1}{2}\beta\mathbf{D}\beta \quad (3.2.2.1.1)$$

con  $\mathbf{d}$  una matriz  $1 \times k$  y  $\mathbf{D}$  una matriz de  $k \times k$ . Si esta aproximación es buena, para pasar de el conjunto de parámetros actuales a otro mejor se calcula:

$$\beta_{\min} \approx \beta_{\text{actual}} + \mathbf{D}^{-1}[\nabla \mathbf{S}(\beta_0)] \quad (3.2.2.1.2)$$

Una aproximación mejor para la actualización anterior será:

$$\beta_{\min} \approx \beta_{\text{actual}} + \text{constante}^*[\nabla \mathbf{S}(\beta_0)] \quad (3.2.2.1.3)$$

El gradiente de  $\mathbf{S}(\beta)$  en su valor mínimo será igual a cero; dicho gradiente tendrá las componentes:

$$\frac{\partial \mathbf{S}}{\partial \beta_i} = 2 \sum_{t=1-Q}^n (a_t | W, \beta) \frac{\partial (a_t | W, \beta)}{\partial \beta_i} \quad i=1,2,\dots,k$$

tomando una derivada parcial adicional:

$$\frac{\partial^2 \mathbf{S}}{\partial \beta_i \partial \beta_j} = 2 \left[ \sum_{t=1-Q}^n (a_t | W, \beta) \frac{\partial^2 (a_t | W, \beta)}{\partial \beta_i \partial \beta_j} + \frac{\partial (a_t | W, \beta)}{\partial \beta_i} \frac{\partial (a_t | W, \beta)}{\partial \beta_j} \right]$$

Se define ahora

$$\varphi_i = \frac{\partial \mathbf{S}}{2 \partial \beta_i}, \quad \alpha_{ij} = \frac{\partial^2 \mathbf{S}}{2 \partial \beta_i \partial \beta_j} \quad (3.2.2.1.4)$$

Tomando la matriz  $\alpha = \frac{1}{2} \mathbf{D}$  (3.2.2.1.2) puede describirse como

$$\varphi_i \approx \sum_{j=1}^k \alpha_{ij} \delta \beta_j \quad (3.2.2.1.5)$$

Este conjunto de ecuaciones es resuelto para los incrementos  $\delta \beta_j$  que agregados a la aproximación actual al valor de los parámetros proporciona los

valores de la siguiente aproximación. Para utilizar la expresión (3.2.2.1.3) se considera que  $1/\lambda\alpha_{ij} = \text{constante}$  tal que  $\delta\beta_j = \varphi_j/\lambda\alpha_j$  esta expresión es equivalente a

$$\delta\beta_j\lambda\alpha_j = \varphi_j \quad (3.2.2.1.6)$$

donde  $\lambda$  es un factor de ajuste adimensional con la posibilidad de ser  $\gg 1$ .

(3.2.2.1.5) y (3.2.2.1.6) pueden ser combinadas si se define una matriz  $\alpha'$  tal que:

$$\alpha'_{jj} = \alpha_j(1+\lambda) \quad \alpha'_{ji} = \alpha_j \quad (j \neq i)$$

y pueden reemplazarse por:

$$\varphi_i \approx \sum_{j=1}^k \alpha'_{ij} \delta\beta_j \quad (3.2.2.1.7)$$

Entonces, cuando  $\lambda$  es muy grande  $\alpha'$  es forzada a convertirse en diagonalmente dominante por lo que (3.2.2.1.5) se aproxima a (3.2.2.1.3) y cuando  $\lambda$  se aproxima a cero,  $\alpha'$  se aproxima a (3.2.2.1.4); así, dados valores iniciales de  $\beta$  el algoritmo de Marquardt consiste en:

- 1 Calcular  $\mathbf{S}(\beta)$  para el valor inicial dado de  $\beta$ .
- 2 Seleccionar  $\lambda$  pequeño (por ejemplo,  $\lambda=0.001$ ).
- 3 Resolver el sistema en (3.2.2.1.7) para  $\delta\beta$  y calcular  $\mathbf{S}(\beta+\delta\beta)$ .
- 4 Determinar si  $\lambda$  se debe incrementar o disminuir.
  - 4.1 Si  $\mathbf{S}(\beta+\delta\beta) \geq \mathbf{S}(\beta)$  aumentar  $\lambda$  por un factor de 10
  - 4.2 Si  $\mathbf{S}(\beta+\delta\beta) < \mathbf{S}(\beta)$  disminuir  $\lambda$  por un factor de 10
- 5 Actualizar la solución de la siguiente iteración, es decir,  $\beta \leftarrow \beta + \delta\beta$  y regresar al paso 3.

- 6 Detenerse cuando  $S(\beta)_{\text{actual}} - S(\beta)_{\text{anterior}} \leq r$  (donde  $r$  es un número pequeño, por ejemplo 0.1). No detenerse después de haber encontrado para algún  $\beta$  que  $S(\beta)$  se incrementa

### 3.2.3 Estimación por momentos: Valores parametrales iniciales

Cuando el modelo identificado es  $AR(p)$  el valor inicial de los  $p$  parámetros puede obtenerse a través de las ecuaciones de Yule-Walker estimadas:

$$\begin{aligned} r_1 &= \phi_1 + \phi_2 r_1 + \dots + \phi_p r_{p-1} \\ r_2 &= \phi_1 r_1 + \phi_2 + \dots + \phi_p r_{p-2} \\ &\dots \dots \dots \\ r_p &= \phi_1 r_{p-1} + \phi_2 r_{p-2} + \dots + \phi_p \end{aligned}$$

es decir,

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \dots \\ \phi_p \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 & \dots & r_{p-2} & r_{p-1} \\ r_1 & 1 & r_1 & \dots & r_{p-3} & r_{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ r_{p-1} & r_{p-2} & r_{p-3} & \dots & r_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ r_2 \\ \dots \\ r_p \end{pmatrix}$$

Cuando el modelo sea  $MA(q)$ , las relaciones entre los parámetros  $\theta$  y las correlaciones  $\rho_1, \rho_2, \dots$  no son lineales por lo que se requiere resolver el sistema de ecuaciones no lineal:

$$\begin{aligned} r_1 &= (-\theta_1 + \theta_1\theta_2 + \dots + \theta_{q-1}\theta_q) / (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \\ r_2 &= (-\theta_2 + \theta_1\theta_3 + \dots + \theta_{q-2}\theta_q) / (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \\ &\dots \dots \dots \\ r_q &= -\theta_q / (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \end{aligned}$$

Este sistema tiene múltiples soluciones pero sólo una de ellas satisfará las condiciones de invertibilidad por lo que esa deberá ser la elegida. Por ejemplo, para un **MA(1)** se tendrá  $r_1 = -\theta_1/(1 + \theta_1^2)$  que tiene las soluciones :

$$\theta = -(2r_1)^{-1} + [(2r_1)^{-2} - 1]^{1/2} \quad \text{y} \quad \theta = -(2r_1)^{-1} - [(2r_1)^{-2} - 1]^{1/2}$$

Para un **MA(2)** se tiene el sistema:

$$r_1 = (-\theta_1 + \theta_1\theta_2)/(1 + \theta_1^2 + \theta_2^2)$$

$$r_2 = -\theta_2/(1 + \theta_1^2 + \theta_2^2)$$

dado que la invertibilidad implica  $\theta_2 - \theta_1 < 1$  y  $\theta_2 + \theta_1 < 1$ , entonces se necesita

$$\theta_1^2 + \theta_2^2 = [(\theta_2 - \theta_1)^2 + (\theta_2 + \theta_1)^2]/2 < 1$$

por lo cual el término  $(1 + \theta_1^2 + \theta_2^2)$  debe tomar algún valor dentro del intervalo (1,2)

que permite la siguiente aproximación  $(1 + \theta_1^2 + \theta_2^2) \cong 1.5$  que da por resultado

$$\theta_1 = -r_1/(0.667 + r_2) \quad \text{y} \quad \theta_2 = -1.5r_2.$$

Para modelos **ARMA(p,q)** se presentan ecuaciones lineales y no lineales que involucran a los parámetros. Para resolver estas ecuaciones se considera el hecho de que  $\phi(B)\rho_k=0$  para  $k>q$  por lo que el sistema:

$$\Gamma_{q+1} = \phi_1\Gamma_q + \phi_2\Gamma_{q-1} + \dots + \phi_p\Gamma_{q-p+1}$$

$$\Gamma_{q+2} = \phi_1\Gamma_{q+1} + \phi_2\Gamma_q + \dots + \phi_p\Gamma_{q-p+2}$$

.....

$$\Gamma_{q+p} = \phi_1\Gamma_{q+p-1} + \phi_2\Gamma_{q+p-2} + \dots + \phi_p\Gamma_q$$

proporciona las estimaciones iniciales para  $\phi_1, \phi_2, \dots, \phi_p$  que deben sustituirse en las primeras  $q$  autocorrelaciones muestrales para poder determinar los valores iniciales  $\theta_1, \theta_2, \theta_3, \dots, \theta_q$ .

En lo tocante al término  $\theta_0$  la estimación de este parámetro por el método de momentos se obtiene a partir de la relación  $E[\phi(B)W_t] = \theta_0$  de forma que

$$\theta_0 = (1 - \phi_1 - \phi_2 - \dots - \phi_p) \underline{W}$$

con  $\underline{W} = \sum_{l=d+p+1}^N W_l / (N - d - p + 1)$ . El hecho de que se presente una tendencia determinista implica que ésta actuará aun fuera del periodo observado; como este supuesto es muy fuerte generalmente se elimina el término determinista y sólo en la etapa de verificación se precisa si dicho término debe incluirse.

### 3.2.4 La varianza del ruido blanco

Como se habrá notado, aunque he incluido a la varianza de los choques aleatorios dentro de los parámetros a determinar, aún no he especificado cómo se calculará dicho valor y es por eso que a continuación explico este aspecto.

Si se revisan las expresiones (3.2.1.1.2) y (3.2.1.2.1) para los logaritmos naturales de las funciones de verosimilitud condicional y exacta se apreciará que los únicos términos donde aparece  $\sigma_\varepsilon^2$  son  $-(n/2)\text{Ln}(\sigma_\varepsilon^2)$  y  $-\mathbf{S}_\varepsilon(\theta, \phi) / 2\sigma_\varepsilon^2$  por lo que si se supone que ya se cuenta con los estimadores correctos de  $\theta$  y  $\phi$  la estimación máximo verosímil de  $\sigma_\varepsilon^2$  se obtiene a partir de.

$$\frac{\partial \text{Ln}[\mathcal{L}(\sigma_\varepsilon^2 | \theta, \phi)]}{\partial \sigma_\varepsilon^2} = \frac{-n}{2\sigma_\varepsilon^2} + \frac{\mathbf{S}(\theta, \phi)}{2\sigma_\varepsilon^4} = 0$$

Donde  $n = N - d$  que produce la estimación

$$S_o^2 = \frac{\mathbf{S}(\theta, \phi)}{n} \quad (3.2.4.1)$$

aunque (3.2.4.1) produce un máximo del logaritmo de la función de verosimilitud, en la práctica se utiliza a (3.2.4.2) que considera la corrección por grados de libertad usados al estimar todos los parámetros del modelo

$$S_o^2 = \frac{\mathbf{S}(\theta, \phi)}{n-p-q-1} \quad (3.2.4.2)$$

### **3.3 Verificación: ¿Es adecuado el modelo seleccionado?**

Una vez que el modelo ha sido identificado y estimado cabe preguntarse si es realmente adecuado. Si es posible encontrar evidencia de que el modelo es seriamente inapropiado, se requerirá saber cómo debe ser modificado para el siguiente ciclo.

En esta sección presento los métodos para diagnosticar qué supuestos cumple el modelo construido. Un método para confirmar si el modelo es adecuado es la sobreparametrización, es decir, calcular los parámetros de un modelo más general que aquél que se había seleccionado lo que nos permitirá saber si el modelo es inadecuado. Adicionalmente es necesario analizar los residuales; en este tipo de análisis los propios datos sugerirán las modificaciones adecuadas.

Es importante recordar que ningún modelo representa absolutamente a la serie verdadera y es por eso que, dados suficientes datos, algunas pruebas estadísticas pueden desacreditar a un modelo aun cuando éste sea suficientemente adecuado para el propósito inmediato, por esto se debe estar preparado para trabajar con modelos que exhiban ligeras fallas en su ajuste.

### 3.3.1 Análisis de residuales

Una vez determinadas las estimaciones paramétricas  $\hat{\theta}$  y  $\hat{\phi}$  del modelo

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

los residuales se estiman a través de

$$\hat{a}_t = \hat{\theta}(B)^{-1}\hat{\phi}(B)W_t$$

considerando  $\hat{\theta}(B)^{-1}\hat{\phi}(B) = \hat{\pi}(B)$  el modelo puede reescribirse como

$$W_t = \hat{a}_t + \hat{\pi}_1 W_{t-1} + \hat{\pi}_2 W_{t-2} + \dots$$

Por otro lado, la estimación de  $W_t$  con base en el polinomio estimado  $\hat{\pi}(B)$  y las observaciones hasta el momento  $t-1$  es

$$\hat{W}_t = \hat{\pi}_1 \hat{W}_{t-1} + \hat{\pi}_2 \hat{W}_{t-2} + \dots$$

A partir de estas relaciones se sigue que

$$\hat{a}_t = W_t - \hat{W}_t \tag{3.3.1.1}$$

esto quiere decir que los errores o residuales se miden como la diferencia entre los valores observados y los estimados por el modelo.

Puede demostrarse que si el modelo es adecuado, los valores  $\hat{a}_t$  se aproximan a los valores  $a_t$  del ruido blanco conforme crece el tamaño de la muestra de la serie. Basándose en lo anterior, los supuestos de  $\{a_t\}$  pueden verificarse y cuando no sean cumplidos podrá corregirse el modelo apoyándose en  $\hat{a}_t$ .



**Supuesto 1:**  $\{a_t\}$  tiene media cero. Para determinar si esto es cierto se calculan primero la media aritmética y la desviación estándar muestral de los residuales utilizando las fórmulas:

$$M(\hat{a}) = \frac{\sum_{t=d+p+1}^N \hat{a}_t}{(N-d-p)} \quad (3.3.1.2)$$

$$\sigma_{\hat{a}} = \left( \frac{\sum_{t=d+p+1}^N (\hat{a}_t - M(\hat{a}))^2}{(N-d-p-q)} \right)^{1/2} \quad (3.3.1.3)$$

posteriormente se evalúa

$$(N-d-p)^{1/2} M(\hat{a}) / \sigma_{\hat{a}} \quad (3.3.1.4)$$

Si el valor absoluto de este cociente es menor que dos se dice que no hay evidencia de que la media del proceso sea distinta de cero y, por tanto, no se rechaza el supuesto; sin embargo, si el valor absoluto es mayor o igual a dos, se dice que la media es significativamente distinta de cero (esto querrá decir que todavía existe una tendencia determinista o semideterminista en la serie) y se ha violado el supuesto.

Para corregir esta violación se prueba si considerando un término autorregresivo o una diferencia más desaparece la tendencia determinista o semideterminista<sup>1</sup>. Si dicha tendencia no desaparece, se procede a incluir en el modelo al término  $\theta_0$ , que deberá ser estimado conjuntamente con los demás

---

<sup>1</sup> Esto surge del hecho de que la diferencia o el parámetro autorregresivo podrían representar a una posible tendencia adaptiva presente en los residuales.

parámetros y cuyo valor inicial será tomado como  $M(\hat{\theta}_t)$ . El término  $\theta_0$  implicará la existencia de una tendencia determinista

**Supuesto 2:**  $\{\hat{\theta}_t\}$  tiene varianza constante. Para establecer si esto es cierto se elabora una gráfica de los residuales contra el tiempo para inspeccionar visualmente si la varianza puede ser considerada constante. Esta revisión tan burda se debe al supuesto de que solamente violaciones muy notorias son las que llegarían a causar problemas.

Para corregir la presencia de un patrón de crecimiento o decrecimiento constante habría que aplicar una transformación potencia a la serie.

**Supuesto 3:** Las variables aleatorias  $\{\hat{\theta}_t\}$  son mutuamente independientes. Para determinar si esto es cierto se analizan las autocorrelaciones residuales mediante la estimación de la FAC muestral de los residuales que, bajo el supuesto de media cero, estará formada por los términos

$$r_k(\hat{\theta}) = \frac{\sum_{t=d+p+1}^{N-k} \hat{\theta}_t \hat{\theta}_{t+k}}{N \sum_{t=d+p+1} \hat{\theta}_t^2} \quad (3.3.1.5)$$

Después se estima la desviación estándar de  $r_k(\hat{\theta})$  a través de

$$\{\text{Var}[r_k(\hat{\theta})]\}^{1/2} \approx 1/(N-d-p)^{1/2} \quad (3.3.1.6)$$

la determinación de la significación estadística individual de las autocorrelaciones de los residuales se realiza a continuación considerando que  $r_k(\hat{\theta}) \sim N(0, [N-d-p]^{-1})$ : Si  $|r_k(\hat{\theta})| \geq 2/[N-d-p]^{1/2}$ , se considera que  $r_k(\hat{\theta})$  es significativamente distinta de cero.

Durbin demostró que, aunque los valores obtenidos con los datos estimados pueden ser una valiosa ayuda para reconocer lo adecuado del modelo, puede ser peligroso utilizar estos valores ya que es posible que se presenten alejamientos importantes respecto de los valores reales. Por su parte Box y Pierce notaron que esta prueba no es completamente válida para autocorrelaciones correspondientes a retrasos pequeños pero sí lo es para retrasos moderados y grandes. Por lo anterior, en estos casos se procede a una prueba conjunta de significación de las primeras  $K$  autocorrelaciones<sup>1</sup> llamada *Prueba Portmanteau* mediante el estadístico  $Q$  de Box y Pierce (1970) definido como:

$$Q = (N-d-p) \sum_{k=1}^K r_k^2(\hat{\alpha})$$

que para  $K > 20$  sigue aproximadamente una distribución  $\chi^2$  con  $K-p-q$  grados de libertad por lo que el valor  $Q$  calculado deberá ser comparado con los valores de las tablas  $\chi^2$  con los grados de libertad correspondientes y la confianza deseada. Una mejor aproximación se obtiene con el estadístico  $Q'$  de Ljung y Box (1978) que se define como.

$$Q' = (N-d-p)(N-d-p+2) \sum_{k=1}^K r_k^2(\hat{\alpha}) / (N-d-p-k) \quad (3.3.1.7)$$

y que tiene aproximadamente la misma distribución que el estadístico  $Q$  por lo que deberá realizarse la misma comparación. Si el valor  $Q'$  (o  $Q$ ) calculado es más

---

<sup>1</sup>  $K$  se selecciona de tal forma que los pesos  $\psi_j$  del modelo  $W_t = \phi(B)^{-1} \theta(B) a_t = \psi(B) a_t$  puedan ser ignorados para  $j > K$ .

grande que el valor correspondiente encontrado en las tablas  $\chi^2$ , se rechaza la hipótesis de que los residuales se comporten como ruido blanco.

Si se llega a que las autocorrelaciones no pertenecen a un proceso de ruido blanco, se debe suponer que los residuales siguen un modelo **ARIMA** y esto implicaría modificaciones al modelo originalmente planteado. Para tratar de identificar el modelo **ARIMA** correspondiente a los residuales se examina la gráfica formada por los valores  $r_k(\hat{\rho})$ . La idea es la siguiente:

Se identificó el modelo

$$\phi'(B)W_t = \theta'(B)b_t \quad \text{donde } \{b_t\} \text{ no es ruido blanco} \quad (3.3.1.8)$$

El modelo **ARIMA** que sigue  $\{b_t\}$  es

$$\phi(B)\nabla^d b_t = \underline{\theta}(B)\underline{\alpha}_t \quad (3.3.1.9)$$

Relacionando las expresiones (3.3.1.8) y (3.3.1.9) para eliminar a  $b_t$  se tiene que

$$\phi'(B)\underline{\phi}(B)\nabla^d W_t = \theta'(B)\underline{\theta}(B)\underline{\alpha}_t$$

es el modelo que nuevamente deberá ser estimado para verificar si es adecuado.

**Supuesto 4:**  $\{\alpha_t\}$  sigue una distribución normal para toda  $t$ . Para determinar si se cumple este supuesto se debe verificar que a lo más  $(N-d-p)/20$  de las observaciones residuales se encuentren fuera del intervalo  $(-2\sigma_\alpha^2, 2\sigma_\alpha^2)$ . En esta verificación también se puede emplear la gráfica de los residuales contra el tiempo. También podría elaborarse un histograma de los residuales para observar la forma de su distribución y si presenta asimetría.

La teoría se basa en el supuesto de que  $\{\alpha_t\}$  debe seguir una distribución normal, pero es obvio que  $\{\alpha_t\}$  no necesariamente sigue una distribución normal por lo que pueden esperarse ligeras desviaciones de este supuesto que no causen mayor problema; sin embargo, si el incumplimiento es muy notorio puede aplicarse la transformación normalizante de Box y Cox descrita en (3.1.1.7).

**Supuesto 5:** *No existen observaciones aberrantes (posiblemente ajenas a la serie en estudio).* Para determinar si se cumple este supuesto también es útil la gráfica de las observaciones residuales pues la existencia de observaciones fuera del intervalo  $(-3\sigma_e^2, 3\sigma_e^2)$  implicaría que, o bien se produjo un evento cuya probabilidad de ocurrencia era de 0.2% aproximadamente (lo que sería muy extraño), o el residual registrado corresponde a una observación que no fue producida por el mismo proceso de la serie en estudio.

Cuando se encuentren observaciones fuera de tal intervalo habrá que revisar si hubo un error en el registro de tales valores o si éstos se debieron a una intervención exógena a la serie; en el primer caso la solución obvia es corregir el dato erróneo y en el segundo caso se debe realizar un análisis específico acerca de la intervención.

### 3.3.2 Análisis independientes de los residuales

**Supuesto 6:** *El modelo considerado es parsimonioso.* Para determinar si todos los parámetros incluidos en el modelo son necesarios y no pueden ser considerados iguales a cero se construyen los intervalos con aproximadamente 95% de confianza del tipo:

$$(\beta - 2[\text{Var}(\beta)]^{1/2}, \beta + 2[\text{Var}(\beta)]^{1/2}) \quad \text{donde } \beta \in \beta = (\phi, \theta) \quad (3.3.2.1)$$

para cada uno de los parámetros y se observa si el valor cero se encuentra dentro del intervalo; si el valor cero es razonable para el parámetro debería de eliminarse éste y volver a estimar el modelo sin incluirlo. De cualquier modo, si se tiene conocimiento de que dicho parámetro debe aparecer en el modelo, es aconsejable mantenerlo aunque la prueba no rechace la posibilidad de que el valor del parámetro sea cero.

**Supuesto 7:** *El modelo considerado es admisible.* Para determinar si el modelo es admisible se requiere que éste sea estacionario e invertible por lo que deberán verificarse las condiciones de estacionaridad y/o invertibilidad correspondientes a los polinomios que intervengan, es decir, trabajar con los determinantes presentados en (2 1.4) y (2 2.4). Para corregir la inadmisibilidad del modelo se emplean los intervalos definidos en (3.3.2.1) ya que las estimaciones que se tienen, al ser puntuales, pueden no ser las que se necesiten. Lo que se requiere entonces es identificar si existe algún o algunos parámetros que al considerar otro valor dentro del intervalo correspondiente vuelvan estacionario e invertible al modelo y, posteriormente, realizar la modificación en este.

**Supuesto 8:** *El modelo considerado es estable en los parámetros.* Lo que se busca en este punto es identificar si existen correlaciones altas (negativas o positivas) entre los parámetros del modelo, es decir, si la modificación de un parámetro no llevará a la modificación de otro parámetro con el que esté fuertemente relacionado. El cálculo de las correlaciones entre parámetros se hará a través de:

$$\text{Corr}(\beta_1, \beta_2) = \frac{\text{Cov}(\beta_1, \beta_2)}{[\text{Var}(\beta_1)\text{Var}(\beta_2)]^{1/2}} \quad \text{donde } \beta_i \in \beta = (\phi, \theta) \quad i = 1, 2, \dots, p+q \quad (3.3.2.2)$$

La principal causa de la inestabilidad de los parámetros es la redundancia, característica que se definirá más adelante. Para evitar esta inestabilidad será necesario eliminar uno de los parámetros con alta correlación, o bien, identificar si existen factores de regresión o de promedios móviles que, por ser muy semejantes, puedan cancelarse. Si la inclusión de dos parámetros con alta correlación entre sí es necesaria para la adecuada descripción del modelo, no quedará si no aceptar la existencia de tal correlación entre los parámetros del modelo.

### 3.3.3 Sobreparametrización

La sobreparametrización implica el conocimiento del tipo de discrepancias que pueden ser esperadas y consiste en ajustar un modelo más elaborado que aquél que fue inicialmente seleccionado. La decisión de si el modelo debe ser aumentado o no en alguna de sus componentes debe ser hecha concienzudamente, recuérdese que en concordancia con la redundancia de parámetros del modelo no se deben añadir simultáneamente factores en ambas componentes de un modelo **ARMA**.

Dado el modelo

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

es idéntico al modelo

$$(1-\nu B)\phi(B)W_t = (1-\nu B)\theta(B)a_t$$

en el cual los dos operadores, el autorregresivo y el de promedios móviles, son afectados por el mismo factor. Esta es la idea que enmarca la redundancia de parámetros. Como la estimación es más problemática conforme el modelo se vuelve más complejo, es necesario evitar estas redundancias o casi-redundancias de parámetros.

Si se detecta que el modelo requiere de los factores que se le incorporaron, es decir, que los parámetros de estos factores son diferentes de cero, obviamente el modelo deberá ser sustituido por el nuevo modelo sugerido.

### 3.3.4 Otras observaciones

Si alguna de las raíces de  $\phi(B)$  es muy próxima a la unidad, la serie original puede estar subdiferenciada por lo que se requerirá de al menos una diferenciación adicional, por el contrario, si alguna de las raíces de  $\theta(B)$  es muy próxima a la unidad, la serie original puede estar sobrediferenciada por lo que se requerirá la *eliminación de alguna diferencia*.

Otra medida que ayuda a verificar el grado de la bondad del ajuste realizado es el coeficiente de determinación  $\mathcal{R}^2$ , el cual se define en este caso como

$$\mathcal{R}^2 = 1 - \frac{\sum_{t=1}^n a_t^2}{\sum_{t=1}^n W_t^2}$$

Este estadístico representa el porcentaje de la varianza de la serie temporal que es explicado por los parámetros (autorregresivos y de promedios móviles) del modelo, exceptuando, obviamente, el nivel de la serie. El valor de este estadístico debiera ser lo más cercano a la unidad para que pudiera considerarse que el modelo es una buena representación. Como  $\mathcal{R}^2$  es muy sensible al grado  $d$  de diferenciación aplicado, no es muy aconsejable decidir si un modelo es adecuado o no con base en este estadístico.



## 4 Aplicación del modelo en pronósticos

A continuación presento la teoría que sustenta la obtención de pronósticos; esta teoría es aplicable directamente sólo a series no estacionales pero su aplicación para modelos estacionales del tipo multiplicativo es casi directa.

El problema que abordaré será el de calcular, para una serie  $\{Z_t\}$  (con  $N$  observaciones) el pronóstico óptimo  $\hat{z}_t(h)$  para el valor de la serie al tiempo  $t+h$  ( $Z_{t+h}$ ) conociendo los valores de la serie hasta el tiempo  $t$  y considerando que a esta serie se le han aplicado una transformación  $T$  y cierta diferenciación  $d$ . Recuérdese que el modelo asignado a la serie es estacionario e invertible y corresponde a un modelo **ARMA**.

El pronóstico de valores futuros de la serie se hace bajo el supuesto de que ya se cuenta con el modelo correcto para representarla y considerando que los errores de estimación en los parámetros no afectarán seriamente los pronósticos, a menos que el número de observaciones utilizado para tal estimación haya sido pequeño.

### 4.1 Pronósticos óptimos para series estacionarias

El criterio que se utiliza para determinar lo óptimo del pronóstico de  $Z_{t+h}$  es el error cuadrático medio mínimo, es decir, si  $\hat{z}_t(h)$  es el pronóstico óptimo de la serie para el punto  $Z_{t+h}$  a partir de un origen  $t$  y  $\tilde{z}_t(h)$  es cualquier otro pronóstico para el mismo punto y con el mismo origen, se satisface la relación

$$E_t [Z_{t+h} - \hat{z}_t(h)]^2 = \text{Mín}_{\tilde{z}_t(h)} E_t [Z_{t+h} - \tilde{z}_t(h)]^2$$

donde  $E_t$  denota la esperanza condicional dada toda la información al momento  $t$ , es decir,

$$E_t [Z_{t+h} - \hat{\varepsilon}_t(h)]^2 = E\{[Z_{t+h} - \hat{\varepsilon}_t(h)]^2 | Z_t, Z_{t-1}, \dots\}$$

Recuérdese que se puede representar al modelo **ARIMA**  $\phi(B)\nabla^d Z_t = \theta(B)a_t$  bajo las dos siguientes formas explícitas:

- En forma directa en términos de una ecuación en diferencia:

$$Z_{t+h} = \varphi_1 Z_{t+h-1} + \varphi_2 Z_{t+h-2} + \varphi_3 Z_{t+h-3} + \dots + \varphi_{p+d} Z_{t+h-p-d} + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \theta_3 a_{t+h-3} - \dots - \theta_q a_{t+h-q} \quad (4.1.1)$$

donde  $\varphi(B) = \phi(B)\nabla^d$  y por tanto si el grado del polinomio  $\phi(B)$  es  $p$ , el grado  $\varphi(B)$  de será  $p+d$

- Como una suma infinita ponderada de choques aleatorios:

$$Z_{t+h} = \sum_{j=0}^{\infty} \psi_j a_{t+h-j} = \sum_{j=0}^{h-1} \psi_j a_{t+h-j} + \sum_{j=h}^{\infty} \psi_j a_{t+h-j} \quad (4.1.2)$$

donde  $\psi(B) = \varphi^{-1}(B)\theta(B)$ ,  $\psi_0=1$  y los valores  $\psi_j$   $j = 1, 2, 3, \dots$  se obtienen igualando los elementos con la misma potencia de  $B$  en la relación

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \dots) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q).$$

Para calcular el pronóstico óptimo  $\hat{\varepsilon}_t(h)$  se toma en cuenta esta última representación en términos de choques aleatorios por lo que resulta obvio que la estructura del pronóstico también deberá presentar una combinación lineal y, por lo tanto, se parte de que  $\hat{\varepsilon}_t(h)$  tiene la forma

$$\bar{z}_t(h) = C_h a_t + C_{h+1} a_{t-1} + C_{h+2} a_{t-2} + \dots = \sum_{j=h}^{\infty} C_j a_{t+h-j}$$

y para determinar el pronóstico óptimo habrá que encontrar los valores de  $C_h$  para los cuales la esperanza del cuadrado de los errores (obtenidos con base en (4.1.2)) sean mínimos, así.

$$Z_{t+h} - \bar{z}_t(h) = \sum_{j=0}^{h-1} \psi_j a_{t+h-j} + \sum_{j=h}^{\infty} \psi_j a_{t+h-j} - \sum_{j=h}^{\infty} C_j a_{t+h-j} = \sum_{j=0}^{h-1} \psi_j a_{t+h-j} + \sum_{j=h}^{\infty} (\psi_j - C_j) a_{t+h-j} \quad (4.1.3)$$

en esta expresión se presenta como primer sumando al conjunto de la información desconocida al tiempo  $t$  necesaria para el pronóstico al tiempo  $t+h$ , como segundo sumando el conjunto de la información conocida al tiempo  $t$  y como tercer sumando al estimador del pronóstico. Por otro lado, considerando que  $E_t(a_i a_j) = 0$  cuando  $i \neq j$  se tiene

$$E_t \{ [Z_{t+h} - \bar{z}_t(h)]^2 \} = \sum_{j=0}^{h-1} \psi_j^2 \sigma_e^2 + \sum_{j=h}^{\infty} (\psi_j - C_j)^2 \sigma_e^2$$

el mínimo de esta expresión se obtiene haciendo  $\psi_j = C_j$  con lo que se obtiene el pronóstico óptimo

$$\bar{z}_t(h) = \sum_{j=h}^{\infty} \psi_j a_{t+h-j} \quad (4.1.4)$$

Un resultado interesante se obtiene si se considera que

$$E_t(\alpha_{t+h_j}) = \begin{cases} \alpha_{t+h_j} & \text{Si } j \geq h \\ 0 & \text{Si } j < h \end{cases} \quad (4.1.5)$$

con lo que

$$E_t [Z_{t+h}] = E_t \left[ \sum_{j=0}^{h-1} \psi_j \alpha_{t+h-j} \right] + E_t \left[ \sum_{j=h}^{\infty} \psi_j \alpha_{t+h-j} \right] = \sum_{j=h}^{\infty} \psi_j \alpha_{t+h-j}$$

esto indica que el error cuadrático medio mínimo se obtiene a través de la esperanza condicional de  $Z_{t+h}$ , es decir,

$$\bar{z}_t(h) = E_t [Z_{t+h}]$$

El error de este pronóstico con base en (4.1.3) está dado por

$$e_t(h) = Z_{t+h} - \bar{z}_t(h) = \sum_{j=0}^{h-1} \psi_j \alpha_{t+h-j} \quad (4.1.6)$$

La esperanza y la varianza de este error están dadas por

$$E_t [e_t(h)] = 0 \quad \text{Var}_t [e_t(h)] = \sum_{j=0}^{h-1} \psi_j^2 \sigma_\alpha^2 \quad (4.1.7)$$

con base en estas expresiones puede deducirse que el pronóstico es insesgado y que entre más alejado se desee el pronóstico mayor será la varianza (menor la precisión) del mismo.

Por (4.1.6) el error a horizonte 1 está dado por

$$e_t(1) = Z_{t+1} - \bar{z}_t(1) = \alpha_{t+1} \quad (4.1.8)$$

lo que implica que este error en particular no está correlacionado pues si lo estuviera este valor podría ser predicho a partir de los choques aleatorios precedentes y, en consecuencia, si dicha predicción estuviera dada por  $a_{t+1}$  se tendría que  $\varepsilon_t(h) + a_{t+1}$  sería un mejor pronóstico para  $Z_{t+h}$ . No ocurre lo mismo para horizontes mayores; para  $h > 1$ , en general, los errores están correlacionados<sup>1</sup>: los errores  $e_t(h)$  y  $e_{t_j}(h)$  (para diferentes orígenes pero con el mismo horizonte) están correlacionados, así como  $e_t(h)$  y  $e_t(h+j)$  (para diferentes horizontes pero con el mismo origen) también lo están. Esta estructura correlacionada de los errores explica la tendencia de los pronósticos de permanecer casi siempre por arriba o casi siempre por debajo de los valores reales cuando estos son observados

Para obtener los pronósticos deseados se parte de la representación del proceso en forma de ecuación en diferencia como en (4.1.1) por lo tanto

$$\begin{aligned}
 E_t [Z_{t+h}] &= E_t [\varphi_1 Z_{t+h-1} + \varphi_2 Z_{t+h-2} + \varphi_3 Z_{t+h-3} + \dots + \varphi_{p+d} Z_{t+h-p-d} \\
 &\quad + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \theta_3 a_{t+h-3} - \dots - \theta_q a_{t+h-q}] \\
 &= \varphi_1 E_t [Z_{t+h-1}] + \varphi_2 E_t [Z_{t+h-2}] + \varphi_3 E_t [Z_{t+h-3}] \\
 &\quad + \dots + \varphi_{p+d} E_t [Z_{t+h-p-d}] + E_t [a_{t+h}] - \theta_1 E_t [a_{t+h-1}] \\
 &\quad - \theta_2 E_t [a_{t+h-2}] - \theta_3 E_t [a_{t+h-3}] - \dots - \theta_q E_t [a_{t+h-q}]
 \end{aligned} \tag{4.1.9}$$

donde, como resultado de (4.1.5) y (4.1.8), se tiene que

$$E_t [Z_{t+h-j}] = \begin{cases} Z_{t+h-j} & \text{Si } j \geq h \\ \varepsilon_t(h-j) & \text{Si } j < h \end{cases} \quad E_t [a_{t+h-j}] = \begin{cases} Z_{t+h-j} - \varepsilon_{t+h-j-1}(1) & \text{Si } j \geq h \\ 0 & \text{Si } j < h \end{cases} \tag{4.1.10}$$

<sup>1</sup> Box y Jenkins, Time Series Analysis. Forecasting and Control, pp. 158 –160.

estas expresiones son importantes ya que en ellas aparecen los valores observados de la serie y los pronósticos estimados para ésta y no los choques aleatorios no observables como en (4.1.5).

Para obtener el pronóstico  $\hat{z}_t(h)$  generalmente se expresa al modelo con el formato (4.1.9) y se siguen estos pasos:

- 1 Los valores ya conocidos se toman sin cambios, es decir, se mantienen los valores  $Z_{t+h-j}$  para  $j \geq h$ .
- 2 Los valores que todavía no se han observado se reemplazan por su estimación óptima, es decir, se mantienen los valores  $Z_{t+h-j} = \hat{z}_t(h-j)$  para  $j < h$ .
3. Los valores  $a_{t+h-j}$  para  $j \geq h$  se calculan a partir de  $Z_{t+h-j} - \hat{z}_{t+h-j-1}(1)$
4. Los valores  $a_{t+h-j}$  para  $j < h$  se toman iguales a cero.

Este método es recursivo y se calcula hasta llegar a  $\hat{z}_0(1)$ , como ya no se cuenta con información para calcular este valor, se suponen los  $q$  valores  $a_t$  iniciales iguales a cero y los valores  $\hat{z}_i(1)$   $i = 0, 1, 2, \dots, p$  se igualan con la media de las observaciones de  $Z_t$ .

También es posible obtener el pronóstico a través de la representación en términos exclusivamente de choques aleatorios; para obtenerlo se siguen los mismos pasos, sólo que la base para obtener el pronóstico será la siguiente:

$$\hat{z}_t(h) = E_t [Z_{t+h}] = \sum_{j=0}^{\infty} \psi_j E_t [a_{t+h-j}] = \sum_{j=0}^{h-1} \psi_j E_t [a_{t+h-j}] + \sum_{j=h}^{\infty} \psi_j E_t [a_{t+h-j}] \quad (4.1.11)$$

## 4.2 Pronósticos óptimos para series transformadas

Recuérdese que generalmente la serie  $\{Z_t\}$  original es no estacionaria en varianza, razón por la cual se recurre a la aplicación de una transformación a fin de lograr esa estacionaridad, de tal modo que el modelo **ARIMA** se ajusta a la serie transformada  $\{T(Z_t)\}$  y no a  $\{Z_t\}$ . Como consecuencia de esto, los pronósticos serán aplicables a la serie  $\{T(Z_t)\}$  bajo las siguientes características.

Se toma como base al modelo **ARIMA**  $\varphi(B)T(Z_t) = \theta(B)a_t$  donde  $\varphi(B) = \phi(B)\nabla^d$  cuya forma explícita será

$$T(Z_{t+h}) = \varphi_1 T(Z_{t+h-1}) + \varphi_2 T(Z_{t+h-2}) + \dots + \varphi_{p+d} T(Z_{t+h-p-d}) + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \dots - \theta_q a_{t+h-q}$$

y el estimador mínimo cuadrático  $\mathcal{I}(Z_t)(h)$  determinado por la esperanza condicional de  $T(Z_{t+h-1})$  dada la información a tiempo  $t$  será

$$\begin{aligned} E_t [T(Z_{t+h})] &= E_t [\varphi_1 T(Z_{t+h-1}) + \varphi_2 T(Z_{t+h-2}) + \varphi_3 T(Z_{t+h-3}) + \dots + \\ &\quad \varphi_{p+d} T(Z_{t+h-p-d}) + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \theta_3 a_{t+h-3} - \dots - \\ &\quad \theta_q a_{t+h-q}] \tag{4.2.1} \\ &= \varphi_1 E_t [T(Z_{t+h-1})] + \varphi_2 E_t [T(Z_{t+h-2})] + \varphi_3 E_t [T(Z_{t+h-3})] \\ &\quad + \dots + \varphi_{p+d} E_t [T(Z_{t+h-p-d})] + E_t [a_{t+h}] - \theta_1 E_t [a_{t+h-1}] - \\ &\quad \theta_2 E_t [a_{t+h-2}] - \theta_3 E_t [a_{t+h-3}] - \dots - \theta_q E_t [a_{t+h-q}] \end{aligned}$$

donde, como se estableció antes,

$$E_t [T(Z_{t+h_j})] = \begin{cases} T(Z_{t+h_j}) & \text{Si } j \geq h \\ \mathcal{I}(Z_t)(h-j) & \text{Si } j < h \end{cases} \quad E_t (a_{t+h_j}) = \begin{cases} T(Z_{t+h_j}) - \mathcal{I}(Z_{t+h_j-1})(1) & \text{Si } j \geq h \\ 0 & \text{Si } j < h \end{cases} \tag{4.2.2}$$





Con lo que se concluye que los pronósticos están determinados por el esquema autorregresivo generalizado del proceso **ARIMA** respectivo, mientras que la parte de promedios móviles sirve exclusivamente para fijar los valores iniciales que especifican la solución particular de la ecuación en diferencia. Mediante el empleo de este método recursivo aplicado sobre la representación en términos de ecuación en diferencia se tiene que el pronóstico depende de los valores de  $Z_{t+h}$  para  $h=1, 2, \dots, q$  pero para  $h>q$  la dependencia se registra sobre los pronósticos previamente calculados.

A partir de los pronósticos obtenidos para  $\{T(Z_t)\}$  se requiere ahora obtener el pronóstico para  $\{Z_t\}$ . La primera aproximación que surge para obtener tales valores es aplicar la función inversa de la transformación con lo que el pronóstico para  $\{Z_t\}$  vendría dado por  $T^{-1}[T(Z_t)(h)]$ ; sin embargo, si la transformación utilizada es no lineal, al utilizar esa función inversa no se preservan necesariamente las propiedades óptimas del pronóstico de  $Z_{t+h}$ .

Como señalé en la sección 3.1, es común que la función transformadora sea la función potencia. Para el caso de esta transformación existe un factor que corrige *aproximadamente el sesgo* causado por la función inversa; este factor puede estimarse como<sup>1</sup>:

$$G_{\lambda}(h) = \begin{cases} \frac{1}{2} + \left[ \frac{[1-2\lambda(\lambda-1)\{1+\lambda T(Z_t)(h)\}^2 \text{Var}_t(e_t(h))]}{2} \right]^{1/\lambda} & \text{si } \lambda \leq 1 \text{ y } \lambda \neq 0 \\ \exp \{ \text{Var}_t(e_t(h))/2 \} & \text{si } \lambda = 0 \end{cases} \quad (4.2.3)$$

<sup>1</sup>Guerrero Víctor M., Análisis estadístico de series de tiempo económicas, pp. 227-230.

cuando se aplique la transformación propuesta por Box – Cox mostrada en (3.1.1.7), o bien,

$$G_{\lambda}(h) = \begin{cases} \frac{1}{2} + \left[ \frac{[1-2\lambda^{-1}(\lambda-1)]\{\mathbf{T}(Z_t)(h)\}^{-2} \mathbf{Var}_t(e_t(h))\}^{1/2}}{2} \right]^{1/\lambda} & \text{si } \lambda \leq 1 \text{ y } \lambda \neq 0 \\ \exp \{ \mathbf{Var}_t(e_t(h))/2 \} & \text{si } \lambda = 0 \end{cases} \quad (4.2.4)$$

si la transformación utilizada fue la propuesta por Bartlett y señalada en (3.1.1.6). En ambas expresiones se considera  $\mathbf{Var}_t(e_t(h))$  como la estimación de (4.1.7) y por tanto el pronóstico aproximadamente insesgado de  $Z_{t+h}$  estaría dado como

$$\hat{\varepsilon}_t(h) = \{\mathbf{T}^{-1}[\mathbf{T}(Z_t)(h)]\} G_{\lambda}(h) \quad (4.2.5)$$

### 4.3 Intervalos de confianza para los pronósticos

En lo referente a los límites de confianza se tiene que, dado que se considera que  $e_t \sim N(0, \sigma_e^2)$ , la distribución de la probabilidad condicional  $f(Z_{t+h} | Z_t, Z_{t-1}, Z_{t-2}, \dots)$  de un valor futuro debe ser  $N[\hat{\varepsilon}_t(h), (\psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2)\sigma_e^2]$ . Tomando la estimación  $S_e^2$  definida en (3.2.4.2) para  $\sigma_e^2$ , los límites al  $(1-\alpha)*100\%$  de confianza estarán dados por

$$\hat{\varepsilon}_t(h) \pm u_{\alpha/2} \left[ \sum_{j=0}^{h-1} \psi_j^2 \right]^{1/2} S_e \quad (4.3.1)$$

donde  $\mathbf{P}(U > u_{\alpha/2}) = \alpha/2$  y  $U \sim N(0, 1)$ .

La interpretación de estos límites es que, dada toda la información al tiempo  $t$ , la probabilidad de que el valor real de  $Z_{t+h}$ , cuando ocurra, esté dentro del intervalo es de  $(1-\alpha)*100\%$

Obviamente, en este punto también debe considerarse la corrección por sesgamiento que se aplica a las series transformadas. En este caso, los intervalos de confianza serán corregidos como se indica a continuación

$$G_{1,\lambda}(h)T^{-1}\{T(Z_t)(h) \pm u_{\alpha/2}(\text{Var}_t(e_t(h)))^{1/2}\}$$

$$G_{1,\lambda}(h)T^{-1}\left\{T(Z_t)(h) \pm u_{\alpha/2} \left( \sum_{j=0}^{h-1} \psi_j^2 \right)^{1/2} S_e \right\} \quad (4.3.2)$$

este será el intervalo al  $(1-\alpha)*100\%$  de confianza aproximadamente.

#### 4.4 Para actualizar los pronósticos

Como anticipé en la sección 4.1, los valores  $\psi_j$  del polinomio  $\psi(B)$  se obtienen igualando los elementos con la misma potencia de B en la relación

$$\varphi(B)\psi(B) = \theta(B) \quad \Rightarrow$$

$$(1 - \varphi_1B - \varphi_2B^2 - \dots - \varphi_{p+d}B^{p+d})(1 + \psi_1B + \psi_2B^2 + \dots) = (1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q) \quad \Rightarrow$$

$$1 - \varphi_1B - \varphi_2B^2 - \dots - \varphi_{p+d}B^{p+d} + \psi_1B - \psi_1\varphi_1B^2 - \psi_1\varphi_2B^3 - \dots - \psi_1\varphi_{p+d}B^{p+d+1} + \psi_2B^2 - \psi_2\varphi_1B^3 - \psi_2\varphi_2B^4 - \dots - \psi_2\varphi_{p+d}B^{p+d+2} + \psi_3B^3 - \dots = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$$

así, siendo conocidos los valores de  $\varphi$  y  $\theta$ , los valores de  $\psi$  se obtienen despejándolos de:

$$-\varphi_1B + \psi_1B = -\theta_1B \quad \Rightarrow \quad \psi_1 = \varphi_1 - \theta_1$$

$$-\varphi_2B^2 - \psi_1\varphi_1B^2 + \psi_2B^2 = -\theta_2B^2 \quad \Rightarrow \quad \psi_2 = \varphi_1\psi_1 + \varphi_2 - \theta_2$$

$$-\varphi_3B^3 - \psi_1\varphi_2B^3 - \psi_2\varphi_1B^3 + \psi_3B^3 = -\theta_3B^3 \quad \Rightarrow \quad \psi_3 = \varphi_1\psi_2 + \varphi_2\psi_1 + \varphi_3 - \theta_3$$

.....

y en general

$$\psi_j = \varphi_1 \psi_{j-1} + \varphi_2 \psi_{j-2} + \varphi_3 \psi_{j-3} + \dots + \varphi_{p+d} \psi_{j-p-d} - \theta_j$$

con  $\psi_0 = 1$ ,  $\psi_j = 0$  para  $j < 0$  y  $\theta_j = 0$  para  $j > q$ . Utilizando estos pesos calculados en forma recursiva es posible actualizar los valores pronosticados una vez que se cuenta con una nueva observación de la serie. Para realizar esta actualización obsérvese que

$$\begin{aligned} \bar{\varepsilon}_{t+1}(h) &= \psi_h \varepsilon_{t+1} + \psi_{h+1} \varepsilon_t + \psi_{h+2} \varepsilon_{t-1} + \psi_{h+3} \varepsilon_{t-2} + \dots \\ \bar{\varepsilon}_t(h+1) &= \psi_{h+1} \varepsilon_t + \psi_{h+2} \varepsilon_{t-1} + \psi_{h+3} \varepsilon_{t-2} + \dots \end{aligned}$$

de donde se sigue que

$$\bar{\varepsilon}_{t+1}(h) = \bar{\varepsilon}_t(h+1) + \psi_h \varepsilon_{t+1} \quad (4.4.1)$$

En otras palabras, si se cuenta con los pronósticos a horizonte 1, 2, ..., H en el origen t, entonces tan pronto como se conoce  $Z_{t+1}$  es posible calcular el valor del choque  $\varepsilon_{t+1} = Z_{t+1} - \bar{\varepsilon}_t(1)$  y con ello obtener los pronósticos  $\bar{\varepsilon}_{t+1}(h)$  conforme (4.2.2) para el origen t+1 y los horizontes  $h=1, 2, \dots, H-1$  que vienen a ser las actualizaciones de  $\bar{\varepsilon}_t(h)$  para el origen t y los horizontes  $h=2, 3, \dots, H$ . El nuevo pronóstico a horizonte H ( $\bar{\varepsilon}_{t+1}(h)$ ) se calcula utilizando el procedimiento descrito en la sección 4.1 y los pronósticos obtenidos

#### **4.5 Capacidad de pronóstico del modelo**

Cuando se ha detectado y estimado cuál será el mejor modelo para la serie en estudio es conveniente comparar éste con algún otro modelo, que en algún momento haya parecido ser también adecuado, con la finalidad de analizar la capacidad de pronóstico del mismo, la cual se medirá a través de la comparación entre valores pronosticados para la serie y valores observados. Si la serie  $\{T(Z_t)\}$  consta de N

valores, se pueden efectuar  $Y \geq 1$  pronósticos para el mismo origen  $v=N-Y$  y horizontes  $h=1, 2, \dots, Y$ ; o bien, para los diferentes orígenes  $v= N-Y, N-Y+1, \dots, N-1$  y a horizonte 1. En el primer caso se habla de una *simulación dinámica* y en el segundo de una *simulación estática*. Las diferencias entre los pronósticos y las observaciones reales permitirán descubrir la existencia de sub o sobreestimación en los pronósticos para lo que se calculan las medias de los errores de pronóstico:

$$e_v(h) = Z_{v+h} - \hat{z}_v(h)$$

o de

$$e_{v+h-1}(1) = T(Z_{v+h}) - \hat{T}(Z_{v+h-1})(1) \quad h=1, 2, \dots, Y$$

como

$$\mathcal{M}_1 = \frac{\sum_{h=1}^Y e_v(h)}{Y} \quad \text{o} \quad \mathcal{M}_2 = \frac{\sum_{h=1}^Y e_{v+h-1}(1)}{Y} \quad (4.5.1)$$

Otra medida de bondad de los pronósticos es el error cuadrático medio que se calcula a través de:

$$\text{ECM}_1 = \frac{\sum_{h=1}^Y [e_v(h)]^2}{Y} \quad \text{o} \quad \text{ECM}_2 = \frac{\sum_{h=1}^Y [e_{v+h-1}(1)]^2}{Y} \quad (4.5.2)$$

Es claro entonces que entre más cercanas sean las medidas  $\mathcal{M}_1$ ,  $\mathcal{M}_2$ ,  $\text{ECM}_1$  y  $\text{ECM}_2$  a cero, mejores serán los pronósticos, en otras palabras, mejor será la capacidad de pronóstico del modelo.

Debe notarse que (4.5.1) y (4.5.2) no son totalmente válidas si los pronósticos empleados en este apartado fueron aplicados para medir observaciones que ya

habían sido previamente consideradas en la etapa de estimación del modelo; esto se debe a que en la etapa de estimación se procuró disminuir el error entre los pronósticos un periodo hacia delante y sus valores reales, en consecuencia, los pronósticos utilizados tenderán a presentar un error pequeño.

Una medida que suele tomarse para evitar esta falta de validez es cortar la serie para identificar y estimar el modelo sin utilizar toda la información disponible, es decir, usar  $N-Y$  observaciones solamente y después de calcular los  $Y$  pronósticos compararlos con las  $Y$  observaciones no utilizadas para determinar, ahora sí con toda la validez, si el modelo tiene una adecuada capacidad de pronóstico.

Cuando en esta sección se pronostican valores para observaciones ya consideradas en la identificación y estimación del modelo se dice que se realiza un *análisis en el periodo muestral*, en tanto que cuando la serie es cortada para identificar y estimar el modelo se habla de un *análisis en periodo postmuestral*.

## 5 Un caso especial: Series de tiempo estacionales

Un tipo especial de no estacionaridad presente en muchas series, sobre todo de carácter económico, es la estacionalidad que, como describí en el apartado 1.2, es un fenómeno que consiste en la presencia en la serie de cualquier comportamiento periódico con duración menor o igual a un año que se registre regularmente a lo largo del tiempo.

No es posible representar este tipo de series como un simple modelo **ARIMA** debido a esa característica estacional y es por eso que a continuación presento la metodología aplicable en estos casos

### 5.1 Operadores y modelos estacionales

Para representar las series estacionales se introduce el *operador de diferencia estacional*  $\nabla^k_E$  definido como

$$\nabla^k_E = (1-B^E)^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-jE} \quad \begin{matrix} k=0,1,2, \dots \\ E=1,2,\dots \end{matrix}$$

y el *polinomio de retraso estacional* de orden  $k$  con coeficientes constantes  $g_1, g_2, \dots, g_k$  que está dado por

$$G(B^E) = 1 - g_1 B^E - g_2 B^{2E} - \dots - g_k B^{kE} = \sum_{j=0}^k g_j B^{jE}$$

Al aplicarse el operador  $\nabla^D_E$  se pierden  $D \times E$  observaciones.

Utilizando estas notaciones se puede describir un modelo **ARIMA(P,D,Q)<sub>E</sub>** puramente estacional como

$$\Phi(B^E)\nabla^D_{E}(Z_t-\mu) = \Theta(B^E)a_t$$

donde  $\mu$  es el nivel de  $Z_t$ ,  $\Phi(B^E)$  es un polinomio autorregresivo estacional de orden  $P$ ,  $\Theta(B^E)$  es un polinomio de promedios móviles estacional de orden  $Q$  y  $a_t$  es ruido blanco. El comportamiento de la **FAC** de un proceso **ARIMA**( $P,D,Q$ ) $_E$  es análogo al de la **FAC** de un modelo **ARIMA**( $p,d,q$ ) no estacional por lo que la metodología de construcción revisada en el capítulo tres se aplica sin mayor modificación a los procesos puramente estacionales. Este hecho puede no ser del todo útil si se toma en cuenta que en la práctica no existen muchas series puramente estacionales, si no que mostrarán tanto efectos estacionales como no estacionales.

## 5.2 Modelo multiplicativo estacional

Muchas series temporales con estructura **ARIMA** estacional también tienen estructura **ARIMA** regular; estas estructuras pueden incorporarse aditivamente, pero para que el modelo resulte más poderoso, las estructuras deben combinarse multiplicativamente. Box y Jenkins propusieron el modelo general de la forma:

$$\Phi(B^E)\nabla^D_{E}(Z_t-\mu) = \Theta(B^E)\alpha_t$$

donde  $\alpha_t$  no es ruido blanco, si no un proceso que sigue un modelo **ARIMA**( $p,d,q$ )

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

con estas dos expresiones se obtiene el modelo **ARIMA** multiplicativo estacional representado por **ARIMA**( $p,d,q$ ) $\times$ ( $P,D,Q$ ) $_E$



$$\phi(B)\Phi(B^E)\nabla^d\nabla^D{}_E(Z_t-\mu) = \theta(B)\Theta(B^E)\varepsilon_t$$

donde los términos  $P, D$  y  $Q$  tienen una interpretación similar a la de  $p, d$  y  $q$  y  $E$  indica la longitud del periodo estacional. Como puede notarse, a través de este modelo se consideran relaciones entre observaciones contiguas dentro del mismo año y entre observaciones presentadas en los mismos periodos de diferentes años.

Finalmente, cabe destacar que el modelo **ARIMA**( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $_E$  deberá tener los parámetros restringidos a las bandas de estacionaridad e invertibilidad para los procesos **ARMA** previamente diferenciados  $d$  veces de forma regular y  $D$  veces de forma estacional, es decir, se deberá cumplir que las raíces de  $\phi(x)=0$ ,  $\Phi(x^E)=0$ ,  $\theta(x)=0$  y  $\Theta(x^E)=0$  se encuentren fuera del círculo unitario. Obviamente, debido a la mayor complejidad del modelo, la estructura de las autocorrelaciones es más complicada. A continuación presento algunas **FAC** de procesos estacionales típicos que se obtienen a partir del hecho de que la función generadora de autocovarianzas de un proceso multiplicativo es el resultado del producto de las funciones generadoras de autocovarianzas de cada uno de los procesos involucrados en el modelo multiplicativo<sup>1</sup>.

**MODELO I: ARIMA(1, d, 1) $\times$ (0, D, 1) $_E$**

Forma explícita:  $(1-\phi B)W_t = (1-\theta B)(1-\Theta B^E)\varepsilon_t \quad E \geq 3$

$$\gamma_0 = \left[ \frac{(1+\Theta^2)(1+\theta^2-2\phi\theta)-2\phi^{E-1}\Theta(1-\phi\theta)(\phi-\theta)}{1-\phi^2} \right] \sigma_\varepsilon^2$$

$$\gamma_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1-\phi^2} [(1+\Theta^2)-\Theta(\phi^{E-2}+\phi^E)] \sigma_\varepsilon^2$$

$$\gamma_k = \phi\gamma_{k-1}-\phi^{E-k-1}\Theta(1-\phi\theta)(\phi-\theta)\sigma_\varepsilon^2 \quad 2 \leq k \leq E-1$$

<sup>1</sup> Guerrero, Víctor M., Análisis estadístico de series de tiempo económicas, pp.195 -197.

$$\gamma_E = \left( \frac{\phi^{E-1}(1+\theta^2-\phi^E)(1-\phi\theta)(\phi-\theta)-\theta(1+\theta^2-2\phi\theta)}{1-\phi^2} \right) \sigma_o^2$$

$$\gamma_{E+1} = \frac{(1-\phi\theta)(\phi-\theta)}{1-\phi^2} [\phi^E(1+\theta^2)-\theta(1+\phi^{2E})] \sigma_o^2$$

$$\gamma_k = \phi \gamma_{k-1} \quad k > E+2$$

**MODELO II: ARIMA(0,d,1)x(1,D,1)<sub>E</sub>**

Forma explícita:  $(1-\Phi B^E)W_t = (1-\theta B)(1-\Theta B^E)\varepsilon_t$

$$\gamma_0 = (1+\theta^2) \left( 1 + \frac{(\Theta-\Phi)^2}{1-\Phi^2} \right) \sigma_o^2$$

$$\gamma_1 = -\theta \left( 1 + \frac{(\Theta-\Phi)^2}{1-\Phi^2} \right) \sigma_o^2$$

$$\gamma_k = 0 \quad 2 \leq k \leq E-2 \text{ si } E \geq 4$$

$$\gamma_{E-1} = \theta \left( \Theta - \Phi - \frac{\Phi(\Theta-\Phi)^2}{1-\Phi^2} \right) \sigma_o^2$$

$$\gamma_E = -(1+\theta^2) \left( \Theta - \Phi - \frac{\Phi(\Theta-\Phi)^2}{1-\Phi^2} \right) \sigma_o^2$$

$$\gamma_{E+1} = \gamma_{E-1}$$

$$\gamma_k = \Phi \gamma_{k-E} \quad k \geq E+2$$

**MODELO III: ARIMA(1,d,E+1)x(0,D,0)<sub>E</sub>**

Forma explícita:  $(1-\phi B)W_t = (1-\theta_1 B - \theta_E B^E + \theta_{E+1} B^{E+1})\varepsilon_t \quad E \geq 3$

$$\gamma_0 = \sigma_o^2 \left( \frac{1+\theta_1^2 + \theta_E^2 + \theta_{E+1}^2}{1-\phi^2} + \frac{2\phi[(\theta_E \theta_{E+1} - \theta_1) + \phi^{E-2} \theta_1 \theta_E + \phi^{E-1}(\theta_1 \theta_{E+1} - \theta_E) - \phi^E \theta_{E+1}]}{1-\phi^2} \right)$$

$$\gamma_1 = \sigma_o^2 \left( \frac{\phi(1+\theta_1^2 + \theta_E^2 + \theta_{E+1}^2)}{1-\phi^2} + \frac{(1+\phi^2)[(\theta_E \theta_{E+1} - \theta_1) + \phi^{E-2} \theta_1 \theta_E + \phi^{E-1}(\theta_1 \theta_{E+1} - \theta_E) - \phi^E \theta_{E+1}]}{1-\phi^2} \right)$$

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

$$2 \leq k \leq E-2 \text{ si } E \geq 4$$

$$\gamma_{E-1} = \sigma_c^2 \left[ \frac{\phi^{E-1}(1+\theta_1^2+\theta_E^2+\theta_{E+1}^2)+\phi^{E-2}(1+\phi^2)(\theta_E\theta_{E+1}-\theta_1)}{1-\phi^2} + \frac{(1+\phi^{2E-2})(\theta_1\theta_{E+1}-\theta_E)+\theta_1\theta_E\phi^2\theta_{E+1}}{1-\phi^2} \right]$$

$$\gamma_E = \sigma_c^2 \left[ \frac{\phi^E(1+\theta_1^2+\theta_E^2+\theta_{E+1}^2)+\phi^{E-1}(1+\phi^2)(\theta_E\theta_{E+1}-\theta_1)}{1-\phi^2} + \frac{(1+\phi^{2E})(\theta_1\theta_{E+1}-\theta_E)+\theta_1(1+\phi^{2E-2})(\theta_1\theta_E-\theta_{E+1})}{1-\phi^2} \right]$$

$$\gamma_{E+1} = \sigma_c^2 \left[ \frac{\phi^{E+1}(1+\theta_1^2+\theta_E^2+\theta_{E+1}^2)+\phi^E(1+\phi^2)(\theta_E\theta_{E+1}-\theta_1)}{1-\phi^2} + \frac{\theta_1(1+\phi^{2E})(\theta_1\theta_{E+1}-\theta_E)(1+\phi^{2E-2})(\theta_1\theta_E-\theta_{E+1})}{1-\phi^2} \right]$$

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

$$k \geq E+2$$

#### MODELO IV. ARIMA(0, d, E+1) x (1, D, 0)\_E

Forma explícita:  $(1-\Phi B^E)W_t = (1-\theta_1 B - \theta_E B^E + \theta_{E+1} B^{E+1})a_t \quad E \geq 3$

$$\gamma_0 = \sigma_c^2 \left[ 1 + \theta_1^2 + \frac{(\theta_E - \Phi)^2 + (\theta_{E+1} - \theta_1 \Phi)^2}{1 - \Phi^2} \right]$$

$$\gamma_1 = \sigma_c^2 \left[ -\theta_1 + \frac{(\theta_E - \Phi)(\theta_{E+1} + \theta_1 \Phi)}{1 - \Phi^2} \right]$$

$$\gamma_k = 0 \quad 2 \leq k \leq E-2 \text{ si } E \geq 4$$

$$\gamma_{E-1} = \sigma_c^2 \left[ (\theta_E - \Phi) \left( \theta_1 + \Phi \frac{(\theta_{E+1} + \theta_1 \Phi)}{1 - \Phi^2} \right) \right]$$

$$\gamma_E = \sigma_c^2 \left\{ -(\theta_E - \Phi) \left[ 1 - \Phi \frac{(\theta_E - \Phi)}{1 - \Phi^2} \right] + (\theta_{E+1} + \theta_1 \Phi) \left[ \theta_1 + \Phi \frac{(\theta_{E+1} + \theta_1 \Phi)}{1 - \Phi^2} \right] \right\}$$

$$\gamma_{E+1} = -(\theta_{E+1} + \theta_1 \Phi) \left[ 1 - \Phi \frac{(\theta_E - \Phi)}{1 - \Phi^2} \right] \sigma_c^2$$

$$\gamma_k = \Phi \gamma_{k-E} \quad k \geq E+2$$

### MODELO V: ARIMA(0,d,2)x(0,D,2)<sub>E</sub>

Forma explícita:  $W_t = (1-\theta_1 B - \theta_2 B^2)(1-\Theta_1 B^E - \Theta_2 B^{2E}) a_t \quad E \geq 5$

$$\gamma_0 = (1+\theta_1^2 + \theta_2^2)(1+\Theta_1^2 + \Theta_2^2) \sigma_a^2$$

$$\gamma_1 = -\theta_1(1-\theta_2)(1+\Theta_1^2 + \Theta_2^2) \sigma_a^2$$

$$\gamma_2 = -\theta_2(1+\Theta_1^2 + \Theta_2^2) \sigma_a^2$$

$$\gamma_k = 0 \quad 3 \leq k \leq E-3 \quad \text{si } E \geq 6$$

$$\gamma_{E-2} = \theta_2 \Theta_1 (1-\Theta_2) \sigma_a^2$$

$$\gamma_{E-1} = \theta_1 \Theta_1 (1-\theta_2)(1-\Theta_2) \sigma_a^2$$

$$\gamma_E = -\Theta_1(1+\theta_1^2 + \theta_2^2)(1-\Theta_2) \sigma_a^2$$

$$\gamma_{E+1} = \gamma_{E-1}$$

$$\gamma_{E+2} = \gamma_{E-2}$$

$$\gamma_k = 0 \quad E+3 \leq k \leq 2E-3 \quad \text{si } E \geq 6$$

$$\gamma_{2E-2} = \theta_2 \Theta_2 \sigma_a^2$$

$$\gamma_{2E-1} = \theta_1 \Theta_2 (1-\theta_2) \sigma_a^2$$

$$\gamma_{2E} = -\Theta_2(1+\theta_1^2 + \theta_2^2) \sigma_a^2$$

$$\gamma_{2E+1} = \gamma_{2E-1}$$

$$\gamma_{2E+2} = \gamma_{2E-2}$$

### **5.3 Identificación de procesos estacionales**

Las series temporales con estructura **ARIMA** estacional también pueden ser detectadas a través de la **FAC** y de la **FACP** correspondientes. La identificación se

vuelve ahora más difícil debido a que se incrementa la cantidad de modelos que se pueden utilizar. Para una estacionalidad de longitud  $E$  estas funciones presentan elevamientos y decaimientos significativos en los retardos  $E, 2E, 3E, \dots$ , pero por lo demás tienen el mismo comportamiento que las correspondientes a un proceso no estacional.

Cabe señalar que en la forma explícita del modelo se utiliza a  $W_t = \nabla^d \nabla^D {}_E T(Z_t)$  y que los valores  $d$  y  $D$  se determinan de tal forma que  $W_t$  sea aproximadamente estacionaria. Para tomar esta decisión puede ser útil graficar la **FAC** muestral de varias diferencias (por ejemplo,  $\{T(Z_t)\}, \{\nabla T(Z_t)\}, \{\nabla^2 T(Z_t)\}, \{\nabla \nabla {}_E T(Z_t)\}, \{\nabla^2 \nabla {}_E T(Z_t)\}$ ) y por inspección visual determinar el grado de diferenciación adecuado. Otra herramienta útil puede ser calcular

$$S^2(i,j) = \frac{1}{N-Ei-j-1} \sum_{t=EI+j+1}^N \nabla^i \nabla \left[ {}_E T(Z_t) - \sum_{t=EI+j+1}^N \frac{\nabla^j \nabla^i {}_E T(Z_t)}{N-Ei-j} \right]^2 \quad \text{Para } j,i = 0,1,2,3$$

bastará con seleccionar  $S^2(d,D) = \min\{S^2(i,j) \mid j,i=0,1,2,3\}$  para obtener a la serie que sea estacionaria.

Para facilitar la especificación de los modelos estacionales, autores como Arnau, McCleary, Hay, Hamilton y Watts recomiendan que se tengan en cuenta las consideraciones que siguen:

- a) La ausencia de estacionaridad estacional se detecta inspeccionando la **FAC**. Si la serie es estacional, la **FAC** presentará valores grandes y casi iguales en los retardos estacionales, es decir,  $\mathbf{FAC}(E) \approx \mathbf{FAC}(2E) \approx \mathbf{FAC}(3E), \dots$

- b) Se puede esperar que los procesos  $ARIMA(P,0,0)_E$  presenten un patrón de decaimiento en la **FAC** donde la tasa de decaimiento estará en función de los valores  $\phi_E, \phi_{2E}, \phi_{3E}, \dots, \phi_P$  y, por el contrario, que su **FACP** presente  $P$  picos en los primeros  $P$  retardos, mientras que los restantes valgan cero.
- c) Es posible que los procesos  $ARIMA(0,0,Q)_E$  presenten  $Q$  picos en los primeros  $Q$  retardos estacionales de su **FAC**, mientras que los restantes valgan cero y, por el contrario, un patrón de decaimiento progresivo en la **FACP** cuya tasa de decaimiento estará en función de los valores  $\theta_E, \theta_{2E}, \theta_{3E}, \dots, \theta_Q$ .
- d) Generalmente, las componentes regular y estacional del modelo  $ARIMA(p,d,q) \times (P,D,Q)_E$  serán de la misma clase, es decir, raramente ocurrirá que la parte regular quede adecuadamente descrita por un proceso **AR** y la estacional por un proceso **MA** o a la inversa

La identificación de un modelo  $ARIMA(p,d,q) \times (P,D,Q)_E$  implica los siguientes pasos: Determinar mediante la **FAC** si la serie es o no estacionaria (si no es estacionaria conviene diferenciarla); después se utilizan la **FAC** y la **FACP** para identificar el modelo correspondiente a la parte regular, posteriormente se inspecciona la **FAC** y la **FACP** para determinar los valores  $P$  y  $Q$  de la parte estacional respectivamente.

#### **5.4 Construcción de modelos estacionales**

La metodología de construcción es igual a la que se sigue para modelos no estacionales pero existen algunos cambios en la verificación; así por ejemplo, en la etapa de *determinación de la significación estadística* de la media residual se utilizan los valores

$$M(\hat{a}_t) = \sum_{t=d+p+(D+P)E+1}^N \hat{a}_t / [N-d-p-q-(D+P)E] \quad y$$

$$\sigma_a = \left( \sum_{t=d+p+(D+P)E+1}^N (\hat{a}_t - M(\hat{a}_t))^2 / [N-d-p-q-(D+P)E] \right)^{1/2}$$

posteriormente se evalúa

$$(N-d-p-(D+P)E)^{1/2} M(\hat{a}_t) / \sigma_a$$

La significación individual de las autocorrelaciones residuales se obtienen a partir de la siguiente expresión

$$r_k(\hat{a}) = \frac{\sum_{t=d+p+(D+P)E+1}^{N-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=d+p+(D+P)E+1}^N \hat{a}_t^2}$$

y  $\text{Var}[r_k(\hat{a}_t)] \approx 1 / (N-d-p-(D+P)E)$ . Por su parte el estadístico  $Q'$  se modifica hacia

$$Q' = [N-d-p-(D+P)E] [N-d-p-(D+P)E+2] \sum_{k=1}^K r_k^2(\hat{a}) / [N-d-p-(D+P)E-k]$$

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

## 6 Serie de tiempo: Ingreso federal

El ingreso federal es una variable económica de gran importancia pues es la base para determinar las partidas presupuestales de los diversos sectores del gobierno federal. Se denomina ingreso federal a los recursos consignados en la ley de ingresos de la federación y que, como a continuación describiré, concretamente se expresan en impuestos, derechos, productos y aprovechamientos.

### Ingresos tributarios:

Impuesto sobre la renta (ISR):

Sociedades mercantiles, personas morales con fines no lucrativos, personas físicas, pagos al extranjero, retención a personas físicas y morales, impuesto al activo, otros.

Impuesto al valor agregado (IVA)

Impuesto especial sobre producción y servicios (IEPS):

Gasolina, tabacos labrados, teléfonos, cervezas, envasamiento de bebidas alcohólicas, aguas envasadas, refrescos, seguros, otros.

Accesorios

Importación

Exportación

Automóviles nuevos

Tenencia o uso de vehículos

Erogación

Adquisición de inmuebles

Azúcar, cacao y otros

Otros impuestos



## Ingresos no tributarios

### Derechos

Por la prestación de servicios que corresponden a funciones de derecho público, por el uso y aprovechamiento de bienes del dominio público, por extracción de hidrocarburos, otros.

### Productos:

Derivados del uso o explotación de bienes del dominio privado, por servicios que no corresponden a funciones de derecho público, otros.

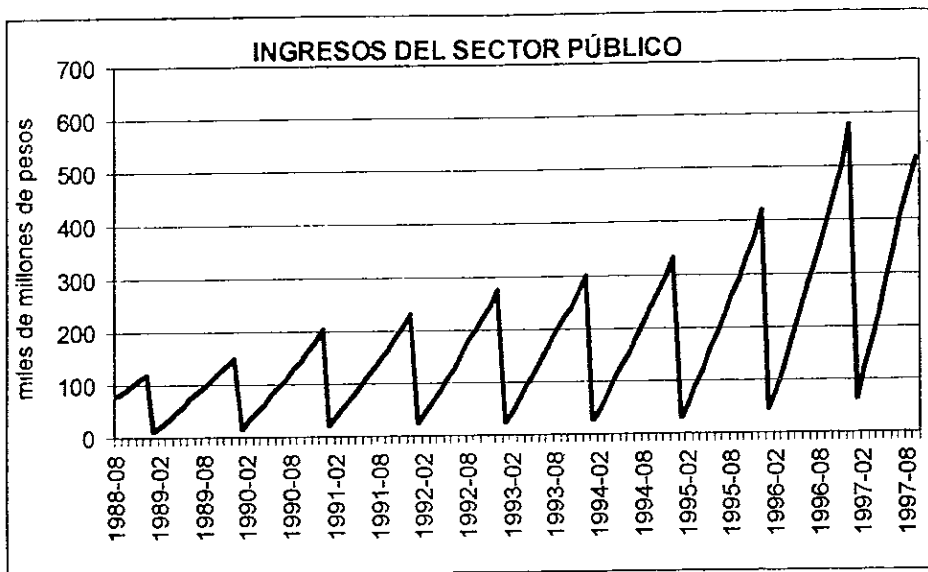
### Aprovechamientos.

Reintegros, recuperación de capital, ingresos por venta de empresas paraestatales, acciones de Telmex y bancos, multas, indemnizaciones, otros.

### Contribuciones de mejoras

Para ajustar un modelo a esta serie utilizaré los valores mensuales registrados desde agosto de 1988 hasta septiembre de 1997, lo que hace un total de ciento diez observaciones a las cuales les será aplicada a continuación la metodología de Box y Jenkins (o del enfoque estocástico).

Empezaré por presentar la gráfica correspondiente a los valores del ingreso federal que utilizaré en el ejercicio pues los datos los mostraré en la tabla A.



### 6.1 Identificación

Como puede notarse en la gráfica, esta serie se obtiene por acumulación y muestra un comportamiento estacional; como presenta un comportamiento con tendencia creciente puedo suponer que requeriré de una transformación estabilizadora de la varianza.

Con el fin de identificar el parámetro de transformación  $\lambda$  que permita la estabilización de la varianza, formo con los datos proporcionados 10 grupos con 11 observaciones cada uno (por lo que no elimino ninguna observación para la prueba) y procedo a calcular los coeficientes de variación para diferentes valores de  $\lambda$ . Los resultados obtenidos son los siguientes:

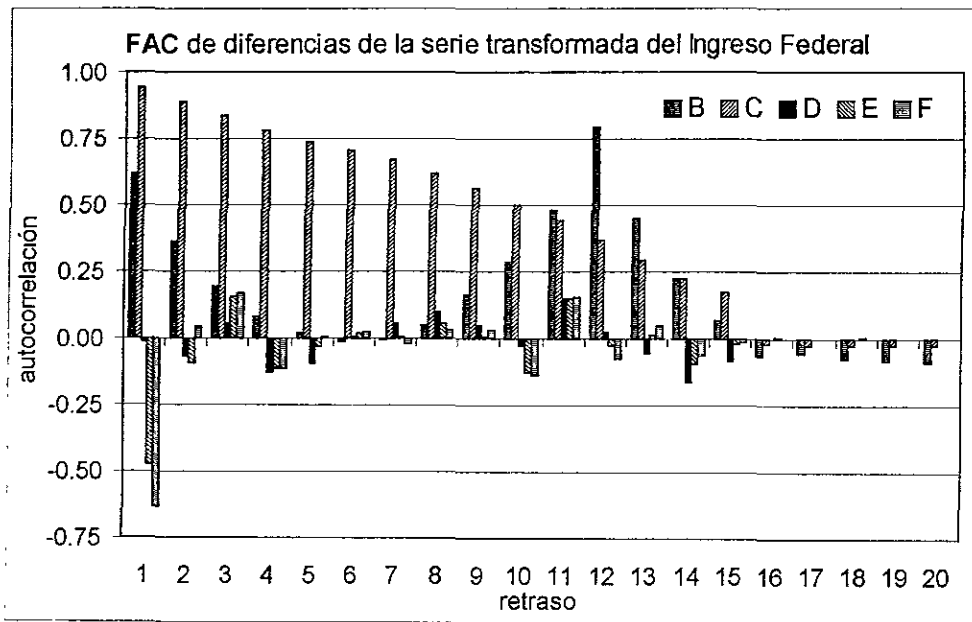
	VALORES DEL GRUPO		VALORES PARAMETRALES DE $\lambda$				
	MEDIA	DESVIACIÓN	-0.1	-0.05	<b>0</b>	0.05	0.1
GRUPO 1	67,029.44	35,413.60	0.1739	0.3031	0.5283	0.9209	1.6052
GRUPO 2	84,439.04	43,285.07	0.1649	0.2907	0.5126	0.9039	1.5939
GRUPO 3	109,128.89	59,360.99	0.1705	0.3046	0.5440	0.9715	1.7352
GRUPO 4	126,163.62	67,592.13	0.1655	0.2978	0.5357	0.9638	1.7340
GRUPO 5	151,326.46	81,983.01	0.1644	0.2984	0.5418	0.9836	1.7857
GRUPO 6	168,387.93	85,687.04	0.1527	0.2788	0.5089	0.9288	1.6953
GRUPO 7	187,905.61	92,852.05	0.1467	0.2692	0.4941	0.9069	1.6644
GRUPO 8	192,603.08	112,076.12	0.1723	0.3167	0.5819	1.0693	1.9648
GRUPO 9	257,835.95	142,121.21	0.1586	0.2956	0.5512	1.0277	1.9162
GRUPO10	334,513.21	175,167.93	0.1468	0.2772	0.5237	0.9891	1.8685
		M( $\lambda$ )	0.1616	0.2932	0.5322	0.9666	1.7563
		S( $\lambda$ )	0.0101	0.0144	0.0248	0.0539	0.1269
		COEFICIENTE DE VARIACIÓN	0.0622	0.0492	<b>0.0466</b>	0.0557	0.0722

Con base en esta tabla determino que el valor de  $\lambda$  que minimiza el coeficiente de variación es cero; así, siguiendo las transformaciones propuestas en las expresiones (3.1.1.6) y (3.1.1.7), debo aplicar la función logaritmo a la serie original con lo que obtengo la serie transformada de la tabla A con 110 observaciones.

Posteriormente estabilizo el nivel de la serie. Dado que se trata de una serie estacional habrá que aplicarle a la serie transformada diferencias regulares y estacionales de periodo 12. Para detectar cuál será la serie estacionaria grafico los primeros valores de la **FAC** de varias diferencias de la serie transformada. Presento a continuación dos gráficas cuyas componentes son las autocorrelaciones de las siguientes diferencias sobre la serie transformada:

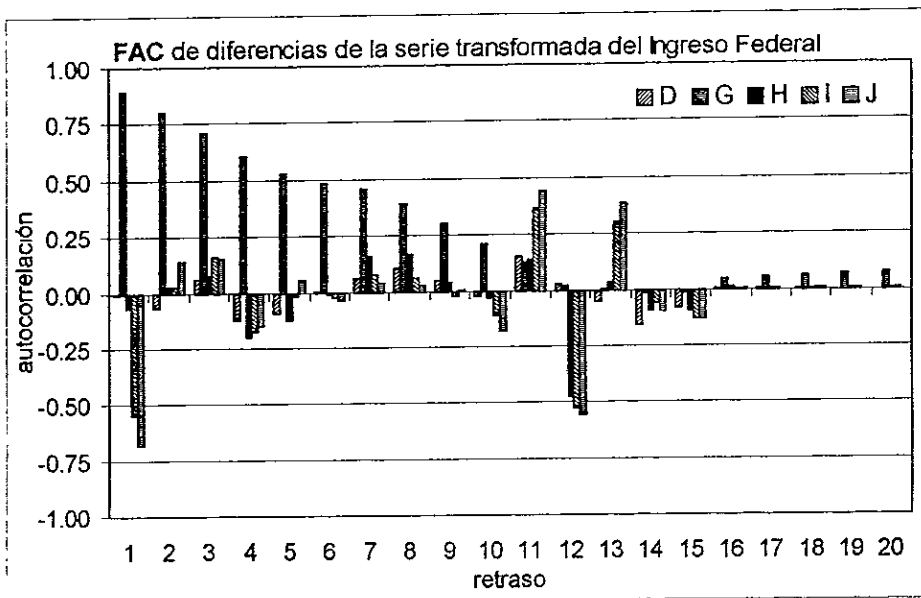
B de la serie transformada.

- C de la primer diferencia estacional de periodo 12 .
- D de la primer diferencia regular de la primer diferencia estacional de periodo 12 .
- E de la segunda diferencia regular de la primer diferencia estacional de periodo 12 .
- F De la tercer diferencia regular de la primer diferencia estacional de periodo 12 .



De esta gráfica las autocorrelaciones más pequeñas corresponden a la serie D por lo que esta es la serie que compararé con las siguientes series:

- G De la segunda diferencia estacional de periodo 12 .
- H De la primer diferencia regular de la segunda diferencia estacional de periodo 12 .
- I De la segunda diferencia regular de la segunda diferencia estacional de periodo 12
- J De la tercer diferencia regular de la segunda diferencia estacional de periodo 12 .



En la segunda gráfica vuelve a ser D la serie con las autocorrelaciones más pequeñas, por lo que pasa a ser la serie estacionaria. Utilizando el método de comparación de las desviaciones estándar de las series llego a los siguientes valores para cada una:

B	0.8281	E	0.0410	H	0.0408
C	0.0864	F	0.0706	I	0.0599
<b>D</b>	<b>0.0287</b>	G	0.0908	J	0.1059

Con esto ratifico a D como la serie que con 97 observaciones ( $110-12-1=97$ ) será considerada como estacionaria. Esto significa que será a la primer diferencia regular de la primer diferencia estacional de periodo 12 de la serie transformada, que presento como serie estacionaria en la tabla A, a la que tendré que ajustar un modelo.

Para identificar el proceso que sigue esta serie analizo a continuación su **FAC** y **FACP** muestrales. En la tabla A muestro los primeros treinta valores de dichas funciones que son calculados con base en la fórmula mostrada en (3.1.3.1) y la sustitución de estos valores en la expresión (1.3.9).

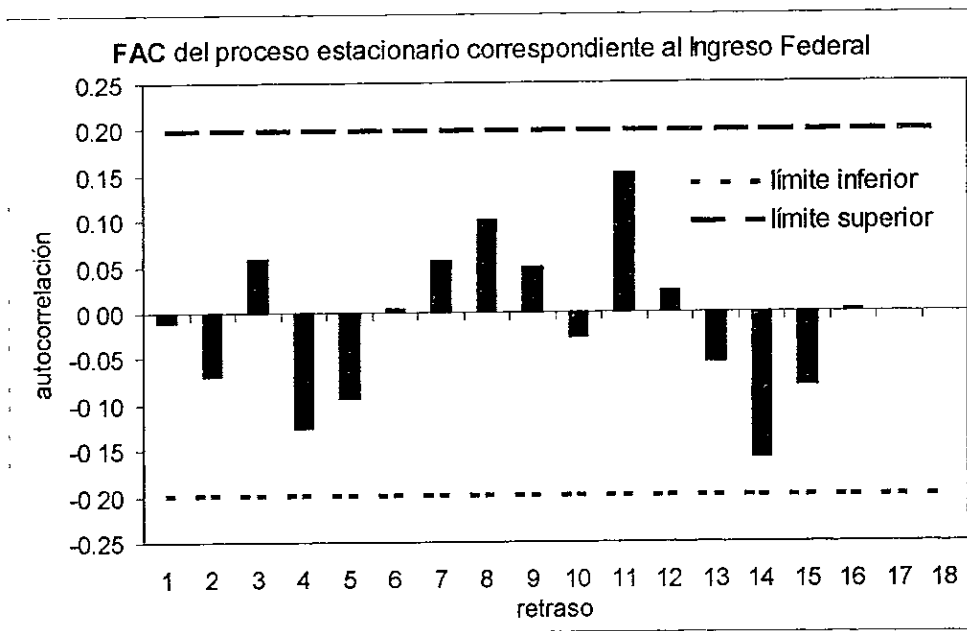
$$M(Z_t) = M(Z_{t+k}) = 0.000344$$

$$\sum_{t=1}^{97} [Z_t - M(Z_t)]^2 = (-0.001112 - 0.000344)^2 + \dots + (-0.012251 - 0.000344)^2 = 0.078929$$

Por ejemplo para  $k=1$

$$\sum_{t=1}^{96} [Z_t - M(Z_t)] * [Z_{t+k} - M(Z_{t+k})] = (-0.001112 - 0.000344) * (0.010883 - 0.000344) + \dots = -0.000854$$

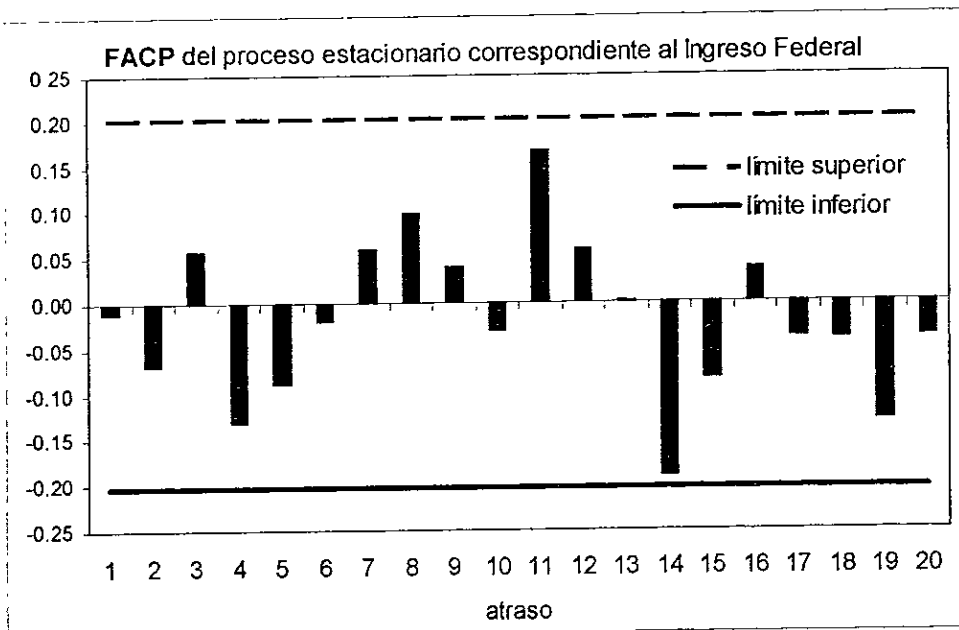
Por lo que el valor de la autocorrelación muestral es  $r_1 = -0.000854 / 0.078929 = -0.01082$  La gráfica correspondiente a la **FAC** muestral es:



Por otro lado, para la **FACP** muestral realizo como ejemplo el cálculo de  $\phi_{33}$  (estimación de  $\phi_{33}$ ) que a continuación presento:

$$\phi_{33} = \frac{\begin{vmatrix} 1 & -0.01082 & -0.01082 \\ -0.01082 & 1 & -0.06874 \\ -0.06874 & -0.01082 & 0.057990 \end{vmatrix} \cdot 0.056443}{\begin{vmatrix} 1 & 0.01082 & -0.06874 \\ 0.01082 & 1 & 0.01082 \\ -0.06874 & 0.01082 & 1 \end{vmatrix} \cdot 0.995025} = 0.05672$$

Finalmente, la gráfica de la **FACP** muestral es



Siguiendo este procedimiento es como obtengo todos los valores de las **FAC** y **FACP** muestrales, los cuales reviso posteriormente para identificar el tipo de proceso que sigue la serie estacionaria.

La primer prueba es para identificar si el proceso estudiado se comporta como ruido blanco. Siguiendo las aproximaciones de Bartlett, si el proceso efectivamente es ruido blanco esperaría que ningún valor de la **FAC** ni de la **FACP** estuviera fuera del intervalo  $(-2/[97]^{1/2}, 2/[97]^{1/2}) = (-2*0.101535, 2*0.101535) = (-0.203069, 0.203069)$ . Este supuesto, como se aprecia en las gráficas, efectivamente se cumple pues los valores más grandes conocidos de la **FAC** y de la **FACP** son -0.159552 y -0.189800, respectivamente y quedan incluidos en el intervalo señalado. Por tanto determino que el proceso correspondiente a la serie transformada y previamente diferenciada es sólo ruido blanco.

## 6.2 Estimación

Debido a que el proceso relacionado resultó representar sólo ruido blanco, el modelo se reduce a la siguiente representación  $Y_t = \alpha_t$  en la que  $Y_t = \nabla \nabla_{\epsilon} T(Z_t)$  es el proceso transformado y diferenciado.

Y por tanto, no necesito estimar parámetros autorregresivos ni de promedios móviles. El valor estimado de la varianza del ruido blanco será la varianza de la serie como tal por lo que:

$$S_{\alpha}^2 = S_Y^2 = \sum_{t=1}^{97} [Y_t - M(Y_t)]^2 / (N - 1) = 0.078929 / 96 = 0.00082218$$

## 6.3 Verificación

Inicio con la revisión de los supuestos para el análisis de residuales que propuse en la sección 3.3.1.

**SUPUESTO 1:**  $\alpha_t$  tiene media cero. Los valores de la media y la desviación estándar muestrales de los errores aleatorios son exactamente los mismos que los que corresponden a la serie en estudio por lo que:



$$M(a_t) = M(Y_t) = 0.000344$$

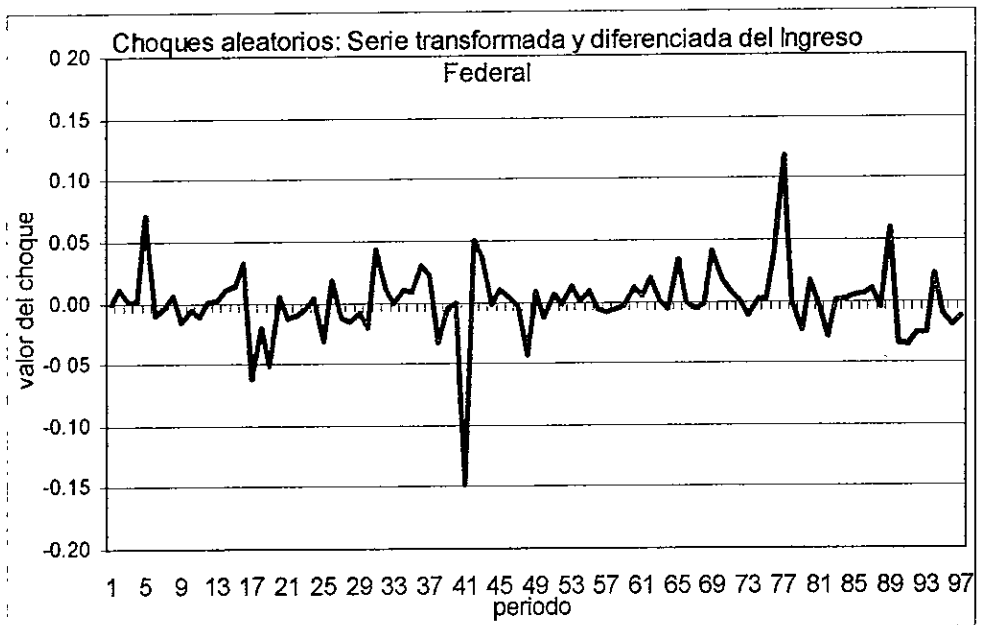
$$S_o = S_Y = [0.00082218]^{1/2} = 0.028674$$

A continuación evalúo el cociente:

$$(N)^{1/2} * M(a_t) / S_o = (97)^{1/2} * 0.000344 / 0.028674 = 0.118156$$

Como se cumple  $0.118156 < 2$  no existe evidencia de que la media del proceso sea diferente de cero y apruebo el supuesto.

**SUPUESTO 2.**  $a_t$  tiene varianza constante. Presento a continuación la gráfica de la serie sobre la que trabajo, misma que coincide con el ruido blanco, ya que a través de su revisión sabré si se cumple con el supuesto.



Como puede verse, la gráfica no presenta una tendencia creciente ni decreciente si no un comportamiento estable por lo que puedo decir que la varianza es constante y en consecuencia se cumple con el supuesto.

**SUPUESTO 3:** Las variables  $\{e_t\}$  son mutuamente independientes. Calculando las primeras 25 estimaciones de  $r_k(a)$  obtengo los valores:

k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$
1	-0.010615	6	0.003686	11	0.152578	16	0.000946	21	-0.001037
2	-0.068492	7	0.057589	12	0.024131	17	0.000080	22	-0.001173
3	0.058267	8	0.101683	13	-0.053891	18	-0.000205	23	-0.001236
4	-0.125526	9	0.049182	14	-0.159330	19	-0.000924	24	-0.001192
5	-0.092825	10	-0.026558	15	-0.079659	20	-0.000862	25	-0.001637

Y la estimación de la varianza de la FAC de los choques aleatorios la obtengo mediante:

$$\underline{Var} [r_k(a)] \approx 1/97 \Rightarrow \{\underline{Var} [r_k(a)]\}^{1/2} \approx [1/97]^{1/2} = 0.101535$$

con este valor formo el intervalo  $(-2*0.101535, 2*0.101535) = (-0.203069, 0.203069)$ . Este intervalo incluye a los valores estimados  $r_k(a)$  por lo que considero que ninguna autocorrelación es significativamente distinta de cero y por tanto el supuesto se cumple.

Realizando la prueba Portmanteau mediante el estadístico  $Q'$  obtengo:

$$Q' = 97 * (97 + 2) \sum_{k=1}^{25} r_k^2(a) / (97 - k) = 97 * 99 * 0.00123403 = 11.850102$$

El valor en tablas de  $\chi^2_{(0.25, 25)} = 40.6465$ , dado que  $\chi^2 > Q'$  no se rechaza el supuesto de independencia de las variables.

**SUPUESTO 4:**  $\{a_t\}$  se distribuye normal. Verifico cuántas observaciones de la serie quedan fuera de  $(-2S_a, 2S_a) = (-2 \cdot 0.028674, 2 \cdot 0.028674) = (-0.057348, 0.057348)$ .

Encuentro que los puntos fuera del intervalo señalado son 0.07, -0.06, -0.15, 0.12 y 0.06 que representan el 5% del total de las observaciones por lo que puede decirse que la serie sigue una distribución normal.

**SUPUESTO 5:** No existen observaciones aberrantes. Encuentro que sólo las observaciones -0.15 (que proviene de las observaciones 41, 42, 53 y 54 de la serie original) y 0.12 (que proviene de las observaciones 77, 78, 89 y 90 de la misma serie original) no se encuentran incluidas en el intervalo formado por  $(-3S_a, 3S_a) = (-0.086022, 0.086022)$ . Para analizar qué pudo ocasionar estos valores de nuestra serie estacionaria realice un comparativo porcentual del comportamiento de la serie original en las observaciones correspondientes a los periodos afectados y obtengo el siguiente cuadro:

<i>Obsv.</i>	<i>Periodo</i>	<i>Valor</i>	<i>Var.</i>	<i>Obsv.</i>	<i>Periodo</i>	<i>Valor</i>	<i>Var.</i>
17	1989-12	148,601.10	25%	66	1994-01	26,755.80	10%
18	1990-01	16,381.80	35%	77	1994-12	332,033.50	11%
29	1990-12	203,486.30	37%	78	1995-01	30,859.20	15%
30	1991-01	21,095.50	29%	89	1995-12	418,882.60	26%
41	1991-12	229,608.30	13%	90	1996-01	43,792.10	42%
42	1992-01	23,603.60	12%	101	1996-12	580,722.00	39%
53	1992-12	274,471.00	20%	102	1997-01	64,379.60	47%
54	1993-01	24,329.00	3%	113	1997-12	730,441.30	26%
65	1993-12	297,951.00	9%				

Como puede apreciarse, efectivamente existe una reducción de los porcentajes de variación en las observaciones 41, 42, 53 y 54 que provocan el valor  $-0.15$ ; sin embargo, no existe en esos periodos ninguna medida económica importante que hubiera podido provocar esta disminución.

Con relación al valor  $0.12$ , encuentro que este valor proviene de observaciones cuyas variaciones porcentuales son significativamente mayores respecto a las correspondientes a los periodos anteriores y esto provoca el valor ya señalado; sin embargo, como en el caso anterior, no existe razón aparente para dicho incremento toda vez que las condiciones económicas y sociales de México en esas fechas eran de aguda crisis consecuencia del "error del diciembre del 1995" y por el estallido del conflicto Chiapaneco. Estas circunstancias lejos de beneficiar al ingreso federal debieron disminuirlo, cosa que, como vemos, no ocurrió.

En resumen, concluyo que no existe razón para pensar que las observaciones fuera del intervalo sean el resultado de una intervención sobre la serie y por tanto acepto el supuesto de que no existen observaciones aberrantes.

Ahora aplico las pruebas independientes de los residuales mostradas en la sección 3.3.2:

**SUPUESTO 6:** *El modelo es parsimonioso.* Dado que no requerí parámetros, el modelo es parsimonioso.

**SUPUESTO 7:** *El modelo es admisible.* Como no existen parámetros no hay necesidad de probar la admisibilidad del modelo.

**SUPUESTO 8:** *El modelo es estable en los parámetros.* Dado que no existen parámetros no puedo verificar correlaciones entre éstos.

En forma independiente de los supuestos anteriores probaré si un modelo más general (con más parámetros) podría describir adecuadamente al proceso, es decir, sobreparametrizaré el modelo.

Iniciaré con un modelo **AR(1)** para el que tendría la relación  $Y_t - \phi Y_{t-1} = a_t$  el valor parametral inicial obtenido por el método de momentos aplicando las ecuaciones de Yule – Walker mostradas en (2.1.7) es el siguiente:

$$r_1 = \phi \quad \Rightarrow \quad \phi = -0.010822$$

Con este valor calculo los que serían los errores aleatorios basados en el supuesto de incondicionalidad para lo cual sigo los pasos marcados en la sección 3.2.1.2 como a continuación lo hago:

a) Expreso al modelo como  $Y_t + 0.010822Y_{t+1} = e_t$

b) Comienzo a calcular  $e_t$  a partir de  $t = N - p = 97 - 1 = 96$  siguiendo este proceso:

$$e_{96} = Y_{96} + 0.010822Y_{97} = -0.019018 + (0.010822)(-0.012251) = -0.019151$$

$$e_{95} = Y_{95} + 0.010822Y_{96} = -0.009602 + (0.010822)(-0.019018) = -0.009808$$

.....

$$e_1 = Y_1 + 0.010822Y_2 = -0.001112 + (0.010822)(0.010883) = -0.000994$$

c) Hago  $e_i = 0$  para  $i = 0, 1, 2, \dots$  y con los valores determinados para  $e_i$  y los valores registrados de  $Y_t$  obtengo los valores para  $Y_i$   $i = 0, 1, 2, \dots$  de tal suerte que:

$$Y_0 = e_0 - 0.010822Y_1 = 0 - 0.010822(-0.001112) = 0.000012$$

$$Y_{-1} = e_{-1} - 0.010822Y_0 = 0 - 0.010822(0.000012) = 0.000000$$

d) Expreso al modelo como  $Y_t + 0.010822Y_{t-1} = a_t$  considerando que para  $Q=1$  se cumple que  $Y_t \approx 0$  para  $t \leq -Q$  por lo que tomo  $a_t = 0$  para  $t \leq -Q$ . Con base en esta última condición y tomando los valores de la serie y los valores calculados para  $Y_0$  y  $Y_{-1}$  es posible calcular los valores de  $a_t$  para  $t > -Q$ . De esta forma, los choques aleatorios vendrían dados por.

$$a_{-1} = 0$$

$$a_1 = Y_1 + 0.010822Y_0 = -0.001112 + 0.010822(0.000012) = -0.001112$$

$$a_2 = Y_2 + 0.010822Y_1 = 0.010883 + 0.010822(-0.001112) = 0.010871$$

.....

$$a_{97} = Y_{97} + 0.010822Y_{96} = -0.012251 + 0.010822(-0.019018) = -0.012457$$

e) Con los datos alcanzados en el punto anterior realizo la suma de cuadrados cuyo valor es 0.07893202448.

A continuación, utilizando la estimación por mínimos cuadrados a través del Método de Marquardt obtengo en el paquete de STATGRAPHICS que el valor que minimiza la suma de cuadrados es  $\hat{\phi} = -0.01070$  con el que obtengo un valor para la suma de cuadrados de 0.07893202207. Dado que la desviación estándar de  $\hat{\phi}$  es 0.10215 deduzco que el cero es un valor permisible y lógico del parámetro por lo cual rechazo que un **AR**(1) sea un mejor modelo para esta serie.

El siguiente modelo a probar será un **MA**(1) para el que tendría la relación  $Y_t = a_t - \theta a_{t-1}$  que tiene como valor parametral inicial por el método de momentos aplicando la **FAC** mostrada en (2.2.7) el siguiente:

$$r_1 = -\theta / (1 + \theta^2) \Rightarrow$$

$$-0.010822 = -\theta / (1 + \theta^2) \Rightarrow$$

$$-0.010822(1 + \theta^2) = -\theta \Rightarrow$$

$$-0.010822 - 0.010822\theta^2 = -\theta \Rightarrow$$

$$-0.010822\theta^2 + \theta - 0.010822 = 0 \Rightarrow$$

$$\theta = \frac{-1 \pm [(1)^2 - 4(-0.010822)(-0.010822)]^{1/2}}{2(-0.010822)}$$

las raíces son  $\theta_1 = 0.010823268$  y  $\theta_2 = 92.39353822$  por lo que, para permitir la invertibilidad del modelo, opto por el valor de  $\theta_1$ .

Con este valor calculo los que serían los errores aleatorios basados en el supuesto de incondicionalidad:

a) Expreso al modelo como  $Y_t = \alpha - 0.010823268a_{t+1}$

b) Comienzo a calcular los valores de  $\alpha$  a partir de  $t=N-p = 97-0=97$ , teniendo en cuenta que  $a_{N-p+i}=0$  para  $i = 1, 2, \dots, q \Rightarrow e_{97-0+1} = e_{98} = 0$  y siguiendo este proceso:

$$e_{97} = Y_{97} + 0.010823268e_{98} = -0.012251 + (0.010823268)(0.000000) = -0.012251$$

$$e_{96} = Y_{96} + 0.010823268e_{97} = -0.019018 + (0.010823268)(-0.012251) = -0.019151$$

.....

$$e_1 = Y_1 + 0.010823268e_2 = -0.001112 + (0.010823)(0.010884) = -0.000994$$

c) Hago  $e_i = 0$  para  $i = 0, 1, 2, \dots$  y con los valores determinados para  $a_t$  y los valores registrados de  $Y_t$  obtengo los valores para  $Y_{-i}$   $i = 0, 1, 2, \dots$  de tal suerte que:

$$Y_0 = e_0 - 0.010823268e_1 = 0 - 0.010823268(-0.000994) = 0.000011$$

$$Y_{-1} = e_{-1} - 0.010823268e_0 = 0 - 0.010823268(0.000000) = 0.000000$$

d) Expreso al modelo como  $Y_t = a_t - 0.010823268a_{t-1}$  considerando que para  $Q=1$  se cumple que  $Y_t \approx 0$  para  $t \leq -Q$  por lo que tomo  $a_t = 0$  para  $t \leq -Q$ . Con base en esta última condición y tomando los valores de la serie y los valores calculados para  $Y_0$  y  $Y_{-1}$  es posible calcular los valores de  $a_t$  para  $t > -Q$ . De esta forma, el proceso es:

$$a_{-1} = 0$$

$$a_0 = Y_0 + 0.0108233a_{-1} = 0.000011 + 0.010823(0.000000) = 0.000011$$

$$a_1 = Y_1 + 0.0108233a_0 = -0.001112 + 0.010823(0.000011) = -0.001112$$

$$a_2 = Y_2 + 0.0108233a_1 = 0.010883 + 0.010823(-0.001112) = 0.010871$$

.....

$$a_{97} = Y_{97} + 0.0108233a_{96} = -0.012251 + 0.010823(-0.019120) = -0.012458$$

e) Con los datos alcanzados en el punto anterior realizo la suma de cuadrados cuyo valor es 0.078930767410857.

A continuación, utilizando el Método de Marquardt, STATGRAPHICS produce un valor que minimiza la suma de cuadrados de  $\theta = 0.01212$  con el que obtengo un valor para la suma de cuadrados de 0.078930620548620. Dado que la desviación estándar de  $\theta$  es 0.10215 deduzco que el cero es un valor permisible y lógico del parámetro por lo cual rechazo que un **MA(1)** sea un mejor modelo para esta serie.



El último modelo a probar será un **ARMA(1,1)** que tiene el formato  $Y_t - \phi Y_{t-1} = a_t - \theta a_{t-1}$  cuyo valor parametral inicial  $\phi$  se obtiene por el método de momentos aplicando la relación mostrada en (2.3.3) para los retrasos  $k \geq q + 1 = 1 + 1 = 2$ . Para  $k = 2$

$$r_2 = \phi r_1 \Rightarrow$$

$$\phi = r_2 / r_1 \Rightarrow$$

$$\phi = -0.068743 / -0.010822 \Rightarrow$$

$$\phi = 6.352467$$

como este valor no permite la estacionaridad resuelvo la misma relación para  $k = 3$  obteniendo.

$$r_3 = \phi r_2 \Rightarrow$$

$$\phi = r_3 / r_2 \Rightarrow$$

$$\phi = 0.057990 / -0.068743 \Rightarrow$$

$$\phi = -0.843577$$

donde  $\phi$  ya permite la estacionaridad. Ahora, el valor de  $\theta$  lo obtengo por la relación:

$$r_1 = (1 - \theta\phi)(\phi - \theta)/(1 + \theta^2 - 2\theta\phi) \Rightarrow$$

$$(r_1 - \phi)\theta^2 + (1 + \phi^2 - 2\phi r_1)\theta + (r_1 - \phi) = 0 \Rightarrow$$

$$(-0.010822 + 0.843577)\theta^2 + (1 - 0.843577^2 - 2 \cdot -0.843577 \cdot -0.010822)\theta +$$

$$(-0.010822 + .843577) = 0 \Rightarrow$$

$$-0.854399\theta^2 + 1.693364\theta - 0.854399 = 0 \Rightarrow$$

$$\theta = \frac{-(1.693364) \pm [(1.693364)^2 - 4(-0.854399)(-0.854399)]^{1/2}}{2 \cdot (-0.854399)}$$

$$\theta = 0.990968 \pm \frac{[-0.052510]^{1/2}}{-1.708798}$$

Con lo que obtengo los números complejos conjugados

$$\theta_1 = 0.990968 + 0.1341001i \quad \text{y} \quad \theta_2 = 0.990968 - 0.1341001i$$

El módulo de estos valores es  $s = [0.990968^2 + 0.1341001^2]^{1/2} = 1.000000414^{1/2} = 1.00000021$  por lo que no permiten invertibilidad del proceso además de que requiero valores reales para los parámetros. Trato ahora la relación para  $k = 5$ :

$$r_5 = \phi r_4 \Rightarrow$$

$$\phi = r_5 / r_4 \Rightarrow$$

$$\phi = -0.092822 / -0.125722 \Rightarrow$$

$$\phi = 0.7383098$$

donde  $\phi$  permite la estacionaridad. Ahora, el valor de  $\theta$  lo obtengo por la relación:

$$r_1 = (1 - \theta\phi)(\phi - \theta) / (1 + \theta^2 - 2\theta\phi) \Rightarrow$$

$$(r_1 - \phi)\theta^2 + (1 + \phi^2 - 2\phi r_1)\theta + (r_1 - \phi) = 0 \Rightarrow$$

$$(-0.010822 - 0.7383098)\theta^2 + (1 + 0.7383098^2 - 2 \cdot 0.7383098 \cdot -0.010822)\theta +$$

$$(-0.010822 - 0.7383098) = 0 \Rightarrow$$

$$-0.7491318\theta^2 + 1.7223298\theta - 0.7491318 = 0 \Rightarrow$$

$$\theta = \frac{-(1.7223298) \pm [(1.7223298)^2 - 4(-0.7491318)(-0.7491318)]^{1/2}}{2 \cdot (-0.7491318)}$$

$$\theta = 1.14955059 \pm \frac{[0.721626]^{1/2}}{-1.4982636} = 1.1495506 \pm 0.5669802$$

Los valores de  $\theta$  son  $\theta_1 = 0.5825704$  y  $\theta_2 = 1.7165308$ , por lo que tomo el valor de  $\theta_1$  como el valor inicial de  $\theta$ .

Con estos parámetros calculo los que serían los errores aleatorios basados en el supuesto de incondicionalidad:

a) Expreso al modelo como  $a_t = Y_t - \phi Y_{t+1} + \theta a_{t+1}$ .

b) Comienzo a calcular los valores de  $a_t$  a partir de  $t = N - p = 97 - 1 = 96$ , sin olvidar que  $a_{N-p+1} = 0$  para  $i = 1, 2, \dots, q \Rightarrow a_{97-1+1} = a_{97} = 0$  con lo que obtengo

$$\begin{aligned} e_{96} &= Y_{96} - 0.7383098Y_{97} + 0.5825704e_{97} \\ &= -0.019018 - 0.7383098(-0.012251) + 0.5825704(0.000000) = -0.009973 \end{aligned}$$

$$\begin{aligned} e_{95} &= Y_{95} - 0.7383098Y_{96} + 0.5825704e_{96} \\ &= -0.009602 - 0.7383098(-0.019018) + 0.5825704(-0.009973) = -0.001370 \end{aligned}$$

.....

$$\begin{aligned} e_1 &= Y_1 - 0.7383098Y_2 + 0.5825704e_2 \\ &= -0.001112 - 0.7383098(0.010883) + 0.5825704(0.007377) = -0.004849 \end{aligned}$$

c) Hago  $e_i = 0$  para  $i = 0, 1, 2, \dots$  y con los valores determinados para  $a_t$  y los valores registrados de  $Y_t$  obtengo los valores para  $Y_{.i}$   $i = 0, 1, 2, \dots$  de forma que

$$\begin{aligned} Y_0 &= e_0 + 0.7383098Y_1 - 0.5825704e_1 \\ &= 0.000000 + 0.7383098(-0.001112) - 0.5825704(-0.004849) = 0.002004 \end{aligned}$$

$$\begin{aligned} Y_{.1} &= e_{.1} + 0.7383098Y_0 - 0.5825704e_0 \\ &= 0.000000 + 0.7383098(0.002004) - 0.5825704(0.000000) = 0.00148 \end{aligned}$$

siguiendo este mismo método obtengo estos valores de  $Y_t$ :

t	Y <sub>t</sub>	t	Y <sub>t</sub>
-2	0.001092	-11	0.000071
-3	0.000806	-12	0.000053
-4	0.000595	-13	0.000039
-5	0.000440	-14	0.000029
-6	0.000325	-15	0.000021
-7	0.000240	-16	0.000016
-8	0.000177	-17	0.000012
-9	0.000131	-18	0.000009
-10	0.000096	-19	0.000006

d) Expreso al modelo como  $Y_t = 0.7383098Y_{t-1} + a_t - 0.5825704a_{t-1}$  considerando que para  $Q = 20$  se cumple que  $Y_t \approx 0$  para  $t \leq -Q$  por lo que tomo  $a_t = 0$  para  $t \leq -Q$ . Con base en esta última condición y tomando los valores observados de la serie y los valores calculados para  $Y_i$   $i = 0, 1, \dots, Q$  es posible calcular los valores de  $a_t$  para  $t > -Q$ . De esta forma, el proceso es:

$$a_{-20} = 0$$

$$a_{-19} = Y_{-19} - 0.7383098Y_{-20} + 0.5825704 a_{-20}$$

$$= 0.000006 - 0.7383098(0.000000) + 0.5825704(0.000000) = 0.000006$$

$$a_{-18} = Y_{-18} - 0.7383098Y_{-19} + 0.5825704 a_{-19}$$

$$= 0.000009 - 0.7383098(0.000006) + 0.5825704(0.000006) = 0.000008$$

.....

$$a_{97} = Y_{97} - 0.7383098Y_{96} + 0.5825704 a_{96}$$

$$=-0.012251-0.7383098(-0.019018)+0.5825704(-0.01679)=-.007991$$

e) Con los datos alcanzados en el punto anterior realizo la suma de cuadrados cuyo valor es 0.083084325596754.

A continuación, utilizando el Método de Marquardt, STATGRAPHICS produce los valores parametrales de  $\phi = -0.00502$  y  $\theta = 0.00926$  que minimizan la suma de cuadrados con el que obtengo un valor para la suma de cuadrados de 0.078930620548620.

Dado que la desviación estándar de  $\phi$  es 3.25147 y la de  $\theta$  es 3.25038 deduzco que el cero es un valor permisible y lógico del parámetro; además, debido a las altas desviaciones pienso que la serie no presenta información suficiente para determinar que el proceso sea efectivamente un **ARMA(1,1)** por lo que rechazo este modelo para representar la serie.

Como resultado de la sobreparametrización, puedo establecer que el mejor modelo para representar a la serie  $Y_t$  en estudio es el ruido blanco.

El cálculo del estadístico  $\mathcal{L}^2$  que aparece en la sección 3.3.4 no tiene sentido ya que los valores de los choques aleatorios son los de la propia serie.

La suposición después de realizar todas las pruebas de verificación es que el proceso de ruido blanco es el mejor modelo para describir a la serie  $Y_t$  y, por tanto, este será el modelo que usaré para realizar los pronósticos correspondientes.

## 6.4 Pronósticos

Para calcular los pronósticos tengo que recurrir a mi serie transformada y para ello debo recordar que apliqué una diferencia estacional de periodo 12 y una diferencia regular por lo que para el cálculo de los pronósticos parto de:

$$Y_t = a_t \quad \text{con } Y_t = \nabla^d \nabla_{12}^D T(Z_t) \Rightarrow$$

$$\nabla \nabla_{12} T(Z_t) = a_t \Rightarrow$$

$$(1-B)(1-B^{12})T(Z_t) = a_t \Rightarrow$$

$$(1-B-B^{12}+B^{13})T(Z_t) = a_t \Rightarrow$$

$$T(Z_t) = T(Z_{t-1}) + T(Z_{t-12}) - T(Z_{t-13}) + a_t$$

Los pronósticos a los diferentes horizontes  $h$  y bajo el origen  $t = 110$  (última observación) considerando las expresiones (4.2.1) y (4.2.2) serán:

$$\mathfrak{T}(Z_t)(h) = E_t [\varphi_1 T(Z_{t+h-1}) + \varphi_2 T(Z_{t+h-2}) + \dots + \varphi_{p+d} T(Z_{t+h-p-d}) + a_{t+h} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \dots - \theta_q a_{t+h-q}]$$

Como  $p = q = 0$ , la expresión anterior se reduce a

$$\mathfrak{T}(Z_t)(h) = E_t [T(Z_{t+h})] = E_t [(T(Z_{t+h-1}))] + E_t [T(Z_{t+h-12})] - E_t [T(Z_{t+h-13})] + E_t (a_{t+h})$$

que, tomando en consideración la última observación al tiempo  $t=110$ , es la expresión que sirve para realizar los pronósticos que a continuación presento:

Para  $T(Z_{111})$ ,  $h = 111 - 110 = 1$

$$\mathfrak{T}(Z_{110})(1) = E_{110} [T(Z_{111})] = E_{110}[T(Z_{110+1-1})] + E_{110}[T(Z_{110+1-12})] - E_{110}[T(Z_{110+1-13})] \\ + E_{110}(a_{110+1})$$

$$\mathfrak{T}(Z_{110})(1) = E_{110} [T(Z_{110})] + E_{110} [T(Z_{99})] - E_{110} [T(Z_{98})] + E_{110} (a_{111})$$

Aquí es donde aplico los resultados de (4.1.10) para obtener.

$$\mathfrak{T}(Z_{110})(1) = T(Z_{110}) + T(Z_{99}) - T(Z_{98}) + 0 = 13.151 + 13.026 - 12.904 + 0$$

$$\mathfrak{T}(Z_{110})(1) = 13.273155$$

Para  $T(Z_{112})$ ,  $h = 112 - 110 = 2$

$$\mathfrak{T}(Z_{110})(2) = E_{110} [T(Z_{112})] = E_{110}[T(Z_{110+2-1})] + E_{110}[T(Z_{110+2-12})] - E_{110}[T(Z_{110+2-13})] \\ + E_{110}(a_{110+2})$$

$$\mathfrak{T}(Z_{110})(2) = E_{110} [T(Z_{111})] + E_{110} [T(Z_{100})] - E_{110} [T(Z_{99})] + E_{110} (a_{112})$$

$$\mathfrak{T}(Z_{110})(2) = \mathfrak{T}(Z_{110})(1) + T(Z_{100}) - T(Z_{99}) + 0 = 13.273 + 13.136 - 13.026 + 0$$

$$\mathfrak{T}(Z_{110})(2) = 13.383038$$

Para  $T(Z_{113})$ ,  $h = 113 - 110 = 3$

$$\mathfrak{T}(Z_{110})(3) = E_{110} [T(Z_{113})] = E_{110}[T(Z_{110+3-1})] + E_{110}[T(Z_{110+3-12})] - E_{110}[T(Z_{110+3-13})] \\ + E_{110}(a_{110+3})$$

$$\mathfrak{T}(Z_{110})(3) = E_{110} [T(Z_{112})] + E_{110} [T(Z_{101})] - E_{110} [T(Z_{100})] + E_{110} (a_{113})$$

$$\mathfrak{T}(Z_{110})(3) = \mathfrak{T}(Z_{110})(2) + T(Z_{101}) - T(Z_{100}) + 0 = 13.383 + 13.272 - 13.136 + 0$$

$$\mathfrak{T}(Z_{110})(3) = 13.519221$$

Para  $T(Z_{114})$ ,  $h = 114 - 110 = 4$

$$\mathfrak{T}(Z_{110})(4) = \mathfrak{T}(Z_{110})(3) + T(Z_{102}) - T(Z_{101}) + 0 = 13.519 + 11.073 - 13.272 + 0$$

$$\mathfrak{T}(Z_{110})(4) = 11.319746$$

Para  $T(Z_{115})$ ,  $h = 115 - 110 = 5$

$$\mathfrak{T}(Z_{110})(5) = \mathfrak{T}(Z_{110})(4) + T(Z_{103}) - T(Z_{102}) + 0 = 11.320 + 11.649 - 11.073 + 0$$

$$\mathfrak{T}(Z_{110})(5) = 11.895870$$

Para  $T(Z_{116})$ ,  $h = 116 - 110 = 6$

$$\mathfrak{T}(Z_{110})(6) = \mathfrak{T}(Z_{110})(5) + T(Z_{104}) - T(Z_{103}) + 0 = 11.896 + 12.041 - 11.649 + 0$$

$$\mathfrak{T}(Z_{110})(6) = 12.288111$$

Para  $T(Z_{117})$ ,  $h = 117 - 110 = 7$

$$\mathfrak{T}(Z_{110})(7) = \mathfrak{T}(Z_{110})(6) + T(Z_{105}) - T(Z_{104}) + 0 = 12.288 + 12.335 - 12.041 + 0$$

$$\mathfrak{T}(Z_{110})(7) = 12.582682$$

Para  $T(Z_{118})$ ,  $h = 118 - 110 = 8$

$$\mathfrak{T}(Z_{110})(8) = \mathfrak{T}(Z_{110})(7) + T(Z_{106}) - T(Z_{105}) + 0 = 12.583 + 12.562 - 12.335 + 0$$

$$\mathfrak{T}(Z_{110})(8) = 12.809027$$

Para  $T(Z_{119})$ ,  $h = 119 - 110 = 9$

$$\mathfrak{T}(Z_{110})(9) = \mathfrak{T}(Z_{110})(8) + T(Z_{107}) - T(Z_{106}) + 0 = 12.809 + 12.757 - 12.562 + 0$$

$$\mathfrak{T}(Z_{110})(9) = 13.003953$$

Para  $T(Z_{120})$ ,  $h = 120 - 110 = 10$

$$\mathfrak{T}(Z_{110})(10) = \mathfrak{T}(Z_{110})(9) + T(Z_{108}) - T(Z_{107}) + 0 = 13.004 + 12.916 - 12.757 + 0$$

$$\mathfrak{T}(Z_{110})(10) = 13.163277$$

Para  $T(Z_{121})$ ,  $h = 121 - 110 = 11$

$$\mathfrak{T}(Z_{110})(11) = \mathfrak{T}(Z_{110})(10) + T(Z_{109}) - T(Z_{108}) + 0 = 13.163 + 13.044 - 12.916 + 0$$

$$\mathfrak{T}(Z_{110})(11) = 13.291110$$

Para  $T(Z_{122})$ ,  $h = 122 - 110 = 12$

$$\mathfrak{T}(Z_{110})(12) = \mathfrak{T}(Z_{110})(11) + T(Z_{110}) - T(Z_{109}) + 0 = 13.291 + 13.151 - 13.044 + 0$$

$$\mathfrak{T}(Z_{110})(12) = 13.398366$$



Para  $T(Z_{123})$ ,  $h = 123 - 110 = 13$

$$\mathbb{T}(Z_{110})(13) = \mathbb{T}(Z_{110})(12) + \mathbb{T}(Z_{110})(1) - T(Z_{110}) + 0 = 13.398 + 13.273 - 13.151 + 0$$

$$\mathbb{T}(Z_{110})(13) = 13.520348$$

A través de este ejercicio he calculado los pronósticos para la serie transformada, pero lo que me interesa es conocer los pronósticos de la serie original y para ello debo aplicar a los pronósticos obtenidos la función inversa de la transformación, que en este caso sería la exponencial dado que apliqué el logaritmo natural a la serie original, y la corrección por sesgo, que está dada por la expresión  $\mathcal{C}_{i,h}(h) = \exp\{\text{Var}_t(e_t(h))/2\}$ .

Para calcular  $\text{Var}_t(e_t(h))$  me baso en la expresión (4.1.7), cuyas ponderaciones  $\psi_i$ ,  $i=1, 2, \dots$  se obtienen a través de los parámetros  $\phi$  y  $\theta$ . Como en este caso no existen parámetros, la estimación de (4.1.7) sólo conserva a  $\psi_0 = 1$  para convertirse en.

$$\text{Var}_t(e_t(h)) = \psi_0 S_0^2 = 1 * 0.00082218$$

Por lo que para toda  $h$

$$\mathcal{C}_{i,h}(h) = \exp[0.00082218 / 2] = \exp[0.0004111] = 1.000411$$

Y así los pronósticos puntuales óptimos finales los obtengo aplicando (4.2.4) Por ejemplo para el primer pronóstico.

$$\mathbb{T}(Z_{110})(1) = [\exp(13.273155)] * 1.000411 = 581377.170552 * 1.000411 \Rightarrow$$

$$\mathbb{T}(Z_{111}) = \mathbb{T}(Z_{110})(1) = \mathbf{581,616.12}$$

Siguiendo esta metodología llego a los pronósticos óptimos que presento en la siguiente tabla:

**ÓPTIMO:**

<b>h</b>	<b><math>\mathcal{T}(Z_{110})(h)</math></b>	<b><math>T^{-1}(\mathcal{T}(Z_{110})(h))</math></b>	<b><math>T^{-1}(\mathcal{T}(Z_{110})(h))\mathcal{G}_{t,\lambda}(h)</math></b>
1	13.273155	581,377.170552	581,616.12
2	13.383038	648,902.660100	649,169.36
3	13.519221	743,572.089625	743,877.70
4	11.319746	82,433.402178	82,467.28
5	11.895870	146,659.668307	146,719.95
6	12.288111	217,099.488542	217,188.72
7	12.582682	291,466.979359	291,586.77
8	12.809027	365,501.988994	365,652.21
9	13.003953	444,165.713329	444,348.27
10	13.163277	520,881.058476	521,095.14
11	13.291110	591,910.073822	592,153.35
12	13.398366	658,925.660107	659,196.48
13	13.520348	744,410.567763	744,716.52

\* Pronóstico óptimo que incluye la corrección por sesgo.

Ahora obtengo los intervalos al 95% de confianza para estos pronósticos aplicando el resultado presentado en (4.3.2), según lo cual el intervalo está dado por.

$$\mathcal{G}_{t,\lambda}(h)T^{-1}\{\mathcal{T}(Z_t)(h) \pm u_{\alpha/2}(\text{Var}_t(e_t(h)))^{1/2}\}$$

Donde

$$[\text{Var}_t(e_t(h))]^{1/2} = [\psi_0 S_{\sigma^2}]^{1/2} = [0.00082218]^{1/2} = 0.028674 \quad \text{y}$$

$$u_{\alpha/2} = u_{0.05/2} = 1.96$$

Así, para el intervalo del primer pronóstico tengo.

$$\mathcal{G}_{t,\lambda}(h)T^{-1}\{\mathcal{T}(Z_t)(h) \pm u_{\alpha/2}(\text{Var}_t(e_t(h)))^{1/2}\} = 1.000411 * \exp\{13.273155 \pm 1.96 * 0.028674\} \Rightarrow$$

$$\begin{aligned} \text{Límite inferior} &= 1.000411 * \exp\{13.273155 - 1.96 * 0.028674\} = 1.000411 * \exp\{13.216954\} \\ &= 1.000411 * 549604.385886 = 549,830.273289 \end{aligned}$$

$$\begin{aligned} \text{Límite superior} &= 1.000411 * \exp\{13.273155 + 1.96 * 0.028674\} = 1.000411 * \exp\{13.329356\} \\ &= 1.000411 * 614986.748868 = 615,239.508422 \end{aligned}$$

La siguiente tabla muestra los intervalos obtenidos a través de esta técnica:

<i>h</i>	<i>ESTIMACIÓN</i>	<i>LÍMITE</i>	
		<i>INFERIOR</i>	<i>SUPERIOR</i>
1	13.273155	549,830.27	615,239.51
2	13.383038	613,691.67	686,698.02
3	13.519221	703,224.11	786,881.48
4	11.319746	77,960.37	87,234.74
5	11.895870	138,701.57	155,201.87
6	12.288111	205,319.16	229,744.46
7	12.582682	275,651.29	308,443.49
8	12.809027	345,668.99	386,790.67
9	13.003953	420,064.23	470,036.16
10	13.163277	492,616.82	551,219.80
11	13.291110	559,791.64	626,385.90
12	13.398366	623,170.80	697,304.81
13	13.520348	704,017.09	787,768.79

Ahora trataré de actualizar pronósticos tomando como base los valores conocidos posteriores a septiembre de 1997 (*t* = 110) Estos valores son:

<i>h</i>	<i>t</i>	<i>PERIODO</i>	<i>INGRESO</i>
1	111	1997-OCT	579,088.00
2	112	1997-NOV	640,291.30
3	113	1997-DIC	730,441.30
4	114	1998-ENE	80,347.40
5	115	1998-FEB	142,996.50
6	116	1998-MAR	208,907.90

Que, como se puede apreciar, están incluidos en los intervalos de confianza calculados previamente. Estas nuevas observaciones de la serie tienen un sesgo respecto de los valores óptimos calculados; dicho sesgo se refleja en que los valores estimados óptimos se encuentran por arriba de los valores reales observados en todos los casos.

Con base en la expresión (4.4.1) concluyo que no es posible actualizar los pronósticos pues, como ya expliqué, no cuento con parámetros  $\psi_1$ ,  $\psi_2$ ,  $\psi_3$ ,  $\psi_4$ ,  $\psi_5$  y  $\psi_6$ . Esta imposibilidad de actualización es también lógica si se recuerda que para la serie estacionaria correspondiente a la serie original del ingreso federal obtuve que sólo representaba ruido blanco, es decir, que no existe influencia de datos previos.

Por último, revisaré la capacidad de pronóstico de mi modelo considerando los pronósticos originales; así, tomando como origen a  $t = 110$  y los horizontes  $h = 1, 2, 3, 4, 5$  y  $6$  calculo los errores de los pronósticos de la serie transformada respecto del valor correspondiente a las observaciones más recientes de la serie una vez que se les ha aplicado la transformación

		serie			error
		original	transformada	estimada	
h	t	$Z_t$	$T(Z_t)$	$\hat{T}(Z_t)$	$T(Z_t) - \hat{T}(Z_t)$
1	111	579,088.00	13.27	13.27	0.00
2	112	640,291.30	13.37	13.38	-0.01
3	113	730,441.30	13.50	13.52	-0.02
4	114	80,347.40	11.29	11.32	-0.03
5	115	142,996.50	11.87	11.90	-0.03
6	116	208,907.90	12.25	12.29	-0.04
<i>Promedio <math>\bar{M}</math></i>					-0.020751
ECM					0.000548

Como puede verse, ambos valores son próximos a cero por lo que es posible concluir que el modelo tiene una muy buena capacidad de pronóstico.

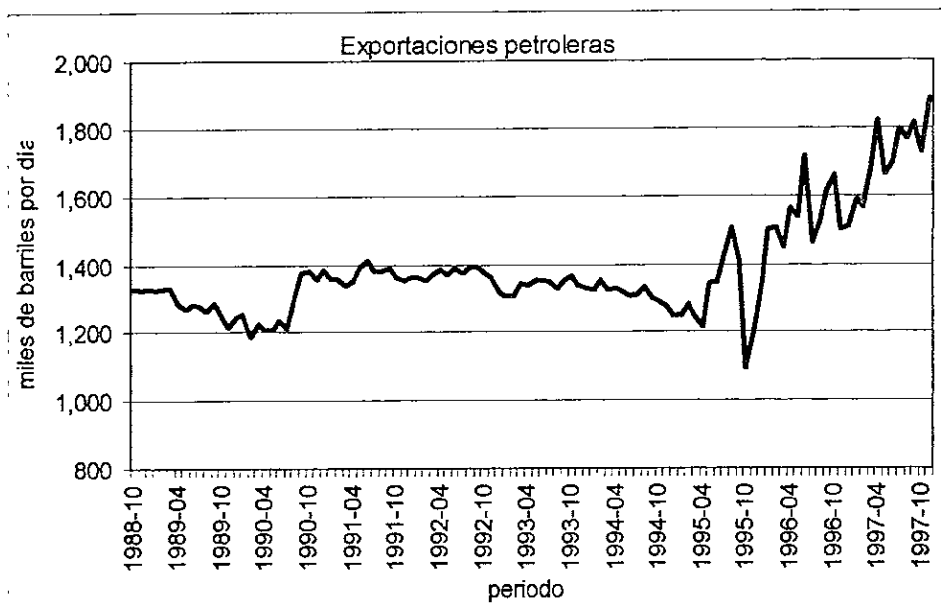
Con esto finalizo el análisis de la primer serie en el que concluyo que, en esencia, la serie de tiempo correspondiente al ingreso federal se comporta como un proceso **ARIMA**(0,1,0)×(0,1,0)<sub>12</sub>, es decir, es un proceso integrado estacional de periodo doce e integrado regular y mediante este modelo se pueden obtener pronósticos confiables de la serie.

## 7 Serie de tiempo: Exportaciones petroleras

De todos es sabido que las exportaciones petroleras son de gran importancia para la economía mexicana, pues gran parte de los ingresos que el gobierno federal obtiene proviene precisamente de éstas. El total de las exportaciones petroleras incluye tres tipos de petróleo: Itsmo, Maya y Olmeca. En este análisis revisaré las exportaciones petroleras con base en los miles de barriles promedio exportados por día en cada mes.

Para ajustar un modelo a esta serie utilizaré los valores mensuales registrados desde octubre de 1988 hasta noviembre de 1997, es decir, ciento diez observaciones a las cuales les aplicaré a ahora la metodología de Box y Jenkins.

La siguiente gráfica muestra los valores de las exportaciones petroleras que utilizaré en el ejercicio, los datos los mostraré en la tabla B:



## 7.1 Identificación

Como puede notarse en la gráfica, esta serie se obtiene por muestreo y presenta un comportamiento que al principio es estable y luego es aparentemente creciente, razón por la cual puedo suponer que requeriré de una transformación estabilizadora de la varianza.

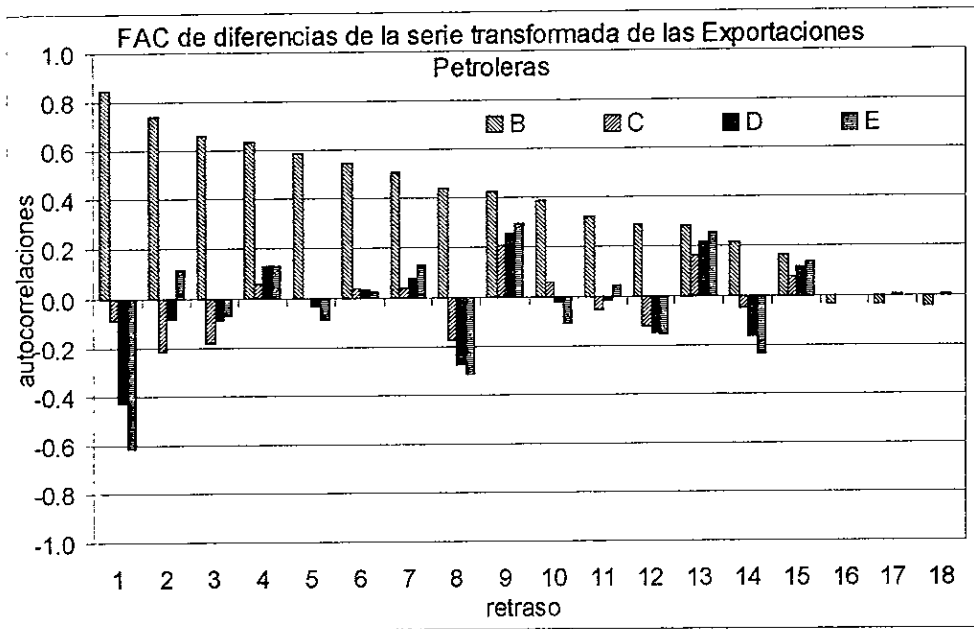
Adicionalmente puedo decir que la serie en estudio no presenta ninguna tendencia estacional por lo que para describirla usaré un modelo **ARIMA** regular.

Con el fin de identificar el parámetro de transformación  $\lambda$  que permita estabilizar la varianza formo 10 grupos con 11 observaciones cada uno (por lo que no elimino ninguna observación para la prueba) y procedo a calcular los coeficientes de variación para diferentes valores de  $\lambda$ . Los resultados obtenidos son los siguientes:

	VALORES DEL GRUPO		VALORES PARAMETRALES DE $\lambda$					
	MEDIA	DESVIACIÓN	-1	-0.5	-0.4	-0.2	0	
GRUPO 1	1,302.86	27.85	0.0000	0.0006	0.0012	0.0051	0.0214	
GRUPO 2	1,227.62	27.20	0.0000	0.0006	0.0013	0.0053	0.0222	
GRUPO 3	1,363.91	28.48	0.0000	0.0006	0.0012	0.0049	0.0209	
GRUPO 4	1,368.27	13.07	0.0000	0.0003	0.0005	0.0023	0.0096	
GRUPO 5	1,351.82	33.36	0.0000	0.0007	0.0014	0.0058	0.0247	
GRUPO 6	1,341.18	14.65	0.0000	0.0003	0.0006	0.0026	0.0109	
GRUPO 7	1,293.64	27.77	0.0000	0.0006	0.0012	0.0051	0.0215	
GRUPO 8	1,330.09	130.97	0.0001	0.0027	0.0055	0.0234	0.0985	
GRUPO 9	1,548.36	82.47	0.0000	0.0014	0.0028	0.0123	0.0533	
GRUPO 10	1,727.64	98.56	0.0000	0.0014	0.0029	0.0128	0.0570	
COEFICIENTE DE VARIACIÓN			$M(\lambda)$	0.0000	0.0009	0.0019	0.0080	0.0340
			$S(\lambda)$	0.0000	0.0007	0.0015	0.0065	0.0277
				0.8164	<b>0.8133</b>	0.8132	0.8137	0.8151

Con base en esta tabla decido que el valor más práctico de  $\lambda$  es  $-0.5$  que transforma a la serie con el inverso de la raíz cuadrada del valor observado, conforme a la transformación propuesta por Bartlett y mostrada en (3.1.1.6). Así, al aplicar el exponente  $-0.5$  a la serie original conforme la serie transformada de la tabla B con 110 observaciones.

A continuación estabilizo el nivel de la serie mediante la aplicación de diferencias regulares a la serie transformada (dado que no percibí ningún comportamiento estacional). Para detectar cuál será la serie estacionaria grafico los primeros valores de la **FAC** de algunas diferencias de la serie transformada.



las **FACs** graficadas corresponden a las siguientes diferencias sobre la serie transformada:



- B De la serie transformada.
- C De la primer diferencia regular de la serie transformada.
- D De la segunda diferencia regular de la serie transformada.
- E De la tercer diferencia regular de la serie transformada.

De esta gráfica las autocorrelaciones más pequeñas corresponden a la serie C, por lo que ésta pasa a ser la serie estacionaria.

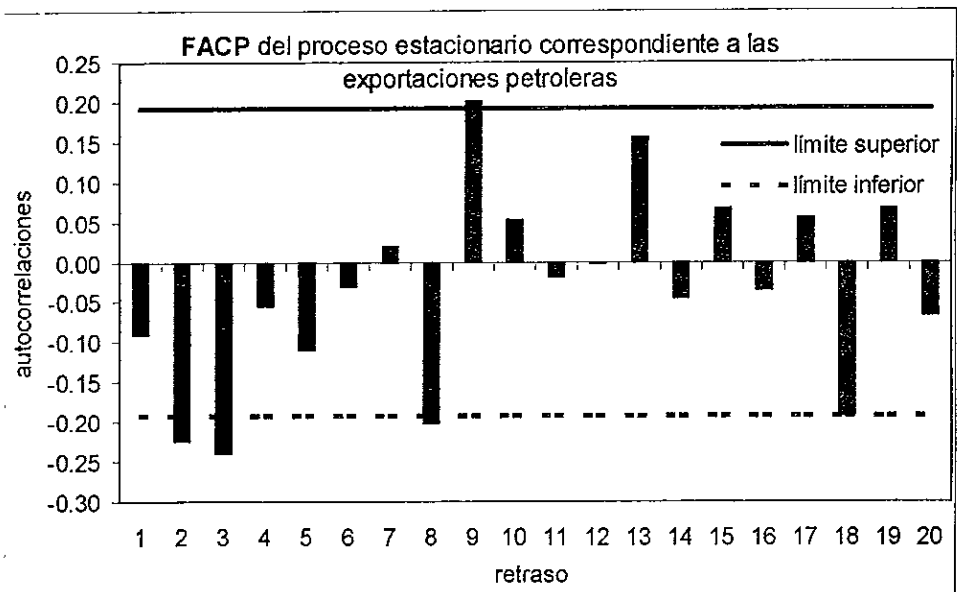
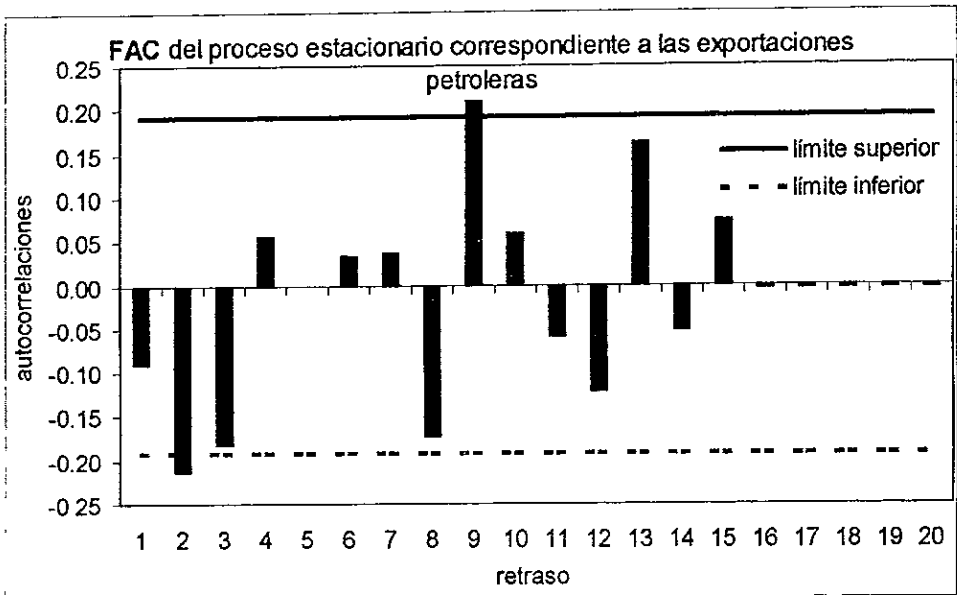
Por otra parte, utilizando el método de comparación de las desviaciones estándar de las series llego a los siguientes valores para cada una:

<i>SERIE</i>	B	C	D	E
<i>DESV. ESTÁNDAR</i>	0.0013	0.0006	0.0009	0.0016

Obtengo que también a través de este análisis la serie de menor desviación estándar es C. Con esto ratifico a C como la serie que con 109 observaciones ( $110-1=109$ ) consideraré como estacionaria. Esto significa que será a la primer diferencia regular de la serie transformada, que presento como serie estacionaria en la tabla B, a la que tendré que ajustar un modelo **ARIMA**.

Para identificar el proceso que sigue esta serie analizo a continuación su **FAC** y **FACP** muestrales. En la tabla B muestro los primeros treinta valores de dichas funciones que cacculo con base en la fórmula mostrada en (3.1.3.1) y la sustitución de estos valores en la expresión (1.3.9) y cuya forma de cálculo revisé en el capítulo 6

Para determinar si el proceso no es ruido blanco utilizo las aproximaciones de Bartlett para formar el intervalo que, en caso de tratarse de ruido blanco, contendrá a todos los valores de la **FAC** y de la **FACP** muestrales  $(-2/[109]^{1/2}, 2/[109]^{1/2}) = (-2*0.095783, 2*0.095783) = (-0.192, 0.192)$ . Las gráficas correspondientes son:



Este supuesto, como se aprecia en las gráficas anteriores, no se cumple pues tanto en la **FAC** como en la **FACP** existen observaciones que quedan fuera del

intervalo construido hasta el retraso  $k = 9$ . Por tanto determino que el proceso no es sólo ruido blanco.

Por el comportamiento de la **FAC** y de la **FACP** muestrales podría pensarse en que el proceso sigue un modelo **MA(9)**, en cuyo caso los valores de la **FAC** para atrasos  $k > 9$  serían diferentes de cero si su valor absoluto fuera mayor al obtenido según el inciso b) de la sección 3.1.3 que en este caso es:

$$1.96 * [(1 + \sum_{j=1}^9 \rho_j^2) / N]^{1/2} \approx 1.96 * [(1 + \sum_{j=1}^9 r_j^2) / 109]^{1/2} = 1.96 * [(1 + (-0.091)^2 + (-0.214)^2 + (-0.182)^2 + (0.057)^2 + (-0.001)^2 + (0.034)^2 + (0.037)^2 + (-0.173)^2 + (0.209)^2) / 109]^{1/2} = 0.2028$$

Como se aprecia en la tabla B, todos los valores de la **FAC** para atrasos  $k > 9$  cumplen con la condición de ser iguales a cero y por tanto podría establecer este modelo como aceptable; sin embargo, la **FACP** no converge a cero tan rápidamente como es de esperarse en un proceso **MA** por lo que puedo pensar en explicar mejor a la serie a través de un modelo **ARMA**

Para determinar los posibles valores de  $p$  y  $q$  calculo la tabla de determinantes que presenté en la sección 3.1.3 del capítulo sobre identificación. Por ejemplo, para  $i = 2$  y  $j = 3$  el cálculo se realiza de la siguiente forma.

$$\Delta(2, 3) = \begin{vmatrix} r_2 & r_1 & r_0 \\ r_3 & r_2 & r_1 \\ r_4 & r_3 & r_2 \end{vmatrix} = \begin{vmatrix} -0.214 & -0.091 & 1.000 \\ -0.182 & -0.214 & -0.091 \\ 0.057 & -0.182 & -0.214 \end{vmatrix} = 0.043$$

Con este procedimiento llego a los siguientes resultados:

ORDEN MA(q)	ORDEN AR(p)								
	1	2	3	4	5	6	7	8	9
1	-0.091	0.222	-0.224	0.045	-0.081	0.019	0.011	0.099	0.083
2	-0.214	0.029	0.043	-0.019	0.008	0.002	-0.003	0.018	0.021
3	-0.182	0.045	-0.011	0.001	-0.001	0.001	-0.002	0.004	0.005
4	0.057	0.003	0.002	0.001	0.000	0.000	0.000	0.002	0.001
5	-0.001	-0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.034	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.037	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000
8	-0.173	0.022	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
9	0.209	0.054	0.012	0.001	0.000	0.000	0.000	0.000	0.000

Con base en esta tabla pudiera pensar que el proceso sigue un modelo **ARMA(3,3)** o bien un **ARMA(4,4)**. También parece lógico pensar en un **MA(9)** o un **AR(9)** que vendrían a incluir más parámetros que los modelos antes señalados.

## 7.2 Estimación

Considerando como el mejor modelo al **ARMA(3,3)** por su menor número de parámetros la representación del modelo sería:

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \phi_3 Z_{t-3} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3}$$

Y por tanto, requiero estimar parámetros autorregresivos y de promedios móviles.

Para encontrar los valores parametrales iniciales de estimación utilizo el método de momentos que presenté en la sección 3.2.3.

Primero debo resolver el siguiente sistema que se basa en la consideración de que cuando  $k > p$ ,  $\phi(B)Z_k = 0$  para determinar los valores iniciales de  $\phi_i$ ,  $i=1, 2, 3$ .

$$r_4 = \phi_1 r_3 + \phi_2 r_2 + \phi_3 r_1 \Rightarrow 0.056514 = -0.182180\phi_1 - 0.213643\phi_2 - 0.091243\phi_3$$

$$r_5 = \phi_1 r_4 + \phi_2 r_3 + \phi_3 r_2 \Rightarrow -0.000806 = 0.056514\phi_1 - 0.182180\phi_2 - 0.213643\phi_3$$

$$r_6 = \phi_1 r_5 + \phi_2 r_4 + \phi_3 r_3 \Rightarrow 0.033695 = -0.000806\phi_1 + 0.056514\phi_2 - 0.182180\phi_3$$

La matriz del sistema es:

$$\begin{vmatrix} -0.182180 & -0.213643 & -0.091243 & \vdots & 0.056514 \\ 0.056514 & -0.182180 & -0.213643 & \vdots & -0.000806 \\ -0.000806 & 0.056514 & -0.182180 & \vdots & 0.033695 \end{vmatrix}$$

que al ser reducida a la identidad proporciona las soluciones:

$$\phi_1 = -0.331899$$

$$\phi_2 = 0.085527$$

$$\phi_3 = -0.156955$$

A continuación estimo los valores iniciales de los parámetros de promedios móviles con base en que

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3}$$

y con base en el cálculo de las autocovarianzas  $\gamma_i$ ,  $i = 0, 1, 2, 3, 4$  que se obtienen a través de

$$\begin{aligned} \gamma_0 &= E(Z_t Z_t) = E[Z_t^*(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3})] \\ &= E[\phi_1 Z_t Z_{t-1} + \phi_2 Z_t Z_{t-2} + \phi_3 Z_t Z_{t-3} + Z_t a_t - \theta_1 Z_t a_{t-1} - \theta_2 Z_t a_{t-2} - \theta_3 Z_t a_{t-3}] \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \phi_3 \gamma_3 + \{1 - \theta_1[-\theta_1 + \phi_1] - \theta_2[-\theta_2 - \phi_1 \theta_1 + \phi_2] - \theta_3[-\theta_3 - \phi_1 \theta_2 - \phi_2 \theta_1 + \phi_3]\} \sigma_a^2 \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E(Z_t Z_{t-1}) = E[Z_{t-1}^*(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3})] \\ &= E[\phi_1 Z_{t-1} Z_{t-1} + \phi_2 Z_{t-1} Z_{t-2} + \phi_3 Z_{t-1} Z_{t-3} + Z_{t-1} a_t - \theta_1 Z_{t-1} a_{t-1} - \theta_2 Z_{t-1} a_{t-2} - \theta_3 Z_{t-1} a_{t-3}] \\ &= \phi_1 \gamma_0 + \phi_2 \gamma_1 + \phi_3 \gamma_2 + \{-\theta_1 - \theta_2[-\theta_1 + \phi_1] - \theta_3[-\theta_2 - \phi_1 \theta_1 + \phi_2]\} \sigma_a^2 \end{aligned}$$

$$\begin{aligned}\gamma_2 &= E(Z_t Z_{t-2}) = E[Z_{t-2}^*(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3})] \\ &= E[\phi_1 Z_{t-2} Z_{t-1} + \phi_2 Z_{t-2} Z_{t-2} + \phi_3 Z_{t-2} Z_{t-3} + Z_{t-2} a_t - \theta_1 Z_{t-2} a_{t-1} - \theta_2 Z_{t-2} a_{t-2} - \theta_3 Z_{t-2} a_{t-3}] \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_0 + \phi_3 \gamma_1 + \{-\theta_2 - \theta_3[-\theta_1 + \phi_1]\} \sigma_a^2\end{aligned}$$

$$\begin{aligned}\gamma_3 &= E(Z_t Z_{t-3}) = E[Z_{t-3}^*(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3})] \\ &= E[\phi_1 Z_{t-3} Z_{t-1} + \phi_2 Z_{t-3} Z_{t-2} + \phi_3 Z_{t-3} Z_{t-3} + Z_{t-3} a_t - \theta_1 Z_{t-3} a_{t-1} - \theta_2 Z_{t-3} a_{t-2} - \theta_3 Z_{t-3} a_{t-3}] \\ &= \phi_1 \gamma_2 + \phi_2 \gamma_1 + \phi_3 \gamma_0 + \{-\theta_3\} \sigma_a^2\end{aligned}$$

Con esto formo el siguiente sistema de ecuaciones donde A, B, C y D representan a los términos que se encuentran multiplicando a  $\sigma_a^2$  en cada una de las expresiones:

$$\begin{aligned}\gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \phi_3 \gamma_3 + \sigma_a^2 A & \Rightarrow & \quad \gamma_0 - \phi_1 \gamma_1 - \phi_2 \gamma_2 - \phi_3 \gamma_3 = \sigma_a^2 A \\ \gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_1 + \phi_3 \gamma_2 + \sigma_a^2 B & \Rightarrow & \quad -\phi_1 \gamma_0 + (1 - \phi_2) \gamma_1 - \phi_3 \gamma_2 = \sigma_a^2 B \\ \gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0 + \phi_3 \gamma_1 + \sigma_a^2 C & \Rightarrow & \quad -\phi_2 \gamma_0 - (\phi_1 + \phi_3) \gamma_1 + \gamma_2 = \sigma_a^2 C \\ \gamma_3 &= \phi_1 \gamma_2 + \phi_2 \gamma_1 + \phi_3 \gamma_0 + \sigma_a^2 D & \Rightarrow & \quad -\phi_3 \gamma_0 - \phi_2 \gamma_1 - \phi_1 \gamma_2 + \gamma_3 = \sigma_a^2 D\end{aligned}$$

resolviendo este sistema llego a que las soluciones iniciales para  $\theta_1, \theta_2$  y  $\theta_3$  son:

$$\theta_1 = -0.01835 \quad \theta_2 = 0.39018 \quad \theta_3 = 0.11259$$

Con estos parámetros calculo los que serían los errores aleatorios basados en el supuesto de incondicionalidad:

a) Expreso al modelo como  $a_t = Z_t - \phi_1 Z_{t+1} - \phi_2 Z_{t+2} - \phi_3 Z_{t+3} + \theta_1 a_{t+1} + \theta_2 a_{t+2} + \theta_3 a_{t+3}$ .

b) Comienzo a calcular los valores de  $a_i$  a partir de  $t = N - p = 109 - 3 = 106$ ,  
 teniendo presente que  $a_{N-p+i} = 0$  para  $i = 1, 2, \dots, q \Rightarrow$

$$e_{109-3+1} = e_{107} = 0 \qquad e_{109-3+2} = e_{108} = 0 \qquad e_{109-3+3} = e_{109} = 0$$

con lo que obtengo

$$\begin{aligned} e_{106} &= Z_{106} + 0.331899Z_{107} - 0.085527Z_{108} + 0.156955Z_{109} - 0.018350e_{107} \\ &+ 0.39018e_{108} + 0.1125900e_{109} \\ &= 0.000193 + 0.331899(-0.000310) - 0.085527(0.000584) + \\ &0.156955(-0.001012) - 0.018350(0.000000) + 0.39018(0.000000) + \\ &0.1125900(0.000000) = -0.000119 \end{aligned}$$

$$\begin{aligned} e_{105} &= Z_{105} + 0.331899Z_{106} - 0.085527Z_{107} + 0.156955Z_{108} - 0.018350e_{106} \\ &+ 0.39018e_{107} + 0.1125900e_{108} = -0.000459 \end{aligned}$$

.....

$$\begin{aligned} e_1 &= Z_1 + 0.331899Z_2 - 0.085527Z_3 + 0.156955Z_4 - 0.01835e_2 + \\ &0.39018e_3 + 0.11259e_4 = 0.000141 \end{aligned}$$

c) Hago  $e_i = 0$  para  $i = 0, 1, 2, \dots$  y con los valores determinados para  $a_i$  y los  
 valores registrados de  $Z_t$  obtengo los valores para  $Z_{-i}$   $i = 0, 1, 2, \dots$  de forma  
 que

$$\begin{aligned} Z_0 &= -0.331899Z_1 + 0.085527Z_2 - 0.156955Z_3 + e_0 + 0.01835e_1 - \\ &0.39018e_2 - 0.11259e_3 = -0.000077 \end{aligned}$$

$$\begin{aligned} Z_{-1} &= -0.331899Z_2 + 0.085527Z_3 - 0.156955Z_4 + e_1 + 0.01835e_2 - \\ &0.39018e_3 - 0.11259e_4 = -0.000027 \end{aligned}$$

siguiendo este mismo método obtengo estos valores de  $Z_t$ :

t	Z <sub>t</sub>	t	Z <sub>t</sub>
-2	-0.000023	-7	0.000003
-3	0.000017	-8	-0.000002
-4	-0.000003	-9	0.000002
-5	0.000006	-10	-0.000001
-6	-0.000005	-11	0.000001

d) Expreso al modelo como  $Z_t = -0.331899Z_{t-1} + 0.085527Z_{t-2} - 0.156955Z_{t-3} + a_t + 0.01835a_{t-1} - 0.39018a_{t-2} - 0.11259a_{t-3}$  considerando que para  $Q = 12$  se cumple que  $Z_t \approx 0$  para  $t \leq -Q$ , por lo que tomo  $a_t = 0$  para  $t \leq -Q$ . Con base en esta última condición y tomando los valores observados de la serie y los valores obtenidos para  $Z_{-i}$   $i = 0, 1, \dots, Q$  es posible calcular los valores de  $a_t$  para  $t > -Q$ . De esta forma, el proceso es:

$$a_{-12} = a_{-13} = a_{-14} = 0$$

$$a_{-11} = Z_{-11} + 0.331899Z_{-12} - 0.085527Z_{-13} + 0.156955Z_{-14} - 0.01835a_{-12} + 0.39018a_{-13} + 0.11259a_{-14}$$

$$= 0.000001 + 0.331899(0.00) - 0.085527(0.00) + 0.156955(0.00) - 0.01835(0.00) + 0.39018(0.00) + 0.11259(0.00) = 0.000001$$

.....

$$a_{109} = Z_{109} + 0.331899Z_{108} - 0.085527Z_{107} + 0.156955Z_{106} - 0.01835a_{108} + 0.39018a_{107} + 0.11259a_{106}$$

$$= -0.001012 + 0.331899(0.000584) - 0.085527(-0.00031) + 0.156955(0.000193) - 0.01835(0.000275) + 0.39018(-0.000657) + 0.11259(-0.000005) = -0.0010221$$



e) Con los datos alcanzados en el punto anterior realizo la suma de cuadrados cuyo valor es 0.00003982.

A continuación, utilizando el Método de Marquardt, STATGRAPHICS produce los valores de

<i>Valor parametral</i>	<i>Desviación estándar</i>
$\phi_1 = -0.18874$	0.62204
$\phi_2 = -0.11728$	0.51557
$\phi_3 = -0.13286$	0.38725
$\theta_1 = -0.01266$	0.62530
$\theta_2 = 0.16506$	0.59590
$\theta_3 = 0.11957$	0.40689

que minimizan la suma de cuadrados. Con estos parámetros la suma de cuadrados se reduce a 0.000038509.

Finalmente el valor estimado de la varianza del ruido blanco conforme (3.2.4.2) será.

$$S_e^2 = \frac{\mathbf{S}(\phi, \theta)}{N-d-p-q-1} = [1/(110-1-3-3-1)] \sum_{i=11}^{109} a_i^2 = [1/(102)] \sum_{i=11}^{109} a_i^2 = \frac{0.000038509}{102}$$

Por lo tanto  $S_e^2 = 0.00000037754$ . La desviación estándar estimada del ruido blanco será entonces de 0.0006144428.

### 7.3 Verificación

Inicio con la revisión de los supuestos para el análisis de residuales que propuse en la sección 3.3.1:

**SUPUESTO 1:**  $\alpha_t$  tiene media cero. Los valores de la media y la desviación estándar muestrales de los  $N-d-p=110-1-3=106$  errores aleatorios son:

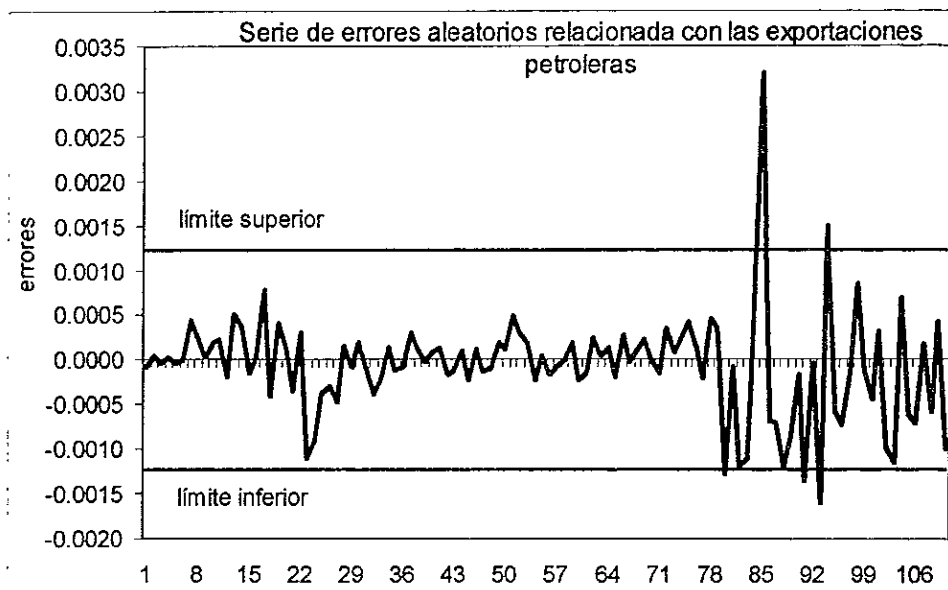
$$M(\alpha_t) = -0.0000755367 \quad S_\alpha = 0.0006144428$$

A continuación calculo el cociente:

$$(N-d-p)^{1/2} * M(\alpha_t) / S_\alpha = (110-1-3)^{1/2} * (-0.0000755367) / 0.0006144428 = -1.26570$$

Como se cumple que  $1.26570 < 2$  puede considerarse que la media del proceso no es diferente de cero y por tanto apruebo el supuesto.

**SUPUESTO 2:**  $\alpha_t$  tiene varianza constante. Presento a continuación la gráfica de la serie de los errores aleatorios calculados previamente.



A través de su revisión puede suponerse que la serie no presenta una tendencia creciente ni decreciente si no un comportamiento estable por lo que puedo decir que la varianza es constante y en consecuencia se cumple con el supuesto.

**SUPUESTO 3:** Las variables  $\{a_t\}$  son mutuamente independientes. Calculando las primeras 25 estimaciones de  $r_k(a)$  obtengo los valores:

k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$
1	-0.014552	6	0.031133	11	-0.032569	16	-0.010187	21	-0.011362
2	-0.018122	7	0.089864	12	-0.048048	17	-0.009225	22	-0.008406
3	-0.016829	8	-0.143184	13	0.159014	18	-0.010556	23	-0.006013
4	0.001679	9	0.208207	14	-0.061650	19	-0.011109	24	-0.005132
5	-0.062660	10	0.073811	15	0.044661	20	-0.010292	25	-0.004487

Y la estimación de la varianza de la **FAC** de los choques aleatorios se obtiene por medio de:

$$\underline{Var} [r_k(a)] = 1/(110-1-3) \Rightarrow$$

$$\{\underline{Var} [r_k(a)]\}^{1/2} = [1/106]^{1/2} = 0.097128586$$

formo el intervalo  $(-2*0.097128586, 2*0.097128586) = (-0.194257172, 0.194257172)$  que incluye a casi todos los valores estimados  $r_k(a)$ , excepto a 0.208207 por lo que podría considerar que el supuesto no se cumple.

Realizando la prueba Portmanteau mediante el estadístico **Q'** obtengo:

$$Q' = 106 * (106 + 2) \sum_{k=1}^{25} r_k^2(a) / (106 - k) = 106 * 108 * 0.00123118 = 14.094590$$

El valor en tablas de  $\chi^2_{(0.25, 25)} = 40.6465$ , dado que  $\chi^2 > Q'$  no se rechaza el supuesto de independencia de las variables.

Como puede observarse, en este supuesto existe una ligera indefinición de si debo aceptarlo o no, y por esta razón es que más adelante probaré con otro modelo.

**SUPUESTO 4:**  $\{\epsilon_t\}$  se distribuye normal. Verifico cuántas observaciones de la serie de choques aleatorios quedan fuera del intervalo:

$$(-2S_{\epsilon}, 2S_{\epsilon}) = (-2 \cdot 0.00061, 2 \cdot 0.00061) = (-0.00122889, 0.00122889)$$

Encuentro que los puntos fuera del intervalo señalado alcanzan un número de 5 que representan el 4.72% del total de las observaciones por lo que puedo decir que la serie sigue una distribución normal. La siguiente tabla muestra los valores de los choques aleatorios que se encuentran fuera del intervalo:

OBSV.	CHOQUE	OBSV.	CHOQUE	OBSV.	CHOQUE
79	0.001280	84	0.003204	90	-0.001358
92	-0.001614	93	0.001499		

**SUPUESTO 5:** No existen observaciones aberrantes. La única observación encontrada fuera del intervalo  $(-3S_{\epsilon}, 3S_{\epsilon}) = (-0.00184333, 0.00184333)$  es la correspondiente al número 84, con un valor de 0.003208. Este choque proviene de las observaciones originales correspondientes a los meses de junio a septiembre de 1995; aunque en este periodo no se presenta ningún hecho que pudiera afectar el comportamiento de la serie, puedo observar que este periodo presenta una breve depresión en el volumen de barriles exportados y que también corresponde a una elevación del nivel que había presentado la serie hasta junio de 1995 aproximadamente. No existe razón aparente para pensar que este valor corresponda a una intervención exógena a la serie y se acepta el supuesto.

Ahora aplico las pruebas independientes de los residuales mostradas en la sección 3.3.2.

**SUPUESTO 6:** El modelo es parsimonioso. Para verificar este supuesto formo los intervalos señalados en la expresión (3.3.2.1).

Parámetro $\beta$	$Var(\beta)$	$[Var(\beta)]^{1/2}$	$(\beta - 2[Var(\beta)]^{1/2}, \beta + 2[Var(\beta)]^{1/2})$
$\phi_1 = -0.18874$	0.38693	0.62204	(-1.432820, 1.055340)
$\phi_2 = -0.11728$	0.26581	0.51557	(-1.148420, 0.913860)
$\phi_3 = -0.13286$	0.14996	0.38725	(-0.907360, 0.641640)
$\theta_1 = -0.01266$	0.39100	0.62530	(-1.263260, 1.237940)
$\theta_2 = 0.16506$	0.35510	0.59590	(-1.026740, 1.356860)
$\theta_3 = 0.11957$	0.16556	0.40689	(-0.694210, 0.933350)

En todos los casos, el valor cero se encuentra dentro del intervalo formado y parece un valor lógico para todos los parámetros pero especialmente para  $\phi_3$  y  $\theta_3$ . Un comentario adicional en este punto es que, dado que los valores de las varianzas para todos los parámetros son muy grandes, puedo deducir que las observaciones de la serie no presentan información suficiente como para especificar que este sea un buen modelo.

**SUPUESTO 7:** El modelo es admisible. Para esta prueba calculo los determinantes en las expresiones (2.1.4) y (2.2.4). En primer término reviso si el modelo es estacionario a través de los parámetros  $\phi_1$ ,  $\phi_2$  y  $\phi_3$  obteniendo, como se puede notar a continuación, que todos los determinantes son positivos y por tanto puedo concluir que el modelo es estacionario:

$$D_1 = \begin{vmatrix} -1 & -0.13286 \\ -0.13286 & -1 \end{vmatrix} = 0.9823$$

$$D_2 = \begin{vmatrix} -1 & 0 & -0.13286 & -0.11728 \\ -0.18874 & -1 & 0 & -0.13286 \\ -0.13286 & 0 & -1 & -0.18874 \\ -0.11728 & -0.13286 & 0 & -1 \end{vmatrix} = 0.9565$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 & -0.13286 & -0.11728 & -0.18874 \\ -0.18874 & -1 & 0 & 0 & -0.13286 & -0.11728 \\ -0.11728 & -0.18874 & -1 & 0 & 0 & -0.13286 \\ -0.13286 & 0 & 0 & -1 & -0.18874 & -0.11728 \\ -0.11728 & -0.13286 & 0 & 0 & -1 & -0.18874 \\ -0.18874 & -0.11728 & -0.13286 & 0 & 0 & -1 \end{vmatrix} = 0.9072$$

A continuación reviso si el modelo es invertible a través de los parámetros  $\theta_1$ ,  $\theta_2$

y  $\theta_3$  para obtener

$$D_1 = \begin{vmatrix} -1 & 0.11957 \\ 0.11957 & -1 \end{vmatrix} = 0.98570$$

$$D_2 = \begin{vmatrix} -1 & 0 & 0.11957 & 0.16506 \\ -0.01266 & -1 & 0 & 0.11957 \\ 0.11957 & 0 & -1 & -0.01266 \\ 0.16506 & 0.11957 & 0 & -1 \end{vmatrix} = 0.94486$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 & 0.11957 & 0.16506 & -0.01266 \\ -0.01266 & -1 & 0 & 0 & 0.11957 & 0.16506 \\ 0.16506 & -0.01266 & -1 & 0 & 0 & 0.11957 \\ 0.11957 & 0 & 0 & -1 & -0.01266 & 0.16506 \\ 0.16506 & 0.11957 & 0 & 0 & -1 & -0.01266 \\ -0.01266 & 0.16506 & 0.11957 & 0 & 0 & -1 \end{vmatrix} = 0.90565$$

Y dado que todos los determinantes son positivos concluyo que el modelo es invertible. Así llego a que se cumple con la admisibilidad.

**SUPUESTO 8:** El modelo es estable en los parámetros. Para definir si se cumple este supuesto calculo las autocorrelaciones entre parámetros.

Para realizar este cálculo me basaré en la aproximación proporcionada por Box que mostré en la sección 3.2.2, calculando  $\delta_i$  como la diferencia entre los valores parametrales iniciales y los óptimos.

i	$\beta_{i,0}$	$\beta_{i,0} + \delta_i$	$\delta_i$
1	$\phi_{1,0} = -0.331899$	$\phi_1 = -0.18874$	0.143159
2	$\phi_{2,0} = 0.085527$	$\phi_2 = -0.11728$	-0.202807
3	$\phi_{3,0} = -0.156955$	$\phi_3 = -0.13286$	0.024095
4	$\theta_{1,0} = -0.01835$	$\theta_1 = -0.01266$	0.005690
5	$\theta_{2,0} = 0.39018$	$\theta_2 = 0.16506$	-0.225120
6	$\theta_{3,0} = 0.11259$	$\theta_3 = 0.11957$	0.006980

Con estos valores calculo los  $\chi_{i,t}$  para obtener la matriz X y a través de ella conocer las aproximaciones de las covarianzas entre parámetros. El método para el cálculo de estos valores es.

$$\chi_{1,0} = [(a_0 | W, \beta_{1,0}, \beta_{2,0}, \beta_{3,0}, \beta_{4,0}, \beta_{5,0}, \beta_{6,0}) - (a_0 | W, \beta_{1,0} + \delta_1, \beta_{2,0}, \beta_{3,0}, \beta_{4,0}, \beta_{5,0}, \beta_{6,0})] / \delta_1$$

$$= [-0.000085 - (-0.000081)] / 0.143159 = -0.00002794$$

$$\chi_{2,5} = [(a_5 | W, \beta_{1,0}, \beta_{2,0}, \beta_{3,0}, \beta_{4,0}, \beta_{5,0}, \beta_{6,0}) - (a_5 | W, \beta_{1,0}, \beta_{2,0} + \delta_2, \beta_{3,0}, \beta_{4,0}, \beta_{5,0}, \beta_{6,0})] / \delta_2$$

$$= [-0.000031 - (-0.000025)] / -0.202807 = -0.00002958$$

recordando que  $a_t$  se calcula con base en  $a_t = Z_t - \beta_1 Z_{t-1} - \beta_2 Z_{t-2} - \beta_3 Z_{t-3} + \beta_4 a_{t-1} + \beta_5 a_{t-2} + \beta_6 a_{t-3}$ . Con base en estos datos formo la matriz X de dimensión  $(N+Q+1) \times k = (109+11+1) \times 6 = 121 \times 6$ ; después obtengo la matriz transpuesta X' para llegar a una

aproximación de la matriz de covarianzas de los parámetros a través de  $(X'X)^{-1}S_e^2$ . El

producto  $X'X$  es:

$$X'X = \begin{vmatrix} 0.00004 & 0.00000 & -0.00001 & -0.00004 & 0.00001 & 0.00001 \\ 0.00000 & 0.00004 & 0.00000 & 0.00000 & -0.00004 & 0.00001 \\ -0.00001 & 0.00000 & 0.00004 & 0.00000 & -0.00001 & -0.00004 \\ -0.00004 & 0.00000 & 0.00000 & 0.00004 & 0.00001 & 0.00000 \\ 0.00001 & -0.00004 & -0.00001 & 0.00001 & 0.00004 & 0.00001 \\ 0.00001 & 0.00001 & -0.00004 & 0.00000 & 0.00001 & 0.00004 \end{vmatrix}$$

La inversa del producto es:

$$(X'X)^{-1} = \begin{vmatrix} 609,914.19 & 7,302.86 & -154,428.10 & 604,724.74 & -172,447.46 & -294,389.70 \\ 7,302.86 & 401,831.42 & -189,592.80 & 9,085.89 & 395,474.37 & -313,276.13 \\ -154,428.10 & -189,592.80 & 248,416.14 & -158,581.16 & -142,678.15 & 333,833.41 \\ 604,724.74 & 9,085.89 & -158,581.16 & 626,112.41 & -173,020.97 & -297,165.56 \\ -172,447.46 & 395,474.37 & -142,678.15 & -173,020.97 & 469,150.65 & -227,871.15 \\ -294,389.70 & -313,276.13 & 333,833.41 & -297,165.56 & -227,871.15 & 515,574.12 \end{vmatrix}$$

La aproximación a la matriz de covarianzas parametrales, considerando el valor  $S_e^2 = 0.00000037754$ , es

$$(X'X)^{-1} S_e^2 = \begin{vmatrix} \phi_1 & \phi_2 & \phi_3 & \theta_1 & \theta_2 & \theta_3 \\ 0.2303 & 0.0028 & -0.0583 & 0.2283 & -0.0651 & -0.1111 & \phi_1 \\ 0.0028 & 0.1517 & -0.0716 & 0.0034 & 0.1493 & -0.1183 & \phi_2 \\ -0.0583 & -0.0716 & 0.0938 & -0.0599 & -0.0539 & 0.1260 & \phi_3 \\ 0.2283 & 0.0034 & -0.0599 & 0.2364 & -0.0653 & -0.1122 & \theta_1 \\ -0.0651 & 0.1493 & -0.0539 & -0.0653 & 0.1771 & -0.0860 & \theta_2 \\ -0.1111 & -0.1183 & 0.1260 & -0.1122 & -0.0860 & 0.1946 & \theta_3 \end{vmatrix}$$



Así llego a que los valores de las autocorrelaciones son.

$$\begin{array}{lll}
 r(\phi_1, \theta_1) = 0.979 & r(\phi_1, \theta_2) = -0.322 & r(\phi_1, \theta_3) = -0.525 \\
 r(\phi_2, \theta_1) = 0.018 & r(\phi_2, \theta_2) = 0.911 & r(\phi_2, \theta_3) = -0.688 \\
 r(\phi_3, \theta_1) = -402 & r(\phi_3, \theta_2) = -0.418 & r(\phi_3, \theta_3) = 0.933 \\
 r(\phi_1, \phi_2) = -0.015 & r(\phi_1, \phi_3) = -0.397 & r(\phi_2, \phi_3) = -0.600 \\
 r(\theta_1, \theta_2) = -0.319 & r(\theta_1, \theta_3) = -0.523 & r(\theta_2, \theta_3) = -0.463
 \end{array}$$

De acuerdo con estos resultados encuentro una autocorrelación fuerte entre las siguientes parejas de parámetros  $(\phi_1, \theta_1)$ ,  $(\phi_1, \theta_3)$ ,  $(\phi_2, \theta_2)$ ,  $(\phi_2, \theta_3)$ ,  $(\phi_3, \theta_3)$ ,  $(\phi_2, \phi_3)$  y  $(\theta_1, \theta_3)$ .

Por los resultados obtenidos hasta este punto es posible determinar que el modelo **ARIMA(3,1,3)** no es adecuado para representar a la serie de las exportaciones petroleras.

Debido a esto estimaré un modelo más general (con más parámetros) para ver si es posible describir adecuadamente al proceso. Recordando que en la etapa de identificación encontré que un buen prospecto para modelo de la serie transformada y diferenciada sería un **ARMA(4,4)**, realizaré ahora las etapas de estimación y verificación para este modelo.

Para el modelo **ARMA(4,4)** tengo la relación

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \phi_3 Z_{t-3} - \phi_4 Z_{t-4} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \theta_4 a_{t-4}$$

y siguiendo las técnicas usadas previamente obtengo los valores parametrales iniciales de

$$\begin{aligned} \phi_1 &= 0.8632 & \theta_1 &= 0.1191 \\ \phi_2 &= 0.9088 & \theta_2 &= 0.3512 \\ \phi_3 &= 0.2937 & \theta_3 &= 0.5327 \\ \phi_4 &= -0.6378 & \theta_4 &= -0.8796 \end{aligned}$$

Y los valores óptimos que a continuación presento:

<i>Valor parametral</i>	<i>Desviación estándar</i>
$\phi_1 = 0.16930$	0.17186
$\phi_2 = 0.45117$	0.16400
$\phi_3 = 0.32462$	0.15958
$\phi_4 = -0.34529$	0.16390
$\theta_1 = 0.33501$	0.13131
$\theta_2 = 0.67971$	0.10547
$\theta_3 = 0.46033$	0.09136
$\theta_4 = -0.75822$	0.11728

que minimizan la suma de cuadrados. Con estos parámetros la suma de cuadrados es de 0.00003548. Finalmente el valor estimado de la varianza del ruido blanco conforme (3.2.4.2) será.

$$S_c^2 = \frac{\mathbf{S}(\hat{\phi}, \hat{\theta})}{N-d-p-q-1} = \frac{1}{[1/(110-1-4-4-1)] \sum_{t=26}^{109} a_t^2} = \frac{1}{[1/(100)] \sum_{t=26}^{109} a_t^2} = \frac{0.000035480}{100}$$

Por lo tanto  $S_c^2 = 0.0000003548$ . El valor estimado de la desviación estándar del ruido blanco será 0.000595651.

Inicio el proceso de verificación con la revisión de los supuestos para el análisis de residuales que propuse en la sección 3.3.1:

**SUPUESTO 1:**  $\alpha_t$  tiene media cero. Los valores de la media y la desviación estándar muestrales de los  $N-d-p=110-1-4=105$  errores aleatorios son:

$$M(\alpha_t) = -0.00006353471$$

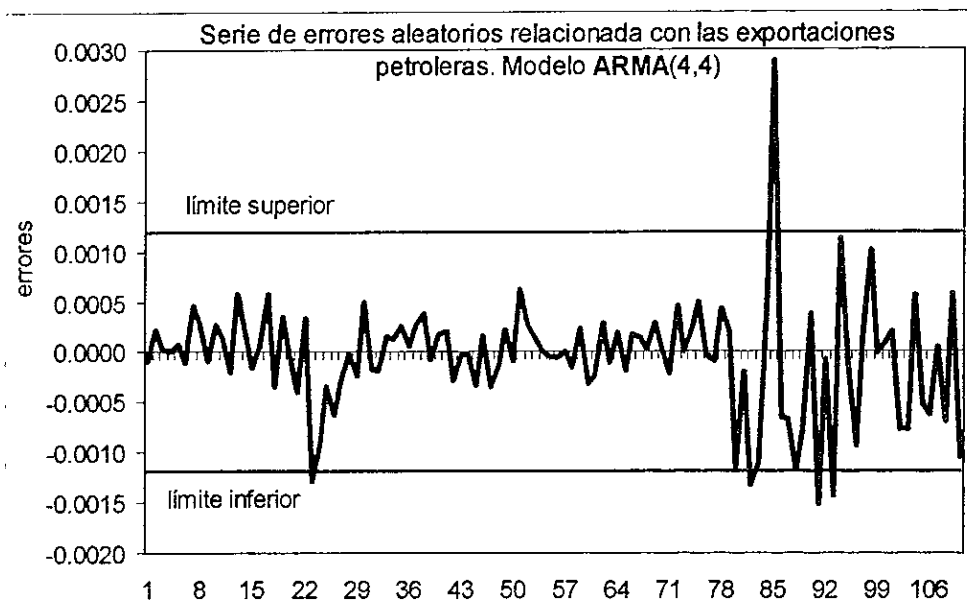
$$S_\alpha = [0.0000003548]^{1/2} = 0.000595651$$

A continuación calculo el cociente:

$$(N-d-p)^{1/2} * M(\alpha_t) / S_\alpha = (110-1-4)^{1/2} * (-0.00006353471) / 0.000595651 = -1.092984055$$

Como se cumple que  $1.092984055 < 2$  se aprueba el supuesto.

**SUPUESTO 2.**  $\alpha_t$  tiene varianza constante Presento a continuación la gráfica de la serie de los errores aleatorios calculados previamente, ya que a través de su revisión sabré si se cumple con el supuesto



Como puede verse, la gráfica no presenta una tendencia creciente ni decreciente si no un comportamiento estable por lo que considero que la varianza es constante y en consecuencia se cumple con el supuesto.

**SUPUESTO 3:** Las variables  $\{a_t\}$  son mutuamente independientes. Calculando las primeras 25 estimaciones de  $r_k(a)$  obtengo los valores:

k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$	k	$r_k(a)$
1	0.000785	6	-0.007822	11	-0.042264	16	-0.024633	21	-0.026317
2	-0.020179	7	-0.021604	12	-0.095565	17	-0.022448	22	-0.016816
3	-0.014081	8	-0.112165	13	0.106952	18	-0.025596	23	-0.010265
4	-0.031945	9	0.110934	14	-0.037282	19	-0.025889	24	-0.008039
5	0.042461	10	0.024549	15	0.037812	20	-0.023276	25	-0.003727

Y la estimación de la varianza de la **FAC** de los choques aleatorios la obtengo por medio de:

$$\underline{Var} [r_k(a)] = 1/(110-1-4) \Rightarrow$$

$$\{\underline{Var} [r_k(a)]\}^{1/2} = [1/105]^{1/2} = 0.097590007$$

el intervalo formado es  $(-2*0.097590007, 2*0.097590007) = (-0.1951800, 0.1951800)$ .

Este intervalo incluye a los valores estimados  $r_k(a)$  por lo que considero que ninguna autocorrelación es significativamente distinta de cero y por tanto el supuesto se cumple.

Realizando la prueba Portmanteau mediante el estadístico  $Q'$  obtengo:

$$Q' = (105) * (107) \frac{\sum_{k=1}^{25} r_k^2(a)}{(105 - k)} = 105 * 107 * 0.0006245 = 7.01626418$$

El valor en tablas de  $\chi^2_{(0.25, 25)} = 40.6465$ , dado que  $\chi^2 > Q'$  no se rechaza el supuesto de independencia de las variables.

**SUPUESTO 4:**  $\{\alpha_i\}$  se distribuye normal Verifico cuántas valores de la serie de choques aleatorios se encuentran fuera del intervalo formado por  $(-2S_{\alpha}, 2S_{\alpha}) = (-2 \cdot 0.000595651, 2 \cdot 0.000595651) = (-0.001191302, 0.001191302)$ .

Encuentro que los puntos fuera del intervalo señalado alcanzan un número de 5 que representan el 4.76% del total de las observaciones por lo que puedo decir que la serie sigue una distribución normal. La siguiente tabla muestra los valores de los choques aleatorios que se encuentran fuera del intervalo

OBSV.	CHOQUE	OBSV.	CHOQUE	OBSV.	CHOQUE
22	-0.001295	81	-0.001320	84	0.002883
90	-0.001517	92	-0.001434		

**SUPUESTO 5:** No existen observaciones aberrantes. La única observación encontrada fuera del intervalo  $(-3S_{\alpha}, 3S_{\alpha}) = (-0.001786953, 0.001786953)$  es la correspondiente al número 84, con un valor de 0.0028883. Este choque proviene de las observaciones originales correspondientes a los meses de mayo a septiembre de 1995. Como expliqué anteriormente, no existe razón aparente para pensar que este valor corresponda a una intervención exógena a la serie y se acepta el supuesto.

Ahora aplico las pruebas independientes de los residuales mostradas en la sección 3.3.2

**SUPUESTO 6:** El modelo es parsimonioso. Para verificar este supuesto formo los intervalos señalados en la expresión (3.3.2 1).

Parámetro $\beta$	$[\text{Var}(\beta)]^{1/2}$	$(\beta - 2[\text{Var}(\beta)]^{1/2}, \beta + 2[\text{Var}(\beta)]^{1/2})$
$\phi_1 = 0.16930$	0.17186	(-0.17442, 0.51302)
$\phi_2 = 0.45117$	0.16400	( 0.12317, 0.77917)
$\phi_3 = 0.32462$	0.15958	( 0.00546, 0.64378)
$\phi_4 = -0.34529$	0.16390	(-0.67309, -0.01749)
$\theta_1 = 0.33501$	0.13131	( 0.07239, 0.59763)
$\theta_2 = 0.67971$	0.10547	( 0.46877, 0.89065)
$\theta_3 = 0.46033$	0.09136	( 0.27761, 0.64305)
$\theta_4 = -0.75822$	0.11728	(-0.99278, -0.52366)

Como puede observarse sólo puedo considerar al primer parámetro como cero, pero al realizar la estimación óptima sin este parámetro vuelve asignársele un valor por lo que considero necesario dicho parámetro para describir el fenómeno.

**SUPUESTO 7:** El modelo es admisible. Para esta prueba calculo los determinantes en las expresiones (2.1 4) y (2.2.4). En primer término reviso si el modelo es estacionario a través de los parámetros  $\phi_1$ ,  $\phi_2$ ,  $\phi_3$  y  $\phi_4$  obteniendo los siguientes valores para los determinantes involucrados:

$$D_1 = 0.8808 \quad D_2 = 0.7049 \quad D_3 = 0.4781 \quad D_4 = 0.3151$$

como todos los determinantes son positivos puedo concluir que el modelo es estacionario. A continuación reviso si el modelo es invertible a través de los parámetros  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  y  $\theta_4$  para obtener los siguientes valores de los determinantes:

$$D_1 = 0.4251 \quad D_2 = 0.1381 \quad D_3 = 0.0343 \quad D_4 = 0.0072$$

Y dado que todos los determinantes son positivos concluyo que el modelo es invertible. Así se tiene que se cumple con la admisibilidad.

**SUPUESTO 8:** El modelo es estable en los parámetros. Para definir si se cumple este supuesto calculo las autocorrelaciones entre parámetros.

Para realizar este cálculo me basaré en la aproximación proporcionada por Box que se mostró en la sección 3.2.2 calculando  $\delta_i$  como la diferencia entre los valores parametrales iniciales y los óptimos.

$i$	$\beta_{i,0}$	$\beta_{i,0} + \delta_i$	$\delta_i$
1	$\phi_1 = 0.8632$	$\phi_1 = 0.16930$	-0.69390
2	$\phi_2 = 0.9088$	$\phi_2 = 0.45117$	-0.45763
3	$\phi_3 = 0.2937$	$\phi_3 = 0.32462$	0.03092
4	$\phi_4 = -0.6378$	$\phi_4 = -0.34529$	0.29251
5	$\theta_1 = 0.1191$	$\theta_1 = 0.33501$	0.21591
6	$\theta_2 = 0.3512$	$\theta_2 = 0.67971$	0.32851
7	$\theta_3 = 0.5327$	$\theta_3 = 0.46033$	-0.07237
8	$\theta_4 = -0.8796$	$\theta_4 = -0.75822$	0.12138

Con estos valores calculo los  $\chi_{it}$  para obtener la matriz X y a través de ella conocer las aproximaciones de las covarianzas entre parámetros. El producto  $X'X$  es.

$$X'X = \begin{bmatrix} 0.0215 & -0.0150 & 0.0001 & 0.0137 & -0.0556 & -0.0181 & 0.0665 & 0.0334 \\ -0.0150 & 0.0211 & -0.0140 & -0.0009 & -0.0164 & -0.0137 & 0.0982 & -0.1023 \\ 0.0001 & -0.0140 & 0.0184 & -0.0114 & 0.0726 & 0.0346 & -0.1908 & 0.1046 \\ 0.0137 & -0.0009 & -0.0114 & 0.0159 & -0.0809 & -0.0331 & 0.1621 & -0.0444 \\ -0.0556 & -0.0164 & 0.0726 & -0.0809 & 0.4350 & 0.1863 & -0.9328 & 0.3299 \\ -0.0181 & -0.0137 & 0.0346 & -0.0331 & 0.1863 & 0.0985 & -0.4099 & 0.1678 \\ 0.0665 & 0.0982 & -0.1908 & 0.1621 & -0.9328 & -0.4099 & 2.2196 & -0.9947 \\ 0.0334 & -0.1023 & 0.1046 & -0.0444 & 0.3299 & 0.1678 & -0.9947 & 0.6508 \end{bmatrix}$$

La inversa de este producto es.

$$(X'X)^{-1} = \begin{vmatrix} 22518.4 & 7330.5 & -1417.3 & -2229.7 & 231.7 & -456.1 & 2572.2 & 5677.8 \\ 7330.5 & 8325.6 & 3387.4 & -1132.4 & 384.3 & 64.8 & -135.5 & -153.1 \\ -1417.3 & 3387.4 & 19505.0 & 7240.8 & -3877.3 & -75.0 & 1911.0 & 4057.2 \\ -2229.7 & -1132.4 & 7240.8 & 18609.7 & -1147.4 & -139.4 & 1620.9 & 4436.3 \\ 231.7 & 384.3 & -3877.3 & -1147.4 & 1607.7 & -59.8 & -620.5 & -1632.8 \\ -456.1 & 64.8 & -75.0 & -139.4 & -59.8 & 94.8 & -128.8 & -219.1 \\ 2572.2 & -135.5 & 1911.0 & 1620.9 & -620.5 & -128.8 & 916.6 & 1978.7 \\ 5677.8 & -153.1 & 4057.2 & 4436.3 & -1632.8 & -219.1 & 1978.7 & 4590.7 \end{vmatrix}$$

La aproximación a la matriz de covarianzas parametrales, considerando el valor de

$S_e^2 = 0.0000003548$ , es:

$$(X'X)^{-1}S_e^2 = \begin{vmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 & \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ 0.00799 & 0.00260 & -0.00050 & -0.00079 & 0.00008 & -0.00016 & 0.00091 & 0.00201 & \phi_1 \\ 0.00260 & 0.00295 & 0.00120 & -0.00040 & 0.00014 & 0.00002 & -0.00005 & -0.00005 & \phi_2 \\ -0.00050 & 0.00120 & 0.00692 & 0.00257 & -0.00138 & -0.00003 & 0.00068 & 0.00144 & \phi_3 \\ -0.00079 & -0.00040 & 0.00257 & 0.00660 & -0.00041 & -0.00005 & 0.00058 & 0.00157 & \phi_4 \\ 0.00008 & 0.00014 & -0.00138 & -0.00041 & 0.00057 & -0.00002 & -0.00022 & -0.00058 & \theta_1 \\ -0.00016 & 0.00002 & -0.00003 & -0.00005 & -0.00002 & 0.00003 & -0.00005 & -0.00008 & \theta_2 \\ 0.00091 & -0.00005 & 0.00068 & 0.00058 & -0.00022 & -0.00005 & 0.00033 & 0.00070 & \theta_3 \\ 0.00201 & -0.00005 & 0.00144 & 0.00157 & -0.00058 & -0.00008 & 0.00070 & 0.00163 & \theta_4 \end{vmatrix}$$

Los valores de las autocorrelaciones entre parámetros son.

$$\begin{aligned} r(\phi_1, \theta_1) &= 0.04 & r(\phi_1, \theta_2) &= -0.31 & r(\phi_1, \theta_3) &= 0.57 & r(\phi_1, \theta_4) &= 0.56 \\ r(\phi_2, \theta_1) &= 0.11 & r(\phi_2, \theta_2) &= 0.07 & r(\phi_2, \theta_3) &= -0.05 & r(\phi_2, \theta_4) &= -0.02 \end{aligned}$$



$$\begin{aligned}
r(\phi_3, \theta_1) &= -0.69 & r(\phi_3, \theta_2) &= -0.06 & r(\phi_3, \theta_3) &= 0.45 & r(\phi_3, \theta_4) &= 0.43 \\
r(\phi_4, \theta_1) &= -0.21 & r(\phi_4, \theta_2) &= -0.10 & r(\phi_4, \theta_3) &= 0.39 & r(\phi_4, \theta_4) &= 0.48 \\
r(\phi_1, \phi_2) &= 0.54 & r(\phi_1, \phi_3) &= -0.07 & r(\phi_1, \phi_4) &= -0.11 & r(\phi_2, \phi_3) &= 0.27 \\
r(\phi_2, \phi_4) &= -0.09 & r(\phi_3, \phi_4) &= 0.38 & r(\theta_1, \theta_2) &= -0.15 & r(\theta_1, \theta_3) &= -0.51 \\
r(\theta_1, \theta_4) &= -0.60 & r(\theta_2, \theta_3) &= -0.44 & r(\theta_2, \theta_4) &= -0.33 & r(\theta_3, \theta_4) &= 0.96
\end{aligned}$$

De acuerdo con estos resultados se presenta una autocorrelación alta entre las siguientes parejas de parámetros  $(\phi_1, \theta_3)$ ,  $(\phi_1, \theta_4)$ ,  $(\phi_3, \theta_1)$ ,  $(\phi_1, \phi_2)$ ,  $(\theta_1, \theta_3)$ ,  $(\theta_1, \theta_4)$  y  $(\theta_3, \theta_4)$ . De entre éstas, la más significativa es  $r(\theta_3, \theta_4) = 0.96$ ; aquí es evidente que no puedo eliminar el parámetro  $\theta_4$  ya que volvería al modelo inicial que ya mostré que no es adecuado.

Después de todas estas pruebas puedo decir que, de los modelos probados, éste último es más adecuado.

## 7.4 Pronósticos

Los pronósticos serán realizados para el modelo **ARIMA(4,1,4)** —recuérdese que para encontrar el modelo **ARMA(4,4)** utilicé la serie resultante de diferenciar una vez a la serie transformada— de la serie resultante de aplicar la transformación de Bartlett a la serie original que tiene la siguiente representación.

$$\begin{aligned}
(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4)(1-B)T(Z_t) &= (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4)a_t \Rightarrow \\
(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4) \cdot (B - \phi_1 B^2 - \phi_2 B^3 - \phi_3 B^4 - \phi_4 B^5)T(Z_t) &= (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4)a_t \Rightarrow \\
(1 - (1 + \phi_1)B + (\phi_1 - \phi_2)B^2 + (\phi_2 - \phi_3)B^3 + (\phi_3 - \phi_4)B^4 + \phi_4 B^5)T(Z_t) &= (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4)a_t \Rightarrow \\
T(Z_t) &= (1 + \phi_1)T(Z_{t-1}) - (\phi_1 - \phi_2)T(Z_{t-2}) - (\phi_2 - \phi_3)T(Z_{t-3}) - (\phi_3 - \phi_4)T(Z_{t-4}) - \phi_4 T(Z_{t-5}) + a_t - \\
&\theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \theta_4 a_{t-4}
\end{aligned}$$

El origen de mis pronósticos será  $t = 110$  que fue el punto donde realicé mi última observación y para calcularlos me basaré en lo revisado en la sección 4.2.

Como ya conozco los valores de  $\phi$  y  $\theta$  la representación de mi modelo será:

$$T(Z_t) = 1.1693*T(Z_{t-1}) + 0.28187*T(Z_{t-2}) - 0.12655*T(Z_{t-3}) - 0.66991*T(Z_{t-4}) + 0.34529*T(Z_{t-5}) + a_t - 0.33501*a_{t-1} - 0.67971*a_{t-2} - 0.46033*a_{t-3} + 0.75822*a_{t-4}$$

Los horizontes para los cuales calcularé los pronósticos serán del uno al diez, es decir, para  $h = 1, 2, \dots, 10$ .

$$\begin{aligned} T_{110}(1) &= E_{110}[T(Z_{110+1})] = E_{110}[1.1693*T(Z_{110+1-1}) + 0.28187*T(Z_{110+1-2}) - \\ & \quad 0.12655*T(Z_{110+1-3}) - 0.66991*T(Z_{110+1-4}) + \\ & \quad 0.34529*T(Z_{110+1-5}) + a_{110+1} - 0.33501*a_{110+1-1} - \\ & \quad 0.67971*a_{110+1-2} - 0.46033*a_{110+1-3} + 0.75822*a_{110+1-4}] \\ &= E_{110}[1.1693*T(Z_{110}) + 0.28187*T(Z_{109}) - 0.12655*T(Z_{108}) - \\ & \quad 0.66991*T(Z_{107}) + 0.34529*T(Z_{106}) + a_{111} - 0.33501*a_{110} - \\ & \quad 0.67971*a_{109} - 0.46033*a_{108} + 0.75822*a_{107}] \\ &= 1.1693*T(Z_{110}) + 0.28187*T(Z_{109}) - 0.12655*T(Z_{108}) - \\ & \quad 0.66991*T(Z_{107}) + 0.34529*T(Z_{106}) + E_{110}[a_{111}] - \\ & \quad 0.33501*E_{110}[a_{110}] - 0.67971*E_{110}[a_{109}] - 0.46033*E_{110}[a_{108}] \\ & \quad + 0.75822*E_{110}[a_{107}] \\ &= 1.1693*T(Z_{110}) + 0.28187*T(Z_{109}) - 0.12655*T(Z_{108}) - \\ & \quad 0.66991*T(Z_{107}) + 0.34529*T(Z_{106}) + 0 - 0.33501*[T(Z_{110+1-1}) - \\ & \quad T_{110+1-1-1}(1)] - 0.67971*[T(Z_{110+1-2}) - T_{110+1-2-1}(1)] - \\ & \quad 0.46033*[T(Z_{110+1-3}) - T_{110+1-3-1}(1)] + 0.75822*[T(Z_{110+1-4}) - \\ & \quad T_{110+1-4-1}(1)] \end{aligned}$$

$$\begin{aligned}
&= 1.1693 \cdot T(Z_{110}) + 0.28187 \cdot T(Z_{109}) - 0.12655 \cdot T(Z_{108}) - \\
&0.66991 \cdot T(Z_{107}) + 0.34529 \cdot T(Z_{106}) + 0 - 0.33501 \cdot [T(Z_{110}) - \\
&T_{109}(1)] - 0.67971 \cdot [T(Z_{109}) - T_{108}(1)] - 0.46033 \cdot [T(Z_{108}) - \\
&T_{107}(1)] + 0.75822 \cdot [T(Z_{107}) - T_{106}(1)]
\end{aligned}$$

Como puede advertirse, y como previamente expliqué, para generar los valores  $E_{110}[a_{110+h-j}]$  necesito en forma directa de los valores estimados  $T_{110-1}(1)$ ,  $T_{110-2}(1)$ ,  $T_{110-3}(1)$  y  $T_{110-4}(1)$  que a su vez requieren de los valores  $T_{110-m}(1)$  para  $m=5, 6, \dots, 110$ .

Para calcular estos valores parto de las siguientes bases para  $T_0(1)$ :

1.  $E_0[a_0] = E_0[a_{-1}] = E_0[a_{-2}] = E_0[a_{-3}] = 0$
2.  $T(Z_0) = T(Z_{-1}) = T(Z_{-2}) = T(Z_{-3}) = T(Z_{-4}) = M(T(Z_t))$  donde  $t=1, 2, \dots, 110$

En la tabla C presento los valores obtenidos para  $T_t(1)$  y  $E_t[a_{-4+m}]$  y utilizando éstos y los valores observados obtengo:

$$\begin{aligned}
T_{110}(1) = E_{110}[T(Z_{110+1})] &= 1.1693 \cdot 0.023051 + 0.28187 \cdot 0.024063 - 0.12655 \cdot 0.023479 - \\
&0.66991 \cdot 0.023789 + 0.34529 \cdot 0.023596 + 0 - \\
&0.33501 \cdot [0.023051 - 0.024097] - 0.67971 \cdot [0.024063 - \\
&0.023491] - 0.46033 \cdot [0.023479 - 0.024168] + \\
&0.75822 \cdot [0.023789 - 0.023757] \\
&= 0.023279
\end{aligned}$$

Siguiendo esta mecánica y sustituyendo los valores  $E_{110}[T(Z_{110+h-j})]$  por las estimaciones previamente calculadas cuando sea necesario obtengo los siguientes resultados:

$$\begin{aligned}
T_{110}(2) = E_{110}[T(Z_{110+2})] &= E_{110}[1.1693*T(Z_{110+2-1}) + 0.28187*T(Z_{110+2-2}) - \\
&0.12655*T(Z_{110+2-3}) - 0.66991*T(Z_{110+2-4}) + \\
&0.34529*T(Z_{110+2-5}) + a_{110+2} - 0.33501*a_{110+2-1} - \\
&0.67971*a_{110+2-2} - 0.46033*a_{110+2-3} + 0.75822*a_{110+2-4}] \\
&= 1.1693*T_{110}(1) + 0.28187*T(Z_{110}) - 0.12655*T(Z_{109}) - \\
&0.66991*T(Z_{108}) + 0.34529*T(Z_{107}) + 0 - 0.33501*0 - \\
&0.67971*[T(Z_{110})-T_{109}(1)] - 0.46033*[T(Z_{109})-T_{108}(1)] + \\
&0.75822*[T(Z_{108}) - T_{107}(1)] \\
&= 0.023083
\end{aligned}$$

$$\begin{aligned}
T_{110}(3) = E_{110}[T(Z_{110+3})] &= E_{110}[1.1693*T(Z_{110+3-1}) + 0.28187*T(Z_{110+3-2}) - \\
&0.12655*T(Z_{110+3-3}) - 0.66991*T(Z_{110+3-4}) + \\
&0.34529*T(Z_{110+3-5}) + a_{110+3} - 0.33501*a_{110+3-1} - \\
&0.67971*a_{110+3-2} - 0.46033*a_{110+3-3} + 0.75822*a_{110+3-4}] \\
&= 1.1693*T_{110}(2) + 0.28187*T_{110}(1) - 0.12655*T(Z_{110}) - \\
&0.66991*T(Z_{109}) + 0.34529*T(Z_{108}) + 0 - 0.33501*0 - \\
&0.67971*0 - 0.46033*[T(Z_{110})-T_{109}(1)] + 0.75822*[T(Z_{109}) - \\
&T_{108}(1)] \\
&= 0.023538
\end{aligned}$$

$$\begin{aligned}
T_{110}(4) = E_{110}[T(Z_{110+4})] &= E_{110}[1.1693*T(Z_{110+4-1}) + 0.28187*T(Z_{110+4-2}) - \\
&0.12655*T(Z_{110+4-3}) - 0.66991*T(Z_{110+4-4}) + \\
&0.34529*T(Z_{110+4-5}) + a_{110+4} - 0.33501*a_{110+4-1} - \\
&0.67971*a_{110+4-2} - 0.46033*a_{110+4-3} + 0.75822*a_{110+4-4}] \\
&= 1.1693*T_{110}(3) + 0.28187*T_{110}(2) - 0.12655*T_{110}(1) -
\end{aligned}$$

$$\begin{aligned}
& 0.66991*T(Z_{110}) + 0.34529*T(Z_{109}) + 0 - 0.33501*0 - \\
& 0.67971*0 - 0.460330 + 0.75822*[T(Z_{110}) - T_{09}(1)] \\
& = 0.023157
\end{aligned}$$

$$\begin{aligned}
T_{110}(5) = E_{110}[T(Z_{110+5})] &= E_{110}[1.1693*T(Z_{110+5-1}) + 0.28187*T(Z_{110+5-2}) - \\
& 0.12655*T(Z_{110+5-3}) - 0.66991*T(Z_{110+5-4}) + \\
& 0.34529*T(Z_{110+5-5}) + a_{110+5} - 0.33501*a_{110+5-1} - \\
& 0.67971*a_{110+5-2} - 0.46033*a_{110+5-3} + 0.75822*a_{110+5-4}] \\
& = 1.1693*T_{110}(4) + 0.28187*T_{110}(3) - 0.12655*T_{110}(2) - \\
& 0.66991*T_{110}(1) + 0.34529*T(Z_{110}) + 0 - 0.33501*0 - \\
& 0.67971*0 - 0.46033*0 + 0.75822*0 \\
& = 0.023155
\end{aligned}$$

$$\begin{aligned}
T_{110}(6) = E_{110}[T(Z_{110+6})] &= E_{110}[1.1693*T(Z_{110+6-1}) + 0.28187*T(Z_{110+6-2}) - \\
& 0.12655*T(Z_{110+6-3}) - 0.66991*T(Z_{110+6-4}) + \\
& 0.34529*T(Z_{110+6-5}) + a_{110+6} - 0.33501*a_{110+6-1} - \\
& 0.67971*a_{110+6-2} - 0.46033*a_{110+6-3} + 0.75822*a_{110+6-4}] \\
& = 1.1693*T_{110}(5) + 0.28187*T_{110}(4) - 0.12655*T_{110}(3) - \\
& 0.66991*T_{110}(2) + 0.34529*T_{110}(1) + 0 - 0.33501*0 - \\
& 0.67971*0 - 0.46033*0 + 0.75822*0 \\
& = 0.023199
\end{aligned}$$

$$\begin{aligned}
T_{110}(7) = E_{110}[T(Z_{110+7})] &= E_{110}[1.1693*T(Z_{110+7-1}) + 0.28187*T(Z_{110+7-2}) - \\
& 0.12655*T(Z_{110+7-3}) - 0.66991*T(Z_{110+7-4}) + \\
& 0.34529*T(Z_{110+7-5}) + a_{110+7} - 0.33501*a_{110+7-1} - \\
& 0.67971*a_{110+7-2} - 0.46033*a_{110+7-3} + 0.75822*a_{110+7-4}]
\end{aligned}$$

$$\begin{aligned}
&= 1.1693 * T_{110}(6) + 0.28187 * T_{110}(5) - 0.12655 * T_{110}(4) - \\
&0.66991 * T_{110}(3) + 0.34529 * T_{110}(2) + 0 - 0.33501 * 0 - \\
&0.67971 * 0 - 0.46033 * 0 + 0.75822 * 0 \\
&= 0.022924
\end{aligned}$$

$$\begin{aligned}
T_{110}(8) = E_{110}[T(Z_{110+8})] &= E_{110}[1.1693 * T(Z_{110+8-1}) + 0.28187 * T(Z_{110+8-2}) - \\
&0.12655 * T(Z_{110+8-3}) - 0.66991 * T(Z_{110+8-4}) + \\
&0.34529 * T(Z_{110+8-5}) + a_{110+8} - 0.33501 * a_{110+8-1} - \\
&0.67971 * a_{110+8-2} - 0.46033 * a_{110+8-3} + 0.75822 * a_{110+8-4}] \\
&= 1.1693 * T_{110}(7) + 0.28187 * T_{110}(6) - 0.12655 * T_{110}(5) - \\
&0.66991 * T_{110}(4) + 0.34529 * T_{110}(3) + 0 - 0.33501 * 0 - \\
&0.67971 * 0 - 0.46033 * 0 + 0.75822 * 0 \\
&= 0.023028
\end{aligned}$$

$$\begin{aligned}
T_{110}(9) = E_{110}[T(Z_{110+9})] &= E_{110}[1.1693 * T(Z_{110+9-1}) + 0.28187 * T(Z_{110+9-2}) - \\
&0.12655 * T(Z_{110+9-3}) - 0.66991 * T(Z_{110+9-4}) + \\
&0.34529 * T(Z_{110+9-5}) + a_{110+9} - 0.33501 * a_{110+9-1} - \\
&0.67971 * a_{110+9-2} - 0.46033 * a_{110+9-3} + 0.75822 * a_{110+9-4}] \\
&= 1.1693 * T_{110}(8) + 0.28187 * T_{110}(7) - 0.12655 * T_{110}(6) - \\
&0.66991 * T_{110}(5) + 0.34529 * T_{110}(4) + 0 - 0.33501 * 0 - \\
&0.67971 * 0 - 0.46033 * 0 + 0.75822 * 0 \\
&= 0.022937
\end{aligned}$$

$$\begin{aligned}
T_{110}(10) = E_{110}[T(Z_{110+10})] &= E_{110}[1.1693 * T(Z_{110+10-1}) + 0.28187 * T(Z_{110+10-2}) - \\
&0.12655 * T(Z_{110+10-3}) - 0.66991 * T(Z_{110+10-4}) + \\
&0.34529 * T(Z_{110+10-5}) + a_{110+10} - 0.33501 * a_{110+10-1} -
\end{aligned}$$

$$\begin{aligned}
& 0.67971*a_{110+10-2} - 0.46033*a_{110+10-3} + 0.75822*a_{110+10-4}] \\
= & 1.1693*T_{110}(9) + 0.28187*T_{110}(8) - 0.12655*T_{110}(7) - \\
& 0.66991* T_{110}(6) + 0.34529*T_{110}(5) + 0 - 0.33501*0 - \\
& 0.67971*0 - 0.46033*0 + 0.75822*0 \\
= & 0.022864
\end{aligned}$$

A través de este ejercicio he calculado los pronósticos para la serie transformada, pero lo que me interesa es conocer los pronósticos de la serie original y para ello debo aplicar a los pronósticos obtenidos la función inversa de la transformación, que en este caso sería  $T(Z_t)^{-2}$  dado que  $T(Z_t) = Z_t^{-1/2}$  y la corrección por sesgo estará dada, conforme a (4.2.2) —dado que utilicé la transformación de Bartlett— por.

$$G_{t,\lambda}(h) = \frac{1}{2} + \left( \frac{[1-2\lambda^{-1}(\lambda-1)\{T(Z_t)(h)\}^{-2} \text{Var}_t(e_t(h))\]^{1/2}}{2} \right)^{1/\lambda}$$

Para calcular  $\text{Var}_t(e_t(h))$  tengo la expresión (4.1 7), cuyas ponderaciones  $\psi_i$ ,  $i=1, 2, \dots$  las obtengo a través de los parámetros  $\phi$  y  $\theta$ . Con base en el desarrollo presentado en la sección 4.4 llego a:

$$\begin{aligned}
T(Z_t) = & 1.1693*T(Z_{t-1}) + 0.28187*T(Z_{t-2}) - 0.12655*T(Z_{t-3}) - 0.66991*T(Z_{t-4}) + \\
& 0.34529*T(Z_{t-5}) + a_t - 0.33501*a_{t-1} - 0.67971*a_{t-2} - 0.46033*a_{t-3} + 0.75822*a_{t-4}
\end{aligned}$$

Por lo que

$$\varphi_1 = 1.16930 \quad \theta_1 = 0.33501$$

$$\varphi_2 = 0.28187 \quad \theta_2 = 0.67971$$

$$\varphi_3 = -0.12655 \quad \theta_3 = 0.46033$$

$$\varphi_4 = -0.66991 \quad \theta_4 = -0.75822$$

$$\varphi_5 = 0.34529$$

Así obtengo:

$$\psi_1 = \varphi_1 - \theta_1$$

$$\psi_1 = 1.1693 - 0.33501 = 0.834290$$

$$\psi_2 = \varphi_1\psi_1 + \varphi_2 - \theta_2$$

$$\psi_2 = 1.1693 * 0.83429 + 0.28187 - 0.67971 = 0.577695$$

$$\psi_3 = \varphi_1\psi_2 + \varphi_2\psi_1 + \varphi_3 - \theta_3$$

$$\psi_3 = 1.1693 * 0.577695 + 0.28187 * 0.83429 - 0.12655 - 0.46033 = 0.323780$$

$$\psi_4 = \varphi_1\psi_3 + \varphi_2\psi_2 + \varphi_3\psi_1 + \varphi_4 - \theta_4$$

$$\psi_4 = 1.1693 * 0.323780 + 0.28187 * 0.577695 - 0.12655 * 0.83429 - 0.66991 = 0.524162$$

$$\psi_5 = \varphi_1\psi_4 + \varphi_2\psi_3 + \varphi_3\psi_2 + \varphi_4\psi_1 + \varphi_5 - \theta_5$$

$$\psi_5 = 1.1693 * 0.524162 + 0.28187 * 0.323780 - 0.12655 * 0.577695 -$$

$$0.66991 * 0.83429 + 0.34529 = 0.41745$$

Continuando con esta mecánica obtengo los siguientes valores:

$$\psi_6 = 0.495964 \quad \psi_7 = 0.613833 \quad \psi_8 = 0.770774 \quad \psi_9 = 0.912857$$

$$\psi_{10} = 1.018871 \quad \psi_{11} = 1.111170 \quad \psi_{12} = 1.166560$$



Aplicando estas ponderaciones y recordando que  $S_0^2 = 0.0000003548$  llego a los siguientes valores estimados de las varianzas de los errores:

h	$Var_t(e_t(h))$	h	$Var_t(e_t(h))$	h	$Var_t(e_t(h))$
1	0.00000035480	5	0.00000085484	9	0.00000134841
2	0.00000060175	6	0.00000091667	10	0.00000164407
3	0.00000072016	7	0.00000100394	11	0.00000201239
4	0.00000075736	8	0.00000113763	12	0.00000245046

Con estos datos y recordando que en la etapa de identificación tomé un valor de  $\lambda = -0.5$  ya es posible calcular la corrección por sesgo, por ejemplo para  $h = 1$

$$C_{110,-0.5}(1) = \frac{1}{2} + \left( \frac{[1 - 2(-0.5)^{-1}(-0.5-1)\{\Phi(Z_{110})(1)\}]^2(0.00000035480)]^{1/2}}{2} \right)^{1/(-0.5)}$$

Realizando la sustitución para todos los pronósticos tengo:

h	$C_{110,0.5}(h)$	h	$C_{110,-0.5}(h)$
1	1.001969	6	1.005143
2	1.003402	7	1.005773
3	1.003919	8	1.006488
4	1.004260	9	1.007764
5	1.004812	10	1.009547

Y así los pronósticos puntuales óptimos para la serie original los obtengo aplicando (4.2.4). Debe tenerse en cuenta que  $T(Z_t) = Z_t^{-0.5} = Z_t^{-1/2} \Rightarrow T^{-1}[T(Z_t)] = T(Z_t)^{-2}$

Por ejemplo para el primer pronóstico:

$$\varepsilon_{110}(1) = (T^{-1}[\Phi(Z_{110})(1)]) * C_{110,0.5}(1) = [\Phi(Z_{110})(1)]^{-2} * (1.001969) = [0.023279]^{-2} * (1.001969)$$

$$\Rightarrow \varepsilon_{110}(1) = 1,848.92$$

Siguiendo esta metodología obtengo los pronósticos óptimos que presento en la siguiente tabla:

			<i>óptimo</i>
h	$\mathcal{F}(Z_{110})(h)$	$T^{-1}[\mathcal{F}(Z_{110})(h)]$	$T^{-1}[\mathcal{F}(Z_{110})(h)]\mathcal{G}_{t,\lambda}(h)$
1	0.023279	1,845.29	1,848.92
2	0.023083	1,876.75	1,883.13
3	0.023538	1,804.89	1,811.96
4	0.023157	1,864.81	1,872.75
5	0.023155	1,865.07	1,874.04
6	0.023199	1,858.14	1,867.70
7	0.022924	1,902.88	1,913.86
8	0.023028	1,885.71	1,897.95
9	0.022937	1,900.80	1,915.56
10	0.022864	1,912.87	1,931.13

Ahora obtengo los intervalos al 95% de confianza para estos pronósticos. Para ello, aplico el resultado presentado en (4.3 2), según lo cual el intervalo está dado por.

$$\mathcal{G}_{t,\lambda}(h)T^{-1}\{\mathcal{F}(Z_t)(h) \pm u_{\alpha/2}(\mathbf{Var}_t(\mathbf{e}_t(h)))^{1/2}\}$$

Donde  $u_{\alpha/2} = u_{0.05/2} = 1.96$ . Así, para el intervalo del primer pronóstico obtengo.

$$\begin{aligned} \mathcal{G}_{t,\lambda}(h)T^{-1}\{\mathcal{F}(Z_t)(h) - u_{\alpha/2}(\mathbf{Var}_t(\mathbf{e}_t(h)))^{1/2}\} &= 1.001969*[0.023279-1.96*(0.0000003548)^{1/2}]^{-2} \\ &= 2,045.29 \end{aligned}$$

$$G_{\lambda}(h)T^{-1}\{T(Z_t)(h) + u_{\alpha/2}(\text{Var}_t(e_t(h)))^{1/2}\}=1.001969*[0.023279+1.96*(0.0000003548)^{1/2}]^2$$

$$= 1,673.25$$

La siguiente tabla muestra los intervalos obtenidos a través de esta técnica:

<i>h</i>	LÍMITE	
	INFERIOR	SUPERIOR
1	1,673.25	2,045.29
2	1,651.96	2,150.74
3	1,574.51	2,089.80
4	1,617.71	2,173.16
5	1,604.16	2,195.22
6	1,590.43	2,199.61
7	1,614.42	2,276.16
8	1,584.90	2,281.07
9	1,573.12	2,342.65
10	1,552.77	2,414.48

Ahora trataré de actualizar pronósticos tomando como base los valores conocidos posteriores a septiembre de 1997 ( $t = 110$ ). Estos valores son:

<i>h</i>	<i>t</i>	PERIODO	EXPORTACIONES	ÓPTIMO MEDIDO	SESGO
1	111	1997-DIC	1,683	1,848.92	-165.92
2	112	1998-ENE	1,958	1,883.13	74.87
3	113	1998-FEB	1,625	1,811.96	-186.96
4	114	1998-MAR	1,927	1,872.75	54.25
5	115	1998-ABR	1,677	1,874.04	-197.04

Que, como se puede apreciar, están incluidos en los intervalos de confianza calculados previamente.

Con base en la expresión (4.4.1) y utilizando las ponderaciones  $\psi_i$  calculadas anteriormente determino los pronósticos actualizados. Para ello tengo:

$$a_{t+1} = a_{110+1} = a_{111} = Z_{110+1} - \varepsilon_{110}(1) = 1,683 - 1,848.92 = -165.92$$

A partir de este resultado tengo para el origen  $t = 111$ :

$h$	$\psi_h$	Pronósticos	
		Óptimo inicial	Óptimo actualizado
1	0.834290	1,883.13	1,744.70
2	0.577695	1,811.96	1,716.11
3	0.323780	1,872.75	1,819.03
4	0.524162	1,874.04	1,787.07
5	0.417450	1,867.70	1,798.44
6	0.495964	1,913.86	1,831.57
7	0.613833	1,897.95	1,796.10
8	0.770774	1,915.56	1,787.67
9	0.912857	1,931.13	1,779.67

Obviamente los pronósticos se reducen debido al valor de  $a_{t+1}$ . Para el origen  $t=112$  repito la operación:

$$a_{t+1} = a_{111+1} = a_{112} = Z_{111+1} - \varepsilon_{111}(1) = 1,958 - 1,744.70 = 213.30$$

$h$	$\psi_h$	Pronósticos	
		Óptimo inicial	Óptimo actualizado
1	0.834290	1,811.96	1,894.06
2	0.577695	1,872.75	1,942.24
3	0.323780	1,874.04	1,856.13
4	0.524162	1,867.70	1,910.24
5	0.417450	1,913.86	1,920.61

6	0.495964	1,897.95	1,901.88
7	0.613833	1,915.56	1,918.60
8	0.770774	1,931.13	1,944.07

Repitiendo la operación para el origen  $t = 113$ :

$$a_{t+1} = a_{112+1} = a_{113} = Z_{112+1} - z_{112}(1) = 1,625 - 1,894.06 = -269.06$$

$h$	$\psi_h$	Pronósticos	
		Óptimo inicial	Óptimo actualizado
1	0.834290	1,872.75	1,717.77
2	0.577695	1,874.04	1,700.70
3	0.323780	1,867.70	1,823.12
4	0.524162	1,913.86	1,779.58
5	0.417450	1,897.95	1,789.57
6	0.495964	1,915.56	1,785.16
7	0.613833	1,931.13	1,778.91

Repitiendo la operación para el origen  $t = 114$ :

$$a_{t+1} = a_{113+1} = a_{114} = Z_{113+1} - z_{113}(1) = 1,927 - 1,717.77 = 209.23$$

$h$	$\psi_h$	Pronósticos	
		Óptimo inicial	Óptimo actualizado
1	0.834290	1,874.04	1,875.25
2	0.577695	1,867.70	1,943.99
3	0.323780	1,913.86	1,847.33
4	0.524162	1,897.95	1,899.24
5	0.417450	1,915.56	1,872.50
6	0.495964	1,931.13	1,882.68

Repitiendo la operación para el origen  $t = 115$ :

$$\hat{\alpha}_{t+1} = \hat{\alpha}_{114+1} = \hat{\alpha}_{115} = Z_{1143+1} - \bar{x}_{143}(1) = 1,677 - 1,875.25 = -198.25$$

		Pronósticos	
$h$	$\psi_h$	Óptimo inicial	Óptimo actualizado
1	0.834290	1,867.70	1,778.59
2	0.577695	1,913.86	1,732.80
3	0.323780	1,897.95	1,835.05
4	0.524162	1,915.56	1,768.58
5	0.417450	1,931.13	1,799.92

Por último, revisaré la capacidad de pronóstico de mi modelo considerando los pronósticos originales; así, tomando como origen a  $t = 110$  y los horizontes  $h = 1, 2, 3, 4$  y  $5$  calculo los errores de los pronósticos de la serie transformada respecto del valor correspondiente a las observaciones más recientes de la serie una vez que les he aplicado la transformación.

		serie			error
$h$	$t$	original	transformada	estimada	
		$Z_t$	$T(Z_t)$	$\hat{T}(Z_t)$	$T(Z_t t) - \hat{T}(Z_t)$
1	111	1,683	0.024376	0.023279	0.001097
2	112	1,958	0.022599	0.023538	-0.000484
3	113	1,625	0.024807	0.023538	0.001269
4	114	1,927	0.022780	0.023157	-0.000377
5	115	1,677	0.024419	0.023155	0.001264
Promedio $\bar{M}$					0.000554
ECM					0.000001

Como puede verse, ambos valores son próximos a cero por lo que es posible concluir que el modelo tienen una muy buena capacidad de pronóstico.

Con esto finalizo el análisis de la segunda serie en el que concluyo que, en esencia, la serie de tiempo correspondiente a las exportaciones petroleras medidas a partir de los miles de barriles exportados por mes se comporta como un proceso **ARIMA(4,1,4)** y mediante este modelo se pueden obtener pronósticos confiables de la serie

## Conclusiones

La primer vez que escuché el término series de tiempo fue en el año de 1992 mientras cursaba el segundo semestre de la carrera de Actuaría, la forma en la que la ayudante de materia habló sobre la importancia de esta metodología para pronosticar inició mi interés por saber acerca de esta teoría. Hoy, puedo decir que tengo algo más de conocimiento acerca de uno de los enfoques para el análisis de series temporales y, aunque soy consciente de que aún no soy una experta, me siento orgullosa de haber logrado el objetivo que me marqué de aprender a analizar series cronológicas

A través de la investigación que desarrollé para obtener la información necesaria para elaborar esta tesis pude darme cuenta de la vigencia que tiene actualmente el uso de las series de tiempo bajo los diferentes enfoques que existen en prácticamente todos los campos del conocimiento, incluso hay revistas especializadas exclusivamente en la difusión de los avances más recientes en la teoría de series cronológicas y encuentros internacionales de personalidades que se han dedicado a la investigación en este campo. Es bueno saber que un conocimiento que fue adquirido con gran esfuerzo sigue vigente y es de gran importancia y de aplicación práctica.

Después de haber trabajado durante varios meses en la elaboración de este escrito puedo decir que los modelos **ARIMA**, presentados por vez primera por Box y Jenkins, son bastante prácticos pues la metodología que se sigue para modelar una serie, verificar la bondad del modelo seleccionado y efectuar el cálculo de pronósticos es relativamente sencilla aunque es cierto que hay algunos casos en los



que la selección del mejor modelo requiere de una amplia experiencia en el manejo de serie temporales.

En particular, para las dos series sobre las que decidí trabajar, la identificación y la construcción del modelo fueron etapas relativamente sencillas como pudo observarse en cada uno de los capítulos en los que se analizan las dos series objeto de esta tesis. Las etapas más laboriosas en ambos ejercicios fueron la estimación de parámetros y la verificación de los modelos; la dificultad se debió en gran medida a los varios pasos que hube de efectuar para llegar a un resultado.

A pesar de las dificultades presentadas por las dos etapas antes mencionadas, éstas fueron, junto con la elaboración de pronósticos, las más interesantes en todo el proceso de análisis. En esta última etapa de pronosticación debo señalar que quedé bastante satisfecha con los resultados obtenidos tanto en la estimación puntual como en la realizada a través de intervalos.

Los valores puntuales obtenidos pueden ser considerados según mi criterio bastante buenos, aunque debo señalar que puede ser mucho mejor guiarnos por la estimación intervalar pues ésta nos da una idea más amplia del comportamiento que podríamos esperar de nuestra serie y, por tanto, nos permitiría prepararnos para los diferentes escenarios resultantes de un cambio en los valores de la serie, claro, siempre y cuando no exista un evento exógeno que intervenga en el comportamiento de la serie pues ese tipo de análisis no fue contemplado dentro del desarrollo de este trabajo.

Considero que para las dos series analizadas los modelos seleccionados fueron los mejores que pude construir y precisamente de ahí viene la bondad de los pronósticos realizados para cada una de ellas. Creo que esos pronósticos son la

mejor muestra de la eficiencia que puede lograrse al aplicar la metodología de modelos **ARIMA** a las series de tiempo.

Intenté realizar este trabajo en el nivel más básico y completo posible y creo haberlo conseguido, espero que el tratamiento que le di a mi texto sea *suficientemente sencillo como para* que otras personas que desconozcan esta metodología la puedan entender como yo a través de este compendio.

Finalmente, es importante destacar que el uso de paquetes estadísticos que cuentan con las rutinas necesarias para las diferentes etapas permiten que el desarrollo de éstas sea más eficiente y rápido. En mi caso el paquete que me auxilió en las diferentes etapas, pero muy especialmente en la estimación de parámetros óptimos, fue **STATGRAPHICS**.

**TABLA A.**

t	INGRESO FEDERAL MILLONES DE PESOS	TRANSFORMADA		FAC	FACP
	Z <sub>t</sub>	$\lambda = 0 \Rightarrow \text{LN}(Z_t)$	ESTACIONARIA	MUESTRAL	MUESTRAL
		T(Z <sub>t</sub> )	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$
1	78,683.30	11.27	-0.001	-0.010822	-0.010820
2	88,897.10	11.40	0.011	-0.068743	-0.068870
3	98,445.70	11.50	0.000	0.057990	0.056720
4	108,180.00	11.59	0 003	-0.125722	-0.130410
5	118,513.80	11.68	0.071	-0.092822	-0.088160
6	12,168.80	9.41	-0.011	0.003546	-0.020030
7	22,463.80	10.02	-0.005	0.057282	0.059990
8	34,333.10	10.44	0.005	0.101258	0.098390
9	45,919.80	10.73	-0.016	0.048935	0.039670
10	58,680.20	10.98	-0.006	-0.026866	-0.029710
11	71,038.20	11.17	-0.011	0.152291	0.166470
12	84,429.70	11.34	0.000	0.023852	0.059970
13	97,444.60	11.49	0 002	-0.054152	0.002490
14	109,971.40	11.61	0.010	-0.159552	-0.189800
15	123,116.20	11.72	0.015	-0.079804	-0.083520
16	135,297.50	11.82	0.032	0.001752	0.038110
17	148,601.10	11.91	-0.061	0.000612	-0.038210
18	16,381.80	9.70	-0.020	0.000232	-0.039140
19	29,924.40	10.31	-0.051	-0.000715	-0.128880
20	45,505.90	10.73	0.004	-0.000641	-0.037630
21	61,193.90	11.02	-0.012	-0.000875	-0.033020
22	76,962.90	11.25	-0.010	-0.001061	0.080170
23	92,640.70	11 44	-0.004	-0.001149	0.024070
24	108,913.70	11.60	0.003	-0 001098	0.040270
25	125,734.50	11.74	-0.032	-0.001687	-0.102370
26	142,248.20	11.87	0.018	-0.001366	0.070410

**TABLA A.**

t	INGRESO FEDERAL MILLONES DE PESOS	TRANSFORMADA $\lambda = 0 \Rightarrow \text{LN}(Z_t)$	ESTACIONARIA	FAC MUESTRAL	FACP MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla \nabla_{12} \bar{T}(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} \bar{T}(Z_t)$
27	160,874.50	11.99	-0.013	-0.001606	0.027360
28	179,391.40	12.10	-0.015	-0.001894	-0.071670
29	203,486.30	12.22	-0.008	-0.002056	-0.073540
30	21,095.50	9.96	-0.020	-0.002441	-0.052080
31	37,762.10	10.54	0.043		
32	54,568.10	10.91	0.011		
33	73,702.80	11.21	0 000		
34	91,554.90	11.42	0 010		
35	109,141.60	11.60	0.009		
36	127,746.00	11.76	0.030		
37	147,929.60	11.90	0.021		
38	162,157.80	12.00	-0.033		
39	186,674.80	12.14	-0.004		
40	205,542.30	12.23	-0.001		
41	229,608.30	12.34	-0.148		
42	23,603.60	10.07	0.050		
43	41,395.10	10.63	0.037		
44	62,445.80	11.04	-0.002		
45	85,310.40	11.35	0.010		
46	105,924.10	11.57	0.003		
47	127,561.80	11.76	-0.003		
48	150,681.30	11.92	-0.043		
49	179,825.80	12.10	0.009		
50	201,365.70	12.21	-0.013		
51	224,312.80	12.32	0.006		
52	245,956.20	12.41	-0.002		

**TABLA A.**

t	INGRESO FEDERAL MILLONES DE PESOS	TRANSFORMADA $\lambda = 0 \Rightarrow \text{LN}(Z_t)$	ESTACIONARIA	FAC MUESTRAL	FACP MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$
53	274,471.00	12.52	0.013		
54	24,329.00	10.10	0.000		
55	44,853.00	10.71	0.009		
56	70,193.50	11.16	-0.006		
57	95,747.70	11.47	-0.009		
58	120,070.20	11.70	-0.006		
59	145,010.90	11.88	-0.004		
60	170,754.50	12.05	0.011		
61	195,140.60	12.18	0.004		
62	220,510.20	12.30	0.019		
63	242,571.40	12.40	0.000		
64	267,561.40	12.50	-0.006		
65	297,951.00	12.60	0.034		
66	26,755.80	10.19	-0.001		
67	49,329.20	10.81	-0.006		
68	77,897.40	11.26	-0.002		
69	105,655.60	11.57	0.040		
70	131,318.80	11.79	0.019		
71	157,694.40	11.97	0.008		
72	184,987.10	12.13	0.000		
73	213,703.30	12.27	-0.012		
74	242,545.50	12.40	0.001		
75	271,891.90	12.51	0.002		
76	299,905.00	12.61	0.040		
77	332,033.50	12.71	0.118		
78	30,859.20	10.34	-0.001		

**TABLA A.**

t	INGRESO FEDERAL	TRANSFORMADA	ESTACIONARIA	FAC	FACP
	MILLONES DE PESOS	$\lambda = 0 \Rightarrow \text{LN}(Z_t)$		MUESTRAL	MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla\nabla_{12}T(Z_t)$	$\nabla\nabla_{12}T(Z_t)$	$\nabla\nabla_{12}T(Z_t)$
79	56,845.60	10.95	-0.024		
80	89,221.10	11.40	0.017		
81	120,795.40	11.70	-0.005		
82	156,286.60	11.96	-0.029		
83	191,184.80	12.16	0.001		
84	226,180.80	12.33	0.002		
85	261,338.70	12.47	0.005		
86	293,176.40	12.59	0.006		
87	329,093.90	12.70	0.010		
88	363,651.40	12.80	-0.005		
89	418,882.60	12.95	0.059		
90	43,792.10	10.69	-0.034		
91	80,567.90	11.30	-0.035		
92	123,493.50	11.72	-0.025		
93	170,036.20	12.04	-0.026		
94	218,835.90	12.30	0.022		
95	260,066.70	12.47	-0.010		
96	307,927.50	12.64	-0.019		
97	356,635.90	12.78	-0.012		
98	401,908.00	12.90			
99	454,049.10	13.03			
100	506,786.00	13.14			
101	580,722.00	13.27			
102	64,379.60	11.07			
103	114,539.60	11.65			
104	169,552.30	12.04			

**TABLA A.**

t	INGRESO FEDERAL MILLONES DE PESOS	TRANSFORMADA $\lambda = 0 \Rightarrow \text{LN}(Z_t)$	ESTACIONARIA	FAC MUESTRAL	FACP MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$	$\nabla \nabla_{12} T(Z_t)$
105	227,632.60	12.34			
106	285,453.00	12.56			
107	346,888.50	12.76			
108	406,802.50	12.92			
109	462,275.20	13.04			
110	514,614.00	13.15			

TABLA B.

EXPORTACIONES PETROLERAS MILES DE BARRILES DIARIOS		TRANSFORMADA $\lambda = -0.5 \Rightarrow (Z_t)^{-0.5}$	ESTACIONARIA	FAC MUESTRAL	FACP MUESTRAL
t	$Z_t$	$T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$
1	1,329	0.027431	0.000062	-0.091243	-0.09124
2	1,323	0.027493	-0.000031	-0.213643	-0.22383
3	1,326	0.027462	0.000031	-0.182180	-0.24041
4	1,323	0.027493	-0.000062	0.056514	-0.05535
5	1,329	0.027431	0.000000	-0.000806	-0.10993
6	1,329	0.027431	0.000444	0.033695	-0.03062
7	1,287	0.027875	0.000208	0.037051	0.01901
8	1,268	0.028083	-0.000132	-0.172984	-0.20168
9	1,280	0.027951	0.000028	0.208809	0.20265
10	1,277	0.027979	0.000178	0.058673	0.05308
11	1,261	0.028157	-0.000243	-0.058691	-0.01802
12	1,283	0.027915	0.000461	-0.122428	-0.00259
13	1,242	0.028376	0.000293	0.162083	0.15513
14	1,217	0.028669	-0.000293	-0.052906	-0.04539
15	1,242	0.028376	-0.000137	0.074703	0.15555
16	1,254	0.028239	0.000786	-0.005261	-0.03444
17	1,187	0.029025	-0.000477	-0.004234	0.05598
18	1,227	0.028548	0.000271	-0.004966	-0.19241
19	1,204	0.028820	0.000000	-0.005061	0.06823
20	1,204	0.028820	-0.000341	-0.004353	-0.06716
21	1,233	0.028479	0.000258	-0.005055	-0.05274
22	1,211	0.028736	-0.001065	-0.002644	-0.00474
23	1,306	0.027671	-0.000713	-0.001062	0.09779
24	1,376	0.026958	-0.000020	-0.001110	-0.03990
25	1,378	0.026939	0.000208	-0.001694	0.03211



TABLA B.

t	EXPORTACIONES PETROLERAS MILES DE BARRILES DIARIOS	TRANSFORMADA	ESTACIONARIA	FAC	FACP
		$\lambda = -0.5 \Rightarrow (Z_t)^{-0.5}$		MUESTRAL	MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$
26	1,357	0.027146	-0.000247	-0.001207	-0.02268
27	1,382	0.026900	0.000257	-0.001906	0.11395
28	1,356	0.027156	-0.000010	-0.001977	-0.05042
29	1,357	0.027146	0.000223	-0.002597	-0.04316
30	1,335	0.027369	-0.000183	-0.002261	-0.00015
31	1,353	0.027186	-0.000393		
32	1,393	0.026793	-0.000162		
33	1,410	0.026631	0.000307		
34	1,378	0.026939	-0.000020		
35	1,380	0.026919	-0.000097		
36	1,390	0.026822	0.000274		
37	1,362	0.027096	0.000110		
38	1,351	0.027206	-0.000100		
39	1,361	0.027106	0.000000		
40	1,361	0.027106	0.000120		
41	1,349	0.027227	-0.000200		
42	1,369	0.027027	-0.000127		
43	1,382	0.026900	0.000137		
44	1,368	0.027037	-0.000176		
45	1,386	0.026861	0.000137		
46	1,372	0.026997	-0.000137		
47	1,386	0.026861	-0.000058		
48	1,392	0.026803	0.000195		
49	1,372	0.026997	0.000129		
50	1,359	0.027126	0.000440		

**TABLA B.**

t	EXPORTACIONES PETROLERAS MILES DE BARRILES DIARIOS	TRANSFORMADA	ESTACIONARIA	FAC	FACP
		$\lambda = -0.5 \Rightarrow (Z_t)^{-0.5}$		MUESTRAL	MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$
51	1,316	0.027566	0.000137		
52	1,303	0.027703	0.000000		
53	1,303	0.027703	-0.000406		
54	1,342	0.027298	0.000031		
55	1,339	0.027328	-0.000112		
56	1,350	0.027217	0.000010		
57	1,349	0.027227	0.000010		
58	1,348	0.027237	0.000225		
59	1,326	0.027462	-0.000255		
60	1,351	0.027206	-0.000160		
61	1,367	0.027047	0.000281		
62	1,339	0.027328	0.000103		
63	1,329	0.027431	0.000073		
64	1,322	0.027503	-0.000287		
65	1,350	0.027217	0.000287		
66	1,322	0.027503	-0.000042		
67	1,326	0.027462	0.000094		
68	1,317	0.027555	0.000148		
69	1,303	0.027703	-0.000064		
70	1,309	0.027639	-0.000229		
71	1,331	0.027410	0.000346		
72	1,298	0.027756	0.000086		
73	1,290	0.027842	0.000152		
74	1,276	0.027995	0.000290		
75	1,250	0.028284	0.000011		

TABLA B.

t	EXPORTACIONES PETROLERAS MILES DE BARRILES DIARIOS	TRANSFORMADA	ESTACIONARIA	FAC	FACP
		$\lambda = -0.5 \Rightarrow (Z_t)^{-0.5}$		MUESTRAL	MUESTRAL
	$Z_t$	$T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$	$\nabla T(Z_t)$
76	1,249	0.028296	-0.000356		
77	1,281	0.027940	0.000412		
78	1,244	0.028352	0.000336		
79	1,215	0.028689	-0.001391		
80	1,342	0.027298	-0.000041		
81	1,346	0.027257	-0.000905		
82	1,440	0.026352	-0.000584		
83	1,506	0.025768	0.000882		
84	1,408	0.026650	0.003556		
85	1,096	0.030206	-0.001290		
86	1,196	0.028916	-0.001588		
87	1,339	0.027328	-0.001500		
88	1,499	0.025828	-0.000052		
89	1,505	0.025777	0.000493		
90	1,449	0.026270	-0.000960		
91	1,561	0.025310	0.000205		
92	1,536	0.025516	-0.001361		
93	1,714	0.024154	0.002008		
94	1,461	0.026162	-0.000546		
95	1,524	0.025616	-0.000740		
96	1,616	0.024876	-0.000310		
97	1,657	0.024566	0.001262		
98	1,499	0.025828	-0.000094		
99	1,510	0.025734	-0.000656		
100	1,590	0.025078	0.000183		

**TABLA B.**

t	EXPORTACIONES PETROLERAS MILES DE BARRILES DIARIOS $Z_t$	TRANSFORMADA	ESTACIONARIA	FAC	FACP
		$\lambda = -0.5 \Rightarrow (Z_t)^{-0.5}$ $T(Z_t)$	$\nabla T(Z_t)$	MUESTRAL $\nabla T(Z_t)$	MUESTRAL $\nabla T(Z_t)$
101	1,567	0.025262	-0.000857		
102	1,679	0.024405	-0.000939		
103	1,816	0.023466	0.001048		
104	1,664	0.024515	-0.000275		
105	1,702	0.024239	-0.000643		
106	1,796	0.023596	0.000193		
107	1,767	0.023789	-0.000310		
108	1,814	0.023479	0.000584		
109	1,727	0.024063	-0.001012		
110	1,882	0.023051			

**TABLA C.**

t	$E_t[\alpha_t]$	$\mathcal{T}_t(1)$	t	$E_t[\alpha_t]$	$\mathcal{T}_t(1)$	t	$E_t[\alpha_t]$	$\mathcal{T}_t(1)$
1	0.000458	0.027355	31	-0.000226	0.026974	61	-0.000240	0.027035
2	0.000138	0.027352	32	-0.000180	0.026447	62	0.000293	0.027561
3	0.000109	0.027292	33	0.000184	0.026889	63	-0.000130	0.027308
4	0.000201	0.027488	34	0.000049	0.026623	64	0.000196	0.027414
5	-0.000058	0.027340	35	0.000296	0.026783	65	-0.000198	0.027353
6	0.000091	0.027423	36	0.000039	0.026855	66	0.000151	0.027306
7	0.000452	0.027885	37	0.000241	0.026780	67	0.000156	0.027551
8	0.000198	0.027880	38	0.000426	0.027224	68	0.000005	0.027418
9	0.000070	0.027869	39	-0.000118	0.026930	69	0.000286	0.027601
10	0.000110	0.028005	40	0.000176	0.027010	70	0.000038	0.027650
11	0.000152	0.028078	41	0.000217	0.027361	71	-0.000239	0.027285
12	-0.000163	0.027963	42	-0.000334	0.026876	72	0.000472	0.027843
13	0.000413	0.028379	43	0.000023	0.027080	73	-0.000001	0.027779
14	0.000290	0.028599	44	-0.000044	0.027213	74	0.000215	0.027781
15	-0.000223	0.028255	45	-0.000352	0.026804	75	0.000504	0.028354
16	-0.000016	0.028356	46	0.000194	0.027242	76	-0.000058	0.028025
17	0.000669	0.029010	47	-0.000381	0.026920	77	-0.000085	0.027926
18	-0.000462	0.028417	48	-0.000117	0.026779	78	0.000427	0.028489
19	0.000402	0.028811	49	0.000219	0.027242	79	0.000200	0.028451
20	0.000008	0.028959	50	-0.000115	0.026919	80	-0.001153	0.027460
21	-0.000481	0.028294	51	0.000647	0.027454	81	-0.000203	0.027673
22	0.000442	0.029012	52	0.000249	0.027595	82	-0.001320	0.026876
23	-0.001341	0.027873	53	0.000108	0.027272	83	-0.001107	0.026216
24	-0.000914	0.027208	54	0.000025	0.027415	84	0.000434	0.027317
25	-0.000269	0.027843	55	-0.000087	0.027254	85	0.002889	0.029575
26	-0.000697	0.027130	56	-0.000038	0.027222	86	-0.000659	0.028007
27	-0.000230	0.027173	57	0.000005	0.027411	87	-0.000679	0.027002
28	-0.000017	0.027441	58	-0.000175	0.027203	88	-0.001173	0.026560
29	-0.000295	0.026796	59	0.000259	0.027548	89	-0.000783	0.025895
30	0.000573	0.027412	60	-0.000341	0.027287	90	0.000376	0.026823

**TABLA C.**

t	$E_i[a_i]$	$T_i(1)$	t	$E_i[a_i]$	$T_i(1)$	t	$E_i[a_i]$	$T_i(1)$
91	-0.001513	0.025594	98	0.001006	0.025745	105	-0.000533	0.024218
92	-0.000078	0.025583	99	-0.000011	0.025000	106	-0.000621	0.023757
93	-0.001429	0.025048	100	0.000078	0.025048	107	0.000032	0.024168
94	0.001114	0.025773	101	0.000213	0.025174	108	-0.000689	0.023491
95	-0.000157	0.025810	102	-0.000769	0.024230	109	0.000572	0.024097
96	-0.000934	0.024449	103	-0.000764	0.023946	110	-0.001046	
97	0.000117	0.024822	104	0.000568	0.024772	<b>PROMEDIO</b>		<b>0.026972</b>

## Bibliografía

- Anderson, O.D., *Time Series Analysis and Forecasting: The Box-Jenkins Approach*, Gran Bretaña, Butterworths, 1979.
- Box, George E. P., y Gwilym M. Jenkins, *Time Series Analysis. Forecasting and Control*, U.S.A. Holden-Day, 1976.
- Guerrero, M. Víctor, *Análisis estadístico de series de tiempo económicas*, México, Universidad Autónoma Metropolitana, 1991.
- Gourieroux, Christian y Alain Monfort, *Cours de series temporelles*, Francia, Economica, 1983
- INEGI, *Cuaderno de información oportuna*, vols. 189, 195, 201, 207, 213, 219, 225, 231, 237, 243, 249, 255, 261, 267, 273, 279, 285, 291, Aguascalientes, México.
- Martín P., F. Javier, *Introducción a la estadística económica y empresarial. Teoría y práctica*, España, A Editorial, 1995.
- Pérez López, César, *Econometría y análisis estadístico multivariable con Statgraphics*, España, RA-MA, 1996.
- Wei, William W. S., *Time Series Analysis. Univariate and Multivariate Methods*, U.S.A., Adison Wesley, 1990