



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

FACULTAD DE CIENCIAS

INTRODUCCION A LOS MODELOS ARIMA  
MULTIPLICATIVOS

**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**

**ALBERTO IVAN MENDOZA MARTINEZ**



FACULTAD DE CIENCIAS  
UNAM

DIRECTORA DE TESIS: M. en A.P. MARIA DEL PILAR ALONSO REYES

2005



m. 344 167



Universidad Nacional  
Autónoma de México



**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.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

Integrado a la Biblioteca General de Bibliotecas de la  
UNAM para su depósito y registro e impreso el  
contenido de la tesis de recepción.

NOMBRE: Alberto Iván Mendoza Martínez

FECHA: 17 mayo - 2005

FIRMA:

**ACT. MAURICIO AGUILAR GONZÁLEZ**  
**Jefe de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

"Introducción a los Modelos ARIMA Multiplicativos"

realizado por Alberto Iván Mendoza Martínez

con número de cuenta 098519259 , quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director  
Propietario

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

Propietario

M. en C. José Antonio Flores Díaz

Propietario

Act. Lucio Gerardo Chávez Heredia

Suplente

Act. María Aurora Valdés Michell

Suplente

Act. Marypaola Janett Maya López

Consejo Departamental de Matemáticas



Act. Jaime Vázquez Alamilla

FACULTAD DE CIENCIAS  
CONSEJO DEPARTAMENTAL  
DE  
MATEMÁTICAS

# ÍNDICE

Agradecimientos

Introducción VII

## **CAPÍTULO 1. Series de Tiempo.** 1

1.1. Introducción . . . . .	1
1.2. Definición y conceptos claves . . . . .	2
1.2.1. Componentes de una serie de tiempo . . . . .	4
1.2.2. Objetivos del análisis de una serie de tiempo . . . . .	7
1.3. Modelos de series de tiempo . . . . .	8
1.4. Exploración de patrones de datos con análisis de autocorrelación . . . . .	9
1.5. Métodos de suavizamiento . . . . .	12
1.5.1. Promedios simples . . . . .	12
1.5.2. Promedios móviles . . . . .	12
1.5.3. Métodos de atenuación exponencial . . . . .	14
1.6. Descomposición de las componente de una serie de tiempo . . . . .	15
1.6.1. Tendencia . . . . .	17
1.6.2. Variación cíclica . . . . .	21
1.6.3. Variación estacional . . . . .	24
1.7. Gráfica de las componentes . . . . .	30

## **CAPÍTULO 2. Modelos ARIMA Multiplicativos.** 33

2.1. Series de tiempo en los procesos estocásticos . . . . .	33
2.2. Procesos estacionarios . . . . .	34

2.2.1. Estacionariedad en media y varianza . . . . .	36
2.2.2. Proceso de ruido blanco . . . . .	37
2.3. Notación de operadores . . . . .	39
2.3.1. El operador retardo . . . . .	39
2.3.2. El operador diferencia . . . . .	41
2.3.3. Operadores inversos . . . . .	43
2.4. Descomposición de Wold . . . . .	44
2.5. Ecuaciones en diferencia . . . . .	45
2.5.1. Ecuaciones en diferencia de orden uno . . . . .	46
2.5.2. Ecuaciones en diferencia de orden dos . . . . .	46
2.5.3. Ecuaciones en diferencia de orden $k$ . . . . .	50
2.6. Modelos Autorregresivos . . . . .	51
2.6.1. El proceso autorregresivo de primer orden $AR(1)$ . . . . .	51
2.6.1.1. Representación del proceso $AR(1)$ como suma de choques aleatorios . . . . .	55
2.6.2. El proceso autorregresivo de segundo orden $AR(2)$ . . . . .	56
2.6.2.1. Representación del proceso $AR(2)$ como suma de choques aleatorios . . . . .	62
2.6.3. El proceso autorregresivo general $AR(p)$ . . . . .	63
2.7. Función de autocorrelación parcial ( $fap$ ) . . . . .	66
2.8. Modelos de promedios móviles . . . . .	74
2.8.1. El proceso de media móvil de primer orden $MA(1)$ . . . . .	75
2.8.2. El proceso de media móvil de segundo orden $MA(2)$ . . . . .	78
2.8.3. El proceso de media móvil de orden general $MA(q)$ . . . . .	81
2.9. Modelos $ARMA$ . . . . .	85
2.9.1. El proceso $ARMA(1,1)$ . . . . .	86
2.9.2. El proceso $ARMA(p,q)$ . . . . .	89
2.10. Procesos no estacionarios . . . . .	91
2.10.1. Procesos Integrados . . . . .	91
2.10.1.1. La caminata aleatoria . . . . .	92

2.10.2. Procesos Integrados <b>ARIMA</b> . . . . .	93
2.10.3. El concepto de estacionalidad . . . . .	95
2.10.4. El modelo <b>ARIMA</b> estacional . . . . .	97
2.10.4.1. La <i>fas</i> y <i>fap</i> de un proceso <b>ARIMA</b> estacional . . . . .	99
2.11. Predicción con modelos <b>ARIMA</b> . . . . .	101
2.12. Transformaciones para estabilizar la varianza . . . . .	105
2.13. Criterios de selección de modelos . . . . .	108
2.13.1. El criterio AIC de Akaike . . . . .	109
2.14. Revisión del modelo . . . . .	110
2.14.1. Contraste de media cero . . . . .	110
2.14.2. Contraste de homocedasticidad . . . . .	110
2.14.3. Contraste de normalidad . . . . .	111
2.14.4. Contraste de Ljung-Box sobre las autocorrelaciones . . . . .	112
2.15. Verificación del modelo ajustado . . . . .	113
<b>CAPÍTULO 3. Aplicaciones del Modelo ARIMA Multiplicativo.</b>	114
3.1. Análisis de la serie del Índice al Valor Agregado (IVA) . . . . .	114
3.2. Análisis de la serie del Agregado Monetario M4 . . . . .	134
<b>CONCLUSIONES</b>	148
<b>APÉNDICES</b>	152
Apéndice A. . . . .	152
Apéndice B. . . . .	155
Apéndice C. . . . .	158
Apéndice D. . . . .	160
<b>BIBLIOGRAFÍA</b>	161

## Agradecimientos.

En el transcurso de toda mi vida, siempre ha existido alguien a quien debo agradecer el permitirme seguir de pie y con la fe en alto en todo momento. Tú, quien sabes cuándo actuar, quien sabes cuándo apretar y cuándo colaborar, lograste en tus tiempos que hoy crea en algo, ese algo que sé se mueve entre todos nosotros. A Tí **Señor**, te doy gracias por acompañarme en esos momentos de soledad y por todos esos obstáculos que me han hecho fuerte.

A Ustedes, mi tesoro lejano, por quienes he entregado todo de mí, les dedico este trabajo, que olvidando todo lo matemático, estadístico, económico, resume el esfuerzo, la conciencia y sacrificio que cada uno de ustedes aportó para que este humilde servidor siguiera con sus estudios y el día de hoy pudiera obtener su título. Sí, a tí Beto, Tere, Ivonne e Irving, les agradezco infinitamente sus cartas, pensamientos, consejos, regaños y sobretodo el apoyo brindado para poder vivir en esta ciudad. Sé que no habrá nada que pueda pagar todo lo que han hecho por mí y mucho menos el cómo devolverles lo que han hecho de mí: los valores y principios que hasta la fecha me han inculcado. Gracias **Familia Mendoza Martínez**, mil gracias por todo. A mis abuelitas Esther Rosales (Q.E.P.D.) y Esthela Reyes no olviden que las quiero mucho y gracias por sus ánimos y bendiciones en cada momento. A mis tíos y primos que tanto quiero y extraño, gracias por hacer más felices mis días de vacaciones en Tuxtla.

Personas a quienes les agradezco mucho el haber compartido parte de su espacio, son aquellas quienes me adoptaron cuando recién llegué y que merecen todo mi reconocimiento. A usted, Sra. Silvia (*Mapía*), gracias por acogerme en su hogar y permitirme comenzar la travesía que hoy consigue un fruto, gracias mil por ayudarme sin siquiera conocerme. A usted **Sra. Rosario Guerrero** y **Sra. Alicia Flores**, gracias por estar el pendiente de mí, gracias por recibirme con los brazos abiertos en sus casas y hospedarme de vez en cuando.

A mis **amigos**, mi segunda familia, tengo que agradecerles el ayudarme a ponerme aquellas caídas, el aconsejarme de distintas formas a cómo sobrevivir en esta ciudad (y que bien saben lo que pienso de ésta). Todas las alegrías, tristezas, aventuras, viajes, fiestas, triunfos y principalmente el poder haber hecho un gran equipo de amistad y de trabajo. A ustedes mis estimados “Chiapas”, que poco a poco fueron entendiendo el idioma con el que yo me desenvolvía, les estaré infinitamente agradecido por su apoyo y comprensión. A ti Armando, Gaby, Memo, Vic, Kinkin y Marisol, Mariana, Peter, Ale, Ray, Raúl, Lilis, gracias por todo. A mis amigos de Tuxtla, Franz, Elí, Silver, Daniela, Tere, gracias por sus consejos y ánimos brindados en su momento para alentarme en la universidad.

Agradecimiento especial merecen personas que con el tiempo han mantenido una gran amistad conmigo, que han sabido aconsejarme y con quienes he pasado momentos muy especiales. Claudia, gracias por todas las experiencias que compartimos, por tus consejos y tu comprensión. Jorge y Jessi, mis amiguitos bimsers, gracias porque ustedes han sido un amuleto importante en mi vida, por su sinceridad y apoyo en todo

momento, personas como ustedes son muy difíciles de encontrar por aquí. Igualmente Citlali, Elodia, Selene, Wendy, Gilberto, Margarita y Samadhi, Karilia, muchas gracias por ser excelentes personas conmigo y por las buenas vibras que me brindaron para terminar este trabajo. Discúlpeme las personas de quienes me estoy olvidando, pero a todos agradezco lo mucho que hicieron por mí.

A mis **profesores** de la facultad de Ciencias, gracias por los conocimientos que por ustedes adquirí y que han hecho de mí un buen Actuario, pues sin ustedes no hubiera sido posible la culminación de esta tesis. Al Dr. Javier Páez, por sus consejos y enseñanzas en Cálculo muchas gracias.

A mi Directora de Tesis, la M. en A. P. María del Pilar Alonso Reyes, muchas gracias por la comprensión y paciencia que tuviste conmigo, pues fuiste tu quien me enseñó las herramientas estadísticas que hicieron posible este trabajo, muchas gracias profe.

A cada uno de mis sinodales, al M. en C. José Antonio Flores, a la Act. MaryPaola Maya, al M. en C. Gerardo Chávez y a la Act. Aurora Valdés, por su tiempo y paciencia para revisar mi trabajo, así como la disponibilidad que desde un principio tuvieron para atenderme, muchas gracias señores.

Finalmente a mi gran casa de estudios, la **UNAM**, por ser la institución que es. Por su pluralidad y apoyo a los estudiantes que de provincia vienen a cumplir sueños. Hoy estoy logrando uno de muchos. Gracias Universidad.

Sinceramente,  
Su amigo Iván



# INTRODUCCIÓN.

Desde tiempos remotos el hombre ha buscado la forma de predecir el futuro, basándose en distintas técnicas que pueden llegar a convertirlos en sabias personas o en simples charlatanes. El uso de técnicas no estadísticas para lograr la predicción, bien pueden explicar la incertidumbre y de algún modo dar una aproximación de lo que sucederá más adelante, sin embargo esta manera empírica de visualizar el futuro no arroja alguna medida de la incertidumbre que el tiempo por sí mismo acarrea. Para esto existen técnicas estadísticas que dan una mejor aproximación de lo que pasará en un horizonte no muy lejano y tienen la ventaja de que poder cuantificar la incertidumbre de nuestra predicción.

El hecho de predecir lo que sucederá más adelante es una habilidad que hoy en día ha tomado fuerza en diversos campos, por ejemplo, es común encontrar que en Medicina se trate de estimar el número de muertes por cierta enfermedad o que economistas traten de pronosticar el valor de cierto índice en México. Estos son sólo dos ejemplos de miles encontrados en la vida diaria, por lo que es necesario analizar a la variable de interés desde diversas perspectivas, sin importar el efecto que otras variables le causen, importando solamente la historia que la variable de estudio traiga consigo.

El análisis de Series de Tiempo brinda una herramienta eficiente para explicar, en gran medida, el comportamiento de esa variable que interesa pronosticar. Esta explicación está basada en la descomposición de un cierto número de mediciones de la variable de estudio, lo que se conoce como *serie* y es de *tiempo* debido a que esas mediciones fueron tomados con el paso de los días, semanas, meses, años, etc. La descomposición de la serie de tiempo involucra herramientas estadísticas, las cuales serán estudiadas en el capítulo 1 del presente trabajo.

El propósito del análisis de la serie es tratar de identificar a ciertos factores que influyen en el comportamiento de la misma, ya que es común encontrar en la mayoría de éstas algún patrón que se repite periódicamente, que los valores de la serie aumentan a medida que avanza el tiempo o que existan cambios en su comportamiento debido a decisiones que terceras personas toman de forma imprevista. El conjunto de estos factores explican el porqué de la serie y el cómo se comportará más adelante.

Aunque esta forma de pronóstico ya es bastante obsoleta (pues se conocen técnicas más modernas, eficientes y sencillas, capaces de pronosticar el futuro de la variable de interés), su estudio será necesario para poder comprender la metodología empleada en la creación de los nuevos modelos. Dichos modelos analizan a la serie como un proceso estocástico, es decir como un conjunto de variables aleatorias que están indexadas en el tiempo.

Estos modelos son muy efectivos, ya que desnudan a la serie con métodos estadísticos tratando de reducirla al más simple de los procesos. Para lograr esto es necesario definir términos y operaciones que permitirán transformar a la serie en otra más fácil de utilizar y con la que se analizarán los factores que explican su comportamiento (por ejemplo, el término de *estacionariedad*, el operador diferencia, etc.). Es precisamente en el capítulo 2 en donde se probarán matemáticamente diversas proposiciones útiles en la construcción de estos modelos, conocidos como *ARIMA* Multiplicativos (*Modelos Autorregresivos Integrados de Media Móvil*), que serán explicados a detalle en el mismo capítulo. Esta construcción se hará en partes, comenzando con los procesos autorregresivos, separadamente de los de media móvil y al final una combinación de ellos. Dichos modelos estudian a su vez la posible estacionalidad de la serie, es decir, patrones que se repiten periódicamente, aunque en ausencia de ese patrón, los modelos también son muy útiles.

Antes de estudiar a estos modelos, será necesario conocer la gráfica que analiza el comportamiento de la serie respecto a su pasado, es decir, el estudio de los rezagos de la serie. El término *rezago* es uno de los más importantes en este trabajo, ya que será esta definición la que proporcionará el grado de los parámetros del modelo. Cabe mencionar que el estudio de estos rezagos será más fácil en una gráfica llamada *correlograma*, donde se plasma la relación existente entre valor de la variable en el período actual y en los anteriores, mediante funciones que traen consigo la correlación de los datos en el tiempo.

Será en este mismo capítulo 2 en donde, además de saber interpretar a los correlogramas, también se aprenderá a identificar las transformaciones necesarias que deben hacerse a la serie original con el propósito de estabilizarla y de mantenerla a un nivel constante respecto al tiempo y de esta forma tener una estimación más sencilla de los parámetros que deben incluirse en el modelo *ARIMA* correspondiente.

No debe olvidarse, que como todo buen modelo es necesario que se cumplan ciertas hipótesis respecto a la variable de estudio y los residuales que el pronóstico arroje; dichos supuestos serán planteados al final del capítulo 2 y además se explicará la importancia de respetarlos para contar con un buen modelo de pronóstico.

En el capítulo 3 se presentan dos aplicaciones a estos modelos, utilizando la serie del Índice del Valor Agregado (IVA) y del agregado monetario M4, con el propósito de mostrar la utilidad y eficiencia de las técnicas de pronóstico actuales. Además de mostrar un Apéndice con la demostración y explicación de teoremas y proposiciones que servirán

para fundamentar matemáticamente el uso de los modelos *ARIMA*.

Será a través de este trabajo en el que el lector podrá contar con una alternativa a las técnicas de pronóstico habituales o modelos de regresión, utilizando únicamente la historia de la variable de interés y no el efecto que otras variables provoquen en la misma. En la medida en que se conozca información acerca de la variable a pronosticar, mejor será el resultado arrojado por los modelos, pues una de las ventajas que tiene es que permite incluir dicha información en la modelación.

Actualmente existen muchos paquetes estadísticos que colaboran en la rápida y precisa estimación de los parámetros del modelo. Para efectos de esta tesis, se utilizará el paquete E-Views versión 4.0, el cual se recomienda para el aplicación de los modelos *ARIMA*.

Por último, cabe mencionar que la lectura de este trabajo no está limitado a personas con bases matemáticas y/o estadísticas, sino que su utilidad puede ampliarse a todos los sectores de la vida laboral y estudiantil en donde sea necesario pronosticar valores para alguna variable de interés. El análisis de las series de tiempo es hoy en día un buen pretexto para presumir que se es capaz predecir el futuro a mediano plazo, sin caer en charlatanerías y lo que es mejor, cobrar por ello.

*Alberto Iván Mendoza Martínez*  
(*El Chiapas*)

# **CAPITULO 1. Series de Tiempo.**

## **1.1. Introducción.**

Toda institución, ya sea la familia, la empresa o el gobierno, tiene que hacer planes para el futuro para sobrevivir y progresar. Hoy en día diversas instituciones requieren conocer el comportamiento futuro de ciertos fenómenos con fin de planificar, prever o prevenir.

En la vida cotidiana siempre se tiene información de algunas variables, por ejemplo, en las noticias se mencionan diariamente los cambios en la bolsa de valores, las condiciones climáticas, las producciones de algún producto y algunas veces el índice de precios al consumidor. Al escuchar estos valores, se aprecia que tienen períodos de tiempo, aunque en muchas ocasiones surgen dudas acerca de lo que éstos indican.

En el mundo de los negocios uno se encuentra constantemente con variables que presentan una variación aleatoria en el tiempo. El tiempo, ya sea como una variable experimental independiente o como una dimensión que se agrega a otras variables de interés, puede resultar una ayuda o un estorbo en el proceso de toma de decisiones. Se sabe que muchas variables son no medibles o no controlables, y pueden hacer que el resultado que se observa varíe en el tiempo y causen así un incremento en el error experimental.

En cambio, hay situaciones de toma de decisiones en las que el tiempo es una de las variables más importantes. Por ejemplo, un investigador que esté interesado en predecir el precio futuro del oro, o el gerente de una tienda interesado en el efecto del tiempo sobre la demanda de sus productos o de un investigador de mercados que está interesado en el patrón de comportamiento de las ventas por temporada.

En cierto sentido, cada persona tiene que hacer planes para el futuro, y en un intento para presupuestar recursos y planear estrategias, se encuentra involucrado con procesos que presentan variaciones aleatorias temporales. Si se piensa que las tasas de interés pueden decaer en un futuro cercano, puede resultar acertado rentar una casa y esperar algún tiempo antes de comprar. Con buenos datos en la mano, en el sentido de que contengan la

información deseada, un pronosticador puede empezar la importante tarea de explorar los patrones de comportamiento. Este paso comprende la observación de los datos, la comprensión de lo que los datos sugieren y el uso de varios métodos gráficos para obtener una mejor visión del proceso que los generó.

Se pueden emplear dos métodos básicos de pronósticos. Las técnicas de pronósticos *cualitativas*, que se basan en el juicio humano y en la intuición, más que en la manipulación de datos históricos anteriores. Las técnicas de pronósticos *cuantitativas*, se usan cuando existen suficientes datos históricos disponibles y cuando se juzga que estos datos son representativos de un futuro desconocido. Estas técnicas se clasifican en dos categorías: estadísticas y determinísticas.

Las técnicas *estadísticas* se enfocan completamente en patrones, cambios en estos y perturbaciones causadas por influencias aleatorias. Éstas a su vez se basan, por un lado, en la suposición de que los datos se pueden descomponer en componentes como la tendencia, ciclo, estacionalidad e irregularidad, y por otro con la metodología de Box-Jenkins para modelos econométricos de series de tiempo. Los fundamentos teóricos se basan principalmente en conceptos estadísticos y no se supone que los datos estén separados en componentes separados. Las técnicas *determinísticas* (causales) comprenden la identificación y determinación de relaciones entre la variable por pronosticar y otras variables de influencia.

## 1.2. Definición y conceptos claves.

Una **Serie de Tiempo** es una colección de observaciones hechas secuencialmente en el tiempo<sup>(1)</sup>. Es común representar a las series de tiempo por medio de una ecuación matemática que describa los valores de la variable observada como una función del tiempo o equivalentemente como una curva en una gráfica en la que la coordenada vertical representa el valor de la variable y la coordenada horizontal representa el tiempo.

En la fig. 1.1 se muestra la gráfica de una serie de tiempo para el tipo de cambio del dólar en nuestro país de 1996 a 2002 <sup>(2)</sup>. La misma información podría ser presentada en forma tabular, pero el patrón de cambio no sería apreciado de la misma manera. Lo que normalmente se usa como instrumento de planeación es el patrón generado por la serie, más que los valores individuales observados. Por ejemplo, un inversionista desearía comprar dólares en tiempos donde se espere un decremento en el precio del dólar y venderlos cuando espere que el precio aumentará. Así, si el inversionista hubiera comprado el 1 de octubre de 1997, cada dólar le hubiera costado \$7.74. Siendo pasivo con su inversión, y decide vender el 1 de octubre de 1998 (es decir, un año después), hubiera tenido una ganancia de \$2.64 por dólar. Para poder proyectar a futuro las fluctuaciones del tipo de cambio, es necesario considerar las tendencias, los ciclos, y otros elementos del

---

<sup>1</sup> Ver referencia 3 de la bibliografía, pp. 1.

<sup>2</sup> Ver referencia 17 de la bibliografía.

patrón indicado por la serie de tiempo de la fig. 1.1.

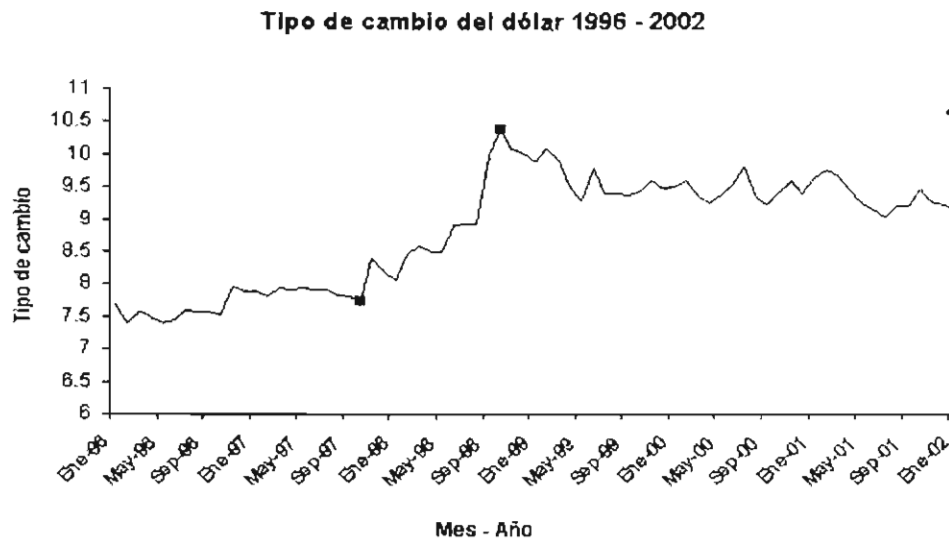


Fig. 1.1. Tipo de cambio del dólar respecto al peso mexicano de 1996 - 2002

La suposición básica que subyace en el **análisis de series de tiempo**, es que los factores que han ocasionado patrones de actividad en el pasado y en el presente continuarán haciéndolo, más o menos de la misma forma en el futuro. Por consiguiente, el principal objetivo del análisis de series de tiempo consiste en aislar dichos factores de influencia con propósitos de hacer predicciones (pronósticos), así como para efectuar una planeación y tratar de tomar un mayor control de la situación que en muchos ámbitos de la vida diaria se presentan.

Otra definición que tiende a ser muy parecida es la que indica que la serie de tiempo es un conjunto de mediciones de cierto fenómeno o experimentos registrados secuencialmente en el tiempo. Estas observaciones serán denotadas por  $\{Y(t_1), Y(t_2), \dots, Y(t_n)\} = \{Y(t) : t \in T \subset R\}$ , con  $Y(t_i)$  el valor de la variable  $x$  en el instante  $t_i$ <sup>(3)</sup>.

Se dice que una serie de tiempo observada es **continua**, cuando las observaciones se toman continuamente en el tiempo, y se denota, por ejemplo por  $Y(t)$ ,  $0 \leq t \leq T$ . Se dice que una serie de tiempo observada es **discreta**, cuando las observaciones se toman sólo en momentos específicos usualmente equiespaciados, y se denota, por ejemplo por  $Y(t)$ ,  $t = 0, 1, 2, 3, \dots, T$ .

<sup>3</sup> Otra definición de serie de tiempo adecuada, es la que proporciona G.S. Maddala., quien define una serie de tiempo como una secuencia de datos numéricos, cada uno de los cuales se asocia a un instante específico de tiempo. Es decir, observaciones de una misma variable en tiempo  $t_1, t_2, \dots, t_T$ .

Como ejemplo de series de tiempo discretas y equiespaciadas se pueden mencionar: precios de acciones en sucesivos días, totales de exportación en sucesivos meses, ventas semanales, mensuales o anuales de un comercio particular, lluvias diarias en cierta ciudad, temperatura del aire medida en sucesivas horas, días o meses, población de un país en forma anual, índice de precio al consumidor y el Producto Interno Bruto (PIB) trimestral o anual.

Las series de tiempo pueden originarse de varias formas:

a) Por muestreo de una serie de tiempo continua, es decir, dada una serie de tiempo continua, se pueden tomar valores de intervalos iguales de tiempo.

b) Por acumulación de una serie de tiempo, ya sea continua o discreta, sobre un período de tiempo dado.

Si los valores futuros de una serie de tiempo pueden ser determinados exactamente por una **función matemática** como por ejemplo:  $Y(t) = \text{sen}(2\pi t)$ , se dice que la serie de tiempo es **determinística**; si los valores futuros pueden ser descritos sólo en términos de una distribución de probabilidad es **no-determinística**. La mayoría de las series de tiempo son **estocásticas**, es decir, que el futuro sólo es determinado en parte por los valores anteriores; en este caso las predicciones exactas son imposibles, pues debe tomarse en cuenta que los valores futuros tienen una distribución de probabilidad condicionada al conocimiento de los valores pasados.

### 1.2.1. Componentes de una serie de tiempo.

Los estadísticos frecuentemente piensan en una serie de tiempo como la adición o producto (o mixtos) de cuatro componentes importantes de la serie:

- 1- Tendencia a largo plazo
- 2- Efecto cíclico
- 3- Efecto estacional
- 4- Variación aleatoria

**Las tendencias a largo plazo ( $T$ )**, se presentan en una serie de tiempo debido a un crecimiento en los valores de variables importantes, como la población, el Producto Nacional Bruto, el efecto de la competencia y otros factores que no llegan a producir cambios violentos en la variable observada, pero que producen un cambio gradual y estable sobre el tiempo. Una serie de tiempo con una tendencia creciente a largo plazo es similar a

la fig. 1.2, donde se muestra el aumento del Índice de Precios y Cotizaciones en el tiempo, de 1993 a 2001 <sup>(4)</sup>. Es ascendente en general, porque hay ciertos períodos en donde se observa que desciende ligeramente.

Al referirse a la tendencia es importante definir antes lo que se entiende por largo plazo, pues se puede confundir con el ciclo que la serie traiga, es decir, se puede concluir que la serie tiene una tendencia ascendente y creer que así seguirá por tiempo indefinido y en realidad es solamente es parte de un ciclo que no es posible visualizar por no tener datos suficientes. Por ejemplo, variables climáticas algunas veces exhiben variaciones cíclicas sobre un período de tiempo largo, tal como 50 años. Si sólo se tienen 20 años de datos, esta oscilación a largo plazo podría parecer parte de una tendencia, pero si se dispusiera de cientos de años, la oscilación a largo plazo sería visible. Sin embargo, en un período corto podría ser más útil pensar en una oscilación a largo plazo como una tendencia. Así, al hablar de tendencia se debe tener en cuenta el número de observaciones disponibles y hacer una evaluación subjetiva del término largo plazo.



Fig. 1.2. Movimiento del Índice de Precios y Cotizaciones de 1993 a 2001

**Los efectos cíclicos (C).** en una serie de tiempo aparecen cuando la serie sube y baja suavemente, a manera de ondas siguiendo la curva de la tendencia a largo plazo, tal como sucede con la variación de los precios del Henry Hub (petróleo) entre 1986 y 1994 <sup>(5)</sup>, representado en la fig. 1.3. Esta gráfica, como se observa, no parece seguir una tendencia creciente o decreciente sobre el tiempo; sino que parece fluctuar dadas diversas condiciones, tales como los conflictos entre los países asiáticos, las crisis financieras en países de Latinoamérica, la inflación en muchos países, principalmente en los mayores exportadores de petróleo, los cambios de gobernantes, etc., lo que provoca oscilaciones en los precios largos. Otro ejemplo serían los ciclos económicos en nuestro país, que se

<sup>4</sup> Ver referencia 18 de la bibliografía.

<sup>5</sup> Ver referencia 19 de la bibliografía.



presentan cada seis años, al término del mandato de un presidente.

En resumen, los efectos cíclicos en una serie de tiempo suelen ser causados por cambios en la demanda, en este caso del petróleo Henry Hub, tal como se muestra en la fig 1.3.

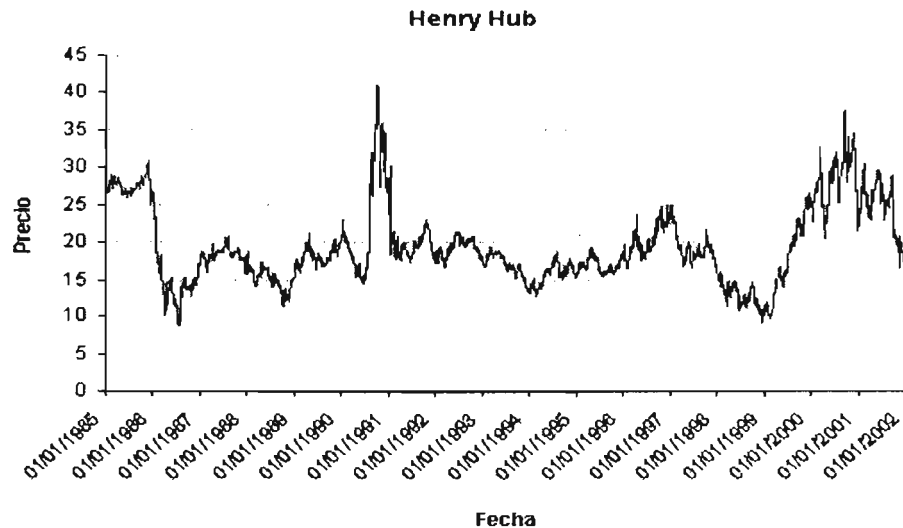


Fig. 1.3. Movimiento de precios del Henry Hub de 1985 - 2002

**Los efectos estacionales ( $S$ )**, son aquellas altas y bajas que ocurren en un tiempo particular del año. Por ejemplo, la cantidad de pasajeros en las líneas aéreas internacionales, durante el período comprendido entre 1949 y 1961 <sup>(6)</sup>, representados en la fig. 1.4. Son totales mensuales, representados en miles de pasajeros. Se puede observar que esta serie tiene una marcada tendencia ascendente, además de que existen oscilaciones dentro de los años, que si bien, tienen la misma forma, la amplitud de las mismas se hace más grande a través de los años. Esto lleva a pensar en que la serie tiene una componente estacional.

Hay que tomar en cuenta que con este gráfico no se puede hacer comparaciones de la cantidad de pasajeros en pares de meses arbitrarios, por ejemplo, en marzo de 1960 con julio del mismo año y concluir que hay mayor cantidad de personas que viajan en julio que en marzo, ya que este aumento se debe precisamente al efecto estacional y no a la realidad. Matemáticamente, se puede decir que la serie representa una variación estacional, si existe un número  $s$  tal que  $Y(t) = Y(t + ks)$ .

<sup>6</sup> Ver referencia 20 de la bibliografía..

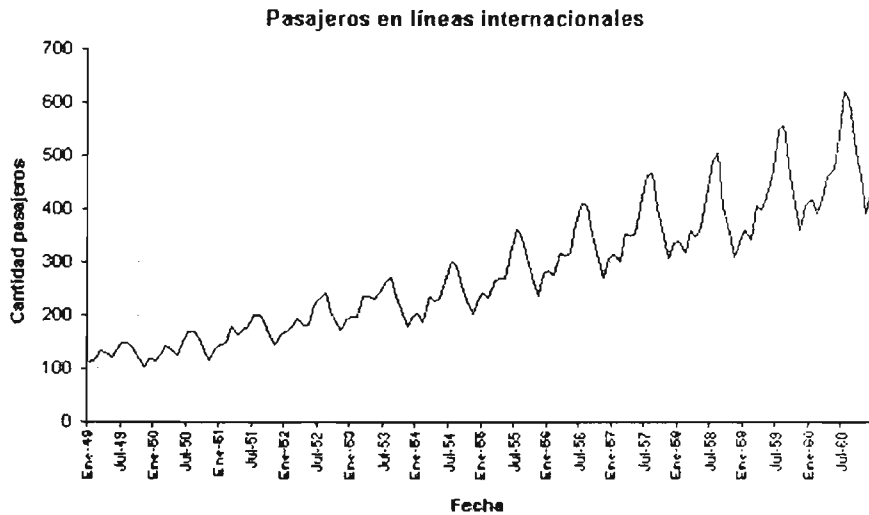


Fig. 1.4. Cantidad de pasajeros que viajaron en el período 1949 - 1960

La diferencia principal entre los efectos cíclicos y los estacionales es que los efectos estacionales pueden predecirse, y ocurren a un intervalo de tiempo fijo de la última ocurrencia, mientras que los efectos cíclicos son componentes impredecibles.

La cuarta componente es **la variación aleatoria ( $I$ )**, la cual representa los movimientos ascendentes y descendentes de la serie después de haber ajustado la tendencia a largo plazo, el efecto cíclico y el efecto estacional. La variación aleatoria son aquellos cambios inexplicables y sacudidas de la serie sobre un período corto de tiempo. Son los eventos políticos, el clima, así como las diversas acciones humanas los que tienden a causar cambios aleatorios e inesperados en una serie de tiempo. Todas las series de tiempo contienen una variación aleatoria. Además una serie de tiempo puede contener todas o ninguna de las otras tres componentes.

En la mayoría de las series de tiempo no resulta nada sencillo distinguir entre las componentes. A menudo la pareja de los efectos cíclico y estacional, o las tres componentes, tendencia a largo plazo, efectos cíclico y estacional, se han integrado tanto que resultan inseparables. Por el contrario si los efectos parecen distinguibles, no es difícil separarlos.

### 1.2.2. Objetivos del análisis de una serie de tiempo.

El primer objetivo en el análisis de series de tiempo es identificar aquellos componentes presentes para conocer sus causas y predecir valores futuros de la serie.

Además se puede clasificar a los objetivos que se persiguen de la siguiente manera:

- *Descripción*: Simplemente enunciar el comportamiento de un proceso.
- *Explicación*: Permitiendo un análisis de tipo causa-efecto al relacionar un fenómeno con el medio circundante.
- *Pronóstico*: En el caso que se desee estimar valores futuros con el fin de tomar decisiones o prever comportamientos; sin embargo, debe resaltarse que el futuro puede ser cierta clase de extensión del pasado, pero difícilmente puede esperarse que sea una réplica exacta; por lo que la predicción puede ser puntual o por intervalos.

### 1.3. Modelos de series de tiempo.

Una serie de tiempo se puede expresar como:

$$Y(t) = f(T_t, S_t, C_t, I_t) + u_t \quad t = 0, 1, 2, \dots, T$$

El término aleatorio  $u\{t\}$  se considera independiente, o al menos no correlacionado, con varianza finita positiva y constante. Surge ahora el problema de estimar cada uno de estos componentes, es decir, se desea estimar que parte del valor de la serie de tiempo  $X(t)$ , para cualquier tiempo dado  $t$ , puede atribuirse a la tendencia, a los factores estacionales, a los factores cíclicos o a una variación aleatoria o irregular. Para poder estimar cada componente, primero se debe hacer supuestos acerca de como están relacionados estos componentes. Estos supuestos conforman el llamado *modelo de la serie de tiempo*.

Los modelos de series de tiempo se clasifican en dos categorías, el modelo aditivo y el multiplicativo. El modelo aditivo se expresa como:

$$Y_t = T_t + C_t + S_t + I_t + u_t \quad t = 0, 1, 2, \dots, T$$

donde se supone que las componentes son independientes entre sí, es decir, por ejemplo, la tendencia no puede afectar a la variación estacional ni a la cíclica, y estos componentes no pueden afectar tampoco a la tendencia. Interacciones representadas por un término tal como  $S\{t\} \cdot T\{t\}$ , por ejemplo, no están contempladas en este modelo.

Para algunas series observadas este modelo aditivo puede resultar demasiado restrictivo, pero una transformación simple de la serie sigue al menos aproximadamente un esquema aditivo. El modelo multiplicativo se expresa como:

$$Y_t = T_t \cdot C_t \cdot S_t \cdot I_t \cdot u_t \quad t = 0, 1, 2, \dots, T$$

Aquí se supone que los cuatro componentes están relacionados entre sí, manteniendo la posibilidad de que las componentes provengan de causas básicas diferentes. Cuando los valores de la serie  $Y_t'$  son mayores que cero, se puede tomar un logaritmo en ambos miembros del modelo multiplicativo y así transformarlo en un modelo aditivo de la forma siguiente:

$$Y_t = \log Y_t' = \log T_t + \log C_t + \log S_t + \log I_t + \log u_t, \quad t = 1, 2, \dots, T$$

Existen otros modelos, que son combinaciones de elementos aditivos y multiplicativos, llamados modelos mixtos, como por ejemplo:

$$Y_t = T_t \cdot C_t + S_t + I_t + u_t, \quad t = 1, 2, \dots, T$$

Una vez que las componentes han sido estimadas, es posible hacer predicciones del valor de la serie en algún tiempo futuro, considerando primero el valor de la componente tendencia en ese punto y modificando su valor mediante un ajuste que tenga en cuenta las componentes estacional y cíclica. Para estudiar los componentes de una serie histórica, el analista debe considerar sus relaciones matemáticas. El enfoque que se utiliza con mayor frecuencia es el de tratar a los datos originales de la serie como un producto de sus componentes; es decir, *una serie anual* es un producto de tendencia y fluctuaciones cíclicas (por definición), y cuando la serie es medida en períodos menores a un año, como las series mensuales y trimestrales, se considera que cada valor original es el producto multiplicativo de las cuatro componentes.

## 1.4. Exploración de patrones de datos con análisis de autocorrelación.

Cuando se mide una variable a través del tiempo, con frecuencia está correlacionada consigo misma cuando se desfasa uno o más períodos. Esta correlación se mide mediante el coeficiente de autocorrelación. Comúnmente se usa la ecuación (1.1) de la página siguiente, para calcular el coeficiente de autocorrelación de orden  $k$  ( $r_k$ ), entre observaciones separadas por  $k$  períodos  $Y_t$  y  $Y_{t-k}$ .

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (1.1)$$

en donde:

$r_k$  = coeficiente de autocorrelación para un desfase de  $k$  períodos

$\bar{Y}$  = media de los valores de la serie

$Y_t$  = observación en el período  $t$

$Y_{t-k}$  = observación en el período  $t - k$

El *correlograma* es una útil herramienta gráfica para exhibir autocorrelaciones para varios desfases de una serie de tiempo. La manera de entender un correlograma varía dependiendo del autor, para efectos de este trabajo se leerá de la manera siguiente: la parte izquierda muestra gráficamente la forma en que se comporta la función de autocorrelación simple y parcial; esta última se explicará más adelante. Las columnas AC y PAC muestran los coeficientes de autocorrelación simple y parcial entre  $Y_t$  y  $Y_{t-k}$ , respectivamente, para el número correspondiente de períodos defasados que aparecen a su izquierda. Además se muestra el estadístico *Q-Stat* que prueba si los AC en conjunto son significativamente distintos de cero o no, lo cual se explicará en el capítulo 2 y su valor-p o significancia asociada se muestra delante de éste. Las bandas que aparecen sobre las barras que representan a los AC y PAC, son bandas de confianza que determinan si individualmente son distintos a cero o no y están dadas según el tamaño de la muestra, ya que cortan en el punto  $\frac{2}{\sqrt{n}}$ .

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.518	-0.518	27.612	0.000
		2	0.284	0.022	36.026	0.000
		3	-0.140	0.021	38.094	0.000
		4	0.030	-0.054	38.192	0.000
		5	-0.032	-0.052	38.300	0.000
		6	-0.030	-0.071	38.397	0.000
		7	-0.104	-0.194	39.575	0.000
		8	0.059	-0.081	39.965	0.000
		9	0.064	0.139	40.420	0.000
		10	-0.082	-0.014	41.183	0.000
		11	0.095	-0.005	42.212	0.000
		12	-0.036	0.033	42.358	0.000
		13	0.057	0.039	42.739	0.000
		14	-0.034	-0.011	42.876	0.000
		15	0.064	0.096	43.365	0.000

Fig. 1.5. Ejemplo de un correlograma

Con esta gráfica se pueden estudiar los patrones de datos, incluyendo tendencia, estacionalidad e irregularidad, mediante el enfoque del análisis de autocorrelación. Los coeficientes de autocorrelación para diferentes períodos desfasados de una variable se pueden utilizar en una serie de tiempo de datos para identificar:

- Si una serie es aleatoria.
- Si los datos tienen una tendencia.
- Si los datos son estacionarios.
- Si los datos son estacionales.

Una serie es aleatoria si la correlación entre  $Y_t$  y  $Y_{t-1}$  es cercana a cero y los valores sucesivos de la serie de tiempo no guardan relación entre sí. Una serie tiene tendencia si  $Y_t$  y  $Y_{t-1}$  están altamente correlacionados y es típico que los coeficientes de autocorrelación sean diferentes de cero de manera significativa para varios de los primeros períodos de desfase y caigan gradualmente hacia cero al incrementarse el número de períodos. El coeficiente de autocorrelación para el período de desfase 1, es por lo regular cercano a 1. El coeficiente de autocorrelación para el período de desfase 2 será también grande, pero no tanto como para el período 1. Y así para cada uno de los demás períodos. Una serie tiene un patrón estacional, si se presenta un coeficiente de autocorrelación significativo en el período de desfase correspondiente: cuatro en los datos trimestrales o doce en los mensuales.

Para saber si un coeficiente de autocorrelación es significativamente diferente de cero, y saber con esto que la serie es aleatoria, se podría desarrollar una distribución muestral de coeficientes de autocorrelación tomando un número infinito de muestras de números aleatorios. Quenouille<sup>(7)</sup> y otros han demostrado que los coeficientes de autocorrelación de datos aleatorios tienen una distribución que se puede aproximar a una curva normal con una media de cero y una desviación estándar de  $\frac{1}{\sqrt{n}}$ . Conociendo lo anterior, se pueden comparar los coeficientes de autocorrelación de la muestra con esta distribución teórica de muestreo y determinar si los datos provienen de una población cuya media sea cero en períodos de desfase en particular.

De hecho si la serie es en efecto aleatoria, la mayoría de los coeficientes de autocorrelación deben ubicarse dentro de un nivel especificado por 0, más o menos un cierto número de errores estándar. A un nivel específico de confianza  $\alpha$ , se puede considerar aleatoria una muestra si los coeficientes de autocorrelación calculados se encuentran todos dentro del intervalo  $0 \pm Z_{\alpha/2} \frac{1}{\sqrt{n}}$

Algunos modelos de pronósticos avanzados, como los modelos de Box-Jenkins están diseñados para utilizarse como series de tiempo estacionarias. Una *serie estacionaria* es aquella cuyas propiedades estadísticas básicas, como la media y la varianza, permanecen

---

<sup>7</sup> Ver referencia 5 de la bibliografía., pp. 561-571..

constantes en el tiempo. Se dice que una serie que no presenta crecimiento o declinación es estacionaria. Una serie que tiene una tendencia, se dice que es no estacionaria.

Los coeficientes de autocorrelación de datos estacionarios caen a cero después del segundo o tercer período de desfase, mientras que las series no estacionarias son significativamente distintas de cero durante varios períodos. En estas series se debe quitar la tendencia antes de realizar cualquier análisis posterior, como su uso en los procedimientos Box-Jenkins. Lo anterior será detallado en el siguiente capítulo.

## 1.5. Métodos de suavizamiento.

Cuando se analizan datos en donde los movimientos de la tendencia en la serie se ven oscurecidos por variaciones de un año a otro (ciclos), y no es fácil darse cuenta de si realmente existe en la serie algún efecto de la tendencia hacia arriba o hacia abajo; se pueden utilizar diversos métodos de suavizamiento para eliminar estas variaciones y en consecuencia ver cuál es el patrón de movimiento de los datos a través del tiempo. Estos son los métodos (ó técnicas) de suavizamiento.

### 1.5.1. Promedios simples.

El objetivo consiste en usar datos anteriores para desarrollar un modelo de pronóstico para períodos futuros. Un *promedio simple* se obtiene encontrando la media de todos los valores pertinentes y usando después esta media para pronosticar el siguiente período.

$$Y_{t+1} = \sum_{i=1}^n \frac{Y_i}{n}$$

Este método debería emplearse cuando los datos son estacionarios: sin tendencia, estacionalidad u otros patrones sistemáticos.

### 1.5.2. Promedios móviles.

Al estar disponible cada nueva observación, se puede calcular una nueva media eliminando el valor más antiguo e incluyendo el más reciente. Entonces, se usa este promedio móvil para pronosticar el siguiente período. Así, pueden promediarse las ventas

mensuales de una compañía sobre un período de cuatro meses y graficar el promedio en el punto medio del intervalo de cuatro meses.

El segundo punto de la serie se obtendría añadiendo al total de los cuatro meses las ventas del siguiente mes de la serie, desechando las ventas correspondientes al primer mes de los cuatro previos, y calculando un nuevo promedio de cuatro meses. Entonces la serie de tiempo de los promedios móviles presenta un punto por cada mes que el promedio de las observaciones calculado para un período específico de tiempo antes y después del mes en cuestión.

El efecto neto es transformar la serie original en la serie de promedios móviles que resulta más suave (menos sujeta a oscilaciones rápidas) y más susceptible de revelar las siguientes tendencias o ciclos en el patrón de la serie sobre el tiempo. Al tiempo  $t$  el promedio móvil  $\bar{y}_t$  de las observaciones de la serie sobre  $M$  períodos de tiempo se encuentra calculando:

$$y_t = \frac{y_{t-(M-1)/2} + y_{t+1-(M-1)/2} + y_{t+2-(M-1)/2} + \dots + y_{t+(M-1)/2}}{M}$$

en donde  $M$  es un número impar y  $y_t$  es la observación del proceso en el tiempo  $t$ ,  $y_{t-1}$  es la observación del proceso al tiempo  $t-1$ , y así sucesivamente. La fórmula de los promedios móviles puede ser simplificada al escribirla de la siguiente manera (forma recursiva):

$$\bar{y}_t = \bar{y}_{t-1} + \frac{\text{siguiente observación} - \text{observación más remota}}{M}$$

puesto que lo que se hace a cada paso es recalcular el promedio, añadiendo la siguiente observación y eliminando a la observación que ocurrió  $M$  períodos en el pasado. Si  $M$  es un número par, los promedios móviles ocurrirán *entre* los puntos de tiempo en lugar de *en* los puntos del tiempo. Entonces resulta constructivo calcular los promedios móviles sobre un número impar de períodos para tener así punto de comparación con los valores originales.

El modelo de promedio móvil funciona mejor con datos estacionarios<sup>(8)</sup>. No maneja muy bien la tendencia o la estacionalidad, aunque lo hace mejor que el método de promedio simple. La principal desventaja al usar un promedio móvil para suavizar la serie es que, al menos en  $M = 1$ , no se tiene un valor suavizado correspondiente a cada valor de la serie original.

Por ejemplo, si se calculan los promedios móviles sobre cada  $M = 5$  períodos



consecutivos de la serie de tiempo, no se tendrá un valor suavizado para los primeros dos y los últimos dos valores. Cuando se cuenta con un número grande de observaciones ésta no resulta una consideración tan seria como cuando se cuenta con un número reducido de observaciones.

Entre mayor sea el orden del promedio móvil, mayor será el efecto de atenuación. Empleado como un pronóstico, un promedio móvil grande presta poca atención a las fluctuaciones en la serie de datos, es decir, al ser más grande el período de tiempo para el cálculo, se dará menor peso a los períodos más recientes. Un número pequeño asigna mayor peso a la historia reciente, lo que permite que el pronóstico se ajuste con mayor rapidez al nivel actual.

### 1.5.3. Métodos de atenuación exponencial.

La atenuación exponencial es un método utilizado para revisar constantemente una estimación a la luz de experiencias más recientes. El método está basado en el promedio de valores anteriores de una serie, haciendo esto de forma decreciente. Las observaciones se ponderan, asignando mayor peso a las más recientes. Las ponderaciones empleadas se designan como  $\alpha$  para la observación más reciente,  $\alpha(1 - \alpha)$  para la siguiente más reciente,  $\alpha(1 - \alpha)^2$  para la siguiente y así sucesivamente. Si se supone que  $0 < \alpha < 1$ , de esta forma:

$$\begin{aligned}\hat{Y}_{t+1} &= \alpha Y_t + (1 - \alpha)\hat{Y}_t \\ &= \alpha Y_t + \hat{Y}_t - \alpha \hat{Y}_t \\ &= \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)\end{aligned}$$

La atenuación exponencial es simplemente el pronóstico anterior ( $\hat{Y}_t$ ) más  $\alpha$  veces el error de  $(Y_t - \hat{Y}_t)$ , en el pronóstico anterior. La constante de atenuación  $\alpha$  sirve como el factor para ponderar.

El valor real de  $\alpha$  determina el grado hasta el cual la observación más reciente puede influir en el valor del pronóstico. Cuando  $\alpha$  es cercana a 1, el nuevo pronóstico incluirá un ajuste sustancial de cualquier error ocurrido en el pronóstico anterior. Inversamente, cuando  $\alpha$  está cercana a 0, el pronóstico es similar al anterior. Si se desea que los pronósticos sean estables y se atenúen las variaciones aleatorias, se requiere de un valor de  $\alpha$  cercano a 0.

Un método para estimar  $\alpha$  consiste en un procedimiento iterativo que minimiza el error medio cuadrático (ECM). Se calculan los pronósticos para  $\alpha$  igual a 0.1, 0.2, 0.3, ..... , 0.9 y

para cada uno se calcula la suma de cuadrados del pronóstico. Para generar pronósticos futuros, se elige el valor de  $\alpha$  que produce el error más pequeño.

En la técnica de atenuación exponencial simple existe la suposición de que los datos son estacionarios. Siempre que exista una tendencia significativa, la atenuación exponencial se ubicará por debajo de los valores reales de pronósticos de la serie de tiempo.

Otras técnicas de suavizamiento que se utilizan con frecuencia para manejar una tendencia lineal son:

- Atenuación exponencial ajustada a la tendencia: *Método de Holt*. Esta técnica atenúa en forma directa la tendencia y la pendiente empleando diferentes constantes de atenuación para cada una de ellas.

- Atenuación exponencial ajustada a la tendencia y a la variación estacional: *Modelo de Winter*. Este modelo es una extensión del modelo de Holt y reduce en gran medida el error de pronóstico. Se utiliza una ecuación adicional para determinar la estacionalidad, la cual va eliminando los efectos estacionales que pudieran existir en el dato original  $Y$ .

## 1.6. Descomposición de las componentes de una serie de tiempo.

El primer paso para un análisis clásico de Series de Tiempo comienza analizando la gráfica de la serie a estudiar; esto permite detectar las componentes esenciales de la serie. Asimismo, el gráfico permitirá detectar *outliers*, es decir, aquellos puntos de la serie que se escapan de lo normal. Un outlier es una observación de la serie que corresponde a un comportamiento anormal del fenómeno (sin incidencias futuras) o a un error de medición.

Un enfoque al análisis de series de tiempo, comprende un intento de identificar los factores que ejercen influencia sobre cada uno de los valores periódicos de una serie. Cada componente: tendencia, variación cíclica, variación estacional, fluctuación irregular, se identifica por separado de tal manera que la serie histórica pueda proyectarse al futuro y utilizarse en pronósticos tanto de corto como largo plazo.

Para un mejor entendimiento de la manera en que se descompone una serie, se utilizará la serie que representa las ventas al mayoreo de vehículos automóviles en el interior de la República Mexicana, desde enero de 1992 a diciembre de 2002 medidos mensualmente <sup>(9)</sup>,

---

<sup>9</sup> Ver referencia 20 de la bibliografía..

datos que se presentan en la tabla 1.1. y así emplear un modelo multiplicativo de series de tiempo de cuatro componentes. De esta forma, el supuesto importante que inicialmente se usa es el que las componentes mantienen una dependencia entre ellas.

Como primer paso se grafican los datos de la serie, y se presentan en la fig. 1.6., en donde se observa que en determinados períodos, se refleja cierta tendencia negativa y el resto del tiempo toma una tendencia positiva. También se pueden ver rasgos de ciclicidad, los cuales se determinarán más adelante.

Para poder observar mejor los efectos de la tendencia, además de que su cálculo será indispensable en la obtención de las demás componentes, se aplica un método de suavizamiento a los datos, el *promedio móvil*, y sus resultados se presentan en la tabla 1.1, junto con los datos originales de la serie en cuestión.

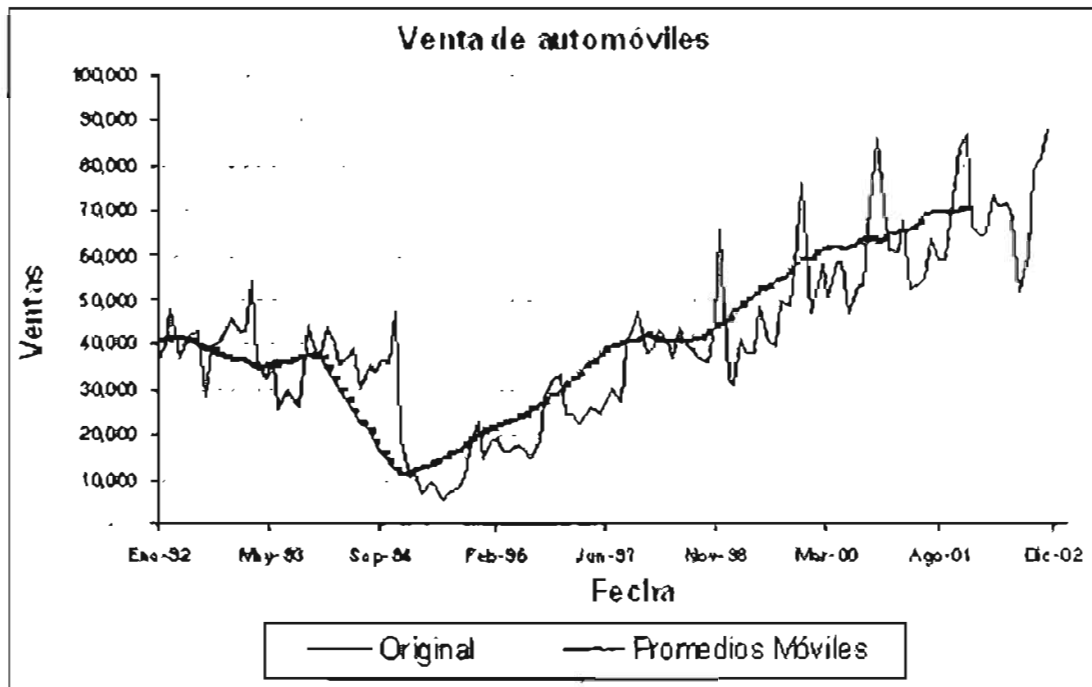


Fig. 1.6. Venta de automóviles de enero-1992 a diciembre-2002 y su promedio móvil

Con el promedio móvil, lo que se logra es suavizar a la serie, lo cual es claro en la gráfica, además de que se observa mejor la tendencia que sigue. Antes de 1994 la tendencia era descendente y después de esa fecha se vuelve positiva, siendo claro la eliminación de los "picos" que la serie original traía consigo.

### 1.6.1. Tendencia.

Antes de medir la tendencia de una serie histórica, se debe primero conocer el propósito de su medición. Este conocimiento servirá en la elección del método y dimensión de la serie histórica a utilizar en la medición. Existen dos propósitos fundamentales: proyectar la tendencia y eliminarla de los datos originales. En el análisis de tendencia, la variable independiente es el tiempo.

El método más ampliamente usado en el análisis de tendencia de línea recta es el método de mínimos cuadrados. Este enfoque calcula la línea que mejor se ajusta a un grupo de puntos; la ecuación de tendencia es:

$$\hat{Y} = b_0 + b_1X$$

donde,  $\hat{Y}$  es el valor de predicción de tendencia de la variable  $Y$ ;  $b_0$  es el valor de la tendencia cuando  $X = 0$  (intercepción);  $b_1$  es el incremento o disminución promedio en  $\hat{Y}$ , para cada incremento en un período de  $X$ , que corresponde a la pendiente de la recta de regresión. Los estimadores de  $b_0$  y  $b_1$  están dados por,

$$b_1 = \frac{\sum_{i=1}^n x_i y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^n x_i^2 - n\bar{X}^2} \quad \text{y} \quad b_0 = \bar{Y} - b_1\bar{X}$$

En la tabla 1.1, se muestran los datos mensuales de los venta de automóviles de enero-1992 a diciembre-2002 y la gráfica de los datos se observa en la fig. 1.6.

Tabla 1.1. Venta al mayoreo de vehículos en el interior de la República Mexicana

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)
01-01-92	1	36233		
01-02-92	2	40215		
01-03-92	3	48440		
01-04-92	4	37447		
01-05-92	5	38962		
01-06-92	6	42266	40315.00	
01-07-92	7	42683	40827.58	40571.29
01-08-92	8	28130	41056.67	40942.13
01-09-92	9	39496	41564.92	41310.79
01-10-92	10	40730	41373.50	41469.21
01-11-92	11	43433	40830.58	41102.04
01-12-92	12	45745	40353.00	40591.79
01-01-93	1	42384	38954.83	39653.92
01-02-93	2	42964	39098.00	39026.42
01-03-93	3	54539	38119.50	38608.75
01-04-93	4	35150	36885.83	37502.67
01-05-93	5	32447	37000.42	36943.13
01-06-93	6	36535	36513.92	36757.17
01-07-93	7	25905	36138.00	36325.96
01-08-93	8	29848	36209.50	36173.75
01-09-93	9	27754	34998.92	35604.21
01-10-93	10	25926	35020.08	35009.50
01-11-93	11	44808	35473.92	35247.00
01-12-93	12	39907	35677.42	35575.67
01-01-94	1	37873	36043.75	35860.58
01-02-94	2	43822	36489.17	36266.46
01-03-94	3	40012	37015.50	36752.33
01-04-94	4	35404	37890.08	37452.79
01-05-94	5	37893	37146.42	37518.25
01-06-94	6	38977	37760.25	37453.33
01-07-94	7	30301	36217.67	36988.96
01-08-94	8	35193	33423.00	34820.33
01-09-94	9	34070	31074.00	32248.50
01-10-94	10	36421	28668.33	29871.17
01-11-94	11	36884	26325.17	27496.75
01-12-94	12	47273	23772.00	25048.58
01-01-95	1	19362	21720.92	22746.46
01-02-95	2	10286	19342.50	20531.71
01-03-95	3	11824	17147.83	18245.17
01-04-95	4	6536	14977.25	16062.54
01-05-95	5	9775	13356.42	14166.83
01-06-95	6	8339	11343.25	12349.83
01-07-95	7	5688	10959.17	11151.21

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)
01-08-95	8	6652	11607.58	11283.36
01-09-95	9	7734	12235.83	11921.71
01-10-95	10	10374	13085.00	12660.42
01-11-95	11	16434	13627.50	13356.25
01-12-95	12	23115	14389.08	14008.29
01-01-96	1	14753	15267.83	14828.46
01-02-96	2	18067	15924.75	15596.29
01-03-96	3	19363	16801.75	16363.25
01-04-96	4	16726	16257.42	17529.58
01-05-96	5	16285	19516.00	18886.71
01-06-96	6	17478	20372.17	19944.08
01-07-96	7	16233	21197.08	20784.63
01-08-96	8	14535	21709.83	21453.46
01-09-96	9	18258	21972.75	21841.29
01-10-96	10	27842	22598.67	22285.71
01-11-96	11	31537	23449.67	23024.17
01-12-96	12	33389	24011.92	23730.79
01-01-97	1	24652	24889.83	24450.88
01-02-97	2	24220	26238.67	25564.25
01-03-97	3	22518	26997.25	26617.96
01-04-97	4	24237	28025.75	27511.50
01-05-97	5	26497	29079.33	28552.54
01-06-97	6	24225	30275.50	29677.42
01-07-97	7	26768	31441.67	30858.58
01-08-97	8	30721	32708.42	32075.04
01-09-97	9	27361	34455.25	33681.83
01-10-97	10	40184	35819.42	35137.33
01-11-97	11	44180	36687.17	36253.29
01-12-97	12	47743	38316.17	37501.67
01-01-98	1	38646	39413.25	38864.71
01-02-98	2	39421	40088.33	39750.79
01-03-98	3	43480	40907.33	40497.83
01-04-98	4	40607	40570.67	40739.00
01-05-98	5	36910	40650.00	40610.33
01-06-98	6	43773	42141.00	41395.50
01-07-98	7	39933	41585.17	41863.08
01-08-98	8	38822	40891.17	41238.17
01-09-98	9	37189	40730.25	40810.71
01-10-98	10	36144	40523.75	40627.00
01-11-98	11	45132	40630.50	40577.13
01-12-98	12	65635	41043.83	40837.17
01-01-99	1	31976	41144.25	41094.04
01-02-99	2	31093	41225.58	41184.92
01-03-99	3	41549	42236.17	41730.88

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)
01-04-99	4	38129	43302.83	42769.50
01-05-99	5	38191	44340.83	43821.83
01-06-99	6	48733	45212.50	44776.67
01-07-99	7	41138	46449.08	45630.79
01-08-99	8	39798	48110.83	47279.96
01-09-99	9	49316	49456.08	48763.46
01-10-99	10	48944	50481.67	49968.88
01-11-99	11	57588	52156.17	51318.92
01-12-99	12	76095	52977.00	52566.58
01-01-00	1	46815	53485.25	53231.13
01-02-00	2	51034	54588.67	54036.96
01-03-00	3	57692	54877.00	54732.83
01-04-00	4	50436	56715.50	55796.25
01-05-00	5	58285	59102.17	57908.83
01-06-00	6	58583	58681.17	58891.67
01-07-00	7	47237	59872.25	59276.71
01-08-00	8	53039	60667.83	60270.04
01-09-00	9	52776	61504.00	61085.92
01-10-00	10	71006	61690.67	61597.33
01-11-00	11	86228	61323.42	61507.04
01-12-00	12	71043	61009.75	61166.58
01-01-01	1	61108	62354.08	61681.92
01-02-01	2	60581	62871.92	62613.00
01-03-01	3	67726	63383.92	63127.92
01-04-01	4	52676	63167.00	63275.46
01-05-01	5	53878	62996.83	63081.92
01-06-01	6	54819	64331.00	63663.92
01-07-01	7	63369	64798.17	64564.58
01-08-01	8	59253	65092.08	64945.13
01-09-01	9	58920	64893.42	64992.75
01-10-01	10	68403	66650.92	65772.17
01-11-01	11	84186	68065.67	67358.29
01-12-01	12	87053	69441.33	68753.50
01-01-02	1	66714	69850.17	69645.75
01-02-02	2	64108	69237.92	69544.04
01-03-02	3	65342	69284.42	69261.17
01-04-02	4	73766	70210.25	69747.33
01-05-02	5	70855	69976.42	70093.33
01-06-02	6	71327	70102.25	70039.33
01-07-02	7	68275		
01-08-02	8	51906		
01-09-02	9	59478		
01-10-02	10	79513		
01-11-02	11	81380		
01-12-02	12	88563		

Se genera la regresión entre el tiempo (variable independiente) y la venta de automóviles (variable dependiente) y según se observa en la tabla 1.2, el coeficiente de determinación es del 0.509, lo que indica que la variación en las ventas es explicada en un 50% por el modelo.

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0.71400766
Coefficiente de determinación R <sup>2</sup>	0.50980693
R <sup>2</sup> ajustado	0.50565276
Error típico	11547.7135
Observaciones	120

Tabla 1.2. Resultados de la regresión de las ventas de automóviles.

Así también, con base en la regresión lineal se obtienen los siguientes regresores con los cuales se puede hacer una estimación de la primer componente de tendencia:

$$\beta_0 = 19,681.71 \quad , \quad \beta_1 = 337.12$$

### 1.6.2. Variación cíclica.

El análisis de la tendencia de la variable dependiente tiene un valor práctico directo para los pronósticos a largo plazo. Sin embargo, el análisis del componente cíclico es de un valor dudoso a largo plazo y por tanto del pronóstico. Como se había mencionado anteriormente, el componente cíclico (*C*) es la fluctuación en forma de onda o ciclo mucho mayor duración, debida a condiciones económicas cambiantes. El componente cíclico de las series de tiempo se identifica eliminando o promediando los efectos de la tendencia (*T*); ya que este componente constituye lo que queda después de dichos ajustes, se le refiere como el método residual.

En forma simbólica, la descomposición de la componente ciclo de una serie (*Y*) se representa como:

$$C = \frac{Y}{T} = \frac{TC}{T}$$

En la determinación del efecto relativo del componente en cada valor, se acepta al valor de  $\hat{Y}$  como una indicación precisa de la tendencia y se trata a la discrepancia (residuo) como el componente cíclico. Los resultados de la estimación de la tendencia y el



componente cíclico se muestran en la tabla 1.3.

Tabla 1.3. Cálculo de las componentes tendencia y cíclico.

X	TxC (Y)	T	C
1	40571.29	20018.84	2.0267
2	40942.13	20355.96	2.0113
3	41310.79	20693.09	1.9964
4	41469.21	21030.21	1.9719
5	41102.04	21367.34	1.9236
6	40591.79	21704.46	1.8702
7	39653.92	22041.58	1.7991
8	39026.42	22378.71	1.7439
9	38608.75	22715.83	1.6996
10	37502.67	23052.96	1.6268
11	36943.13	23390.08	1.5794
12	36757.17	23727.21	1.5492
13	36325.96	24064.33	1.5095
14	36173.75	24401.46	1.4824
15	35604.21	24738.58	1.4392
16	35009.50	25075.70	1.3962
17	35247.00	25412.83	1.3870
18	35575.67	25749.95	1.3816
19	35860.58	26087.08	1.3746
20	36266.46	26424.20	1.3725
21	36752.33	26761.33	1.3733
22	37452.79	27098.45	1.3821
23	37518.25	27435.57	1.3675
24	37453.33	27772.70	1.3486
25	36968.96	28109.82	1.3159
26	34820.33	28446.95	1.2240
27	32248.50	28784.07	1.1204
28	29871.17	29121.20	1.0258
29	27496.75	29458.32	0.9334
30	25048.58	29796.45	0.8407
31	22746.46	30132.57	0.7549
32	20531.71	30469.69	0.6738
33	18245.17	30806.82	0.5922
34	16062.54	31143.94	0.5158
35	14166.83	31481.07	0.4500
36	12349.83	31818.19	0.3861
37	11151.21	32155.32	0.3468
38	11283.38	32492.44	0.3473
39	11921.71	32829.56	0.3631
40	12660.42	33166.69	0.3817
41	13356.25	33503.81	0.3986
42	14008.29	33840.94	0.4139
43	14828.46	34178.06	0.4339

X	TxC (M)	T	C
44	15596.29	34515.19	0.4519
45	16363.25	34852.31	0.4695
46	17529.58	35189.43	0.4981
47	18886.71	35526.56	0.5316
48	19944.08	35863.68	0.5661
49	20784.63	36200.81	0.5741
50	21453.46	36537.93	0.5872
51	21841.29	36875.06	0.5923
52	22285.71	37212.18	0.5989
53	23024.17	37549.31	0.6132
54	23730.79	37886.43	0.6264
55	24450.88	38223.55	0.6397
56	25564.25	38560.68	0.6630
57	26617.96	38897.80	0.6843
58	27511.50	39234.93	0.7012
59	28552.54	39572.05	0.7215
60	29677.42	39909.18	0.7436
61	30858.58	40246.30	0.7667
62	32075.04	40583.42	0.7903
63	33581.83	40920.55	0.8207
64	35137.33	41257.67	0.8517
65	36253.29	41594.80	0.8716
66	37501.67	41931.92	0.8943
67	38864.71	42269.05	0.9195
68	39750.79	42606.17	0.9330
69	40497.83	42943.29	0.9431
70	40739.00	43280.42	0.9413
71	40610.33	43617.54	0.9311
72	41395.50	43954.67	0.9418
73	41863.08	44291.79	0.9452
74	41238.17	44628.92	0.9240
75	40810.71	44966.04	0.9076
76	40627.00	45303.17	0.8968
77	40577.13	45640.29	0.8891
78	40837.17	45977.41	0.8882
79	41094.04	46314.54	0.8873
80	41184.92	46651.66	0.8828
81	41730.88	46988.79	0.8881
82	42769.50	47325.91	0.9037
83	43821.83	47663.04	0.9194
84	44776.67	48000.16	0.9328
85	45830.79	48337.28	0.9481
86	47279.96	48674.41	0.9714
87	48783.46	49011.53	0.9953

X	TxC (Y)	T	C
88	49968.88	49348.66	1.0126
89	51318.92	49685.78	1.0329
90	52566.58	50022.91	1.0509
91	53231.13	50360.03	1.0570
92	54036.96	50697.15	1.0659
93	54732.83	51034.28	1.0725
94	55796.25	51371.40	1.0861
95	57908.83	51708.53	1.1199
96	58891.67	52045.65	1.1315
97	59276.71	52382.78	1.1316
98	60270.04	52719.90	1.1432
99	61085.92	53057.03	1.1513
100	61597.33	53394.15	1.1536
101	61507.04	53731.27	1.1447
102	61166.58	54068.40	1.1313
103	61681.92	54405.52	1.1337
104	62613.00	54742.65	1.1438
105	63127.92	55079.77	1.1461
106	63275.46	55416.90	1.1418
107	63081.92	55754.02	1.1314
108	63663.92	56091.14	1.1350
109	64564.58	56428.27	1.1442
110	64945.13	56765.39	1.1441
111	64992.75	57102.52	1.1382
112	65772.17	57439.64	1.1451
113	67358.29	57776.77	1.1658
114	68753.50	58113.89	1.1831
115	69645.75	58451.02	1.1915
116	69544.04	58788.14	1.1830
117	69261.17	59125.26	1.1714
118	69747.33	59462.39	1.1730
119	70093.33	59799.51	1.1721
120	70039.33	60136.64	1.1647

### 1.6.3. Variación Estacional.

A diferencia del análisis de la tendencia, el análisis del componente estacional de una serie histórica, tiene implicaciones más inmediatas de corto plazo. Por otro lado, ambos análisis se diferencian en por lo menos dos formas:

1.- Mientras que la tendencia se determina en forma directa a partir de los datos disponibles, el componente estacional se determina eliminando los otros componentes, de modo que al final sólo quede el estacional.

2.- Mientras que la tendencia se determina mediante una ecuación o líneas de mejor ajuste, se debe calcular un valor estacional por separado para cada mes (o trimestre) del año, por lo regular en la forma de un número índice. Al igual que en el análisis de tendencia, se han desarrollado diversos métodos para medir la variación estacional; la mayoría de los cálculos del índice estacional que ahora se emplean, son variaciones del *método de la proporción al promedio móvil*.

Los promedios móviles calculados anteriormente (ver tabla 1.1) son estimaciones de la componente tendencia-ciclo, por lo que si se divide la serie original entre los resultados de los promedios móviles, se obtendrá una estimación (ver tabla 1.4) de la componente estacional-irregular, como se muestra en la siguiente ecuación:

$$\frac{Y_i}{\text{promedio móvil centrado}} = \frac{T_i S_i C_i I_i}{T_i C_i} = S_i I_i$$

Para obtener el índice estacional, primero se deben eliminar las variaciones irregulares. Esto se logra obteniendo un promedio, en este caso, la mediana ( $Me_i$ ) de los valores obtenidos en la ecuación anterior para cada mes. La forma habitual de obtener la mediana de un conjunto de datos es ordenándolos y tomar el dato que está en medio de todos, y en caso de ser dos, se toma un promedio de ellos, de tal forma que se tome el dato en donde se acumula el 50% de la distribución.

Estos valores de la mediana se ajustan de tal manera que la suma de los valores del índice estacional para un año sea 4 y el valor promedio igual a 1. Se pueden observar los resultados del ajuste en la tabla 1.5.

Mes	Mediana mensual	Factor ajuste	Índice estacional
1	0.9925	0.9932	0.9858
2	0.9575	0.9932	0.9510
3	1.0635	0.9932	1.0562
4	0.9206	0.9932	0.9144
5	0.8936	0.9932	0.8875
6	0.9944	0.9932	0.9876
7	0.8192	0.9932	0.8136
8	0.8800	0.9932	0.8740
9	0.8640	0.9932	0.8581
10	1.0400	0.9932	1.0329
11	1.2498	0.9932	1.2413
12	1.4070	0.9932	1.3974

Tabla 1.4. Cálculo del factor de ajuste para la componente estacional.

El componente estacional en una serie, se mide en la forma de un número índice. Su cálculo, que representa el grado de influencia estacional para un segmento del año en particular, comprende la comparación de los valores medidos o esperados para ese segmento (mes), con el promedio general de todos los segmentos del año.

De este modo, un índice estacional de 1 para un mes en particular, indica que el valor esperado para ese mes es de 1/12 del total del período anual centrado en ese mes. Un índice estacional de 1.0562, indica que el valor esperado para ese mes es 5.62% mayor que el promedio mensual. Un índice de 0.9858 indica que el precio esperado para ese mes es sólo el 98.58% del promedio mensual.

Así, un índice indica las altas y bajas esperadas en los niveles de actividad dependiendo del período manejado, eliminando los efectos ocasionales por los componentes de tendencia, cíclico e irregular.

A continuación, se utiliza el índice estacional como ajuste para una proyección de la tendencia, que se muestra en la Tabla 1.5. Esto es, se multiplica el valor de la tendencia proyectada para un mes determinado por el índice estacional correspondiente a ese mes. Si el propósito es encontrar una serie desestacionalizada, entonces se debe eliminar los efectos de la componente estacional de los datos. Esto se logra dividiendo cada valor observado de la serie de tiempo mensual por el índice estacional de ese mes.

Tabla 1.5. Datos desestacionalizados de las ventas de automóviles.

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)	E x I	I	Valor ajustado estacionalmente
01-01-92	1	36233					36755.01
01-02-92	2	40215					42287.82
01-03-92	3	48440					45861.06
01-04-92	4	37447					40954.74
01-05-92	5	38962					43899.66
01-06-92	6	42266	40315.00				42796.32
01-07-92	7	42683	40827.58	40571.29	1.0520	1.2930	52459.94
01-08-92	8	28130	41056.67	40942.13	0.6871	0.7861	32183.52
01-09-92	9	39496	41564.92	41310.79	0.9561	1.1142	46027.29
01-10-92	10	40730	41373.50	41469.21	0.9822	0.9509	39431.12
01-11-92	11	43433	40830.58	41102.04	1.0567	0.8513	34988.77
01-12-92	12	45745	40353.00	40591.79	1.1270	0.8064	32734.83
01-01-93	1	42384	38954.83	39653.92	1.0688	1.0842	42994.63
01-02-93	2	42964	39098.00	39026.42	1.1009	1.1576	45178.51
01-03-93	3	54539	38119.50	38608.75	1.4126	1.3374	51635.35
01-04-93	4	35150	36885.83	37502.67	0.9373	1.0251	38442.57
01-05-93	5	32447	37000.42	36943.13	0.8783	0.9896	36559.01
01-06-93	6	36535	36513.92	36757.17	0.9940	1.0064	36993.41
01-07-93	7	25905	36138.00	36325.96	0.7131	0.8765	31838.78
01-08-93	8	29848	36209.50	36173.75	0.8251	0.9440	34149.09
01-09-93	9	27754	34998.92	35604.21	0.7795	0.9084	32343.57
01-10-93	10	25926	35020.08	35009.50	0.7405	0.7169	25099.22
01-11-93	11	44808	35473.92	35247.00	1.2713	1.0241	36096.45
01-12-93	12	39907	35677.42	35575.67	1.1217	0.8027	28557.20
01-01-94	1	37873	36043.75	35860.58	1.0561	1.0713	38418.64
01-02-94	2	43822	36489.17	36266.46	1.2083	1.2706	46080.73
01-03-94	3	40012	37015.50	36752.33	1.0887	1.0307	37881.77
01-04-94	4	35404	37890.08	37452.79	0.9453	1.0338	38720.36
01-05-94	5	37893	37146.42	37518.25	1.0100	1.1380	42695.18
01-06-94	6	38977	37760.25	37453.33	1.0407	1.0537	39466.05
01-07-94	7	30301	36217.67	36988.96	0.8192	1.0068	37241.73
01-08-94	8	35193	33423.00	34820.33	1.0107	1.1563	40264.30
01-09-94	9	34070	31074.00	32248.50	1.0565	1.2312	39704.02
01-10-94	10	36421	28668.33	29871.17	1.2193	1.1804	35259.53
01-11-94	11	35884	26325.17	27496.75	1.3050	1.0513	28907.45
01-12-94	12	47273	23772.00	25048.58	1.8873	1.3505	33828.26
01-01-95	1	19362	21720.92	22746.46	0.8512	0.8635	19640.95
01-02-95	2	10286	19342.50	20531.71	0.5010	0.5268	10816.18
01-03-95	3	11824	17147.83	18245.17	0.6481	0.6136	11194.49
01-04-95	4	8536	14977.25	16062.54	0.4069	0.4450	7148.24
01-05-95	5	9775	13356.42	14166.83	0.6900	0.7774	11013.79
01-06-95	6	8339	11343.25	12349.83	0.6752	0.6837	8443.63
01-07-95	7	5688	10959.17	11151.21	0.5101	0.6269	6990.89

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)	E x I	I	Valor ajustado estacionalmente
01-08-95	8	6652	11607.58	11283.38	0.5895	0.6745	7610.55
01-09-95	9	7734	12235.83	11921.71	0.6487	0.7560	9012.94
01-10-95	10	10374	13085.00	12660.42	0.8194	0.7933	10043.17
01-11-95	11	16434	13627.50	13356.25	1.2304	0.9912	13238.91
01-12-95	12	23115	14389.08	14008.29	1.6501	1.1808	16540.95
01-01-96	1	14753	15267.83	14828.46	0.9949	1.0092	14965.55
01-02-96	2	18067	15924.75	15596.29	1.1584	1.2181	18998.23
01-03-96	3	19363	16801.75	16363.25	1.1833	1.1203	18332.12
01-04-96	4	16726	18257.42	17529.58	0.9542	1.0435	18292.76
01-05-96	5	16285	19516.00	18886.71	0.8622	0.9715	18348.80
01-06-96	6	17478	20372.17	19944.08	0.8764	0.8873	17697.30
01-07-96	7	16233	21197.08	20784.63	0.7810	0.9599	19951.32
01-08-96	8	14535	21709.83	21453.46	0.6775	0.7751	16629.49
01-09-96	9	18258	21972.75	21841.29	0.8359	0.9742	21277.25
01-10-96	10	27842	22598.67	22285.71	1.2493	1.2095	26954.12
01-11-96	11	31537	23449.67	23024.17	1.3697	1.1034	25405.59
01-12-96	12	33389	24011.92	23730.79	1.4070	1.0068	23892.96
01-01-97	1	24652	24889.83	24450.88	1.0082	1.0228	25007.16
01-02-97	2	24220	26238.67	25564.25	0.9474	0.9962	25468.38
01-03-97	3	22518	26997.25	26617.96	0.8460	0.8009	21319.14
01-04-97	4	24237	28025.75	27511.50	0.8810	0.9635	26507.33
01-05-97	5	26497	29079.33	28552.54	0.9280	1.0456	29854.97
01-06-97	6	24225	30275.50	29677.42	0.8163	0.8265	24528.95
01-07-97	7	26768	31441.67	30858.58	0.8674	1.0661	32899.46
01-08-97	8	30721	32708.42	32075.04	0.9578	1.0958	35147.89
01-09-97	9	27361	34455.25	33581.83	0.8148	0.9495	31885.58
01-10-97	10	40184	35819.42	35137.33	1.1436	1.1072	38902.53
01-11-97	11	44180	36687.17	36253.29	1.2186	0.9817	35590.54
01-12-97	12	47743	38316.17	37501.67	1.2731	0.9110	34164.59
01-01-98	1	38646	39413.25	38864.71	0.9944	1.0087	39202.78
01-02-98	2	39421	40088.33	39750.79	0.9917	1.0428	41452.89
01-03-98	3	43480	40907.33	40497.83	1.0736	1.0165	41165.13
01-04-98	4	40607	40570.67	40739.00	0.9968	1.0901	44410.74
01-05-98	5	36910	40650.00	40610.33	0.9089	1.0241	41587.61
01-06-98	6	43773	42141.00	41395.50	1.0574	1.0707	44322.22
01-07-98	7	39933	41585.17	41863.08	0.9539	1.1724	49080.02
01-08-98	8	38822	40891.17	41238.17	0.9414	1.0771	44416.24
01-09-98	9	37189	40730.25	40810.71	0.9113	1.0619	43338.79
01-10-98	10	36144	40523.75	40627.00	0.8897	0.8613	34991.37
01-11-98	11	45132	40630.50	40577.13	1.1123	0.8960	36357.45
01-12-98	12	65635	41043.83	40837.17	1.6072	1.1501	46967.99
01-01-99	1	31976	41144.25	41094.04	0.7781	0.7893	32436.68
01-02-99	2	31093	41225.58	41184.92	0.7550	0.7939	32695.64
01-03-99	3	41549	42236.17	41730.88	0.9956	0.9426	39336.94

Fecha	Mes	Venta	Promedio móvil	Promedio móvil centrado (TxC)	E x I	I	Valor ajustado estacionalmente
01-04-99	4	38129	43302.83	42769.50	0.8915	0.9750	41700.62
01-05-99	5	38191	44340.83	43821.83	0.8715	0.9820	43030.95
01-06-99	6	48733	45212.50	44776.67	1.0884	1.1020	49344.46
01-07-99	7	41138	46449.08	45830.79	0.8976	1.1032	50561.04
01-08-99	8	39798	48110.83	47279.96	0.8418	0.9630	45532.88
01-09-99	9	49316	49456.08	48783.46	1.0109	1.1781	57471.19
01-10-99	10	48944	50481.67	49968.88	0.9795	0.9483	47383.17
01-11-99	11	57588	52156.17	51318.92	1.1222	0.9040	46391.76
01-12-99	12	76095	52977.00	52566.58	1.4476	1.0359	54453.10
01-01-00	1	46815	53485.25	53231.13	0.8795	0.8921	47489.47
01-02-00	2	51034	54588.67	54036.96	0.9444	0.9931	53664.46
01-03-00	3	57692	54877.00	54732.83	1.0541	0.9979	54620.49
01-04-00	4	50436	56715.50	55796.25	0.9039	0.9886	55160.44
01-05-00	5	58285	59102.17	57908.83	1.0065	1.1340	65671.46
01-06-00	6	58583	58681.17	58891.67	0.9948	1.0072	59318.05
01-07-00	7	47237	59872.25	59276.71	0.7969	0.9794	58057.07
01-08-00	8	53039	60667.83	60270.04	0.8800	1.0068	60681.90
01-09-00	9	52776	61504.00	61085.92	0.8640	1.0068	61503.35
01-10-00	10	71006	61890.67	61597.33	1.1527	1.1160	68741.62
01-11-00	11	66228	61323.42	61507.04	1.4019	1.1294	69463.59
01-12-00	12	71043	61009.75	61166.58	1.1615	0.8311	50837.92
01-01-01	1	61108	62354.08	61681.92	0.9907	1.0050	61968.39
01-02-01	2	60581	62871.92	62613.00	0.9675	1.0174	63703.55
01-03-01	3	67726	63383.92	63127.92	1.0728	1.0157	64120.28
01-04-01	4	52676	63167.00	63275.46	0.8325	0.9105	57610.27
01-05-01	5	53878	62996.83	63081.92	0.8541	0.9623	60705.96
01-06-01	6	54819	64331.00	63663.92	0.8611	0.8719	55506.82
01-07-01	7	63369	64798.17	64564.58	0.9815	1.2063	77884.26
01-08-01	8	59253	65092.08	64945.13	0.9124	1.0438	67791.34
01-09-01	9	58920	64893.42	64992.75	0.9066	1.0565	68663.36
01-10-01	10	68403	66650.92	65772.17	1.0400	1.0068	66221.63
01-11-01	11	84186	68065.67	67358.29	1.2498	1.0068	67818.59
01-12-01	12	87053	69441.33	68753.50	1.2662	0.9061	62294.58
01-01-02	1	66714	69860.17	69645.75	0.9579	0.9717	67675.15
01-02-02	2	64108	69237.92	69544.04	0.9218	0.9693	67412.34
01-03-02	3	65342	69284.42	69261.17	0.9434	0.8932	61863.20
01-04-02	4	73766	70210.25	69747.33	1.0576	1.1567	80675.81
01-05-02	5	70855	69976.42	70093.33	1.0109	1.1390	79834.46
01-06-02	6	71327	70102.25	70039.33	1.0184	1.0312	72221.95
01-07-02	7	68275					83914.02
01-08-02	8	51906					59385.64
01-09-02	9	59478					69313.63
01-10-02	10	79513					76977.33
01-11-02	11	81380					65558.13
01-12-02	12	68563					63375.12



## 1.7. Gráficas de las componentes.

A continuación se presentan los gráficos de cada componente para entender un poco más el comportamiento de la serie de automóviles.

Separando el componente de tendencia, se observa que de 1992 a 2002 la venta de autos ha crecido sostenidamente y esto se confirma revisando la serie original y darse cuenta que de aproximadamente 20 millones aumentó en 10 años a 90 millones aproximadamente. No podría decirse que el crecimiento irá creciendo conforme el tiempo lo haga, pues como se dijo anteriormente, la serie podría estar envuelta en algún ciclo desconocido, que provoque un descenso más adelante.

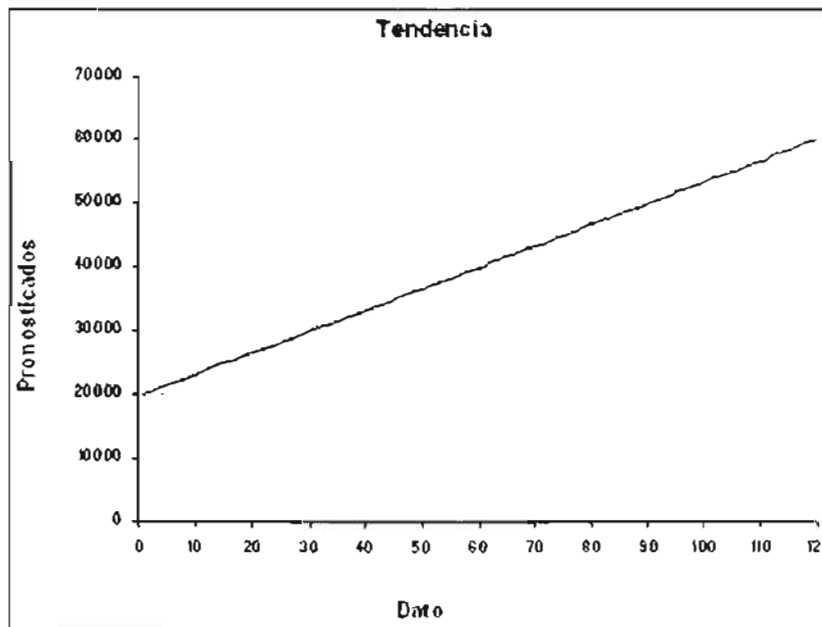


Fig. 1.7.1. Gráfico de la componente Tendencia de las ventas de automóviles.

Lo anterior no es posible definirlo con la siguiente gráfica, pues no se observa algún patrón repetitivo que confirme que la tendencia está formando parte del ciclo. Para poder afirmarlo necesitaríamos más datos, y de esa forma se sabría en que mes/año el ciclo comienza y en cuál termina.

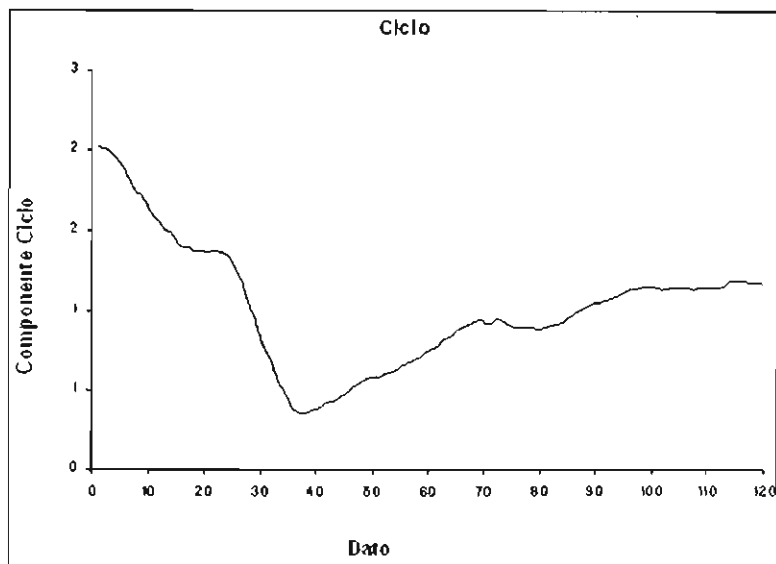


Fig. 1.7.2. Gráfico de la componente Ciclo de las ventas de automóviles.

Respecto al componente estacional se observa muy claramente la repetición de los "picos" cada 12 meses y de manera muy parecida año tras año.

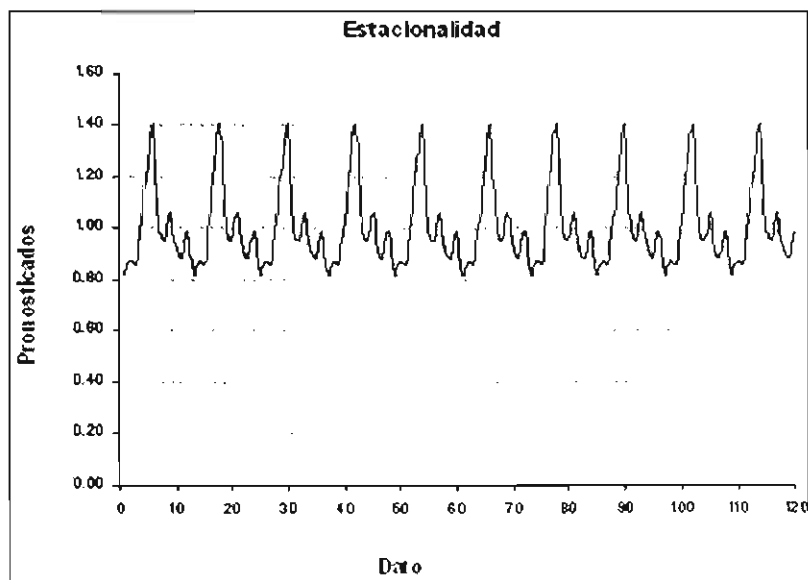


Fig. 1.7.3. Gráfico de la componente Estacionalidad de las ventas de automóviles.

Se observa que en general el componente irregular se mantiene constante en el tiempo, es decir, que no supone tener información importante de la serie, que pudiera indicar cierta dependencia con los datos. Se puede concluir que el componente irregular (ó fluctuaciones locales) en realidad es aleatorio y no es posible modelarlo.

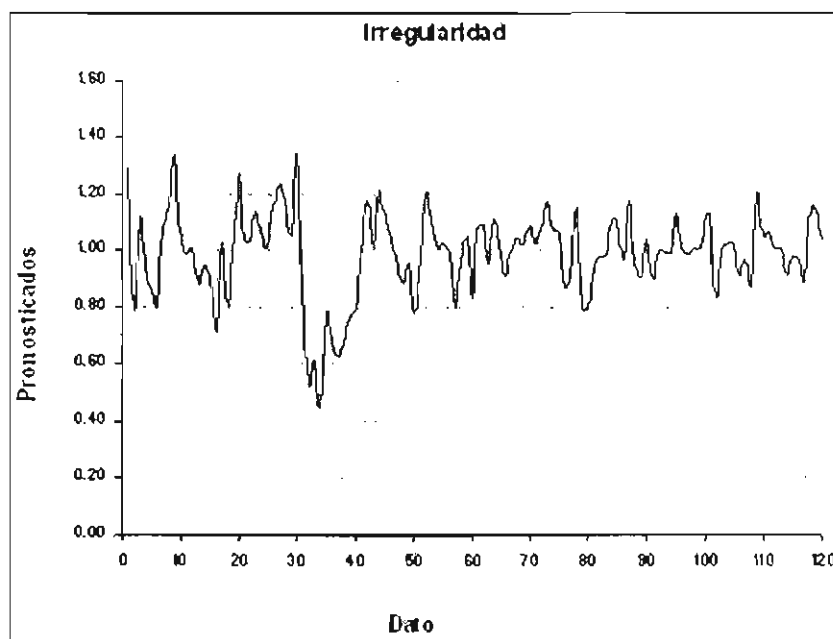


Fig. 1.7.4. Gráfico de la componente Irregular de las ventas de automóviles.

De esta forma, se puede concluir que el análisis de series que dependen del tiempo, es una herramienta muy útil para el estudio del comportamiento de la misma y se basa principalmente en la descomposición de los factores que influyen en la variable de interés. Básicamente lo que se pretende con el caso univariante de las series temporales es el conocimiento de de una variable a lo largo del tiempo para que, a partir de este conocimiento, y bajo el supuesto de que no van a producirse cambios estructurales, se podrán realizar predicciones. Hasta el primer cuarto del siglo XX, los métodos de análisis de series temporales, se basaban en los modelos de descomposición, es decir, en donde se supone que las series resultan de la agregación de cuatro componentes básicas. Posteriormente surgió la concepción moderna o estocástica, fundamentada en la teoría de los procesos aleatorios, los cuales representan a la serie mediante relaciones analíticas que conectan los valores de la variable con combinaciones lineales o no de parámetros y variables. En el capítulo siguiente se analizarán a las series mediante esta nueva concepción, tratando de eliminar los efectos que no permitan explicar ni pronosticar adecuadamente a la variable de interés.

# CAPITULO 2. MODELOS ARIMA MULTIPLICATIVOS.

## 2.1. Series de tiempo en los procesos estocásticos.

Un *proceso estocástico* es un conjunto de variables aleatorias  $\{Z_t\}$ , donde el índice  $t$  toma valores en un cierto conjunto  $C$ . En este caso, el conjunto es ordenado y corresponde a los instantes temporales (días, meses, años, etc.). Para cada valor  $t$  del conjunto  $C$  (para cada instante temporal) está definida una variable aleatoria,  $Z_t$ , y los valores observados de las variables aleatorias en distintos instantes forman una serie temporal. Es decir, una serie de  $T$  datos,  $(Z_1, \dots, Z_t, \dots, Z_T)$ , es una muestra de tamaño uno del vector de  $T$  variables aleatorias ordenadas en el tiempo correspondiente a los momentos  $t = 1, \dots, T$ , y la serie observada se considera una realización o trayectoria del proceso estocástico.

El proceso queda caracterizado si se define la distribución de probabilidad conjunta de las variables aleatorias  $(Z_1, \dots, Z_t, \dots, Z_T)$ , para cualquier valor de  $T$ . Estas distribuciones se denominan la distribuciones finito-dimensionales del proceso. Se diría que se conoce la estructura probabilística de un proceso estocástico, cuando se tenga conocimiento de estas distribuciones, que determinan la distribución de cualquier subconjunto de variables  $Z$ , en particular, las distribuciones marginales de cada variable.

Por ejemplo, se considera la variable climática lluvia a las 12 horas de cada día del año en alguna ciudad de Chiapas. Las 365 observaciones de cada año de esta variable constituyen una realización del proceso estocástico. Al comenzar otro año, se obtiene una nueva realización del proceso y así sucesivamente. Si se estudia el valor 26 de agosto de los distintos años en que se hayan levantado las observaciones, se tendrá la distribución de probabilidad de una variable aleatoria. Si se contempla el levantamiento de observaciones del período 1991-2002, se tienen 12 años, es decir, 12 realizaciones de la serie de lluvia en, por ejemplo, Tuxtla Gutiérrez en los 12 meses del año. Se tiene por tanto 12 variables aleatorias, una para cada mes y la trayectoria de los 12 valores en un año dado, representa una realización del proceso estocástico, Si se considera un mes cualquiera, los 12 valores disponibles forman una muestra de tamaño 12 de esa variable.

Se pueden observar varias realizaciones de un proceso estocástico en fenómenos estables repetitivos a lo largo del tiempo. Por ejemplo, con las variables climáticas, astronómicas o ambientales que siguen cada año un patrón similar como consecuencia de la rotación de la tierra alrededor del Sol. También en variables que se suponen se pueden repetir en las mismas condiciones; por ejemplo, el número de clientes en un puesto de servicio de un supermercado, gasolinera, etc. en la primera semana del mes en distintos meses que se pueden suponer similares, o el ritmo cardiovascular de una persona a lo largo del día o la cotización horaria de una acción en distintos días.

Determinar la distribución conjunta del proceso, requiere observar un gran número de realizaciones, para estimar la probabilidad de los distintos intervalos. Esta tarea se puede simplificar cuando se supone que la distribución conjunta es normal multivariada, pues así, quedará determinada por el vector de medias y la matriz de varianzas y covarianzas entre las variables.

Para un proceso  $\{Z_t\}$ , su función media está dada por

$$\mu_t = E(Z_t),$$

su función varianza por

$$\sigma_t^2 = Var(Z_t) = E[(Z_t - \mu_t)^2],$$

y su función de autocovarianza<sup>(1)</sup> entre los instantes de tiempo  $t_1$  y  $t_2$  por

$$\gamma(t_1, t_2) = Cov(Z_{t_1}, Z_{t_2}) = E[(Z_{t_1} - \mu_{t_1})(Z_{t_2} - \mu_{t_2})]$$

## 2.2. Procesos estacionarios.

La obtención de las distribuciones de probabilidad del proceso es posible en ciertas situaciones, por ejemplo, con variables climáticas, donde se supone que en cada año se observa una realización del mismo proceso o valores que pueden generarse en un laboratorio. Sin embargo, en muchas situaciones de interés, como ocurre con variables económicas y sociales, sólo se puede observar una realización del proceso. Por ejemplo, si se observa la serie de crecimientos anuales de la riqueza de un país, no es posible volver atrás en el tiempo para generar otra realización. El proceso estocástico existe conceptualmente, pero no es posible obtener muestras sucesivas o realizaciones independientes del mismo. Para poder estimar las características “transversales” del

<sup>1</sup> O también, para simplificar la notación de la función de autocovarianza, es posible definirla respecto al retardo establecido como sigue  $\gamma(t_i, t_{i+k}) = \gamma_k$ .

proceso como medias, varianzas, etc. a partir de su evolución “longitudinal”, es necesario suponer que las propiedades “transversales” o la distribución de las variables en cada instante) son estables a lo largo del tiempo. Esto conduce al concepto de estacionariedad, que se define a continuación:

Se dice que un proceso estocástico es *estacionario en sentido estricto* si:

1. Las distribuciones marginales de todas las variables son idénticas.
2. Las distribuciones finito-dimensionales de cualquier conjunto de variable sólo dependen de los retardos ( $k$ ) entre las variables, entendiendo por retardos a los rezagos en las variables, es decir, tomando como punto de partida el tiempo  $t$ , aquellos valores de las variables en los tiempos  $t - 1, t - 2, \dots, t - k$ .

Estas dos condiciones pueden resumirse estableciendo que la distribución conjunta de cualquier conjunto de variables no se modifica si se trasladan las variables en el tiempo, es decir:

$$F(x_i, x_j, \dots, x_k) = F(x_i + h, x_j + h, \dots, x_k + h)$$

El problema de esta definición es que no es empíricamente contrastable, por que la serie está formada por una sola observación de cada una de las distribuciones de probabilidad que componen el proceso en cada período de tiempo. Además, el número de condiciones a comprobar podría ser infinito, es decir, es necesario disponer de las distribuciones conjuntas para cualquier selección de variables del proceso. Una propiedad más débil, pero más fácil de contrastar en la práctica, es la *estacionariedad en sentido débil*, que implica la estabilidad de la media, la varianza y la estructura de covarianzas a lo largo del tiempo. Un proceso es estacionario en sentido débil si su media es constante e independiente del tiempo, su varianza finita y constante, y el valor de la covarianza entre dos períodos no depende del tiempo en el cual se ha calculado, sino de la distancia o desfase entre ellos. Todos los momentos de primer y segundo orden son independientes del tiempo, incluyendo la función de autocorrelación parcial. Es decir, si para toda  $t$ :

1.  $E(Z_t) = \mu_t = \mu$  (constante)
2.  $Var(Z_t) = \sigma_t^2 = \sigma^2$  (constante)
3.  $\gamma_{-k} = \gamma(t - k, t) = \gamma(t, t + k) = \gamma_k$

Ahora, si se supone que el proceso  $Z_{t_1}, Z_{t_2}, Z_{t_3}, \dots, Z_{t_n}$  sigue una distribución normal multivariada, obviamente la estacionariedad débil coincide con la estricta, esto debido a que la función normal multivariada está caracterizada por el primero y segundo momento y basta con conocer la media y función de autocovarianza<sup>(2)</sup> del proceso para poder completamente describir a la serie estacionaria.

---

<sup>2</sup> Nótese que  $\gamma(0) = E[(Z_t - \mu)(Z_t - \mu)] = \sigma^2$ .

Se define el coeficiente de autocorrelación como

$$\rho_k = \frac{\text{Cov}(Z_t, Z_{t+k})}{\sqrt{\text{Var}(Z_t)\text{Var}(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

el cual, para diferentes valores de  $k$ , genera la *función de autocorrelación (fas)*, o *correlograma*. Esta función, tiene las siguientes propiedades:

- $\rho_0 = 1$
- $|\rho_k| \leq 1$
- $\rho_k = \rho_{-k}$ , para todo  $k$ , es decir, la función de autocorrelación es una función par y simétrica en  $k = 0$ .

La estacionariedad débil no garantiza la estabilidad completa del proceso. Por ejemplo, la distribución de las variables  $Z_t$  puede estar cambiando en el tiempo. Sin embargo, si se supone que estas variables tienen conjuntamente una distribución normal  $n$ -dimensional, como ésta queda determinada por las medias y las varianzas y covarianzas, todas las distribuciones marginales serán idénticas, y también lo serán las distribuciones de subconjuntos de hasta  $n - 1$  variables del tipo  $(Z_{t_1}, Z_{t_1+j}, Z_{t_1+k})$  y  $(Z_{t_2}, Z_{t_2+j}, Z_{t_2+k})$  para cada  $t$ . Por tanto, para procesos con distribuciones normales la estacionariedad débil coincide con la estricta.

### 2.2.1. Estacionariedad en media y varianza.

Supóngase una serie temporal con tendencia creciente. Como cada observación de la serie proviene de la distribución de probabilidad del período correspondiente, es razonable creer que la esperanza matemática de dichas distribuciones crece también con el tiempo.

Para conocer si la serie temporal es estacionaria en varianza, puede realizarse la observación de los datos originales. Si alcanzan oscilaciones en torno a la tendencia, se puede defender la no estacionariedad de la serie. La fig. 2.1 presenta una serie temporal generada por un proceso no estacionario en varianza, siendo ésta creciente con el tiempo.

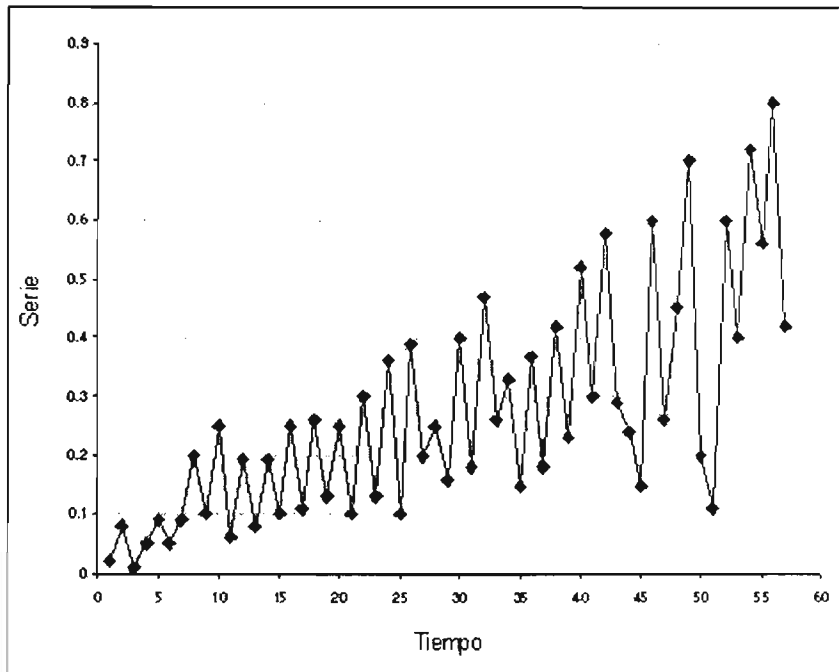


Fig. 2.1. Serie no estacionaria en varianza.

Los modelos de predicción de series temporales están diseñados para procesos estacionarios. Si las características del proceso cambian a lo largo del tiempo, resultará difícil representar a la serie para intervalos de tiempos pasados y futuros, mediante un modelo lineal sencillo. Sin embargo, por regla general, las series económicas no son series que proceden de procesos estacionarios, sino que suelen tener una tendencia creciente o decreciente, y variabilidad no constante. Esta limitación no es tan importante porque, en la práctica, se pueden transformar las series no estacionarias en otras que sí lo son.

Se dice que un proceso estocástico es *ergódico*, cuando los valores de la serie alejados en el tiempo, están poco correlacionados, es decir, cuando  $\rho_k$  decrece al aumentar el retardo  $k$ .

## 2.2.2. Proceso de ruido blanco

Un proceso estacionario muy importante es el definido por:

1.  $E[Z_t] = 0$
2.  $Var(Z_t) = \sigma^2$



3.  $Cov(Z_t, Z_{t-k}) = 0, \quad k = \pm 1, \pm 2, \dots$

Este proceso se denomina proceso de ruido blanco. Por ejemplo, se tira una moneda en instantes  $t = 1, 2, \dots$  y se define  $Z_t = -1$  si se obtiene águila y  $Z_t = +1$  si se obtiene sol, se tendrá un proceso de ruido blanco: la esperanza es cero, la varianza es constante (igual a la unidad) y las covarianzas son nulas. Como segundo ejemplo, si se extraen independientemente elementos al azar de una población cualquiera con media conocida  $\mu$  y se considera la secuencia de valores  $x_i - \mu$ , llamando  $x_i$  al valor observado en la  $i$ -ésima extracción, se obtiene un proceso de ruido. En estos procesos conocer los valores pasados no proporciona ninguna información sobre el futuro, ya que el proceso no tiene memoria.

Si se supone que en un proceso de ruido todas las variables tienen distribución normal, la no correlación garantiza la independencia y se llamará al proceso resultante *proceso de ruido blanco*, representándose por  $a_t$ . Normalmente se supone que las perturbaciones de un modelo de regresión considerados como un proceso de índice  $i = 1, \dots$ , forman un proceso de ruido blanco. La fig. 2.2. presenta una simulación de este proceso. Se observa que debido a la variabilidad muestral, habrá siempre períodos cortos donde la variabilidad sea mayor que en otros, o donde la media local sea mayor; lo importante es que el nivel medio de la serie es constante, igual a cero en este caso, que la variabilidad es estable y no tiene tendencia a crecer o decrecer con el tiempo y que las autocorrelaciones son todas muy pequeñas.

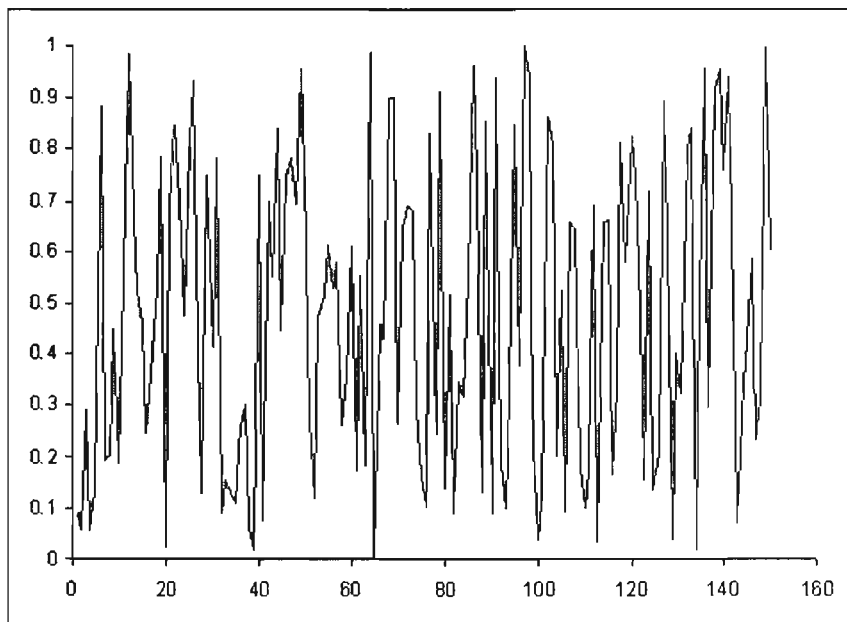


Fig. 2.2. Simulación de 150 observaciones de un proceso de ruido blanco

La fig. 2.3. presenta los coeficientes de autocorrelación para retardos mayores que cero de los datos simulados en la fig. 2.2. Además a la gráfica se han incorporado dos bandas de confianza del 95% para las autocorrelaciones, las cuales están a una distancia de  $\frac{2}{\sqrt{T}}$  y representan el intervalo de confianza aproximado al 95% donde deben estar los coeficientes estimados si la serie se ha generado por un proceso de ruido blanco. Un intervalo de confianza del 95% implica que una de cada 20 autocorrelaciones, se espera que salga fuera de las bandas en promedio, cuando el proceso es verdaderamente ruido blanco, lo cual ocurre en la fig. 2.3.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.109	0.109	1.2164	0.270
		2 0.069	0.048	1.5809	0.454
		3 0.113	0.103	2.9227	0.404
		4 0.086	0.063	3.7137	0.446
		5 -0.036	-0.062	3.8502	0.571
		6 0.001	-0.007	3.8504	0.697
		7 -0.084	-0.098	4.6308	0.705
		8 -0.127	-0.109	6.4145	0.601
		9 -0.065	-0.029	6.8860	0.649
		10 0.070	0.111	7.4432	0.683
		11 0.104	0.143	8.6925	0.650
		12 -0.135	-0.150	10.810	0.545
		13 -0.052	-0.070	11.123	0.601
		14 -0.018	-0.055	11.163	0.673
		15 -0.056	-0.053	11.532	0.714

Fig. 2.3. Función de autocorrelación de un proceso de ruido blanco.

## 2.3. Notación de operadores.

### 2.3.1. El operador retardo.

Se define el *operador retardo*, B, como un operador lineal que aplicado a una función temporal proporciona esa misma función retardada un período. Por definición:

$$Bf(t) \equiv f(t - 1)$$

donde, para simplificar la notación, en lugar de utilizar la notación más precisa

$B[f(t)] \equiv f(t - 1)$ , se escribe el operador como un producto sobre la función; pero esta notación debe interpretarse como una operación aplicada a una función,  $f(t)$ , que proporciona otra función  $f(t - 1)$ . En particular, si se aplica el operador de retardo a una serie temporal, se obtiene la misma serie temporal retardada un período y se escribe:

$$BZ_t = Z_{t-1}$$

Una constante se interpreta como una función que es constante en el tiempo, y se establece que al aplicar el operador de retardo a una constante ésta no se modifica:

$$B\mu = \mu$$

De acuerdo con la definición, si se aplica este operador al producto de una constante  $a$ , por una serie temporal se tiene que:

$$BaZ_t = aBZ_t = aZ_{t-1}$$

El operador de retardo es lineal, es decir:

$$B(af(t) + bg(t)) = af(t - 1) + bg(t - 1)$$

y puede aplicarse sucesivamente,

$$B^k Z_t = \overbrace{B \dots B}^{\text{k-veces}} Z_t = Z_{t-k}$$

Se definen los polinomios en el operador de retardo como nuevos operadores que, aplicados a una función proporciona otra función, de acuerdo con las propiedades anteriores del operador  $B$ . Por ejemplo, el operador  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2$  aplicado a una serie temporal proporciona otra serie temporal, mediante la operación

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2)Z_t &= Z_t - \phi_1 BZ_t - \phi_2 B^2 Z_t \\ &= Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} \end{aligned}$$

y al aplicar  $\phi(B)$  a una serie,  $Z_t$ , se obtiene otra serie,  $y_t$ , que se calcula como  $y_t = Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2}$ .

### 2.3.2. El operador diferencia.

Un caso particular importante de operador polinómico es el operador  $(1 - B)$ , que se conoce como *operador diferencia*, y se denota por  $\nabla = (1 - B)$ . Entonces:

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

El resultado de aplicar el operador diferencia a una serie  $z_t$  con  $T$  observaciones es obtener una nueva serie de  $T - 1$  observaciones mediante  $y_t = Z_t - Z_{t-1}$ , para  $t = 2, \dots, T$ . El operador diferencia puede aplicarse más de una vez:

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

El análisis de series temporales, no puede estar limitado a las series de tiempo estacionarias, pues en la realidad la mayoría de las series que existen son no estacionarias. Pero como se había comentado, existen varias formas de convertir en estacionarias las series que no lo son. En muchos casos la primera diferenciación del proceso no estacionario  $\{Z_t\}$  da como resultado un nuevo proceso  $\{W_t\}$ , que sí cumple con las características de estacionariedad.

$$W_t = Z_t - Z_{t-1}$$

Si  $\{W_t\}$  todavía no es estacionario, se repite la operación, obteniendo un nuevo proceso  $\{V_t\}$ , donde

$$V_t = W_t - W_{t-1}$$

es decir, se sigue diferenciando para encontrar un proceso estacionario, que se denominará *proceso integrado de orden 2*, por ser resultado de dos diferenciaciones. Para este tipo de relaciones, se utilizará el operador  $\nabla$ , el cual está definido como:

$$\nabla Z_t = Z_t - Z_{t-1}$$

y por lo tanto se tiene que  $W_t$  puede escribirse como  $W_t = \nabla Z_t$  y  $Z_t$  como

$$V_t = \nabla \nabla Z_t = \nabla^2 Z_t.$$

Como ejemplo, considérese el proceso  $Z_t$  que sigue una tendencia polinomial de orden uno,

$$Z_t = a + bt + u_t$$

donde  $\{u_t\}$  es un proceso ruido blanco, el cual es estacionario. Al aplicar el operador diferencia una vez, se obtiene,

$$\begin{aligned} \nabla Z_t &= Z_t - Z_{t-1} \\ &= a + bt + u_t - [a + b(t-1) + u_{t-1}] \\ &= b + u_t - u_{t-1} \\ &= b + \nabla u_t \end{aligned}$$

Para ver que la diferencia sigue siendo un proceso estacionario, se debe demostrar que la diferencia de dos procesos estacionarios sigue siendo estacionario, ya que  $b$  es una constante.

Sean  $u_t$  y  $u_{t-1}$  dos procesos estacionarios, con media cero y varianza  $\sigma^2$ . Considérese la diferencia  $\nabla u_t = u_t - u_{t-1}$ . Tomando la función esperanza y varianza, se tiene

$$\begin{aligned} E(\nabla u_t) &= E(u_t - u_{t-1}) \\ &= E(u_t) - E(u_{t-1}) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{Var}(\nabla u_t) &= \text{Var}(u_t - u_{t-1}) \\ &= \text{Var}(u_t) + \text{Var}(u_{t-1}) - 2\text{Cov}(u_t, u_{t-1}) \\ &= 2\sigma^2 - 2\gamma_1 \\ &= 2(\sigma^2 - \gamma_1) \end{aligned}$$

$$\begin{aligned} \gamma(t, t+k) &= \text{Cov}(\nabla u_t, \nabla u_{t+k}) = E[(u_t - u_{t-1})(u_{t+k} - u_{t+k-1})] \\ &= E[u_t u_{t+k} - u_t u_{t+k-1} - u_{t-1} u_{t+k} + u_{t-1} u_{t+k-1}] \\ &= E(u_t u_{t+k}) - E(u_t u_{t+k-1}) - E(u_{t-1} u_{t+k}) + E(u_{t-1} u_{t+k-1}) \\ &= 2\gamma_k - \gamma_{k+1} - \gamma_{k-1} \end{aligned}$$

de donde se obtiene que la esperanza es siempre cero, la varianza constante, y que la autocovarianza entre  $\nabla u_t$  y  $\nabla u_{t+k}$  sólo depende del retardo y no del tiempo. De esta forma,

se demuestra que la tendencia polinomial de primer orden, puede ser eliminada mediante la aplicación del operador  $\nabla$ . A este tipo de no estacionariedad se le conoce como *no estacionariedad homogénea*.

En general, un proceso que sigue una tendencia polinomial de orden  $h$ , será un proceso integrado de orden  $h$  (véase demostración en el Apéndice A).

$$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_h t^h + u_t = \sum_{i=0}^h \beta_i t^i + u_t$$

donde  $\{u_t\}$  es un proceso estacionario.

### 2.3.3. Operadores inversos.

Estos operadores tienen sus operadores inversos que verifican la propiedad de que su producto por el operador inicial es la unidad. El inverso del operador de retardo, es el operador de adelanto  $F = B^{-1}$ , que verifica  $FB = 1$  y está definido por:

$$FZ_t = B^{-1}Z_t = Z_{t+1}$$

y se verifica que  $FBZ_t = Z_t = BFZ_t$ . Si se aplica un operador y su universo, en cualquier orden, la serie no se modifica. Generalizando esta definición, se puede definir el inverso de cualquier operador  $\phi(B)$  como un operador  $\phi(B)^{-1}$  tal que  $\phi(B)\phi(B)^{-1} = 1$ . Por ejemplo, se encontrará frecuentemente con el operador  $\phi(B) = (1 - \phi B)$ , cuyo inverso se define para  $|\phi| < 1$  como:

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

La definición anterior es coherente, ya que  $(1 - \phi B)(1 - \phi B)^{-1} = 1$ , y como  $|\phi| < 1$  la suma es convergente y la definición del operador inverso es consistente con la expresión para la suma de una progresión geométrica indefinida. En general, el operador inverso puede obtenerse a partir de la condición  $\phi(B)\phi(B)^{-1} = 1$ , y será una serie infinita. Las condiciones para que esa suma sea convergente determinan las situaciones en las que se puede definir sin ambigüedad a este operador.

Si  $\phi = 1$  el operador  $(1 - \phi B)$  es el operador diferencia y al invertirlo, la suma  $S = (1 - B)^{-1} = 1 + B + B^2 + \dots$  no es convergente. Sin embargo, se puede todavía definir el inverso del operador diferencia como el operador suma:

$$SZ_t = Z_t + Z_{t-1} + Z_{t-2} + \dots$$

ya que es fácil comprobar que  $(1 - B)(1 + B + B^2 + \dots) = 1$ .

## 2.4. Descomposición de Wold.

Los modelos para series que se están estudiando, están basados en la idea de que una serie de tiempo estacionaria, cuyos valores sucesivos pueden estar altamente dependientes, puede considerarse generada a partir de variables aleatorias no correlacionadas  $a_t$ , mediante la relación

$$\tilde{Z}_t = a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots \quad (2.1)$$

en donde  $\tilde{Z}_t = Z_t - \mu$ . Esta idea fue fundamentada en la representación general de procesos estacionarios, obtenida por Wold en 1938. Este autor demostró que todo proceso débilmente estacionario  $Z_t$ , que no contenga componentes deterministas puede expresarse en función de  $a_t$ , mediante (2.1) donde  $\{a_t\}$  es un proceso de ruido blanco. Los coeficientes de  $\Psi_i$  tienen que cumplir con ciertas condiciones para que el proceso sea estacionario. La esperanza de  $\tilde{Z}_t$  tiene que ser constante al igual que la varianza, y la función de autocovarianza sólo debe depender del retardo.

Debido a que  $a_t$  es un proceso de ruido blanco, se puede verificar de inmediato que la esperanza de la *representación lineal general*, es constante:

$$\begin{aligned} E(\tilde{Z}_t) &= E(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots) \\ &= E(a_t) + \Psi_1 E(a_{t-1}) + \Psi_2 E(a_{t-2}) + \dots \\ &= 0 \end{aligned}$$

y por otro lado,

$$\begin{aligned} Var(\tilde{Z}_t) &= Var(a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots) \\ &= Var\left[\sum_{i=0}^{\infty} \Psi_i a_{t-i}\right] \\ &= \sigma_a^2 \sum_{i=0}^{\infty} \Psi_i^2 \end{aligned}$$

entonces, para que el proceso sea estacionario, la serie  $\{\Psi_i^2\}$ , deberá ser convergente. En el caso de la autocovarianza del proceso  $\tilde{Z}_t$ , ésta puede calcularse multiplicando de ambos lados de (2.1) por  $\tilde{Z}_{t-k}$  y aplicando el operador esperanza, se obtiene:

$$\gamma_k = E(\tilde{Z}_t \tilde{Z}_{t-k}) = E\left[\sum_{i=0}^{\infty} \Psi_i a_{t-i} \tilde{Z}_{t-k}\right]$$

sustituyendo  $\tilde{Z}_{t-k}$  en términos del proceso  $a_t$ , se obtiene:

$$\gamma_k = E\left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Psi_i \Psi_j a_{t-i} a_{t-k+j}\right] = \sigma_a^2 \sum_{i=0}^{\infty} \Psi_i \Psi_{i+k}$$

ya que los últimos términos que serán diferentes de cero, serán aquellos en donde  $i = j - k$ . La expresión final sólo depende del rezago. Por tanto, un proceso de la forma (2.1) se dirá que es estacionario, si  $\sum_{i=0}^{\infty} \Psi_i^2 < \infty$ . Dentro de los casos particulares de la representación general de Wold se encuentran los modelos de media móvil, los cuales se representan mediante la expresión,

$$Z_t = \theta_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

en donde  $\theta_0$  denota la media del proceso  $Z_t$ ,  $a_t$  un proceso de ruido blanco, y  $\theta_1, \theta_2, \dots, \theta_q$ , son constantes que sirven para relacionar a las sucesiones  $\{Z_t\}$  y  $\{a_t\}$ .

Asímismo, los modelos autorregresivos, que se definen como:

$$Z_t = \theta_0 + a_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}$$

en donde  $\theta_0$  y  $\phi_i$ ,  $i = 1, \dots, p$  son constantes a determinar y  $a_t$  un proceso de ruido blanco. Las características más importantes de los procesos de media móvil y autorregresivos se presentarán posteriormente.

## 2.5. Ecuaciones en diferencia.

Las ecuaciones en diferencia son los análogos discretos de las ecuaciones diferenciales (en donde el tiempo es continuo) y juegan un papel importante en el estudio de las series



temporales, ya que algunas de sus propiedades dependen de las características de las raíces de una ecuación de este tipo. Es por eso que se hará un paréntesis para dar una breve introducción enfocada básicamente a los métodos para obtener sus soluciones.

Una ecuación en diferencia de orden  $k$  (que es el máximo retardo de la variable que aparece en la ecuación) está dada por la ecuación,

$$f_t = x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_k x_{t-k}$$

donde  $\phi_i$ ,  $i = 1, \dots, n$ , son constantes a determinar. Si la función  $f_t$  es idénticamente cero, se dice que la ecuación es homogénea y no-homogénea si  $f_t \neq 0$  <sup>(3)</sup>. A menos que se indique otra situación, las ecuaciones que se estudiarán son las ecuaciones lineales homogéneas en diferencia con coeficientes constantes.

### 2.5.1. Ecuaciones en diferencia de orden uno.

El caso más simple y por ello el más fácil de resolver, es el caso de la ecuación en diferencia de orden uno,

$$x_t - x_{t-1} = 0 \tag{2.2}$$

la cual puede ser reescrita como  $x_t = \phi x_{t-1}$  e iterando sucesivamente se obtiene la solución general de la ecuación (2.2)

$$x_t = \phi x_{t-1} = \phi(\phi x_{t-2}) = \dots = \phi^t x_0$$

Dada la solución inicial  $x_0$ , ésta se vuelve única.

### 2.5.2. Ecuaciones en diferencia de orden dos.

#### CASO 1.

Considérese la ecuación lineal homogénea en diferencia de orden dos:

<sup>3</sup> En general, las ecuaciones diferenciales lineales son de la forma:

$$a_0(x) y^{(n)} + a_1(x) y^{(n-1)} + \dots + a_{n-1}(x) y' + a_n(x) y = h(x)$$

Si  $h(x) = 0$ , la ecuación es homogénea.

Si  $h(x) \neq 0$ , la ecuación es lineal no homogénea.

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = 0 \quad (2.3)$$

Ahora, se demostrará que su solución general es

$$x_t = A_1 G_1^t + A_2 G_2^t \quad (2.4)$$

en donde  $A_1$  y  $A_2$  son constantes que dependen de las condiciones iniciales y  $G_i^t$ ,  $i = 1, 2$ , son las raíces de una ecuación del siguiente estilo:

$$y^2 - \phi_1 y - \phi_2 = 0 \quad (2.5)$$

Para demostrarlo, sustitúyase (2.4.) en (2.3),

$$\begin{aligned} x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} &= (A_1 G_1^t + A_2 G_2^t) - \phi_1 (A_1 G_1^{t-1} + A_2 G_2^{t-1}) - \phi_2 (A_1 G_1^{t-2} + A_2 G_2^{t-2}) \\ &= A_1 G_1^{t-2} (G_1^2 - \phi_1 G_1 - \phi_2) + A_2 G_2^{t-2} (G_2^2 - \phi_1 G_2 - \phi_2) \\ &= 0 \end{aligned}$$

Por lo que se observa de la última igualdad, que  $G_1$  y  $G_2$ , son raíces de la ecuación (2.5).

En términos del operador de retardo  $B$ , (2.3) se expresa como:

$$(1 - \phi_1 B - \phi_2 B^2)x_t = 0$$

Además, si se factoriza  $B^2$  en el primer factor del producto, se observa que éste tiene la forma de (2.5),

$$B^2(B^{-2} - \phi_1 B^{-1} - \phi_2)x_t = 0$$

De esta manera, se dice que (2.4.) es la solución general de la ecuación en diferencia  $x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = 0$ , en donde  $G_i^{-1}$ ,  $i = 1, 2$  son las raíces de la ecuación  $1 - \phi_1 B - \phi_2 B^2 = 0$ ; ya que si  $G_i$  satisface  $B^{-2} - \phi_1 B^{-1} - \phi_2 = 0$  entonces  $G_i^{-1}$  satisface

$$1 - \phi_1 B - \phi_2 B^2 = 0.$$

A  $1 - \phi_1 B - \phi_2 B^2 = 0$  se le conoce como ecuación auxiliar asociada a la ecuación en diferencia (o también, ecuación característica), y sus raíces están dadas por:

$$G_1^{-1} = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}, \quad G_2^{-1} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Hasta ahora se ha supuesto que  $G_1$  y  $G_2$  son raíces reales y diferentes, es decir que  $\phi_1^2 + 4\phi_2 > 0$ . Ahora se debe analizar los dos casos restantes, en donde  $\phi_1^2 + 4\phi_2 = 0$  y  $\phi_1^2 + 4\phi_2 < 0$ .

### CASO 2.

Cuando  $\phi_1^2 + 4\phi_2 = 0$ , la ecuación auxiliar sólo tiene una raíz,  $G^{-1} = \frac{\phi_1}{2}$ . Por lo que se tiene que demostrar que la solución general de (2.4.) puede escribirse como  $(A_1 + tA_2)G^t$ , es decir, que si  $G^t$  es solución de (2.3), entonces  $tG^t$  también lo es. Para comprobarlo, sustitúyase  $tG^t$  en (2.3).

$$\begin{aligned} tG^t - \phi_1(t-1)G^{t-1} - \phi_2(t-2)G^{t-2} &= tG^t - \phi_1 B G^t - \phi_2 B^2 G^t \\ &= tG^t(1 - \phi_1 B - \phi_2 B^2) \end{aligned}$$

Ahora, se escribe a  $1 - \phi_1 B - \phi_2 B^2$  en términos de su raíz doble  $G^{-1}$ , es decir como  $(1 - GB)^2$ , por lo que

$$\begin{aligned} tG^t(1 - \phi_1 B - \phi_2 B^2) &= tG^t(1 - GB)^2 \\ &= tG^t(1 - GB)(1 - GB) \\ &= (tG^t - tG^{t+1}B)(1 - GB) \\ &= [tG^t - (t-1)G^t](1 - GB) \\ &= G^t(1 - GB) \\ &= G^t - G^{t+1}B = G^t - G^t \\ &= 0 \end{aligned}$$

es decir  $tG^t$  satisface (2.3) y con esto la solución general de (2.4.) puede escribirse como  $(A_1 + tA_2)G^t$ , en donde  $A_1$  y  $A_2$  quedarán determinadas por las condiciones iniciales.

### CASO 3.

Si  $\phi_1^2 + 4\phi_2 < 0$ , las raíces de  $1 - \phi_1 B - \phi_2 B^2 = 0$  son complejas y  $G_1$  y  $G_2$  son de la forma

$$G_1 = \alpha + i\beta \quad \text{y} \quad G_2 = \alpha - i\beta$$

$$\text{donde } \alpha = \frac{\phi_1}{2} \text{ y } \beta = \frac{\sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Como las soluciones de la ecuación en diferencia son  $G_1^t$  y  $G_2^t$  se debe describir como calcular las potencias de números complejos. Esto es más fácil cuando los números están escritos en forma polar. Es decir,

$$\begin{aligned} r &= \sqrt{\alpha^2 + \beta^2} & \theta &= \tan^{-1} \frac{\beta}{\alpha} \\ \alpha &= r \cos \theta & \beta &= r \operatorname{sen} \theta \end{aligned}$$

Usando las dos últimas relaciones, se tiene,

$$G_1 = r(\cos \theta + i \operatorname{sen} \theta) \quad \text{y} \quad G_2 = r(\cos \theta - i \operatorname{sen} \theta)$$

La fórmula de Euler dice que  $e^{i\theta} = \cos \theta + i \operatorname{sen} \theta$  con lo cual se puede expresar a  $G_1$  y  $G_2$  como:

$$G_1 = r e^{i\theta} \quad \text{y} \quad G_2 = r e^{-i\theta}$$

Entonces

$$G_1^t = r^t e^{it\theta} = r^t [\cos(t\theta) + i \operatorname{sen}(t\theta)]$$

y

$$G_2^t = r^t e^{-it\theta} = r^t [\cos(t\theta) - i \operatorname{sen}(t\theta)]$$

son las dos soluciones independientes de (2.3). Por lo tanto,  $\frac{1}{2}(G_1^t + G_2^t)$  es solución, lo mismo que  $\frac{1}{2i}(G_1^t - G_2^t)$ , las cuales pueden ser reescritas como  $G_{1*}^t = r^t \cos(t\theta)$  y  $G_{2*}^t = r^t \operatorname{sen}(t\theta)$ , respectivamente. quedando la solución general como,

$$x_t = A_1 G_1^t + A_2 G_2^t$$

### 2.5.3. Ecuaciones en diferencia de orden k.

Se llamará ecuación en diferencia de orden  $k$ , a una ecuación del tipo

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_k x_{t-k} = 0 \quad (2.6)$$

Cuando las raíces de la ecuación auxiliar asociada

$$1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_k B^k = 0 \quad (2.7)$$

son de multiplicidad uno, la solución general de la ecuación en diferencia de orden  $k$  es:

$$x_t = A_1 G_1^t + \dots + A_k G_k^t = \sum_{i=1}^k A_i G_i^t \quad (2.8)$$

para demostrarlo sustitúyase (2.8) en (2.6),

$$\sum_{i=1}^k A_i G_i^t - \phi_1 \sum_{i=1}^k A_i G_i^{t-1} - \dots - \phi_k \sum_{i=1}^k A_i G_i^{t-k} = \sum_{i=1}^k A_i G_i^{t-k} (G_i^k - \phi_1 G_i^{k-1} - \dots - \phi_k) = 0$$

Pero cuando la raíz  $G_i^{-1}$ ,  $i = 1, \dots, k$  es de multiplicidad  $m_i$ , la solución está dada por,

$$x_t = \sum_{i=1}^N G_i^t \sum_{j=1}^{m_i-1} j^t A_{ij} \quad (2.9)$$

en donde  $m_i$  indica la multiplicidad de la raíz  $G_i^{-1}$  y  $N$  el número de raíces diferente de la ecuación auxiliar. Es así como:  $\sum_{i=1}^N m_i = k$  y  $\prod_{i=1}^N (1 - G_i B)^{m_i} = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_k B^k$ . (Véase demostración en el Apéndice B).

## 2.6. Modelos Autorregresivos.

A continuación se presentan algunos modelos que tienen la particularidad de ser estables en el tiempo y sin tendencia a crecer o decrecer a largo plazo, es decir, procesos estacionarios cuya dinámica se caracteriza por sus covarianzas. Éstas miden la dependencia entre el valor de la variable hoy y sus valores pasados. Posteriormente se analizarán los modelos que tienen involucrados algún factor de tendencia y de estacionalidad, es decir, los no estacionarios.

El modelo más simple de dependencia entre dos variables aleatorias es el modelo de regresión simple que explica la evolución de una variable,  $y$ , como función lineal de otra variable,  $x$ , mediante la ecuación

$$y = c + bx + a$$

donde  $c$  y  $b$  son constantes a determinar y  $a$  es una variable aleatoria normal, con media nula y varianza constante. Si se aplica esta estructura de dependencia a las variables aleatorias  $y = Z_t$  y  $x = Z_{t-1}$ , se obtiene el proceso autorregresivo de primer orden. Esta dependencia puede generalizarse para permitir la relación actual de la serie con otros retardos,  $Z_{t-2}, \dots, Z_{t-p}$ , con lo que se obtiene un proceso autorregresivo de orden  $p$ .

### 2.6.1. El proceso autorregresivo de primer orden $AR(1)$ .

Se dirá que una serie  $Z_t$  sigue un proceso autorregresivo de primer orden si ha sido generada por:

$$Z_t = \theta_0 + \phi Z_{t-1} + a_t \quad (2.10)$$

donde  $\theta_0$  y  $-1 < \phi < 1$  son constantes a determinar y  $a_t$  es un proceso de ruido blanco independiente de  $Z_{t-k}$  para todo  $k$  positivo. El proceso (2.10) se denomina proceso autorregresivo de primer orden ó  $AR(1)$ .

Como ejemplo de una situación donde es esperable la aparición de este proceso, supóngase que  $Z_t$  es la cantidad de agua en un embalse a fin de mes. Cada mes llega al embalse una cantidad  $\theta_0 + a_t$ , donde  $\theta_0$  es el valor medio de la cantidad que entra y  $a_t$  una variable aleatoria de media cero, que hace que esa cantidad pueda variar de un período al siguiente. Cada mes se gasta una proporción fija de las existencias iniciales,  $(1 - \phi)Z_{t-1}$  y se mantiene la proporción  $\phi Z_{t-1}$ . Esta descripción conduce a (2.10).

La condición  $-1 < \phi < 1$  es necesaria para que el proceso sea estacionario. Para comprobarlo, supóngase que el proceso comienza con  $Z_0 = h$ , siendo  $h$  un valor cualquiera fijo. Entonces, el valor siguiente será  $Z_1 = \theta_0 + \phi h + a_1$ , el siguiente,  $Z_2 = \theta_0 + \phi Z_1 + a_2 = \theta_0 + \phi(\theta_0 + \phi h + a_1) + a_2$  y sustituyendo sucesivamente se puede escribir:

$$\begin{aligned} Z_1 &= \theta_0 + \phi h + a_1 \\ Z_2 &= \theta_0(1 + \phi) + \phi^2 h + \phi a_1 + a_2 \\ Z_3 &= \theta_0(1 + \phi + \phi^2) + \phi^3 h + \phi^2 a_1 + \phi a_2 + a_3 \\ &\dots \dots \dots \\ Z_t &= \theta_0 \sum_{i=0}^{t-1} \phi^i + \phi^t h + \sum_{i=0}^{t-1} \phi^i a_{t-i} \end{aligned}$$

Si se calcula la esperanza de  $Z_t$  y como  $E[a_t] = 0$ , entonces

$$E[Z_t] = \theta_0 \sum_{i=0}^{t-1} \phi^i + \phi^t h$$

Para que el proceso sea estacionario en media, es necesario que esta función no dependa de  $t$ . Esto ocurrirá si el primer término converge a una constante y el segundo a cero, de manera que para  $t$  grande el proceso tenga media constante. Ambas condiciones se verifican si  $|\phi| < 1$ , y de esta manera la suma forma una progresión geométrica indefinida de razón  $\phi$  y converge a  $\frac{\theta_0}{1-\phi}$  y el segundo término a cero. Con esta condición, después de un período transitorio inicial, cuando  $t \rightarrow \infty$ ,  $\phi^t$  tenderá a cero y todas las variables  $Z_t$  tendrán la misma esperanza,  $\mu = \frac{\theta_0}{1-\phi}$ , independientemente de las condiciones iniciales. Lo anterior lleva a que si se toma la esperanza en (2.10) y suponiendo que  $|\phi| < 1$  (de manera que  $E[Z_t] = E[Z_{t-1}] = \mu$ ), se tiene que:

$$\begin{aligned} \mu &= \theta_0 + \phi \mu \\ \Rightarrow \mu &= \frac{\theta_0}{1 - \phi} \end{aligned}$$

Ahora, sustituyendo  $\theta_0$  por  $\mu(1 - \phi)$  y llamando  $\tilde{Z}_t = Z_t - \mu$ , el proceso puede escribirse en desviaciones a su media:

$$\tilde{Z}_t = \phi\tilde{Z}_{t-1} + a_t \quad (2.11)$$

que es la expresión más utilizada del  $AR(1)$ .

Elevando al cuadrado (2.11) y tomando esperanzas, se tiene que:

$$\sigma_Z^2 = E(\tilde{Z}_t^2) = \phi^2 E(\tilde{Z}_{t-1}^2) + 2\phi E(\tilde{Z}_{t-1}a_t) + \sigma_a^2$$

Llamando  $\sigma_Z^2$  la varianza del proceso, y como  $a_t$  es un proceso de ruido blanco, se tiene que  $\tilde{Z}_{t-1}$  y  $a_t$  son independientes y la esperanza de su producto es el producto de las esperanzas y ambas variables tienen esperanza nula (recordar que  $\tilde{Z}_{t-1} = Z_{t-1} - \mu$ ). Por lo que, suponiendo estacionariedad, se concluye que:

$$\sigma_Z^2 = \phi^2 \sigma_Z^2 + \sigma_a^2$$

de donde se obtiene que la varianza del proceso es,

$$\sigma_Z^2 = \frac{\sigma_a^2}{1 - \phi^2} \quad (2.12)$$

Obsérvese que en esta ecuación, aparece de nuevo la condición  $|\phi| < 1$  para que  $\sigma_Z^2$  sea finita y positiva. Por otro lado, se tiene que, para un valor fijo de  $\sigma_a^2$ , la varianza es tanto mayor cuanto mayor sea  $\phi^2 \leq 1$ .

Por otro lado, utilizando (2.11) y escribiendo:

$$Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t$$

multiplicando por  $Z_{t-k}$ , tomando esperanzas y llamando  $\gamma_k$  a la covarianza entre observaciones separadas por  $k$  períodos, o autocovarianza de orden  $k$ :



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

Por tanto, se obtiene:

$$\gamma_k = \phi \gamma_{k-1} \quad k = 1, \dots, p$$

entonces, como  $|\phi| < 1$  la dependencia entre observaciones se amortiguan al aumentar el retardo. En particular, sabiendo que  $\gamma_0 = \sigma^2$ , utilizando (2.12),

$$\gamma_1 = \frac{\phi \sigma_a^2}{1 - \phi^2} \quad (2.13)$$

Llamando  $\rho_k$  a la correlación entre observaciones separadas por  $k$  períodos, o autocorrelación de orden  $k$ , definida por:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi \gamma_{k-1}}{\gamma_0} = \phi \rho_{k-1}$$

y como  $\rho_1 = \phi$ , utilizando (2.12) y (2.13), se tiene que esta función será  $\rho_k = \phi^k$

Los valores de la función de autocorrelación son las sucesivas potencias de  $\phi$ . La condición  $|\phi| < 1$  garantiza que los sucesivos valores de  $\rho_k$  converjan a cero, por lo que la función de autocorrelación o correlograma puede tener dos aspectos distintos, dependiendo del signo de  $\phi$ .

A continuación se presenta la simulación de este proceso y su correspondiente correlograma, para observar que su comportamiento depende del signo de  $\phi$ .

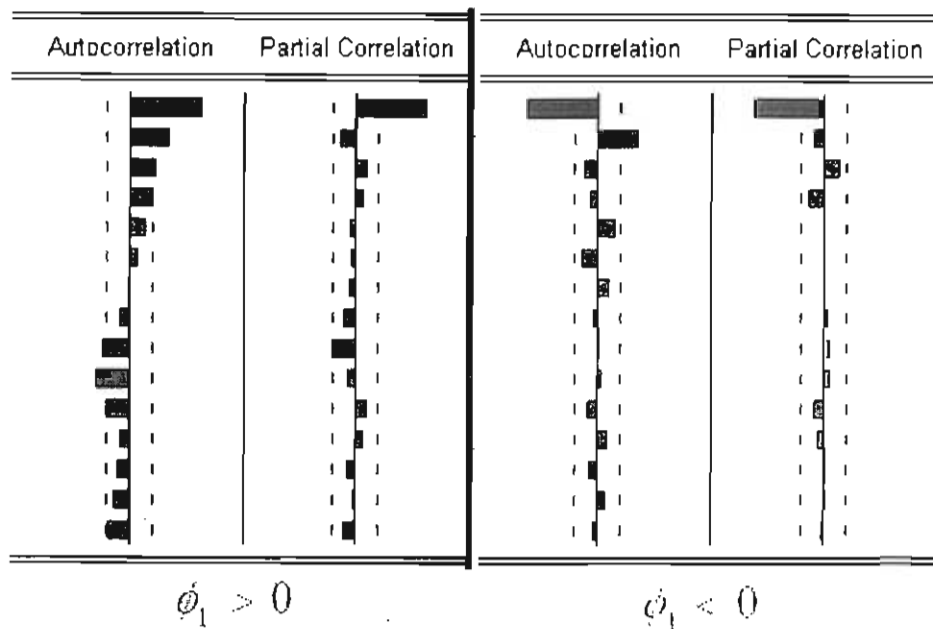


Fig. 2.4. Proceso Autorregresivo de orden 1.

### 2.6.1.1. Representación del proceso $AR(1)$ como suma de choques aleatorios <sup>(4)</sup>.

El proceso  $AR(1)$  puede escribirse utilizando la notación del operador de retardo como:

$$(1 - \phi B)\tilde{Z}_t = a_t \quad (2.14)$$

Esta condición indica que una serie sigue un proceso  $AR(1)$  si al aplicarle el operador  $(1 - \phi B)$  se obtiene un proceso de ruido blanco. El operador  $(1 - \phi B)$  puede interpretarse como un filtro que, aplicado a la serie, la convierte en ruido blanco. Si se considera el operador como una ecuación en  $B$  el coeficiente  $\phi$  se denomina factor de la ecuación. La condición de estacionaridad es que este factor sea menor que la unidad en valor absoluto. Alternativamente se puede hablar de la raíz de la ecuación del operador, que se obtiene igualando el operador a cero y resolviendo la ecuación con  $B$  como incógnita:

$$1 - \phi B = 0$$

con el resultado  $B = \frac{1}{\phi}$ . La condición de estacionaridad requiere entonces que la raíz del operador sea en valor absoluto mayor que uno. El proceso  $AR(1)$  puede expresarse en

<sup>4</sup> Se entiende por choques aleatorios, aquellas fluctuaciones alrededor de un punto de equilibrio, que son causadas por eventos inesperados, los cuales pueden causar efectos después de transcurridos un cierto número de periodos y que se reflejan según sea el valor de su ponderación en el modelo.

función de los valores pasados de choques aleatorios. Sustituyendo en la expresión del proceso  $\tilde{Z}_{t-1}$  por su expresión, se tiene

$$\tilde{Z}_t = \phi(\phi\tilde{Z}_{t-2} + a_{t-1}) + a_t = a_t + \phi a_{t-1} + \phi^2 \tilde{Z}_{t-2}$$

y aplicando repetidamente esta sustitución se tiene que

$$\tilde{Z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots$$

donde la serie se representa como función de todos los choques aleatorios con pesos que decrecen geoméricamente. Esta representación se denomina la forma de Media Móvil de orden infinito ( $MA(\infty)$ ), la cual se verá más adelante. Se observa que los coeficientes de estos choques son precisamente los coeficientes de la función de autocorrelación simple. La expresión  $MA(\infty)$  puede obtenerse también directamente multiplicando la ecuación (2.14) por el operador  $(1 - \phi B)^{-1}$

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

## 2.6.2. El proceso autorregresivo de segundo orden $AR(2)$ .

La dependencia que establece un proceso  $AR(1)$  puede generalizarse permitiendo que  $Z_t$  dependa linealmente, no sólo de  $Z_{t-1}$  sino también de  $Z_{t-2}$ . Se obtiene entonces el proceso autorregresivo de segundo orden o proceso  $AR(2)$ :

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t$$

donde ahora  $\theta_0, \phi_1, \phi_2$  son constantes a determinar y  $a_t$  es un proceso de ruido blanco independiente de la historia de  $Z_t$ . Tomando esperanzas en la ecuación del proceso e imponiendo que la media sea constante, se obtiene:

$$\mu = \theta_0 + \phi_1 \mu + \phi_2 \mu$$

que implica

$$\mu = \frac{\theta_0}{1 - \phi_1 - \phi_2}$$

y la condición para que el proceso tenga media finita es que  $1 - \phi_1 - \phi_2 \neq 0$ . Despejando  $\theta_0$  de la ecuación anterior y llamando  $\tilde{Z}_t = Z_t - \mu$  al proceso en desviaciones a su media, se puede escribir al proceso  $AR(2)$  de la siguiente manera:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t \quad (2.15)$$

y utilizando la notación de operadores:

$$(1 - \phi_1 B - \phi_2 B^2) \tilde{Z}_t = a_t$$

Como anteriormente se expuso, el operador  $(1 - \phi_1 B - \phi_2 B^2)$  puede siempre escribirse como  $(1 - G_1 B)(1 - G_2 B)$ , donde  $G_1^{-1}$  y  $G_2^{-1}$  son las raíces de la ecuación del operador; considerando  $B$  como variable y resolviendo  $1 - \phi_1 B - \phi_2 B^2 = 0$  (ecuación característica). También se dice que  $G_1$  y  $G_2$  son los factores del polinomio característico del proceso.

Estas raíces pueden ser reales o complejas conjugadas y para que se cumpla la condición de estacionaridad se debe satisfacer que  $|G_i| < 1$ ,  $i = 1, 2$ . Esta condición es análoga a la vista en el  $AR(1)$ . Se observa que este resultado es coherente con la condición encontrada para que la media sea finita. Si la ecuación  $1 - \phi_1 B - \phi_2 B^2 = 0$  tiene una raíz unitaria, se verifica que  $1 - \phi_1 - \phi_2 = 0$  y con esto el proceso no es estacionario (pues depende de  $t$ ) y no tiene media finita.

Por otro lado, elevando al cuadrado la expresión del proceso (2.15), y aplicando las mismas propiedades para el proceso  $AR(1)$ , se obtiene que varianza debe satisfacer (recordando que  $\sigma^2 = \gamma_0$ ).

$$\gamma_0 = \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma_a^2 \quad (2.16)$$

Para calcular las autocovarianzas de este proceso, multiplicando la ecuación del proceso (2.16) por  $\tilde{Z}_{t-k}$  y tomando esperanzas, resulta la expresión general de las autocovarianzas:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \quad k \geq 1 \quad (2.17)$$

Particularizando esta ecuación para  $k = 1$ , como  $\gamma_{-1} = \gamma_1$ , se obtiene que

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$$

que proporciona,

$$\gamma_1 = \frac{\phi_1 \gamma_0}{1 - \phi_2}$$

Sustituyendo esta expresión en (2.16), se obtiene la expresión para la varianza:

$$\gamma_0 = \frac{(1 - \phi_2) \sigma_a^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)}$$

Para que la varianza sea positiva, los parámetros del proceso deben ser tales que el numerador y el denominador tengan el mismo signo. Puede comprobarse que los valores de los parámetros que hacen la varianza siempre positiva son los incluidos en la región,

$$\begin{aligned} -1 &< \phi_2 < 1 \\ \phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1 \end{aligned}$$

y la figura 2.5 representa estos valores admisibles de los parámetros para que el proceso sea estacionario. El resto de las autocovarianzas se calculan recursivamente con la ecuación (2.17).

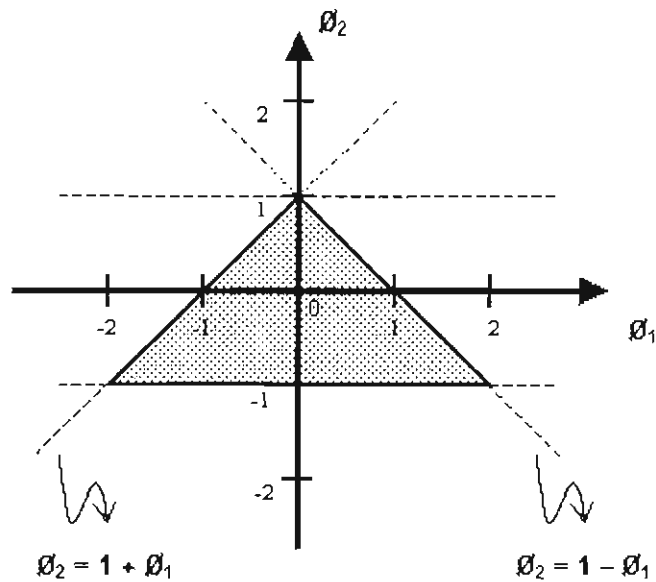


Fig. 2.5. Región de admisibilidad para los parámetros de un modelo AR(2).

Ahora, dividiendo la ecuación (2.17) por la varianza, se obtiene que las relaciones entre los coeficientes de autocorrelación son:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad k \geq 1 \quad (2.18)$$

particularizando (2.18), para  $k = 1$  y como en un proceso estacionario  $\rho_1 = \rho_{-1}$  se obtiene que:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \quad (2.19)$$

Para  $k = 2$  se obtiene,

$$\rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$$

Para  $k \geq 3$ , los coeficientes de autocorrelación pueden obtenerse recursivamente a partir de la ecuación en diferencias (2.18). En la sección 2.5.2. se demostró que la solución general de una ecuación de este tipo es:

$$\rho_k = A_1 G_1^k + A_2 G_2^k \quad (2.20)$$

donde  $G_1$  y  $G_2$  son los factores del polinomio característico del proceso y  $A_1$  y  $A_2$  son constantes a determinar a partir de las condiciones iniciales  $\rho_0 = 1$ , (que implica que  $A_1 + A_2 = 1$ ) y  $\rho_1 = \frac{\phi_1}{1-\phi_2}$  que se determinaron anteriormente. Según (2.20) los coeficientes  $\rho_k$  serán menores o iguales a la unidad si  $|G_1| < 1$  y  $|G_2| < 1$ , que son las condiciones de estacionariedad del proceso. Si los factores  $G_1$  y  $G_2$  son complejos del tipo  $a \pm bi$ , siendo  $i = \sqrt{-1}$ , entonces la condición se convierte en  $\sqrt{a^2 + b^2} < 1$ . Se puede entonces, encontrarse con los casos siguientes:

1. Los dos factores  $G_1$  y  $G_2$  son reales. El decrecimiento de (2.20) será la suma de dos exponenciales y la forma del correlograma dependerá de que  $G_1$  y  $G_2$  tengan signo igual u opuesto.
2. Los dos factores  $G_1$  y  $G_2$  son complejos y conjugados. Entonces se puede demostrar que la función  $\rho_k$  decrecerá de forma senosoidal (pues en la sección 2.5.2., en el caso 3, se vio que si los factores son complejos, es mejor tratar a las soluciones en sus expresiones polares, la cuales involucran términos senosoidales).

Los cuatro tipos de correlogramas posibles para un  $AR(2)$  se presentan en la fig 2.6., tanto para la *función de autocorrelación simple (fas)*, como para la *función de autocorrelación parcial (fap)*.

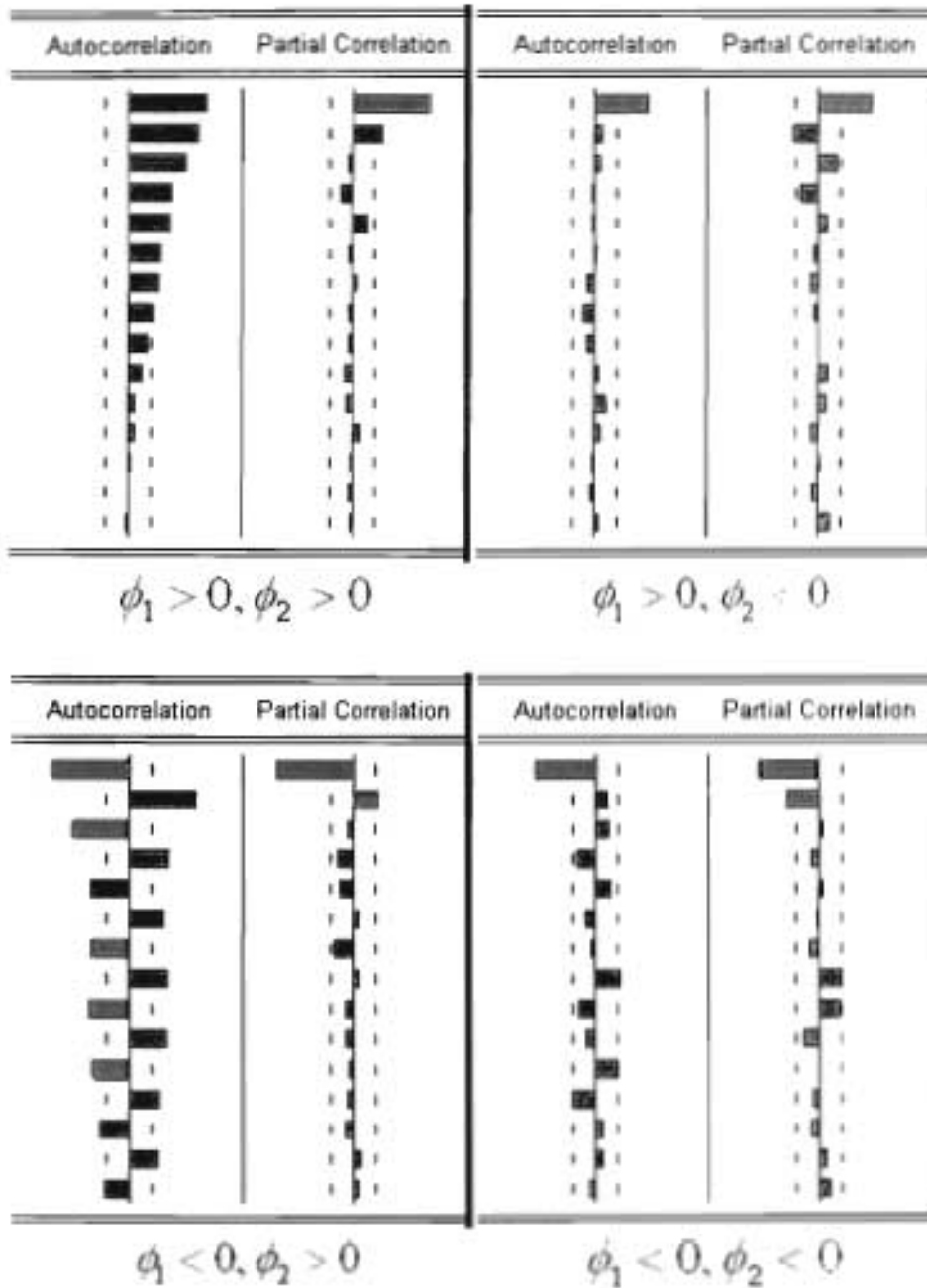


Fig. 2.6. Proceso Autorregresivo de orden 2.



### 2.6.2.1. Representación del proceso AR(2) como suma de choques aleatorios.

El proceso  $AR(2)$  puede representarse, de forma análoga al  $AR(1)$ , como combinación lineal de choques aleatorios, escribiendo el proceso como,

$$(1 - G_1B)(1 - G_2B)\tilde{Z}_t = a_t \quad (2.21)$$

e invirtiendo estos operadores, se tiene que:

$$\tilde{Z}_t = (1 + G_1B + G_1^2B^2 + \dots)(1 + G_2B + G_2^2B^2 + \dots)a_t,$$

que conducirá a una expresión de la forma

$$\tilde{Z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (2.22)$$

que, como más adelante se detallará, se denomina la forma  $MA(\infty)$  del proceso. Es posible obtener los coeficientes  $\psi_i$  como función de las raíces, igualando potencias de  $B$  en (2.21) y (2.22), es decir, ya sabiendo que  $(1 - \phi_1B - \phi_2B^2) = (1 - G_1B)(1 - G_2B)$  y utilizando la notación de operadores  $B$  en (2.22), se pueden igualar las ecuaciones para obtener:

$$(1 - \phi_1B - \phi_2B^2)(1 + \psi_1B + \psi_2B^2 + \dots) = 1 \quad (2.23)$$

Esta ecuación permite obtener los coeficientes  $\psi_i$  como función de los parámetros  $\phi_1$  y  $\phi_2$ , imponiendo la condición de que todos los coeficientes de las potencias de  $B$  en (2.23) sean nulos. El coeficiente de  $B$  en esta ecuación es  $\psi_1 - \phi_1$ . El coeficiente de  $B^2$  es  $\psi_2 - \phi_1\psi_1 - \phi_2$ , que generalizando se tiene la relación general:

$$\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2} \quad (2.24)$$

Los coeficientes de  $B^k$  para  $k \geq 2$  verifican la ecuación (2.24) que es similar a la que deben verificar los coeficientes de autocorrelación. Por lo que se concluye que la forma de los coeficientes  $\psi_i$  será similar a los de autocorrelación, aunque los valores iniciales no

serán idénticos porque los valores iniciales de la secuencia son distintos: para los coeficientes de autocorrelación  $\rho_0 = 1$  y  $\rho_1 = \frac{\phi_1}{1-\phi_2}$ , mientras que para los coeficientes  $\psi_i$ ,  $\psi_0 = 1$  y  $\psi_1 = \phi_1$ .

### 2.6.3. El proceso autorregresivo general $AR(p)$ .

Un proceso autorregresivo de orden  $p$ , denotado por  $AR(p)$ , está dado por la ecuación:

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.25)$$

en donde  $\phi_i$ ,  $i = 1, \dots, p$  son constantes a determinar y  $a_t$  un proceso de ruido blanco independiente de  $Z_{t-k}$  para todo  $k$  positivo. A primera vista podría parecer un modelo de regresión múltiple, sin embargo  $Z_t$  no está explicado en términos de variables independientes, sino en términos de ella misma, de sus propios valores pasados.

Si  $Z_t$  es un proceso estable en la media, al aplicar la función de medias a (2.25), se obtiene el valor de  $\theta_0$ ,

$$\begin{aligned} \mu &= \theta_0 + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu \\ \theta_0 &= (1 - \phi_1 - \phi_2 - \dots - \phi_p) \mu \end{aligned} \quad (2.26)$$

Sustituyendo (2.26) en (2.25) y de nuevo llamando  $\tilde{Z}_t = Z_t - \mu$ , se tiene el proceso expresado en desviaciones a su media, la cual es la expresión más usada del  $AR(p)$ .

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

La función de autocorrelación entre  $Z_{t-k}$  y  $Z_t$  está dada por:

$$\begin{aligned} E[(Z_{t-k} - \mu)(Z_t - \mu)] &= E(\tilde{Z}_{t-k} \tilde{Z}_t) \\ &= E[\tilde{Z}_{t-k} (\phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t)] \\ &= \phi_1 E(\tilde{Z}_{t-k} \tilde{Z}_{t-1}) + \phi_2 E(\tilde{Z}_{t-k} \tilde{Z}_{t-2}) + \dots + \phi_p E(\tilde{Z}_{t-k} \tilde{Z}_{t-p}) + E(\tilde{Z}_{t-k} a_t) \end{aligned}$$

y por ser  $a_t$  independiente de  $Z_{t-k}$  para  $k > 0$  y  $\gamma_k$  una función par, se tiene:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k > 0$$

y dividiendo por la varianza del proceso, se llega a:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k > 0 \quad (2.27)$$

cuya solución, según (2.8), es:

$$\rho_k = \sum_{i=1}^p A_i G_i^k \quad (2.28)$$

en donde  $A_i$  son constantes a determinar a partir de las condiciones iniciales y los  $G_i$  son los factores de la ecuación característica. Para que el proceso sea estacionario, el módulo de  $G_i$  debe ser menor que uno, o lo que es lo mismo, las raíces de la ecuación característica deben ser mayores que uno en módulo. Para comprobarlo, obsérvese que la condición  $|\rho_k| < 1$  exige que no exista en (2.28) ningún  $G_i$  mayor que la unidad, pues entonces, cuando  $k$  aumenta, el término  $G_i^k$  crecería sin límite. Se observa además, que para que el proceso sea estacionario no puede existir una raíz  $G_i$  igual a la unidad, ya que entonces su componente  $G_i^k$  no decrecería y los coeficientes  $\rho_k$  no tenderían a cero para ningún retardo<sup>(5)</sup>.

Para poder determinar las constantes  $A_1, \dots, A_p$  se requerirá de  $p$  condiciones iniciales que se pueden obtener mediante el conocimiento de  $\phi_i$ ,  $i = 1, \dots, p$ , además de resolver el siguiente sistema de ecuaciones, que resulta de particularizar (2.27) para  $k = 1, \dots, p$ .

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

A este sistema de ecuaciones se le conoce como **ecuaciones de Yule - Walker**. De igual manera, si de antemano se conocen los coeficientes de autocorrelación  $\rho_k$  para  $k = 1, \dots, p$ , utilizando el mismo sistema, quedarán determinados los coeficientes  $\phi_i$ ,  $i = 1, \dots, p$ .

---

<sup>5</sup> La ecuación (2.28) muestra que la función de autocorrelación de un proceso  $AR(p)$  es una mezcla de funciones exponenciales (debido a los términos con raíces reales) y senosoidales (por las raíces complejas conjugadas que presenta). Su estructura puede ser, en consecuencia, muy compleja.

Para ilustrar estos procedimientos considérese el proceso  $AR(2)$ ,

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

con función de autocorrelación

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad (2.29)$$

cuya solución general está dada por

$$\rho_k = A_1 G_1^k + A_2 G_2^k \quad (2.30)$$

en donde  $A_1$  y  $A_2$  serán determinadas como sigue: particularizando (2.30) para  $k = 0, 1, 2$ , se tiene

$$\rho_0 = A_1 + A_2 = 1$$

$$\rho_1 = A_1 G_1 + A_2 G_2$$

y de la primera ecuación de Yule - Walker

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$

lo que implica que

$$\frac{\phi_1}{1-\phi_2} = A_1 G_1 + (1 - A_1) G_2$$

despejando  $A_1$  de la ecuación anterior y  $A_2$  de la igualdad de  $\rho_0$ , se obtiene:

$$A_1 = \frac{\left[ \frac{\phi_1}{1-\phi_2} - G_2 \right]}{(G_1 - G_2)}$$

$$A_2 = 1 - A_1$$

Tomando en cuenta todo lo anterior, se concluye que  $\rho_k$  decrecerá conforme el retardo  $k$  aumente; lo hará de forma exponencial si  $G_1$  y  $G_2$  son reales y en forma senoidal si  $G_1$  y  $G_2$  son complejas.

## 2.7. Función de autocorrelación parcial (*fap*).

Determinar el orden de un proceso autorregresivo a partir de la función de autocorrelación simple es difícil. En general esta función es una mezcla de decrecimientos exponenciales y senosoidales, que se amortiguan al avanzar el retardo, y no presenta rasgos fácilmente identificables para determinar el orden del proceso. Para resolver este problema se introduce la función de autocorrelación parcial.

Si se compara un  $AR(1)$  con un  $AR(2)$  se observa que aunque en ambos procesos cada observación está relacionada con las anteriores, el tipo de relación entre observaciones separadas por más de un retardo es distinto en ambos procesos. En el  $AR(1)$  el efecto de  $Z_{t-2}$  sobre  $Z_t$  es siempre a través de  $Z_{t-1}$ , y dado  $Z_{t-1}$  es irrelevante el valor de  $Z_{t-2}$  para prever  $Z_t$ . Sin embargo, en un  $AR(2)$  además del efecto de  $Z_{t-2}$  que se transmite a  $Z_t$  a través de  $Z_{t-1}$ , existe un efecto de  $Z_{t-2}$  sobre  $Z_t$ . En general, un  $AR(p)$  presenta efectos *directos* de observaciones separadas por  $1, 2, \dots, p$  retardos y los efectos *directos* de las observaciones separadas por  $p + 1, \dots$  son nulos. Esta idea es la clave para la utilización de la función de autocorrelación parcial.

La labor de la función de autocorrelación parcial consiste en medir la relación lineal que existe entre  $Z_t$  y  $Z_{t-k}$ ,  $k > 0$ , eliminando de antemano los efectos de las variables intermedias  $Z_{t-1}, \dots, Z_{t-k+1}$ . Cuando la función es diferente de cero indica que existe un efecto directo entre  $Z_t$  y  $Z_{t-k}$  que no se transmite a través de las variables  $Z_{t-1}, \dots, Z_{t-k+1}$ , y lo más adecuado será considerar a esta variable en el modelo.

De esta manera, un proceso  $AR(p)$  tendrá los primeros  $p$  coeficientes de autocorrelación parcial distintos de cero, y serán cero los coeficientes asociados a  $p + 1, p + 2, \dots$

El coeficiente de correlación parcial entre  $\tilde{Z}_{t-1}$  y  $\tilde{Z}_{t-k}$  se obtiene de la siguiente manera:

1. Se calcula de  $\tilde{Z}_t$  el efecto de  $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots, \tilde{Z}_{t-k+1}$  mediante la regresión

$$\tilde{Z}_t = \alpha_1 \tilde{Z}_{t-1} + \alpha_2 \tilde{Z}_{t-2} + \dots + \alpha_{k-1} \tilde{Z}_{t-k+1} + u_t \quad (2.31)$$

en donde  $u_t$  recoge la parte de  $\tilde{Z}_t$  no común con  $\tilde{Z}_{t-1}, \dots, \tilde{Z}_{t-k+1}$ . Procediendo de igual manera que en la sección anterior, se multiplica (2.31) de ambos lados por  $\tilde{Z}_{t-i}$ ,  $i > 0$  y aplicando el operador esperanza se obtiene:

$$\gamma_i = \alpha_1 \gamma_{i-1} + \alpha_2 \gamma_{i-2} + \dots + \alpha_{k-1} \gamma_{i-k+1} \quad (2.32)$$

Por lo que las ecuaciones de Yule - Walker serán:

$$\rho_i = \alpha_1 \rho_{i-1} + \alpha_2 \rho_{i-2} + \dots + \alpha_{k-1} \rho_{i-k+1}$$

que se obtienen de dividir (2.32) por  $\gamma_0$ . Particularizando para  $i = 1, 2, \dots, k-1$  se obtiene el siguiente sistema de ecuaciones:

$$\begin{aligned} \rho_1 &= \alpha_1 \rho_0 + \alpha_2 \rho_1 + \dots + \alpha_{k-1} \rho_{k-2} \\ \rho_2 &= \alpha_1 \rho_1 + \alpha_2 \rho_0 + \dots + \alpha_{k-1} \rho_{k-3} \\ &\vdots \\ \rho_{k-1} &= \alpha_1 \rho_{k-2} + \alpha_2 \rho_{k-3} + \dots + \alpha_{k-1} \rho_0 \end{aligned}$$

que en términos matriciales se puede escribir de la siguiente manera:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \dots & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{k-1} \end{bmatrix} \quad (2.33)$$

A través de este sistema de ecuaciones y conociendo los valores de  $\rho_i$  se pueden obtener los parámetros  $\alpha_i$ ,  $i = 1, \dots, k-1$ . Los valores  $\rho_i$  pueden conocerse resolviendo el sistema de ecuaciones Yule Walker, descrito anteriormente.

2. De igual manera, mediante la regresión

$$\tilde{Z}_{t-k} = \beta_1 \tilde{Z}_{t-k+1} + \beta_2 \tilde{Z}_{t-k+2} + \dots + \beta_{k-1} \tilde{Z}_{t-1} + v_t$$

se eliminará de  $\tilde{Z}_{t-k}$  el efecto de  $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots, \tilde{Z}_{t-k+1}$ . De nuevo  $v_t$  contendrá la parte de  $\tilde{Z}_{t-k}$  no común con  $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots, \tilde{Z}_{t-k+1}$ . Y resolviendo el sistema:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \dots & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{k-1} \end{bmatrix} \quad (2.34)$$

se pueden conocer los parámetros  $\beta_i$ ,  $i = 1, \dots, k$ . Obsérvese que si  $\alpha_i$  satisface (2.33), entonces también lo hace con (2.34), con lo cual  $\alpha_i = \beta_i$  ( $1 \leq i \leq k-1$ ).

3. El coeficiente de correlación simple entre  $u_t$  y  $v_t$  será el coeficiente de autocorrelación parcial de orden  $k$ , que se denotará por  $P_k$ .

$$P_k = \frac{Cov(u_t, v_t)}{\sqrt{Var(u_t)Var(v_t)}}$$

y de (2.31) se obtiene  $Var(u_t)$ ,

$$\begin{aligned}
Var(u_t) &= Var(\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \\
&= E\left[ (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] \\
&= E\left[ \tilde{Z}_t (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] - \alpha_1 E\left[ \tilde{Z}_{t-1} (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] \\
&\quad - \dots - \alpha_{k-1} E\left[ \tilde{Z}_{t-k+1} (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] \\
&= \gamma_0 - \alpha_1 \gamma_1 - \alpha_2 \gamma_2 - \dots - \alpha_{k-1} \gamma_{k-1} - \alpha_1 (\gamma_1 - \alpha_1 \gamma_0 - \alpha_2 \gamma_1 - \dots - \alpha_{k-1} \gamma_{k-2}) - \dots \\
&\quad - \alpha_{k-1} (\gamma_{k-1} - \alpha_1 \gamma_k - \alpha_2 \gamma_{k-1} - \dots - \alpha_{k-1} \gamma_0)
\end{aligned}$$

el término  $\alpha_1 (\gamma_1 - \alpha_1 \gamma_0 - \alpha_2 \gamma_1 - \dots - \alpha_{k-1} \gamma_{k-2})$  de esta última expresión será igual a cero, ya que para  $i = 1$  se obtiene de (2.32) la igualdad a cero, y así se particulariza para  $i = 2, 3, \dots, k-1$ , se obtiene finalmente la expresión para  $Var(u_t)$ .

$$Var(u_t) = \gamma_0 - \alpha_1 \gamma_1 - \alpha_2 \gamma_2 - \dots - \alpha_{k-1} \gamma_{k-1}$$

Y como  $\alpha_i = \beta_i$ , se tiene que  $Var(u_t) = Var(v_t)$ ,  $i = 1, 2, \dots, k-1$ . A continuación se desarrolla la  $Cov(u_t, v_t)$ :

$$Cov(u_t, v_t) = E\left[ (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) (\tilde{Z}_{t-k} - \beta_1 \tilde{Z}_{t-k+1} - \beta_2 \tilde{Z}_{t-k+2} - \dots) \right]$$

haciendo uso de la igualdad  $\alpha_i = \beta_i$ , ( $1 \leq i \leq k-1$ ), se tiene:

$$\begin{aligned}
Cov(u_t, v_t) &= E\left[ (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) (\tilde{Z}_{t-k} - \alpha_1 \tilde{Z}_{t-k+1} - \alpha_2 \tilde{Z}_{t-k+2} - \dots) \right] \\
&= E\left[ \tilde{Z}_{t-k} (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] - \alpha_1 E\left[ \tilde{Z}_{t-k+1} (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] \\
&\quad - \dots - \alpha_{k-1} E\left[ \tilde{Z}_{t-1} (\tilde{Z}_t - \alpha_1 \tilde{Z}_{t-1} - \alpha_2 \tilde{Z}_{t-2} - \dots - \alpha_{k-1} \tilde{Z}_{t-k+1}) \right] \\
&= \gamma_k - \alpha_1 \gamma_{k-1} - \alpha_2 \gamma_{k-2} - \dots - \alpha_{k-1} \gamma_1 - \alpha_1 (\gamma_{k-1} - \alpha_1 \gamma_{k-2} - \dots - \alpha_{k-1} \gamma_0) - \dots \\
&\quad - \alpha_{k-1} (\gamma_1 - \alpha_1 \gamma_0 - \dots - \alpha_{k-1} \gamma_{k-2})
\end{aligned}$$

utilizando de nuevo la ecuación (2.32), se concluye:

$$Cov(u_t, v_t) = \gamma_k - \alpha_1 \gamma_{k-1} - \alpha_2 \gamma_{k-2} - \dots - \alpha_{k-1} \gamma_1$$

Por tanto,



$$\rho_k = \frac{\text{Cov}(u_t, v_t)}{\sqrt{\text{Var}(u_t)\text{Var}(v_t)}} = \frac{\text{Cov}(u_t, v_t)}{\text{Var}(u_t)}$$

$$\rho_k = \frac{\rho_k - \alpha_1 \rho_{k-1} - \alpha_2 \rho_{k-2} - \dots - \alpha_{k-1} \rho_1}{1 - \alpha_1 \rho_1 - \alpha_2 \rho_2 - \dots - \alpha_{k-1} \rho_{k-1}} \quad (2.35)$$

Para resolver el sistema de ecuaciones (2.33), se puede utilizar la regla de Cramer y así conocer los valores de  $\alpha_i$ , en términos de los coeficientes de autocorrelación  $\rho_j$ .

$$\alpha_1 = \frac{\begin{bmatrix} \rho_1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & 1 & \dots & \rho_{k-3} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho_{k-1} & \rho_{k-3} & \dots & 1 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_1 & 1 & \dots & \rho_{k-3} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix}}$$

$$\alpha_2 = \frac{\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_1 & \rho_2 & \dots & \rho_{k-3} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho_{k-2} & \rho_{k-1} & \dots & 1 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_1 & 1 & \dots & \rho_{k-3} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix}}$$

y en general:

$$\alpha_i = \frac{\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 & \rho_i & \dots & \rho_{k-2} \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_2 & \rho_{i-1} & \dots & \rho_{k-3} \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \rho_{k-2} & \rho_{k-3} & \dots & 1 & \rho_{k-1} & \rho_{k-i-2} & \dots & 1 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{i-2} & \rho_{i-1} & \rho_i & \dots & \rho_{k-2} \\ \rho_1 & 1 & \dots & \rho_{i-3} & \rho_{i-2} & \rho_{i-1} & \dots & \rho_{k-3} \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & & \cdot \\ \rho_{k-2} & \rho_{k-3} & \dots & \rho_{k-i} & \rho_{k-i-1} & \rho_{k-i-2} & \dots & 1 \end{bmatrix}}$$

sustituyendo  $\alpha_i$ ,  $i = 1, \dots, k-1$  en (2.35), se puede ver a  $P_k$  como la siguiente expresión:

$$P_k = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{bmatrix}} \quad (2.36)$$

Ahora se demostrará que  $P_k$  es igual a  $\phi_{kk}$ , en donde este último es el  $k$ -ésimo coeficiente de regresión del siguiente modelo.

$$\tilde{Z}_{t+k} = \phi_{k1}\tilde{Z}_{t+k-1} + \phi_{k2}\tilde{Z}_{t+k-2} + \dots + \phi_{kk}\tilde{Z}_t + e_{t+k}$$

en donde  $e_{t+k}$  es un término de error no correlacionado con  $\tilde{Z}_{t+k-j}$  con  $j \geq 1$ .

Nuevamente utilizando las ecuaciones de Yule-Walker y aplicando la regla de Cramer sucesivamente se obtienen los valores de  $\phi_{kk}$ ,

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

·  
·  
·

$$\phi_{kk} = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{bmatrix}} \quad (2.37)$$

Así comparando (2.36) con (2.37) se concluye que  $\phi_{kk}$  es igual a  $P_k$ .

En la práctica se requiere estimar a las autocorrelaciones parciales con los datos observados; para este fin se sustituyen en (2.37) los coeficientes teóricos  $\rho_k$  por sus estimaciones ( $r_k$ ), que no es más que el coeficiente obtenido con los valores de la muestra. Por esto mismo, debe tomarse muy en cuenta que la *fap* estimada está sujeta a errores de estimación, por lo que para decidir si la *fap* puede ser o no considerada como cero, se debe verificar mediante una aproximación sugerida por Quenouille<sup>(6)</sup>, para conocer la variación muestral respecto a su estimador  $\hat{\phi}_{ii}$ .

Esto sugiere que si el proceso es un  $AR(p)$ , entonces las autocorrelaciones se distribuyen como una normal, con media  $E(\hat{\phi}_{ii}) = \phi_{ii}$  y con

$$Var(\hat{\phi}_{ii}) = \frac{1}{N-d} \quad \text{para } i > p$$

Así, se puede definir el intervalo de confianza, para el cual se podrá concluir que  $\phi_{ii}$  es distinto de cero (con cierto nivel de significancia, sea éste 5%). Dicho intervalo se define por,

$$\pm 2\sqrt{Var(\hat{\phi}_{ii})} = \pm 2\sqrt{N-d} \quad \text{para } i > p$$

<sup>6</sup> Ver referencia 1 de la bibliografía, pp. 76.

y si  $\hat{\phi}_{ii}$  se encuentra fuera de este intervalo, se concluye que  $\phi_{ii}$  es distinto de cero.

Finalmente, es importante decir que así como un  $AR(p)$  tiene las  $p$  primeras autocorrelaciones parciales distintas de cero, de la misma forma, un  $MA(q)$  tendrá todas sus autocorrelaciones parciales distintas de cero, con convergencia a cero. A continuación se resume en el cuadro (2.1) los comportamientos de la  $FAS$  y  $FAP$  para una serie estacionaria, para efectos de una mejor identificación del proceso en cuestión. El proceso definido como  $ARMA(p, q)$  será abordado más adelante.

Proceso	FAS	FAP
$AR(p)$	Convergencia a cero, con comportamiento dado por la ecuación $\phi(B)\rho_k = 0$ , para $k \geq p$	Solamente las primeras $p$ autocorrelaciones parciales son distintas de cero.
$MA(q)$	Sólo las primeras $q$ autocorrelaciones son distintas de cero.	Sucesión infinita convergente a cero.
$ARMA(p, q)$	Comportamiento irregular de las primeras $q$ autocorrelaciones y después convergen a cero, de acuerdo con $\phi(B)\rho_k = 0$ , para $k \geq p$	Sucesión infinita convergente a cero.

Cuadro 2.1. Comportamiento de la  $FAS$  y  $FAP$  para distintos procesos.

## 2.8. Modelos de promedios móviles.

Anteriormente se estudiaron los procesos autorregresivos, que se caracterizan por tener muchos coeficientes de autocorrelación distintos de cero y que decrecen con el retardo. Es por esto que a los procesos  $AR$  se les conoce como procesos con memoria larga, ya que incorporan todos los choques aleatorios (innovaciones) que han generado, aunque con coeficientes decrecientes. Estos procesos no pueden representar series con memoria muy corta, que sólo tengan pocas autocorrelaciones distintas de cero. Una familia de procesos que tienen esta propiedad de “memoria corta” son los procesos de media móvil, los cuales se generan a partir de un número finito de choques aleatorios.

El término de promedios móviles sugeriría que el modelo se obtiene como un

promedio de choques aleatorios, y no es así, ya que para ser cierto se necesita que los parámetros fueran todos positivos y que su suma debiera ser uno, para ser un promedio. Como más adelante se verá, todo proceso  $MA$  es estacionario, lo cual sugiere que para este tipo de procesos que se encuentran en equilibrio, las fluctuaciones alrededor del punto de equilibrio,  $\{\tilde{Z}_t\}$ , son causadas por choques asociados con eventos inesperados; tales choques no necesariamente se asimilan de manera instantánea, sino que pueden seguir causando efectos aun después de haber transcurrido un cierto número de períodos y además, la intensidad del choque se refleja en el valor de su ponderación<sup>(5)</sup>.

### 2.8.1. El proceso de media móvil de primer orden $MA(1)$ .

Se define el proceso de media móvil de orden uno,  $MA(1)$ , como el generado por una combinación lineal de los dos últimos choques aleatorios, de acuerdo con la ecuación:

$$\tilde{Z}_t = a_t - \theta a_{t-1} \quad (2.38)$$

donde  $\tilde{Z}_t = Z_t - \mu$ , siendo  $\mu$  la media del proceso y  $a_t$  es un proceso de ruido blanco.

Este proceso puede escribirse con la notación de operadores:

$$\tilde{Z}_t = (1 - \theta B)a_t$$

El proceso  $MA(1)$  es la suma de los dos procesos estacionarios,  $a_t$  y  $a_{t-1}$  y por tanto, siempre será estacionario, para cualquier valor del parámetro, a diferencia de los procesos  $AR$ .

En las aplicaciones de este proceso se supondrá que  $|\theta| < 1$ , de manera que el choque aleatorio pasado tenga menos peso que la presente. Entonces, se dice que el proceso es *invertible* y tiene la propiedad de que el efecto de los valores pasados de la serie decrece con el tiempo. Para justificar esta propiedad, se sustituye  $a_{t-1}$  en (2.31) como función de  $z_{t-1}$ , el resultado es:

$$\tilde{Z}_t = a_t - \theta(\tilde{Z}_{t-1} + \theta a_{t-2}) = -\theta\tilde{Z}_{t-1} - \theta^2 a_{t-2} + a_t$$

y haciendo iterativamente la operación, se obtiene:

$$\tilde{Z}_t = - \sum_{i=1}^{t-1} \theta^i \tilde{Z}_{t-i} - \theta^t a_0 + a_t$$

Se observa que si  $|\theta| < 1$ , el efecto de  $\tilde{Z}_{t-k}$  tiende hacia cero con  $k$  y el proceso se denomina invertible. Si  $|\theta| \geq 1$  se daría la situación paradójica de que el efecto de las observaciones pasadas aumentaría con la distancia y, aunque el proceso seguiría siendo estacionario, parece poco adecuado para representar series reales. En adelante, se supondrá que el proceso es invertible. Entonces, como  $|\theta| < 1$ , existe el operador inverso  $(1 - \theta B)^{-1}$  y se puede escribir:

$$(1 + \theta B + \theta^2 B^2 + \dots) \tilde{Z}_t = a_t$$

y se ve que un proceso  $MA(1)$  con  $|\theta| < 1$  equivale a un  $AR(\infty)$  con coeficientes que decrecen en progresión geométrica.

La varianza del proceso se calcula a partir de (2.38), elevando al cuadrado y tomando esperanzas se obtiene:

$$\sigma_z^2 = \sigma^2(1 + \theta^2) \tag{2.39}$$

y se observa que la varianza del proceso aumenta con  $\theta^2$ .

Para obtener la función de autocorrelación simple de un  $MA(1)$  se calculan las covarianzas, multiplicando (2.38) por  $\tilde{Z}_{t-1}$  y tomando esperanzas, se tiene

$$\gamma_1 = E(\tilde{Z}_t \tilde{Z}_{t-1}) = E(a_t \tilde{Z}_{t-1}) - \theta E(a_{t-1} \tilde{Z}_{t-1})$$

En esta expresión el término  $E(a_t \tilde{Z}_{t-1})$  es cero, ya que  $\tilde{Z}_{t-1}$  depende de  $a_{t-1}$ ,  $a_{t-2}$ , pero no de  $a_t$ . Ahora, para calcular el otro sumando se emplea (2.38), es decir

$$E(a_{t-1}\tilde{Z}_{t-1}) = E(a_{t-1}(a_{t-1} - \theta a_{t-2})) = \sigma^2$$

$$\Rightarrow \gamma_1 = -\theta\sigma^2$$

Para  $j > 1$  se procede de igual forma, sólo que ahora se multiplica (2.38) por  $\tilde{Z}_{t-j}$ , se toman esperanzas y,

$$\gamma_j = 0, \quad j > 1$$

Por otro lado, utilizando (2.39), se concluye que los coeficientes de autocorrelación son:

$$\rho_1 = \frac{-\theta}{1 + \theta^2}, \quad \rho_k = 0, \quad k > 1$$

Con eso la *fas* tendrá únicamente un valor distinto de cero en el primer retardo. Además, teniendo la restricción de que  $|\theta| < 1$ , el coeficiente de autocorrelación será de 0.5, pues de otra forma indicaría que existiría una fuerte relación dependencia de la observación anterior con la actual, por lo cual sería mejor pensar en un modelo autorregresivo para esta situación.

Como anteriormente se mencionó la *fas* de un  $MA(1)$  tiene propiedades similares a la *fap* de un  $AR(1)$ , pues sólo existe un coeficiente de autocorrelación distinto de cero. Esta dualidad entre el  $AR(1)$  y  $MA(1)$  se presenta también en la *fap*. En efecto, si se observa la ecuación

$$(1 + \theta B + \theta^2 B^2 + \dots)\tilde{Z}_t = a_t$$

se puede ver que al escribir un  $MA(1)$  en forma autorregresiva, existe un efecto directo de  $\tilde{Z}_{t-k}$  sobre  $\tilde{Z}_t$  de magnitud  $\theta^k$ , para todo  $k$ . Por tanto, la función de autocorrelación parcial tenderá todos los coeficientes no nulos y que decrecen geométricamente con  $k$ . Ésta es precisamente la estructura de la *fas* de un  $AR(1)$  y por tanto, se concluye que la *fap* de un  $MA(1)$  tiene una estructura similar a la *fas* de un  $AR(1)$ .

A continuación se muestran la *fas* y *fap* del proceso de medias móviles de orden 1.



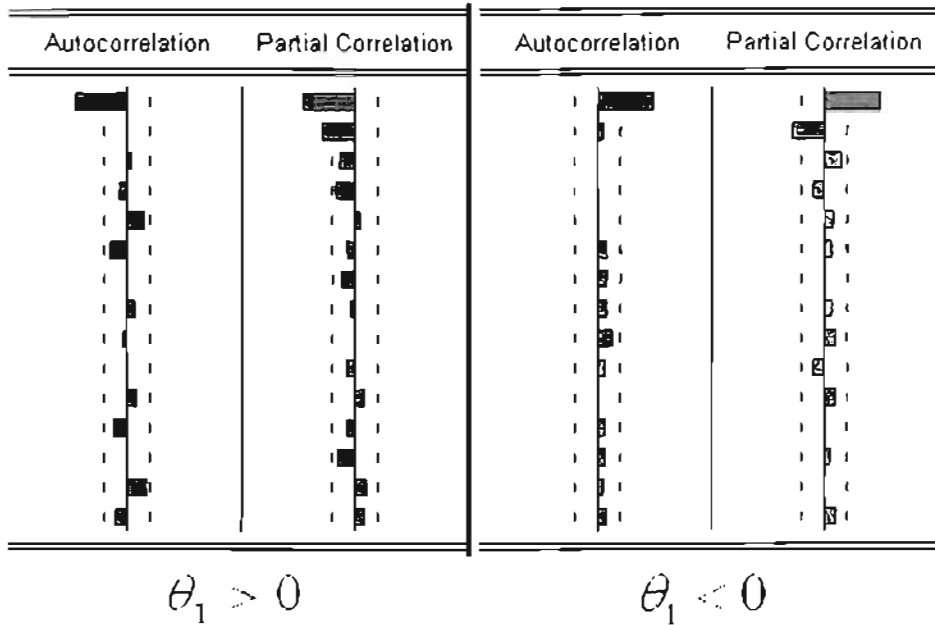


Fig. 2.7. Proceso de media móvil MA de orden 1, con coeficiente positivo y negativo.

### 2.8.2. El proceso de media móvil de segundo orden MA(2).

Se define un proceso de promedios móviles de orden 2 con el modelo

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2)a_t \tag{2.40}$$

y dicho proceso tiene como media y varianza a

$$E(\tilde{Z}_t) = E(Z_t) - \mu = 0 \quad \text{y} \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_a^2$$

y como función de autocovarianza a

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

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

por lo cual, la correspondiente *fas* viene a ser

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

De esta forma se observa (y se confirma) que el modelo es estacionario. Para que también sea invertible, *es necesario que las raíces de la siguiente ecuación estén fuera del círculo unitario.*

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

Y ya se había visto que para que esto suceda se debe cumplir que,

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

Se puede demostrar que las restricciones para los valores de las autocorrelaciones de un proceso  $MA(2)$  son que  $\rho_1^2 \leq 0.5$  y  $|\rho_2| \leq 0.5$ , lo cual puede observarse en las siguientes gráficas que presentan los distintos comportamientos de la *fas* y *fap* de estos procesos.

**ESTA TESIS NO SALE  
DE LA BIBLIOTECA**

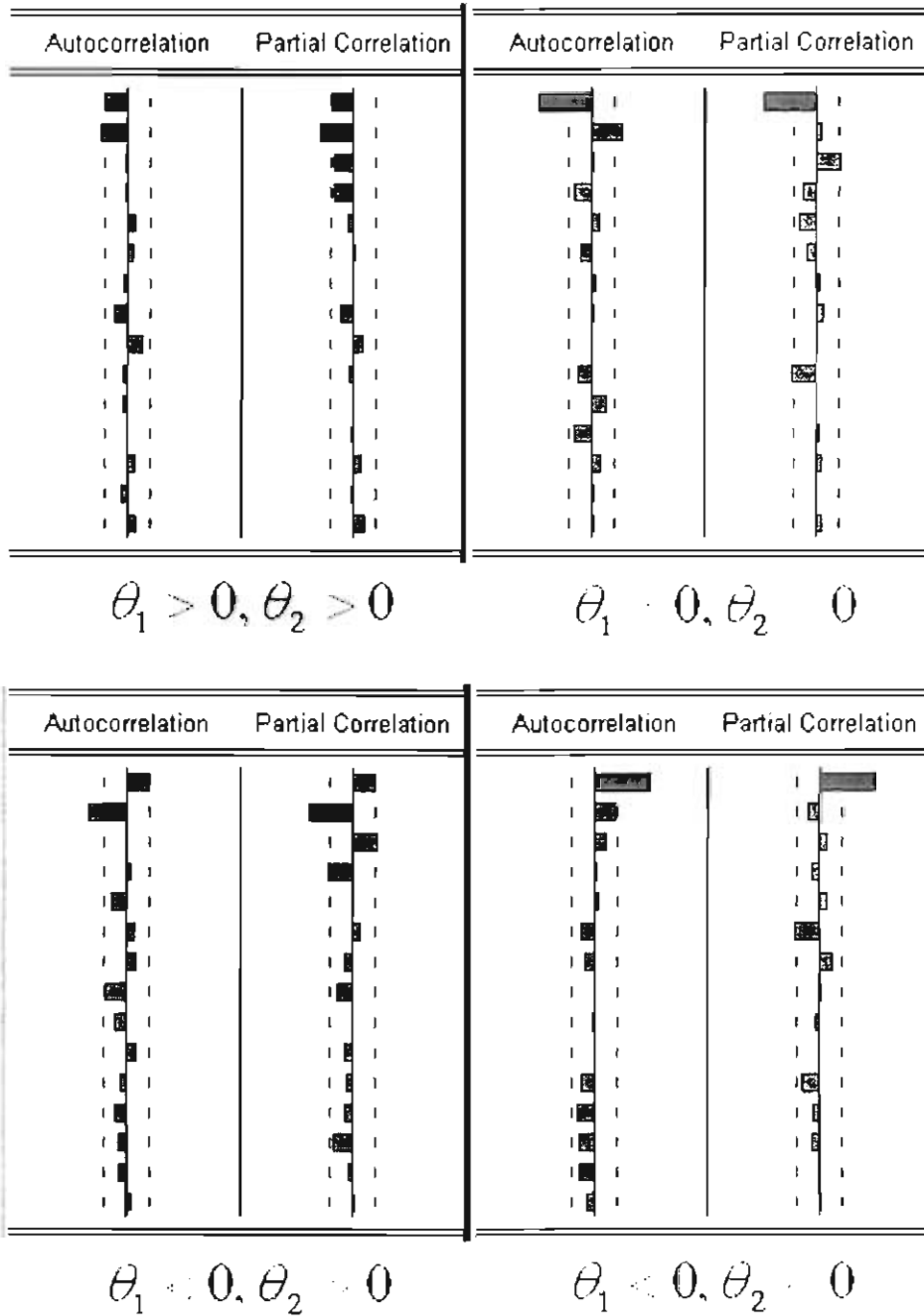


Fig 2.8. Simulación de procesos de medias móviles de orden 2, con coeficientes positivos y negativos.

### 2.8.3. El proceso de media móvil de orden general $MA(q)$ .

Un proceso de media móvil de orden  $q$  se define como:

$$Z_t = \theta_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.41)$$

en donde  $a_t$  es un proceso de ruido blanco y  $\theta_i$ ,  $i = 1, 2, \dots, q$  son constantes a determinar.

El valor de  $\theta_0$  se obtiene aplicando la función esperanza a (2.41)

$$\mu = E(Z_t) = E(\theta_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) = \theta_0$$

por tanto  $\theta_0$  representa a la media del proceso  $Z_t$  y así se puede representar a (2.41) en términos de sus desviaciones a la media por

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.42)$$

con  $\tilde{Z}_t = Z_t - \mu$

Debido a la descomposición de Wold, la cual se mencionó en el apartado 2.4, se puede ver también que todo proceso *MA* es estacionario, ya que es un caso particular de dicha representación, en donde

$$\Psi_0 = 1, \Psi_1 = -\theta_1, \dots, \Psi_q = -\theta_q, \Psi_{q+1} = 0, \Psi_{q+2} = 0, \dots,$$

cumpliéndose la condición de estacionariedad  $\sum_{i=0}^{\infty} \Psi_i^2 < \infty$ . Por otro lado, utilizando el operador de retardo en (2.42) se tiene que

$$\tilde{Z}_t = (1 - B\theta_1 - B^2\theta_2 - \dots - B^q\theta_q)a_t$$

Si se aplica la función esperanza y la función varianza, se obtiene

$$E(\tilde{Z}_t) = 0 \quad \text{y} \quad \gamma_0 = \text{Var}(\tilde{Z}_t) = \sigma_a^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)$$

que no dependen del tiempo. Para calcular la función de autocovarianzas, se multiplica (2.42) por  $\tilde{Z}_{t-k}$

$$\tilde{Z}_t \tilde{Z}_{t-k} = a_t \tilde{Z}_{t-k} - \theta_1 a_{t-1} \tilde{Z}_{t-k} - \dots - \theta_q a_{t-q} \tilde{Z}_{t-k}$$

se sustituye a  $\tilde{Z}_{t-k}$  en términos de  $a_t$

$$\begin{aligned} \tilde{Z}_t \tilde{Z}_{t-k} &= a_t (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q}) - \theta_1 a_{t-1} (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q}) - \dots \\ &\quad - \theta_q a_{t-q} (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q}) \end{aligned}$$

aplicando el operador esperanza,

$$\begin{aligned} E(\tilde{Z}_t \tilde{Z}_{t-k}) &= E[a_t (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q})] - \theta_1 E[a_{t-1} (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q})] \\ &\quad - \theta_q E[a_{t-q} (a_{t-k} - \theta_1 a_{t-k+1} - \dots - \theta_q a_{t-k+q})] \end{aligned}$$

y debido a que el proceso  $a_t$  es un proceso de ruido blanco, se tiene que  $E(a_t) = 0$ , para toda  $t$ . Entonces la función de autocovarianzas se define:

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

que dependen únicamente del retardo y de donde se sigue que la función de autocorrelación,

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

Anteriormente se hizo notar la dualidad que existe entre la estacionariedad de un *AR* y la invertibilidad para un *MA*; mientras que todo proceso *MA* es estacionario, todo proceso *AR* es invertible. Además, las condiciones para invertibilidad de un proceso *MA* pueden obtenerse de manera similar a las condiciones de estacionariedad de un *AR*. Por tanto, un proceso de media móvil será invertible si éste puede reescribirse como un proceso autorregresivo, esto debido a que la influencia de la variable retrasada disminuye con el

paso del tiempo.

La importancia del concepto de invertibilidad está en el hecho de que **todo proceso invertible está determinado de manera única por su FAS**, lo cual no ocurre con procesos que no son invertibles. Para aclarar esta propiedad tómense como ejemplo a dos procesos de media móvil de orden uno:

$$\begin{aligned} \text{a) } \tilde{Z}_t &= a_t - \theta a_{t-1} \\ \text{b) } \tilde{Z}_t &= a_t - \frac{1}{\theta} a_{t-1} \end{aligned}$$

Ya se había desarrollado el procedimiento para determinar la *fas* de a), la cual resultó ser

$$\rho_1 = \frac{-\theta}{1+\theta^2}, \quad \rho_k = 0, \quad k > 1$$

Y siguiendo el mismo procedimiento, se puede encontrar que la *fas* de b) es:

$$\rho_1 = \frac{-\frac{1}{\theta}}{1+\frac{1}{\theta^2}}, \quad \rho_k = 0, \quad k > 1$$

que es igual a la *fas* de a) y sin embargo solamente a) es un proceso invertible. Sustituyendo sucesivamente  $a_{t-1}, a_{t-2}, \dots$ , en función de  $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$ , se tiene que:

$$\begin{aligned} \text{a) } \tilde{Z}_t &= a_t - \theta \tilde{Z}_{t-1} + \theta^2 \tilde{Z}_{t-2} - \dots \\ \text{b) } \tilde{Z}_t &= a_t - \frac{1}{\theta} \tilde{Z}_{t-1} + \frac{1}{\theta^2} \tilde{Z}_{t-2} - \dots \end{aligned}$$

Se observa que si  $|\theta| < 1$ , la influencia de la variable retrasada sobre la variable actual será menor, conforme aumente el retardo, lo cual ocurre en a), por lo cual el proceso es invertible, situación que no se presenta en b).

Por otro lado, se había hecho mención en el apartado 2.8.2. que un modelo de media móvil será invertible sólo si las raíces de su ecuación característica están fuera del círculo unitario. Para demostrar esto, considérese un proceso *MA* de orden  $q$ :

$$\tilde{Z}_t = \theta_q(B)a_t$$

el cual se puede escribir como

$$\pi(B)\tilde{Z}_t = \frac{1}{\theta_q(B)}\tilde{Z}_t = a_t \quad (2.43)$$

Se sabe que a  $\theta_q(B)$  se le puede escribir en términos de sus raíces, como sigue,

$$\frac{1}{\theta_q(B)} = \prod_{i=1}^q \frac{1}{(1 - G_i B)}$$

y desarrollando en fracciones parciales

$$\frac{1}{\theta_q(B)} = \sum_{i=1}^q \frac{\lambda_i}{(1 - G_i B)}$$

si  $|G_i| < 1$  (es decir, la serie será convergente), se define,

$$(1 - G_i B)^{-1} = \sum_{j=0}^{\infty} G_i^j B^j$$

De esta forma (2.43) puede ser expresado por

$$\sum_{j=0}^{\infty} \left( \sum_{i=1}^q \lambda_i G_i^j \right) \tilde{Z}_{t-j} = a_t$$

Obsérvese que  $\sum_{i=1}^q \lambda_i G_i^j$  decrece conforme  $k$  aumenta, es decir, un proceso *MA* es invertible siempre y cuando  $\pi(B)$  sea convergente y esto se cumplirá sólo si las raíces de  $\theta_q(B) = 0$  viven fuera del círculo unitario.

Los coeficientes del polinomio  $\pi(B)$  se obtienen de (2.43), mediante la relación

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

de donde se pueden igualar las potencias de  $B$ , para obtener las ecuaciones:

$$\begin{aligned}
\pi_1 &= -\theta_1 \\
\pi_2 &= \pi_1\theta_1 - \theta_2 \\
\pi_3 &= \pi_2\theta_1 - \pi_1\theta_2 - \theta_3 \\
&\vdots \\
&\vdots \\
&\vdots \\
\pi_j &= \pi_{j-1}\theta_1 + \pi_{j-2}\theta_2 + \dots + \pi_1\theta_{j-1} - \theta_j \quad j < q
\end{aligned}$$

en donde  $\theta_j = 0$  para  $j > q$ .

Una vez que se obtienen los coeficientes  $\pi_j$  de la representación de un  $AR(\infty)$ , se puede escribir el proceso  $MA$  como:

$$\tilde{Z}_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t$$

Esta expresión permite concluir que la *fap* de un  $MA$  será no nula para todo retardo, ya que existe un efecto directo de  $\tilde{Z}_{t-j}$  sobre  $\tilde{Z}_t$  para toda  $j$ . De esta forma, la *fap* de un proceso  $MA$  tendrá la misma estructura que la *fap* de un  $AR$  del mismo orden.

## 2.9. Modelos $ARMA$ .

En la práctica regularmente se encuentran series de tiempo que presentan características tanto de procesos  $AR$  como de procesos  $MA$ . Existen modelos en los cuales es posible hallar una combinación de parámetros que permitan explicar la mayoría de estos fenómenos, tratando de utilizar el menor número de parámetros posibles. Estos modelos son los  $ARMA$  los cuales, según Granger y Morris, surgen al considerar: (i) series obtenidas por agregación de componentes, por ejemplo el PIB que se obtiene al agregar series correspondientes a los distintos sectores de la economía y (ii) series en donde los datos contienen errores de observación, lo cual es común en series macroeconómicas.

Los procesos  $AR$  y  $MA$  son aproximaciones al proceso lineal general  $MA(\infty)$  desde puntos de vista complementarios: los  $AR$  suponen estructura  $MA(\infty)$ , pero imponen restricciones sobre las pautas de decrecimiento de los coeficientes; los  $MA$  suponen un número de términos finitos, pero a cambio no imponen restricciones sobre sus valores.



Desde el punto de vista de la estructura de autocorrelación los *AR* permiten muchos coeficientes distintos de cero, pero con una pauta de decrecimiento fija, mientras que los *MA* permiten unos pocos coeficientes distintos de cero con valores arbitrarios. Los procesos *ARMA* intentan combinar estas ventajas y permite representar de forma escueta, utilizando pocos parámetros, lo que se conoce como *principio de parsimonia*, procesos cuyos primeros  $q$  coeficientes son cualesquiera y los siguientes decrecen.

### 2.9.1. El proceso *ARMA*(1, 1).

Un proceso *ARMA*(1, 1) se define como sigue

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + a_t - \theta_1 a_{t-1} \quad (2.44)$$

o con notación de operadores

$$(1 - \phi_1 B) \tilde{Z}_t = (1 - \theta_1 B) a_t \quad (2.45)$$

donde  $|\phi_1| < 1$  para que el proceso sea estacionario y  $|\theta_1| < 1$  para que sea invertible. Además se supone que  $\phi_1 \neq \theta_1$ , ya que si ambos parámetros son iguales, se tiene que  $\tilde{Z}_t = a_t$  y así el proceso sería ruido blanco. En la formulación de los modelos *ARMA* se supondrá siempre que no existen raíces comunes en los operadores *AR* y *MA*.

Para obtener la función de autocorrelación de un *ARMA*(1, 1), se multiplica (2.44) y tomando esperanzas:

$$E(\tilde{Z}_t \tilde{Z}_{t-k}) = E(\phi_1 \tilde{Z}_{t-1} \tilde{Z}_{t-k} + a_t \tilde{Z}_{t-k} - \theta_1 a_{t-1} \tilde{Z}_{t-k})$$

$$\gamma_k = \phi_1 \gamma_{k-1} + a_t \tilde{Z}_{t-k} - \theta_1 a_{t-1} \tilde{Z}_{t-k} \quad (2.46)$$

Para  $k > 1$ , el ruido  $a_t$  es independiente de la historia de la serie y resulta

$$\gamma_k = \phi_1 \gamma_{k-1} \quad k > 1 \quad (2.47)$$

Para  $k = 0$ ,  $E(a_t \tilde{Z}_t) = E[a_t(\phi_1 \tilde{Z}_{t-1} + a_t - \theta_1 a_{t-1})] = \sigma^2$  y

$$E(a_{t-1} \tilde{Z}_t) = E[a_{t-1}(\phi_1 \tilde{Z}_{t-1} + a_t - \theta_1 a_{t-1})] = \sigma^2(\phi_1 - \theta_1)$$

sustituyendo estos resultados en (2.46) para  $k = 0$  se tiene

$$\gamma_0 = \phi_1 \gamma_1 + \sigma^2 - \theta_1 \sigma^2 (\phi_1 - \theta_1) \quad (2.48)$$

En (2.46) para  $k = 1$  resulta que  $E(a_t \tilde{Z}_{t-1}) = 0$ ,  $E(a_{t-1} \tilde{Z}_{t-1}) = \sigma^2$  y así se obtiene que

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma^2 \quad (2.49)$$

y resolviendo (2.48) y (2.49) se hallan los valores de  $\gamma_0$  y  $\gamma_1$

$$\gamma_0 = \sigma^2 \frac{1 - 2\phi_1 \theta_1 + \theta_1^2}{1 - \phi_1^2}$$

$$\gamma_1 = \sigma^2 \frac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{1 - \phi_1^2}$$

De esta forma, la *fas* de un modelo  $ARMA(1, 1)$  es:

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{1 - 2\phi_1 \theta_1 + \theta_1^2} & \text{si } k = 1 \\ \phi_1 \rho_{k-1} & \text{si } k > 1 \end{cases}$$

En el caso de que  $k = 1$ , obsérvese que si  $\phi_1 = \theta_1$ , la función de autocorrelación es cero y el proceso resulta ser de ruido blanco. Para las demás autocorrelaciones se observa que tienen un decrecimiento exponencial determinado por  $\phi_1$  de la parte  $AR$ . La diferencia con un  $AR(1)$  es que este decrecimiento comienza a partir de  $\rho_1$ , y no de  $\rho_0 = 1$  y este

primer valor de autocorrelación de primer orden depende de la diferencia que exista entre  $\phi_1$  y  $\theta_1$ . Además si  $\phi_1 \approx 1$  y  $\phi_1 - \theta_1 = \varepsilon$ , se pueden tener muchos coeficientes distintos de cero pero todos de tamaño pequeño.

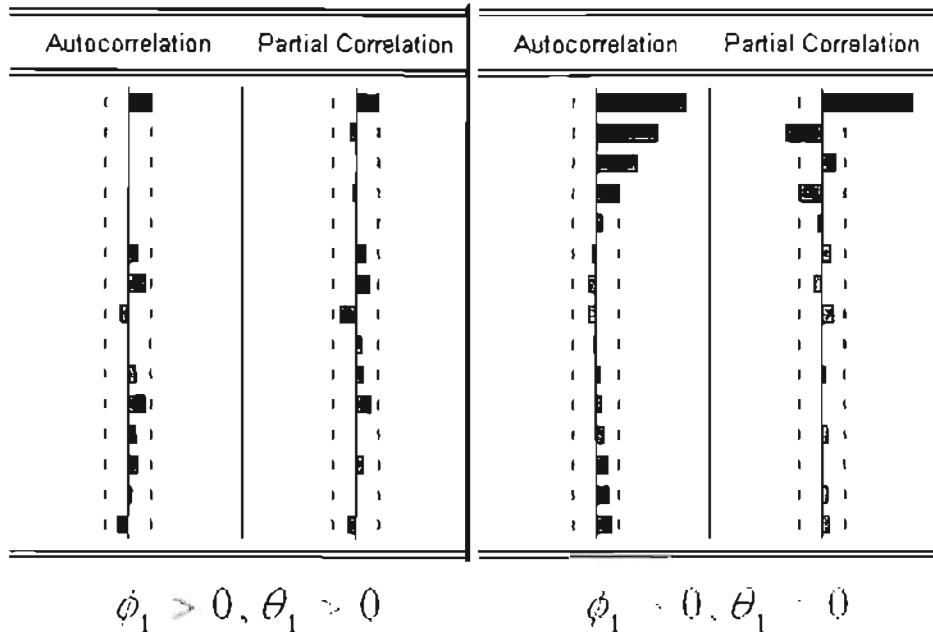
Para explicar la función de autocorrelación parcial, se escribe el  $ARMA(1,1)$  en la forma de un  $AR(\infty)$

$$(1 - \theta_1 B)^{-1}(1 - \phi_1 B)\tilde{Z}_t = a_t$$

y utilizando que  $(1 - \theta_1 B)^{-1} = 1 + \theta_1 B + \theta_1^2 B^2 + \dots$  y operando se tiene

$$\tilde{Z}_t = (\phi_1 - \theta_1)\tilde{Z}_{t-1} + \theta_1(\phi_1 - \theta_1)\tilde{Z}_{t-2} + \theta_1^2(\phi_1 - \theta_1)\tilde{Z}_{t-3} + \dots + a_t$$

Se puede concluir entonces, que en un proceso  $ARMA(1,1)$  la *fas* y la *fap* tienen una estructura similar, es decir, un valor inicial cuya magnitud depende de la diferencia entre  $\phi_1$  y  $\theta_1$  y después un decrecimiento geométrico. La tasa de este decrecimiento en la *fas*<sup>(7)</sup> depende de  $\theta_1$ .



<sup>7</sup> Se observa que el efecto directo de  $Z_{t-k}$  sobre  $Z_t$  decrece geoméricamente con  $\theta_1^k$ , por tanto la *fap* tendrá un decrecimiento geométrico a partir de un valor inicial

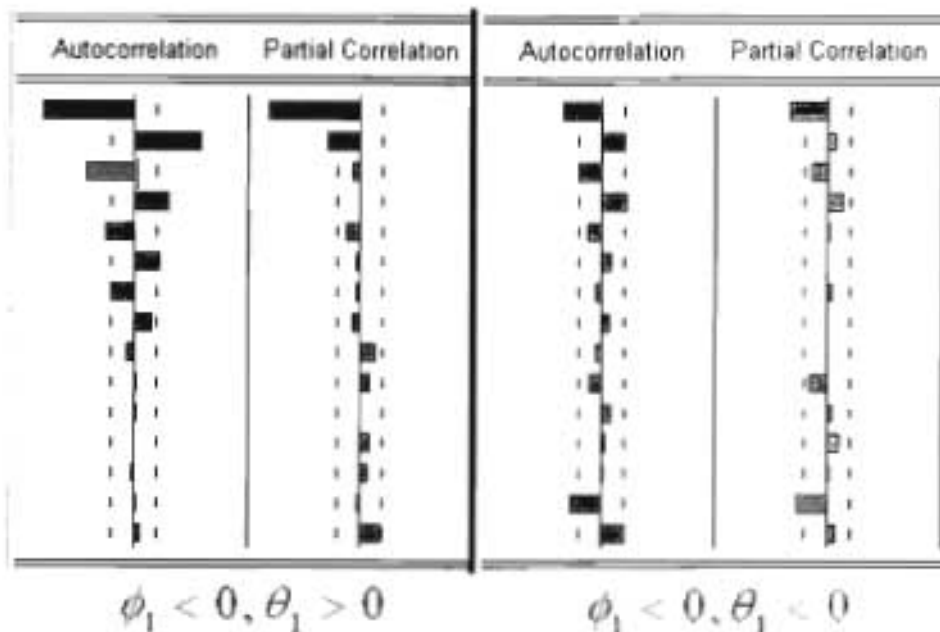


Fig 2.9. Simulación de procesos de medias móviles de orden 2, con coeficientes positivos y negativos.

### 2.9.2. El proceso $ARMA(p, q)$ .

El proceso  $ARMA(p, q)$ , será:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \bar{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.50)$$

o en notación compacta

$$\phi_p(B) \bar{Z}_t = \theta_q(B) a_t$$

El proceso será estacionario si las raíces de la ecuación  $\phi_p(B) = 0$  están fuera del círculo unitario e invertible si las raíces de  $\theta_q(B) = 0$  lo están. De nuevo, se supondrá que no existen raíces comunes para evitar las cancelaciones de los operadores  $AR$  y  $MA$ .

Debido a la composición mixta de este modelo, es posible expresarlo como un  $AR(\infty)$  o como un  $MA(\infty)$ , es decir:

$$\tilde{Z}_t = \phi_p(B)^{-1} \theta_q(B) a_t = \psi(B) a_t \quad \text{o} \quad \theta_q(B)^{-1} \phi_p(B) \tilde{Z}_t = \pi(B) \tilde{Z}_t = a_t$$

en donde  $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$  y  $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$ . Si se reescribe la ecuación (2.50), multiplicando por  $\tilde{Z}_{t-k}$  y tomando esperanzas se tiene

$$\begin{aligned} E(\tilde{Z}_t \tilde{Z}_{t-k}) &= E(\phi_1 \tilde{Z}_{t-1} \tilde{Z}_{t-k} + \dots + \phi_p \tilde{Z}_{t-p} \tilde{Z}_{t-k} + a_t \tilde{Z}_{t-k} - \theta_1 a_{t-1} \tilde{Z}_{t-k} - \dots - \theta_q a_{t-q} \tilde{Z}_{t-k}) \\ &= \phi_1 E(\tilde{Z}_{t-1} \tilde{Z}_{t-k}) + \dots + \phi_p E(\tilde{Z}_{t-p} \tilde{Z}_{t-k}) + E(a_t \tilde{Z}_{t-k}) - \theta_1 E(a_{t-1} \tilde{Z}_{t-k}) - \dots - \theta_q E(a_{t-q} \tilde{Z}_{t-k}) \end{aligned}$$

De esta última expresión debe tomarse en cuenta que  $\tilde{Z}_{t-k}$  está afectada por  $a_{t-k}, a_{t-k-1}, \dots$ , pero es independiente de  $a_{t-k+1}, a_{t-k+2}, \dots$ , por lo que

$$E(a_{t-i} \tilde{Z}_{t-k}) = 0 \quad \text{para } k > i$$

por tanto

$$\gamma_k = \phi_1 \gamma_1 + \dots + \phi_p \gamma_{k-p} \quad \text{para } k < q$$

De esto mismo, si se divide por  $\gamma_0$ , se obtiene que

$$\rho_k = \phi_1 \rho_1 + \dots + \phi_p \rho_{k-p} \quad \text{para } k < q \quad (2.51)$$

Como se vio en la sección 2.5.3, la solución de esta ecuación en diferencia está dado por:

$$\rho_k = \sum_{i=1}^p G_i^k \sum_{j=1}^{m_i-1} t^j A_{ij}$$

en donde  $A_{ij}$  son constantes a determinar a partir de las  $p$  condiciones iniciales y  $G_i^{-1}$  son las raíces de multiplicidad  $m_i$  para  $\phi_p(B) = 0$ . Por tanto las primeras  $q$  autocorrelaciones tendrán cualquier estructura, pues involucran tanto a parámetros autorregresivos como de media móvil y a partir de  $q + 1$  la *fas* se comportará a (2.51) y a las  $p$  condiciones iniciales, decreciendo de manera exponencial o senoidal, dependiendo del signo de los operadores.

## 2.10. Procesos no estacionarios.

Hasta el momento se ha tratado con procesos estacionarios, los cuales deben cumplir ciertas condiciones para que esta propiedad se pueda dar. En la práctica es más común encontrar series de datos que no cumplen alguna de esas condiciones que las hace ser no estacionarias, ya sea porque muestran alguna tendencia, porque su varianza no es constante o porque están influenciadas por algún factor estacional.

Un proceso puede ser no estacionario en media, varianza, covarianzas o en otras características de la distribución de las variables. Si el nivel de la serie no es estable en el tiempo, ya sea porque lleva una tendencia creciente o decreciente, se dirá que la serie no es estacionaria en la media. Si la varianza o covarianzas varían con el tiempo, se dice que la serie no es estacionaria en las covarianzas. Si además, la distribución de la variable es distinta en cada instante del tiempo, manteniendo o no la media y covarianzas, se dice que la serie no es estacionaria en distribución.

Los procesos no estacionarios más importantes son los *procesos integrados*, que son procesos no estacionarios únicamente en la media y que tienen la particularidad de convertirse en estacionarios tomando diferencias. Debe tomarse en cuenta que existen procesos estacionarios que tienen características similares a los integrados, lo cual puede llegar a confundir a quien intente analizar alguna serie de tiempo.

### 2.10.1. Procesos Integrados.

La mayoría de las series reales son no estacionarias, y su nivel medio varía con el tiempo. Es frecuente que las series estacionarias tengan un nivel de estabilidad, por ejemplo, su crecimiento puede variar en torno a un valor medio y aunque la serie observada  $Z_t$  sea no estacionaria, es posible que la serie de crecimientos  $W_t$  sí lo sea, como se había mencionado en la sección 2.3.2. mediante diferencias:

$$W_t = Z_t - Z_{t-1} = \nabla Z_t$$

A la serie  $W_t$  se le denomina *serie de primera diferencia* de  $Z_t$ , y si  $W_t$  es estacionaria, se dice que el proceso  $Z_t$  es integrado de orden uno, o también, con incrementos estacionarios. Aún con esta primera diferencia, es posible que los incrementos no sean estacionarios pero sí las variaciones de éstos, por lo que es necesario tomar una segunda diferencia para hacer al proceso  $Z_t$  estacionario y se dice que  $Z_t$  es integrada de orden dos o con incrementos de segundo orden estacionarios. Generalizando, se dirá que un proceso es integrado de orden  $h \geq 0$ , si al diferenciarlo  $h$  veces se obtiene un proceso estacionario. En

la misma sección 2.3.2. se demostró que cualquier proceso generado como suma de una tendencia polinómica de orden  $h$  más un proceso estacionario cualquiera  $u_t$ , será integrado de orden  $h$ .

$$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_h t^h + u_t = \sum_{i=0}^h \beta_i t^i + u_t$$

### 2.10.1.1. La caminata aleatoria.

En secciones anteriores se ha visto que los procesos  $MA$  son siempre estacionarios y que los  $AR$  lo son si las raíces de la ecuación característica están fuera del círculo unitario. Sea el proceso  $AR(1)$ :

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t$$

Si  $|\phi| > 1$  el proceso es explosivo, pues esto diría que la historia de la serie tiene más peso que el dato actual y con esto la función de autocorrelación iría en aumento. Si  $|\phi| = 1$  el proceso no es estacionario, pero tampoco es explosivo y pertenece a la clase de procesos integrados de orden uno, pues su primera diferencia  $\tilde{Z}_t - \tilde{Z}_{t-1} = a_t$  sí es un proceso estacionario. Este proceso se conoce como *caminata (o paseo) aleatorio*. Utilizando la notación del operador diferencia, la caminata aleatoria puede escribirse como:

$$\nabla \tilde{Z}_t = a_t$$

Se calcula la función de este proceso, suponiendo que comienza en  $t = 0$ . Sustituyendo sucesivamente  $\tilde{Z}_t$  por  $\tilde{Z}_{t-1}$  se tiene:

$$\tilde{Z}_t = a_t + a_{t-1} + a_{t-2} + \dots + a_1$$

Por lo que la varianza del proceso  $\sigma^2$  será

$$\text{Var}(\tilde{Z}_t) = \gamma_0(t) = \sigma^2 t$$

es decir, la varianza tiende a infinito cuando aumenta el tiempo. Para obtener las covarianzas se multiplican los procesos  $\tilde{Z}_t$  y  $\tilde{Z}_{t+k}$ , tomando esperanzas se tiene:

$$Cov(t, t+k) = \sigma^2 t$$

Obsérvese que la covarianza depende de los dos subíndices  $(t, t+k)$  porque al ser el proceso no estacionario, la covarianza no sólo es función del retardo, sino también del instante en que se calcula. Además se puede comprobar que  $Cov(t, t-k) = \sigma^2(t-k) \neq Cov(t, t+k)$ .

La función de autocorrelación es:

$$\rho(t, t+k) = \frac{t}{\sqrt{t(t+k)}} = \frac{\sqrt{t} \sqrt{t}}{\sqrt{t(t+k)}} = \sqrt{\frac{t}{t+k}} = \left(1 + \frac{k}{t}\right)^{-1/2}$$

y si  $t$  es grande, los coeficientes de la función de autocorrelación serán próximos a uno y decrecerán aproximadamente de forma lineal con  $k$ . Tomando logaritmos y utilizando la expansión de Taylor para esa función, se tiene

$$\log \rho(t, t+k) = -\frac{1}{2} \left(1 + \frac{k}{t}\right) \approx -\frac{k}{2t}$$

y con esto, los logaritmos de las correlaciones decrecerán linealmente con  $k$ . Este decrecimiento lineal se puede ver también con los coeficientes de autocorrelación, ya que empleando de nuevo la expansión de Taylor, alrededor del 0, se tiene que  $f(k) = \left(1 + \frac{k}{t}\right)^{-1/2}$ , por lo que  $f'(k) = -\frac{1}{2t} \left(1 + \frac{k}{t}\right)^{-3/2}$ . Entonces:

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= -\frac{1}{2t} \end{aligned}$$

Por tanto  $f(k) = 1 - \frac{k}{2t}$ . Así, los coeficientes de autocorrelación tienen un comportamiento en decrecimiento lineal.

## 2.10.2. Procesos Integrados ARIMA.

Como se vio en la sección anterior, se obtuvieron procesos admitiendo que la raíz de la parte AR es la unidad, igualmente se puede suponer como unidad la raíz de un proceso ARMA(1, 1) (resultando los procesos de alisado exponencial simple), y de esta manera, los procesos anteriores se convierten en no estacionarios. Esta idea puede generalizarse para cualquier proceso ARMA, permitiendo una o varias raíces unitarias en el operador AR, obteniéndose entonces procesos del tipo:



$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d \tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

que se conocen como procesos *ARIMA*( $p, d, q$ ). Bajo esta notación,  $p$  es el orden de la parte autorregresiva estacionaria;  $d$ , es el número de raíces unitarias (orden de integración del proceso) y  $q$  es el orden de la parte de media móvil. Utilizando el operador diferencia  $\nabla = 1 - b$ , el proceso se escribe como:

$$\phi_p(B) \nabla^d \tilde{Z}_t = \theta_q(B) a_t$$

El nombre *ARIMA* proviene de las iniciales en inglés de los procesos autorregresivos integrados de media móvil (autoregressive integrated moving average), donde la palabra "integrado" indica que el proceso  $\tilde{Z}_t$  se puede obtener como suma (integración) de  $w_t$ , donde  $w_t = \nabla^d \tilde{Z}_t$ . Esto es,

$$w_t = (1 - B) \tilde{Z}_t$$

y como

$$(1 - B)^{-1} = 1 + B + B^2 + B^3 + \dots$$

entonces:

$$\tilde{Z}_t = (1 - B)^{-1} w_t = \sum_{j=-\infty}^t w_j$$

En la sección anterior se vio un ejemplo de proceso *ARIMA*, la caminata aleatoria, el cual es el modelo *ARIMA*(0, 1, 0). Estos procesos se caracterizan porque su función de autocorrelación simple tiene coeficientes que decrecen lentamente. Todos los procesos *ARIMA* no estacionarios tienen esta propiedad general. Para demostrarlo, recuérdese que el correlograma de un *ARMA*( $p, q$ ) satisface para  $k > q$  la ecuación:

$$\phi_p(B) \rho_k = 0$$

cuya solución es del tipo

$$\rho_k = \sum_{i=1}^p A_i G_i^k$$

donde  $G_i^{-1}$  son las raíces de  $\phi_p(B) = 0$  y  $|G_i| < 1$ . Si una de estas raíces  $G_i$  es muy próxima a la unidad, se puede escribir  $G_i = 1 - \varepsilon$ , con  $\varepsilon$  muy pequeño. Para  $k$  grande los términos  $A_j G_j^k$  para  $j \neq i$ , debida a las otras raíces serán cero (pues  $|G_j| < 1$ ) y se tendrá que,

$$\rho_k = A_i(1 - \varepsilon)^k \approx A_i(1 - k\varepsilon) \quad \text{para } k \text{ grande}$$

Es decir, la *fas* tendrá coeficientes positivos que se amortiguarán aproximadamente de forma lineal y pueden ser distintos de cero, incluso para valores muy altos de  $k$ . Esta propiedad de tener valores positivos en el correlograma para retardos grandes y el decrecimiento lineal es lo que caracteriza a los procesos no estacionarios.

Es importante mencionar que existe una clase de procesos estacionarios que pueden confundirse con los procesos no estacionarios integrados. Estos se conocen como procesos de memoria larga, que se caracterizan por tener muchos coeficientes de autocorrelación pequeños y que se amortiguan lentamente. Este lento decrecimiento puede llevar a modelar estas series suponiendo que son integradas de orden uno, y así tomar una diferencia para hacerlos estacionarios, cuando en realidad ya lo eran. Los procesos de memoria larga ocupan un lugar intermedio entre los procesos *ARMA*, donde las correlaciones decrecen rápidamente a cero y los *ARIMA* en que decrecen linealmente.

### 2.10.3. El concepto de estacionalidad.

Cuando se habla de series de tiempo lo ideal es que todas fueran estacionarias, pues de esta forma no se tendrían problemas con la estabilidad de la media, de la varianza, de distribución, pero en la realidad es más común encontrarse con series que no cumplen con estas condiciones. Uno de los factores que influyen en la falta de estacionariedad es la *estacionalidad*. Este factor hace que el valor de la serie varíe con una pauta cíclica. Por ejemplo, en una serie trimestral la estacionalidad se presentará si toma valores similares en el trimestre 1 de distintos años y lo mismo para el trimestre 2, 3 o 4, siendo además distintos entre cada trimestre.

El efecto de estacionalidad hace que las series sean no estacionarias, ya que este factor hace que el valor medio de la serie varíe de un período a otro, pudiendo ser estos períodos semanas, meses, trimestres, años, etc. Se llamará período estacional,  $s$ , al número de observaciones que formen el ciclo estacional. Así, si la serie es mensual, entonces  $s = 12$ ; para series trimestrales,  $s = 4$ , etc. Puede suceder que la estacionalidad no sea fija y que no sea sólo una para una misma serie, porque puede pasar que con una serie diaria, se tengan

estacionalidad semanal, otra mensual o anual. Para efectos de este texto se supondrá que existe sólo un tipo de estacionalidad.

Lo anterior implica que las observaciones  $Z_t$  y  $Z_{t+s}$  tienen el mismo componente estacional  $S_t^{(s)}$ , por lo que el modelo de la serie se puede escribir:

$$Z_t = S_t^{(s)} + n_t \quad (2.52)$$

donde  $n_t$  es un proceso estacionario. Por tener un componente estacional, la serie (2.52) es no estacionaria, pero ya se demostró que tomando diferencias es posible volverla estacionaria. Para esto se debe aplicar el operador  $\nabla_s = 1 - B^s$  (tomando en cuenta que  $\nabla_s \neq \nabla^s$ ) a (2.52), por lo que se tiene:

$$\nabla_s Z_t = \nabla_s S_t^{(s)} + \nabla_s n_t$$

Cabe mencionar que el componente estacional puede comportarse de dos formas posibles, las cuales hay que considerarlas. El primer caso es que sea determinista, es decir, constante para el mismo mes en distintos años. Para este caso se puede expresar a la componente de la siguiente forma:

$$S_t^{(s)} = S_{t+ks}^{(s)} \quad k = \pm 1, \pm 2, \dots$$

de donde se tiene que  $\nabla_s S_t^{(s)} = S_t^{(s)} - S_{t-s}^{(s)} = 0$ , con lo que se obtiene:

$$\nabla_s Z_t = \nabla_s n_t$$

es decir, la serie diferenciada  $s$  períodos se transforma en un proceso estacionario. Ahora, el segundo caso es que la componente estacional no sea constante, sino que evoluciona con el tiempo, y se supondrá que esta evolución es simple, es decir de la forma:

$$S_t^{(s)} = S_{t-s}^{(s)} + v_t$$

siendo  $v_t$  un proceso estacionario. Así, se toman diferencias estacionales en (2.52) y se tiene

$$\nabla_s Z_t = v_t + \nabla_s n_t$$

y de nuevo se obtiene un proceso estacionario. Se concluye entonces que en condiciones generales, el operador  $\nabla_s$  convierte un proceso estacional en estacionario.

#### 2.10.4. El modelo *ARIMA* estacional.

Para explicar este tipo de procesos se utilizarán series estacionales  $Z_t$  con período  $s$  conocido, por ejemplo, considérese una serie con datos mensuales y de la que se disponen de  $h$  años completos, es decir la serie tiene un total de  $n = 12h$  observaciones. Como ya se había mencionado, se tienen dos tipos de dependencia, uno en donde las observaciones del mismo mes se parecen entre sí, y en general, que las observaciones separadas por  $s$  períodos se parecen entre sí.

Para modelar la dependencia estacional de una serie mensual, se divide la serie total en 12 series de  $h$  datos, donde  $h$  es el número de años disponibles. Se denotarán a estas series por  $y_t^{(j)}$ , ( $j = 1, \dots, 12$ ) donde ahora el índice temporal de la serie,  $\tau$ , es el año que varía entre  $\tau = 1, \dots, h$  y el índice  $j$  indica el mes dentro del año. De esta forma es posible establecer una relación entre estas series y las primitivas, ya que  $t = j + 12(\tau - 1)$ , por lo que

$$y_t^{(j)} = Z_{j+12(\tau-1)} \quad \tau = 1, \dots, h$$

Si se supone que cada una de estas series sigue un modelo *ARIMA*:

$$\Phi^{(j)}(B)y_t^{(j)} = \Theta^{(j)}(B)u_t^{(j)} \quad \tau = 1, \dots, h$$

que en principio será distinto para unos meses que para otros. Si se supone, el caso más simple, en que el modelo es el mismo para todas las series  $y_t^{(j)}$ , entonces el modelo será del tipo:

$$(1 - \Phi_1 B - \dots - \Phi_p B^p)(1 - B)^D y_t^{(j)} = (1 - \Theta_1 B - \dots - \Theta_q B^q)u_t^{(j)} \quad \tau = 1, \dots, h$$

Este modelo será no estacionario si  $D \geq 1$ , ya que si  $D = 0$  y las series fueran estacionarias, entonces el modelo podría escribirse de la siguiente manera:

$$y_t^{(j)} = \mu_j + \Psi_j(B)u_t^{(j)} \quad \tau = 1, \dots, h \quad (2.53)$$

donde  $\mu_j$  es la media del mes  $j$  y  $\Psi_j(B)$  es la representación del modelo en la forma  $MA(\infty)$ . Si hay estacionalidad, las medias  $\mu_j$  son, claramente, diferentes, por lo que las 12 series no pueden tener un mismo modelo estacionario común. Aunque se sabe que al tomar diferencias en (2.53) se logra que las constantes  $\mu_j$  desaparezcan y así la serie estacionaria  $(1-B)^D y_t^{(j)}$  tendrá media cero y además, si los parámetros  $\Psi_j(B)$  son iguales para todas las series, entonces las series diferenciadas seguirán todas, el mismo modelo. De esta manera las 12 series podrán escribirse bajo un mismo modelo común, utilizando que:

$$y_t^{(j)} = Z_{j+12(\tau-1)} \quad \tau = 1, \dots, h$$

Aplicando el operador retardo,

$$B y_t^{(j)} = y_{t-1}^{(j)} = Z_{j+12(\tau-2)} = Z_{j-12+12(\tau-1)} = B^{12} Z_{j+12(\tau-1)}$$

Es decir, aplicar el operador de retardo regular  $B$  a la serie de los meses  $y_t^{(j)}$ , es relacionar un mes con el mismo mes del año anterior y esto es finalmente aplicar el operador de retardo estacional  $B^{12}$  a la serie original  $Z_t$ .

Por otro lado, debido a que en este proceso no se consideró a la dependencia entre los choques aleatorios, el modelo estacional no se considerará como proceso de ruido blanco a la serie  $\alpha_t$  generada por éstos. Entonces, la serie original puede escribirse, llamando  $s$  al período estacional, como:

$$(1 - \Phi_1 B^s - \dots - \Phi_p B^{p s})(1 - B^s)^D \tilde{Z}_t = (1 - \Theta_1 B^s - \dots - \Theta_q B^{q s}) \alpha_t \quad t = 1, \dots, n \quad (2.54)$$

y ahora el modelo *ARIMA* se formula en  $B^s$ , pues se está relacionando períodos de tamaño  $s$  de distintos años. Se definirá como *estructura regular* a la relacionada a los intervalos naturales de la serie, para diferenciarla de la *estructura estacional*, que es la que se asocia a los intervalos de amplitud  $s$ . Suponiendo que  $\alpha_t$  sigue un proceso *ARIMA* regular:

$$\phi_p(B) \nabla^d \alpha_t = \theta_q(B) \alpha_t \quad (2.55)$$

sustituyendo este modelo regular (2.54) en el estacional (2.55), se obtiene el *modelo*

completo para el proceso observado:

$$\Phi_P(B^s)\phi_P(B)\nabla^d\nabla_s^D\tilde{Z}_t = \theta_q(B)\Theta_Q(B^s)a_t \quad (2.56)$$

el cual se denomina modelo  $ARIMA(P,D,Q)_s \times (p,d,q)$ . Estos modelos fueron introducidos por Box y Jenkins, y representan de forma simple muchos fenómenos reales que se encuentran en la práctica. Aunque se debe recordar que *estos modelos se basan en la hipótesis de que la relación de dependencia estacional es la misma para todos los períodos*.

#### 2.10.4.1. La *fas* y la *fap* de un proceso $ARIMA$ estacional.

Sea  $w_t = \nabla^d\nabla_s^D Z_t$ , el proceso estacionario que se obtiene de diferenciar  $d$  veces de forma regular y  $D$  veces de forma estacional, por lo que el proceso  $w_t$  sigue el proceso  $ARIMA$  estacional multiplicativo:

$$\Phi_P(B^s)\phi_P(B)w_t = \theta_q(B)\Theta_Q(B^s)a_t$$

La *fas* de este proceso es una mezcla de funciones de autocorrelación correspondientes a la parte regular y estacional y se demuestra que el coeficiente de autocorrelación del modelo completo es:

$$\rho_j = \frac{r_j + \sum_{i=1}^{\infty} R_{si}(r_{si+j} + r_{si-j})}{1 + 2 \sum_{i=1}^{\infty} r_{si}R_{si}} \quad (2.57)$$

Para poder dar una explicación acerca de este coeficiente, se supondrá que  $s = 12$  y que  $r_j \simeq 0$  para retardos altos, por ejemplo para  $j \geq 8$ , por lo que el denominador será la unidad y la *fas* estará determinada según el número de retardos, de la siguiente manera:

1.- En los primeros retardos se observa únicamente la estructura generada por la parte regular, esto porque:

$$\rho_j \simeq r_j \quad j < 8$$

2.- En los retardos estacionales se observa básicamente la estructura generada por la parte estacional, esto porque:

$$\rho_{12i} \approx R_{12i}(r_{24i} + r_0) + R_{24i}(r_{36i} + r_{12i})$$

y como  $r_{12i} \approx 0$  para  $i \geq 1$ , la expresión anterior se reduce a:

$$\rho_{12i} \approx R_{12i}$$

3.- Alrededor de los retardos estacionales, se observará la interacción que hay entre la parte estacional y la regular, que se observará como una repetición a ambos lados de cada retardo estacional de la *fas* simple de la parte regular. Es decir, si por ejemplo, la parte regular es media móvil de orden  $q$  habrá a ambos lados de cada retardo estacional distinto de cero,  $q$  coeficientes distintos de cero. Si la parte regular es autorregresiva a ambos lados se observará el decrecimiento exponencial que impone la parte *AR*.

En la práctica lo recomendable es fijarse en los retardos iniciales 1,2,3,4... para identificar la estructura regular y en los retardos estacionales  $s, 2s, 3s, \dots$  para poder identificar la estructura estacional. La interacción alrededor de los coeficientes estacionales puede entonces usarse como confirmación de la identificación realizada.

Por otro lado, la *fap* del proceso  $w_t$  es compleja, pues no es solamente una interacción entre la *fap* de la parte regular y de la estacional, sino que en la interacción alrededor de los retardos estacionales, aparece a un lado la *fap* de la parte regular y al otro la *fas*. Según Hamilton y Watts (1978), la *fap* de un proceso estacional tiene la estructura siguiente:

1.- En los primeros retardos aparece la *fap* de la estructura regular y en los estacionales la *fap* de la estacional.

2.- A la derecha de cada coeficiente estacional (retardos  $js + 1, js + 2, \dots$ ) aparece la *fap* de la parte regular. Si el coeficiente estacional es positivo la *fap* regular aparece invertida, mientras que si es negativo, la *fap* aparece con su signo.

3.- A la izquierda de los coeficientes estacionales (retardos  $js - 1, js - 2, \dots$ ) se observa la *fas* de la parte regular.

Se comprueba que en modelos estacionales, los coeficientes de correlación parcial para los retardos estacionales suelen ser de magnitud mucho menor que la que correspondería a una estructura estacional pura. Esta complejidad recomienda fijarse sólo en los retardos iniciales  $1, 2, 3, 4, \dots$  para identificar la estructura regular, y en los retardos estacionales  $s, 2s, 3s, \dots$  para identificar la estructura estacional.

## 2.11. Predicción con modelos *ARIMA*.

Después de haber establecido las condiciones y criterios para elegir el posible modelo *ARIMA*, sea estacional o no, con el que se explicará a la serie de tiempo en cuestión, es necesario estudiar la predicción que el modelo realizará. Esta predicción debe cumplir ciertas condiciones, de tal forma que sea la mejor en el sentido de que minimiza el error cuadrático medio de los errores de predicción, como sucede en cualquier modelo de estimación.

Sea  $Z_T = (Z_1, \dots, Z_T)$  una serie temporal de longitud  $T$ , de la cual se desea obtener una predicción de un valor futuro  $Z_{T+k}$ ,  $k > 0$ . Esta predicción claramente será una función de los valores pasados, es decir  $f(Z_T) = \hat{Z}_{T+k}$ . Al tratar obtener esta predicción, posiblemente se encontrarán una infinidad, que tratarán de estimar al verdadero valor futuro, pero sólo aquella que minimice el error de predicción será la óptima para considerar.

Se sabe que el valor  $Z_{T+k}$  es desconocido (de otra forma no se estimaría ningún valor), por lo que es imposible conocer el error que se tendrá al estimarlo; pero si se conoce su distribución de probabilidad se puede calcular el error promedio que se cometerá con un estimador fijo. Es decir, si se desea minimizar el error cuadrático medio de  $Z_{T+k}$ , dado por:

$$ECMP(Z_{T+k}) = E\left[(Z_{T+k} - \hat{Z}_{T+k}|Z_T)^2\right]$$

es necesario conocer esa distribución condicional a los valores observados  $Z_T$ . Ahora, para conocer el predictor que minimiza el error cuadrático medio, no es necesario conocer la distribución condicional completa, ya que si,

$$\mu_{T+k|T} = E[Z_{T+k}|Z_T]$$

es la media de esta distribución condicional, se puede restar y sumar  $\mu_{T+k|T}$  a la expresión de *ECMP* y desarrollando, se llega a que:



$$ECMP(Z_{T+k}) = \text{var}(Z_{T+k}|Z_T) + [\mu_{T+k|T} - \hat{Z}_{T+k}|Z_T]^2$$

Dado que el primer término es constante, se concluye que para minimizar a  $ECMP(Z_{T+k})$  es necesario hacer cero al segundo, y esto sucederá siempre y cuando:

$$\hat{Z}_{T+k} = E[Z_{T+k}|Z_T] = \mu_{T+k|T}$$

Es decir, el predictor que minimiza el error cuadrático medio de predicción es la esperanza de la observación que se quiere prever condicionada con los datos observados.

Supóngase que se dispone de una realización de tamaño  $T$ ,  $Z_T = (Z_1, \dots, Z_T)$ , de un proceso  $ARIMA(p, d, q)$  con parámetros conocidos  $\phi(B)\nabla^d Z_t = c + \theta(B)a_t$ . Se tratará de encontrar una relación en como aplicar el resultado anterior para calcular las predicciones. Como los parámetros ya son conocidos, se pueden obtener todas las innovaciones  $a_t$  fijando valores iniciales. Por ejemplo, si el proceso es  $ARMA(1, 1)$  tomando  $a_1 = 0$  las innovaciones  $a_2, \dots, a_T$ , se calculan recursivamente mediante la ecuación:

$$a_t = Z_t - c - \phi Z_{t-1} + \theta a_{t-1}$$

Así, se supondrá en adelante que tanto las observaciones como las innovaciones son conocidas hasta el instante  $T$ . De acuerdo con el resultado anterior, la predicción que minimiza el error cuadrático medio de  $Z_{T+k}$ , que de ahora en adelante se llamará *predicción óptima de  $Z_{T+k}$* , es la esperanza de la variable condicionada a los valores observados.

A continuación se define lo siguiente:

$$\begin{aligned} \hat{Z}_T(j) &= E[Z_{T+j}|Z_T] & j &= 1, 2, \dots \\ \hat{a}_T(j) &= E[a_{T+j}|Z_T] & j &= 1, 2, \dots \end{aligned}$$

donde el subíndice  $T$  representa el origen de la predicción, que se supone fijo y el  $j$  el horizonte de la misma, que irá cambiando para generar predicciones de distintas variables futuras desde el origen  $T$ . Para calcular las predicciones recursivamente, comenzando para  $j = 1, 2, 3, \dots$  hasta el horizonte deseado, se tiene la ecuación:

$$\hat{Z}_T(k) = c + \varphi_1 \hat{Z}_T(k-1) + \dots + \varphi_h \hat{Z}_T(k-h) - \theta_1 \hat{a}_T(k-1) - \dots - \theta_q \hat{a}_T(k-q) \quad (2.58)$$

donde  $\phi_h = \phi(B)\nabla^d$  es el operador de orden  $h = p + d$ . Con la ecuación anterior algunas esperanzas se calculan para variables observadas y otras para variables no observadas. Cuando  $i > 0$  la expresión  $\hat{Z}_T(i)$  es la esperanza condicionada de la variable  $Z_{T+i}$  que no se ha observado todavía. Sin embargo, cuando  $i \leq 0$ ,  $\hat{Z}_T(i)$  es la esperanza condicionada de la variable  $Z_{T-|i|}$ , que ya se ha observado y es conocida, por lo que esta esperanza coincidirá con la observación y  $\hat{Z}_T(-|i|) = Z_{T-|i|}$ . Ahora, con relación a los choques aleatorios, la esperanza de estos choques condicionada a la historia de la serie es igual a su esperanza absoluta, ya que la variable  $a_{T+i}$  es independiente de  $Z_T$ . Como los choques aleatorios tienen esperanza nula, se concluye que para  $i > 0$ , los choques futuros  $\hat{a}_T(i)$  serán cero. Cuando  $i \leq 0$  las innovaciones  $a_{T-|i|}$  son conocidas, por lo que sus esperanzas coinciden con los valores observados y  $\hat{a}_T(-|i|) = a_{T-|i|}$ .

Por otro lado, se tiene que:

$$a_{T+1} = Z_{T+1} - \hat{Z}_T(1)$$

lo cual indica que los choques aleatorios pueden interpretarse como los errores de predicción a un período por delante cuando se conocen los parámetros del modelo. La ecuación (2.58) indica que después de  $q$  valores iniciales los términos de media móvil desaparecerán y la predicción quedará determinada exclusivamente por la parte autorregresiva del modelo; es decir, para  $j > q$  todos los choques aleatorios desaparecerán de la ecuación, entonces las predicciones de (2.58) para  $j > q$  satisfacen la ecuación

$$(1 - \phi_1 B - \dots - \phi_h B^h) \hat{Z}_T(j) = \phi(B) \nabla^d \hat{Z}_T(j) = c \quad j > q$$

donde en esta ecuación  $B\hat{Z}_T(j) = \hat{Z}_T(j-1)$  ya que el operador  $B$  actúa sobre el horizonte de la previsión  $j$ , siendo  $T$ , el origen de la predicción, un valor fijo. Esta ecuación se denomina la *ecuación de predicción final* porque establece como están ligadas las predicciones para horizontes altos cuando desaparece la parte de media móvil. Se observa que la relación entre las predicciones es similar a la que existe entre las autocorrelaciones de un proceso estacionario ARMA, aunque en las predicciones además de estar el operador estacionario  $\phi(B)$  aparece también el operador no estacionario  $\nabla^d$ , y la ecuación no está igualada a cero, sino a la constante  $c$ .

Para dejar más claro el uso de esta ecuación de predicción, supóngase el proceso  $AR(1)Z_t = c + \phi_1 Z_{t-1} + a_t$ . La ecuación de predicción un período adelante es

$$\hat{Z}_T(1) = c + \phi_1 Z_T$$

y para dos períodos:

$$\hat{Z}_T(2) = c + \phi_1 \hat{Z}_T(1) = c(1 + \phi_1) + \phi_1^2 Z_T$$

Generalizando para cualquier período  $k > 0$  se tiene:

$$\hat{Z}_T(k) = c + \phi_1 \hat{Z}_T(k-1) = c(1 + \phi_1 + \dots + \phi_1^k) + \phi_1^k Z_T$$

Para  $k$  grande (sabiendo que  $|\phi_1| < 1$ ), el término  $\phi_1^k Z_T$  será próximo a cero y la predicción será el término  $c(1 + \phi_1 + \dots + \phi_1^k) = c/(1 - \phi_1)$ , que es la media del proceso. Finalmente, para cualquier proceso  $ARMA(p, q)$  la predicción para un horizonte  $k$  grande es la media del proceso,  $\mu = c/(1 - \phi_1 - \dots - \phi_p)$ .

Para un paseo aleatorio  $\nabla Z_T = c + a_t$ , la predicción a un paso es:

$$\hat{Z}_T(1) = c + Z_T$$

y para dos:

$$\hat{Z}_T(2) = c + \hat{Z}_T(1) = 2c + Z_T$$

y para un horizonte  $k$  cualquiera:

$$\hat{Z}_T(k) = c + \hat{Z}_T(k-1) = kc + Z_T$$

Obsérvese que la predicción del siguiente período se obtiene sumando siempre la misma constante  $c$ , por lo que se puede concluir que las predicciones seguirán una línea recta con pendiente  $c$ .

## 2.12. Transformaciones para estabilizar la varianza.

Muchas de las series económicas hoy en día presentan mayor variabilidad cuanto

mayor sea el nivel de la serie, es decir, que su valor promedio va aumentando con el tiempo, lo que hace que la variabilidad no sea constante, lo que se conoce como *heterocedasticidad*. Por tal motivo, se tiene que estabilizar la varianza, encontrando alguna función que transforme la serie en otra que cumpla el supuesto de varianza constante (*homocedasticidad*) y así no afectar al pronóstico a realizar. Existen diversos métodos para hallar esa función estabilizadora de la varianza, uno de ellos se presenta con el siguiente ejemplo:

Se tiene la serie de pasajeros que viajan en avión del período 1949-1960, presentada anteriormente, en donde se observa claramente como el nivel de variabilidad aumenta conforme aumenta el nivel a través del tiempo.

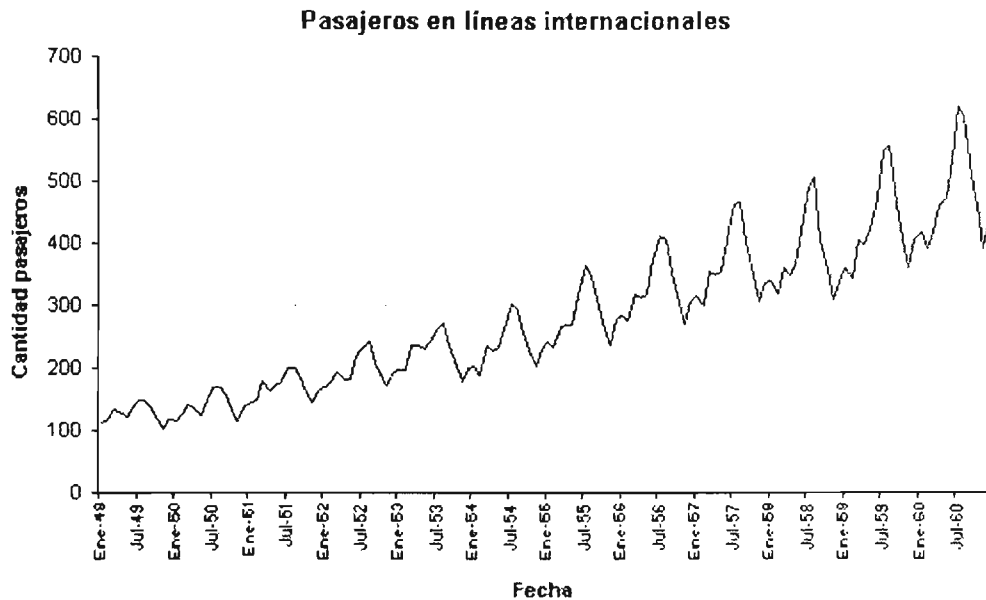


Fig. 2.10. Número de pasajeros en líneas internacionales de 1949 a 1960.

Esta impresión visual que se tiene con esta gráfica se puede apreciar mucho mejor si se grafica la media contra la desviación estándar anual de esta misma serie. En la siguiente figura se puede ver esta relación.

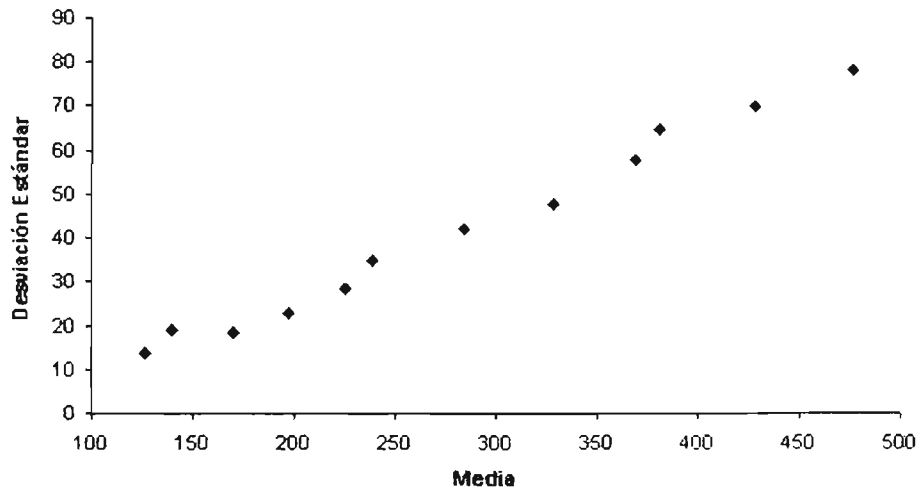


Fig. 2.11. Relación entre la media y la desviación estándar de la serie de pasajeros.

Como se observa, la relación es lineal respecto al tiempo, es decir, entre más tiempo pase y la media crezca, la variación de la serie lo hará también. Cuando se tiene este tipo de casos, el tomar logaritmos a la serie ayuda a tener variabilidad constante. A continuación se presenta la gráfica del logaritmo.

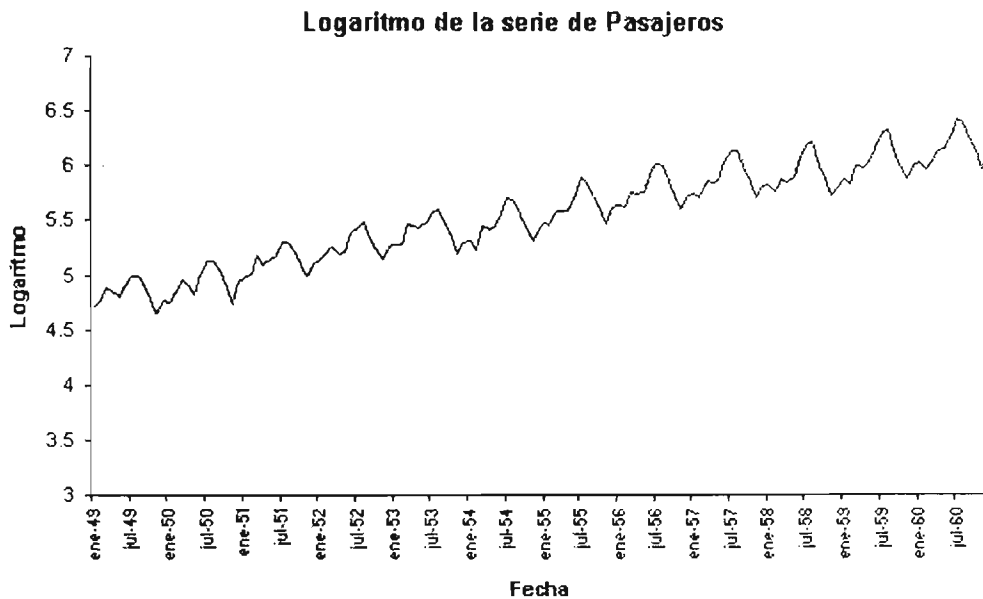


Fig. 2.12. Logaritmo del número de pasajeros en líneas internacionales de 1949 a 1960.

En general, se puede demostrar que cuando existe una relación entre las medias,  $\mu_i$ , de un conjunto de variables aleatorias  $x_i$  y sus desviaciones estándar  $\sigma_i$ , del tipo

$$\sigma_i = k\mu_i^\alpha \quad (2.59)$$

es posible tener variabilidad constante en las variables  $x_i$  si se transforman mediante la función

$$y_i = \frac{x_i^{1-\alpha} - 1}{1-\alpha} \quad (2.60)$$

Esta es la llamada familia Box-Cox, las cuales son transformaciones con el parámetro  $\lambda = 1 - \alpha$ , que incluyen a las funciones potencia cuando  $\alpha \neq 1$ , y la función logaritmo si  $\alpha = 1$ , pues se puede demostrar que al tomar límites cuando  $\alpha \rightarrow 1$  (es decir,  $\lambda \rightarrow 0$ ) se obtiene el logaritmo de la variable.

Es posible obtener el valor aproximado de  $\alpha$ ; si en la serie a transformar, se toman grupos de tamaño 5 ó 6 de observaciones consecutivas, o si la serie tiene período estacional  $s$ , se toman grupos de  $s$  datos para tener comparaciones homogéneas. Así en cada grupo se calcula la media y la desviación estándar y se estima mediante una regresión lineal el valor del parámetro  $\alpha$ . Esto es, si de (2.59) se toman logaritmos, entonces la regresión de

$$\log s_i = c + \alpha \log \bar{x}_i$$

estimaré el valor de  $\alpha$ . Para el ejemplo de los pasajeros de avión, se tiene que el valor del coeficiente de determinación es de 98.11%, lo cual refleja la linealidad entre la media y la desviación estándar de cada subconjunto de datos. Además se estima el valor de  $c = -3.70$  y de  $\alpha = 1.3125$  (por lo que  $\lambda = -0.3125$ , es decir  $\lambda \rightarrow 0$ ).

Otro enfoque utilizado para corregir heterocedasticidad es el propuesto por Guerrero<sup>(8)</sup>. Este método también propone el uso de transformaciones de la familia Box-Cox, en la cual Guerrero dice que la transformación óptima será aquella que minimice un coeficiente de variación que se calcula con una serie construida para determinar la posible heterocedasticidad. Esta nueva serie depende de las estadísticas de tendencia central y de dispersión provenientes de los  $s$  grupos formados (como en la técnica anterior). El método se describe de la siguiente forma:

---

<sup>8</sup> Ver referencia 1 de la bibliografía, pp. 102-104.

**Paso 1.** Llevar a cabo las transformaciones de la familia Box-Cox sobre la serie original utilizando valores  $\lambda_1 < \lambda_2 < \dots < 1 < \dots < \lambda_{p-1} < \lambda_p$ .

**Paso 2.** Dividir la serie transformada  $y(\lambda_i)$  en  $s$  grupos consecutivos respetando al período estacional de la serie.

**Paso 3.** Calcular las medias  $m\{y(\lambda_i)_j\}$  y desviaciones estándar  $\sigma\{y(\lambda_i)_j\}$  por cada bloque  $j = 1, \dots, s$ , para cada una de las  $p$  series transformadas y se construyen las  $p$  series que se muestran a continuación:

$\lambda_1$	...	1	...	$p$
$\frac{\sigma\{y(\lambda_1)_1\}}{[m\{y(\lambda_1)_1\}]^{1-\lambda_1}}$	...	$\sigma\{y(1)_1\}$	...	$\frac{\sigma\{y(\lambda_p)_1\}}{[m\{y(\lambda_p)_1\}]^{1-\lambda_p}}$
...	...	...	...	...
$\frac{\sigma\{y(\lambda_1)_j\}}{[m\{y(\lambda_1)_j\}]^{1-\lambda_1}}$	...	$\sigma\{y(1)_j\}$	...	$\frac{\sigma\{y(\lambda_p)_j\}}{[m\{y(\lambda_p)_j\}]^{1-\lambda_p}}$
...	...	...	...	...
$\frac{\sigma\{y(\lambda_1)_r\}}{[m\{y(\lambda_1)_r\}]^{1-\lambda_1}}$	...	$\sigma\{y(1)_r\}$	...	$\frac{\sigma\{y(\lambda_p)_r\}}{[m\{y(\lambda_p)_r\}]^{1-\lambda_p}}$
$CV\left(\frac{\sigma\{y(\lambda_1)_j\}}{[m\{y(\lambda_1)_j\}]^{1-\lambda_1}}\right)$	...	$CV(\sigma\{y(1)_j\})$	...	$CV\left(\frac{\sigma\{y(\lambda_p)_j\}}{[m\{y(\lambda_p)_j\}]^{1-\lambda_p}}\right)$

De esta forma, se calculan los coeficientes de variación para cada una de las series construidas con la media y la desviación estándar. La transformación óptima  $\lambda_i$  para reducir el posible problema de heteroscedasticidad será aquella que minimice el coeficiente de variación  $CV\left(\frac{\sigma\{y(\lambda_i)_j\}}{[m\{y(\lambda_i)_j\}]^{1-\lambda_i}}\right)$ .

## 2.13. Criterios de selección de modelos.

Suponiendo que se ha estimado un conjunto de modelos  $M_1, M_2, \dots, M_m$  y se debe escoger el que mejor explica la serie observada. Existe un criterio de ajuste dentro de la muestra, el cual es aplicable a modelos que utilicen el mismo número de parámetros y se toma como el mejor modelo, aquél que tenga la menor suma de cuadrados de los errores

dentro de la muestra, pero para modelos con número distinto de parámetros es necesario acudir a otros principios.

El problema puede verse como discriminatorio, en donde se tienen distintos modelos  $M_i$  y una serie  $Z_t$  y puede ser abordado, inicialmente, con el enfoque clásico, aunque no es útil comparar la verosimilitud de distintos modelos porque siempre el modelo con más parámetros tendrá mayor verosimilitud. Sin embargo es posible calcular el valor esperado de la verosimilitud para cada uno de los modelos, es decir, el valor que se espera obtener para la verosimilitud sobre muchas realizaciones del proceso si en cada una se estiman los parámetros por máxima verosimilitud y seleccionar el modelo que produzca un valor esperado más alto de esta verosimilitud esperada. Otro enfoque sería utilizar la misma probabilidad para cada modelo  $M_i$  en lugar de tomar la máxima.

### 2.13.1. El criterio AIC de Akaike.

Este criterio está basado en la idea de tomar la probabilidad máxima del modelo, pues sabiendo que la función de verosimilitud de un modelo ARIMA es la siguiente:

$$L(\beta) = -\frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \log v_t | I_{t-1} - \frac{1}{2\sigma^2} \sum_{t=1}^T \frac{e_t^2}{v_t | I_{t-1}}$$

se multiplica por -2 y tomando esperanzas, se obtiene que:

$$E[-2L(\beta)] = T \log \sigma^2 + \sum_{t=1}^T \log v_t | I_{t-1} + E \left[ \sum_{t=1}^T \frac{e_t^2}{\sigma^2 v_t | I_{t-1}} \right]$$

Se demuestra que si se supone que los parámetros no son conocidos, es decir, que se estima con los datos, se obtiene el criterio AIC (debido a Akaike), calculando la esperanza de la expresión anterior, llamando  $k$  al número de parámetros estimados:

$$AIC = E[-2L(\beta)] = T \log \hat{\sigma}_{MV}^2 + 2k$$

donde  $T$  es el tamaño de la muestra utilizado para estimar el modelo,  $\hat{\sigma}_{MV}^2$  el estimador máximo verosímil de la varianza de los errores. Por tanto, para seleccionar el modelo que tenga la verosimilitud esperada máxima equivale a escoger el que minimiza la verosimilitud con signo negativo, que es el que minimiza el criterio AIC dado en la expresión anterior<sup>(9)</sup>.

<sup>9</sup> Un criterio alternativo, conocido por BIC (propuesto por Schwarz) supone que las probabilidades son las mismas para cada modelo supuesto y también selecciona el modelo que minimiza la verosimilitud con signo negativo. A diferencia del criterio de Akaike, el BIC tiene una penalización mayor por la inclusión de nuevos parámetros, con lo que tiende a elegir modelos más parsimoniosos. La diferencia entre el AIC y el BIC es más grande entre más grande sea la muestra.



## 2.14. Revisión del modelo.

Para revisar el modelo, es necesario comprobar que las hipótesis básicas realizadas respecto a los residuos se cumplan, es decir, se debe comprobar que los residuos tengan media cero, varianza constante, no correlación y distribución normal.

### 2.14.1. Contraste de media cero.

Para contrastar la hipótesis de que las perturbaciones tienen esperanza nula, en el caso general de un modelo *ARIMA*, suponiendo  $T$  residuos y  $p + q$  parámetros, se calcula la media de éstos.

$$\bar{a} = \frac{\sum_{i=1}^T a_i}{T}$$

y su varianza

$$\hat{S}_a^2 = \frac{\sum_{i=1}^T (a_i - \bar{a})^2}{T - p - q}$$

y se concluye que  $E[a_t] \neq 0$ , si

$$\frac{\bar{a}}{\hat{S}_a / \sqrt{T}}$$

es significativamente grande con relación a la distribución  $N(0, 1)$ . Este contraste debe de aplicarse una vez revisado el supuesto de no correlación, pues de esa forma se asegura que  $\hat{S}_a$  es un estimador razonable para la varianza de los errores.

### 2.14.2. Contraste de homocedasticidad.

Para revisar la estabilidad de la varianza de la serie, es necesario estudiar el gráfico de los residuos a lo largo del tiempo. Lo que se hace es dividir la serie en dos partes y aplicar un contraste de varianzas en cada tramo (es conveniente que cada tramo disponga de al menos 10 observaciones), si se observa algún cambio de varianza a partir de un punto

$t = n_1$ . En la hipótesis de que ambos tramos tienen la misma varianza, el estadístico

$$F = \frac{\sum_{t=1}^{n_1} \hat{a}_t^2/n_1}{\sum_{t=n_1+1}^T \hat{a}_t^2/(T-n_1)} = \frac{S_1^2}{S_2^2}$$

se distribuirá aproximadamente como una  $F$  con  $n_1$  y  $T - n_1$  grados de libertad. De la misma forma, si se sospecha que hay  $h$  cambios de varianza en los períodos  $n_1, \dots, n_h$ , el contraste de igualdad de varianzas es

$$\lambda = T \log \hat{\sigma}^2 - \sum_{i=1}^h n_i \log S_i^2$$

donde  $\hat{\sigma}^2$  es la varianza de los residuos en toda la muestra y  $S_i^2$  la varianza en el tramo  $i$  de longitud  $n_i$  observaciones. Bajo la hipótesis de que la varianza es la misma en todos los tramos se demuestra que este estadístico será, asintóticamente, una Ji-cuadrada con  $h - 1$  grados de libertad.

### 2.14.3. Contraste de normalidad.

Dado que las perturbaciones del modelo no son observables, el estudio de la normalidad se lleva a cabo con los residuos. Si la hipótesis se cumple, la distribución empírica de dichos residuos debería presentar características similares a las de la distribución normal. Para este contraste se trabaja con el histograma de los residuos y el cálculo de ciertos estadísticos descriptivos para dicha serie, entre los que se encuentra el estadístico Jarque-Bera que contrasta la normalidad de una variable. En el primer caso, si aproximadamente la distribución de los residuos es una normal, la representación de los residuos debería parecer a una campana de Gauss simétrica y con un apuntamiento similar al de la distribución normal. En el segundo, como los momentos poblacionales de orden impar de una variable normal son cero (incluyendo a su coeficiente de asimetría) y su coeficiente de kurtosis o apuntamiento es tres, se tendría que encontrar valores similares a éstos al calcular los momentos muestrales de los residuos.

El contraste de Jarque-Bera plantea en la hipótesis nula la existencia de normalidad de las perturbaciones y en la alternativa la no normalidad. El estadístico del contraste es el de los multiplicadores de Lagrange y se construye a partir de los coeficientes de asimetría y de kurtosis mencionados anteriormente. Su expresión es:

$$JB = T \left[ \frac{m_3^2}{6m_2^3} + \frac{1}{24} \left( \frac{m_4}{m_2^2} - 3 \right)^2 \right] \rightarrow \chi^2_{(2)}$$

La distribución asintótica que sigue el estadístico, bajo la hipótesis nula de normalidad, es una Ji-cuadrada con 2 de grados de libertad, donde los términos  $m_i$  son los momentos muestrales centrados de orden  $i$  de los residuos, es decir,

$$m_i = \frac{\sum_{i=1}^T (\epsilon_i - \bar{\epsilon})^i}{T}$$

#### 2.14.4. Contraste de Ljung - Box sobre las autocorrelaciones.

Uno de los contrastes más importantes a considerar para la validación del modelo, es verificar que los residuos estimados están no correlacionados. Para ello se calculan sus coeficientes de autocorrelación simples ( $\hat{r}_k$ ). Si los residuos son independientes los  $\hat{r}_k$ , para  $k$  no muy pequeño, serán aproximadamente variables aleatorias con media cero, varianza asintótica  $\frac{1}{T}$  y distribución normal. La varianza asintótica es válida para  $k$  grande, pero no para los primeros retardos. Por ejemplo, se puede demostrar que si la serie sigue un proceso  $AR(1)$ , la desviación típica asintótica de la autocorrelación de orden uno ( $\hat{r}_1$ ) para los residuos es  $\frac{\sqrt{(1-\phi^2)}}{T}$ , que puede ser mucho menor que  $\frac{1}{\sqrt{T}}$ . En consecuencia, el valor  $\frac{1}{\sqrt{T}}$  debe considerarse como un límite máximo de la desviación típica de las autocorrelaciones residuales.

El procedimiento habitual de verificar la no correlación de los residuos, consiste en dibujar a sus coeficientes de autocorrelación simple o parcial y trazar dos líneas paralelas a una distancia de  $\frac{2}{\sqrt{T}}$  del origen y comprobar si todos los coeficientes  $\hat{r}_k$  están dentro de los límites de confianza y al ser estos límites al 95% de confianza, se espera que en promedio uno de cada veinte coeficientes estimados, caerán fuera de estas bandas. Por tanto un valor próximo a los límites de confianza  $\pm \frac{2}{\sqrt{T}}$  en los retardos iniciales, debe considerarse un indicio de que el modelo es inadecuado.

Un contraste global de que los primeros  $h$  coeficientes son cero (donde  $h$  debe ser grande) es el contraste de Ljung-Box. Si los residuos son realmente ruido blanco, se sabe que los coeficientes de autocorrelación estimados son asintóticamente normales, con media cero y varianza  $\frac{T-k}{T(T+2)}$ . Por tanto el estadístico,

$$Q(h) = T(T+2) \sum_{j=1}^h \frac{\hat{r}_j^2}{T-j} \quad (2.61)$$

se distribuye como una  $\chi^2_{(h-n)}$ , donde los grados de libertad se refieren al número de coeficientes en la suma ( $h$ ) menos el número de parámetros estimados ( $n$ ). Para modelos no estacionales  $n = p + q + 1$ , ó  $n = p + q$ , según que el modelo tenga o no constante, y para los estacionales (que habitualmente no tienen constante),  $n = P + p + Q + q$ . Se concluye entonces, que el modelo es inadecuado si el valor de  $Q(h)$  obtenido con (2.61) es mayor que el percentil 0.95 de la distribución  $\chi^2$  con  $h - n$  grados de libertad.

## 2.15. Verificación del modelo ajustado.

Los siguientes criterios pueden utilizarse como guía de verificación del ajuste del modelo de serie de tiempo:

- Comparar la *fas* y la *fap* muestrales con los teóricos y observar el parecido de uno con otro.
- Verificar que los residuales, producto del modelo ajustado, sean ruido blanco, lo que indicará que el modelo extrajo toda la información expresada a través de las autocorrelaciones y que fue correctamente ajustado.
- Comparar el criterio Akaike de diversos modelos ajustados por:

$$\begin{aligned} AIC &= -2 * \ln(\text{verosimilitud maximizada}) + 2 * (\text{número de parámetros estimados}) \\ &= -2 \ln f(y/\theta) + 2p \end{aligned}$$

y elegir el modelo que proporcione el menor valor *AIC*.

- Verificar que los residuales se distribuyan de manera normal y como un proceso de ruido blanco.

## **CAPITULO 3. APLICACIONES DEL MODELO ARIMA MULTIPLICATIVO.**

Para aplicar lo expuesto anteriormente se utilizará la serie del Índice al Valor Agregado (IVA), con la cual se analizarán las partes que componen a una serie de tiempo y se hará un pronóstico de dos años. Son 168 datos medidos mensualmente que se toman a partir de enero de 1990 a diciembre de 2003, y la idea es pronosticar el 2004 y hacer la comparación con lo publicado por el INEGI en ese año. Todo dependerá del horizonte de pronóstico que pueda tomarse y eso a su vez está ligado a la estacionalidad que la serie presente.

### **3.1. Análisis de la Serie del Impuesto al Valor Agregado (IVA).**

El Impuesto al Valor Agregado (IVA) es un impuesto indirecto sobre el consumo<sup>(1)</sup>. Un impuesto indirecto es el que no es percibido por el fisco directamente de la persona que soporta la carga del tributo. Se aplica a compra y venta de bienes y servicios vendidos a usuarios finales o consumidores, sean estos personas naturales o jurídicas. Cada eslabón en la cadena de añadido de valor paga a su antecesor en la cadena el IVA correspondiente al precio facturado por este, y a su vez percibe de su sucesor en la cadena el monto correspondiente al impuesto asociado al precio que facturó. Cada actor, excepto el consumidor final, es responsable ante la autoridad tributaria por liquidar y pagar la diferencia entre el IVA pagado o crédito fiscal y el IVA cobrado o débito fiscal.

Otra definición del IVA es la siguiente<sup>(2)</sup>: Es un impuesto indirecto, en virtud de que los contribuyentes del mismo no lo pagan directamente sino que lo trasladan o cobran a una tercera persona y quien lo absorbe es el consumidor final del bien o servicio. Los pagos que se hacen por este impuesto son definitivos, lo cual significa que a diferencia de los pagos provisionales, no se tiene la obligación de presentar declaración anual.

A continuación se presenta la gráfica de la serie original del IVA, publicada en la página de la Secretaría de Hacienda y de Crédito Público ([www.shcp.gob.mx](http://www.shcp.gob.mx)):

---

<sup>1</sup> Definición encontrada en "La Enciclopedia Libre". Ver referencia 21 de la bibliografía.

<sup>2</sup> Definición del Servicio de Administración Tributaria. Ver referencia 17 de la bibliografía.

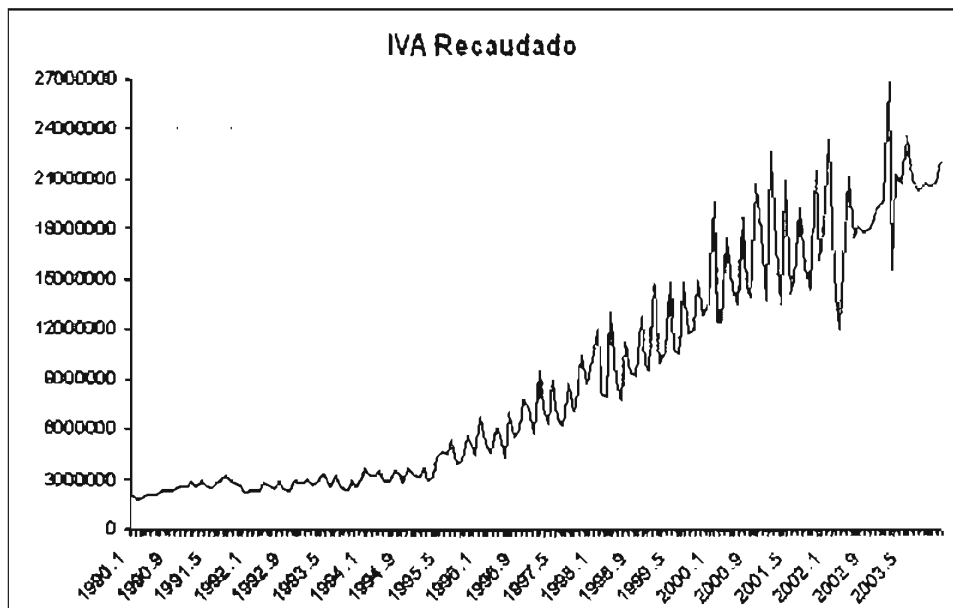


Fig. 3.1. Serie original del IVA de enero-1990 a diciembre-2003

En el gráfico anterior se observa a primera vista que la serie presenta tendencia y estacionalidad, quizá escondidas se encuentren las fluctuaciones locales, por lo que es conveniente hacer el análisis de cada componente por separado. Sin embargo, es importante mencionar que la serie presentada anteriormente está en su forma natural, lo que provoca que el cambio de nivel de la primera fecha a la última sea muy grande, pues es sabido que los precios de hace 13 años no son los mismos que los de ahora por el efecto de la inflación, en general este efecto se presenta en las series involucradas con precios.

Por tal motivo se utilizará la serie del Índice Nacional de Precios al Consumidor (INPC) para indizar a la serie del IVA, es decir, en lugar de tomar los precios crudos se tomarán índices que reflejen el cambio que ha habido de 1990 al año 2003. La gráfica de serie indizada se presenta en la fig. 3.2. A dicha serie se le denominó como IVA Real. Cabe mencionar que este índice tuvo cambios entre 1991 y 1992 cuando se aumentó del 10% al 15%, además de que la recaudación que antes de 1992 controlaba cada estado, a partir de ese año comenzó a controlarla la federación, lo que hizo que el sistema fuera un poco más eficiente. Estos cambios hacen que la serie cambie de nivel observándose un "pulso" a finales de 1991, que se modelaría con un *Análisis de Intervenciones* (ver Apéndice C), lo cual se verá un poco más adelante.

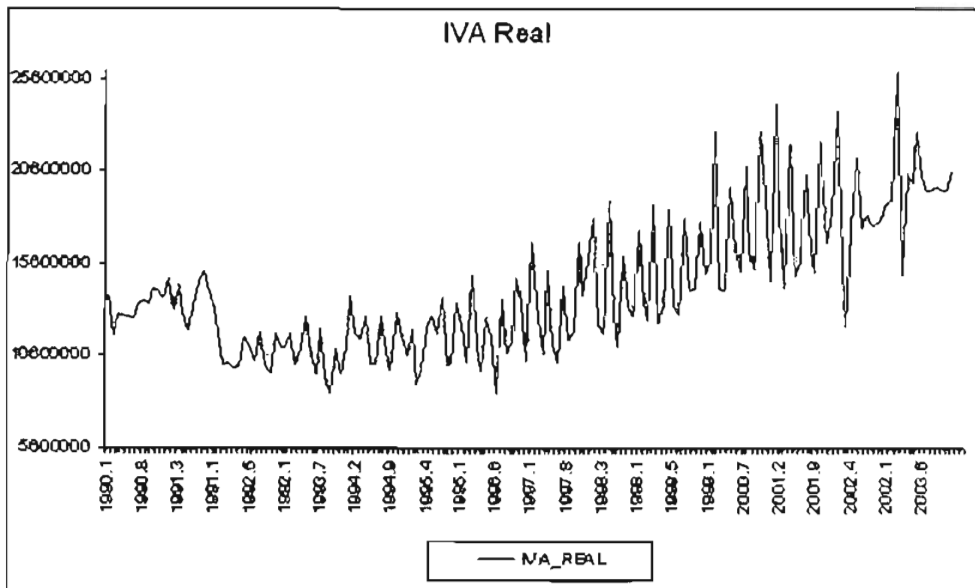


Fig. 3.2. Serie del IVA Real de enero-1990 a diciembre-2003

Para esta serie se puede observar que la tendencia es menos pronunciada que con la serie original aunque se sigue presentando cierto cambio de nivel, el cual hay que analizar, para posteriormente eliminar y así estudiar la estacionalidad de una forma más clara, pues es posible que esta componente se esconda con la tendencia. Esto es importante, pues el tener a la serie a nivel constante (es decir, sin tendencia) permite que sea más fácil identificar los periodos en que se presentan los “picos”. Si no se tiene a esta serie sin tendencia, el nivel sería creciente, por lo que no se tendría algún nivel de comparación. De ahora en adelante cuando se nombre a la serie del IVA se hará referencia a la serie del IVA Real (IVAR).

De acuerdo a lo explicado en el Capítulo I, la tendencia se analiza con base en un análisis de regresión, pues de esa forma es posible analizar la constante de crecimiento. Con base en la regresión aplicada a esta serie, se obtuvieron los siguientes resultados:

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0.691984032
Coefficiente de determinación R <sup>2</sup>	0.478841901
R <sup>2</sup> ajustado	0.475702394
Error típico	2662931.832
Observaciones	168

Tabla 3.1. Resultados de la regresión aplicada la serie del Iva Real

De acuerdo al análisis de regresión aplicado a la serie, el coeficiente de determinación es del 47.88% lo que indica que cerca de la mitad de la varianza de la serie es explicada por el modelo, teniendo como variable independiente al tiempo ( $t = 1, 2, \dots, 168$ ). Además los coeficientes de la regresión resultaron ser significativos a un 95% de confianza y los valores fueron:

	Coefficientes	Error típico	Estadístico t	Probabilidad
Beta <sub>0</sub>	9888361.253	412740.5334	23.95781479	9.3221E-56
Beta <sub>1</sub>	52319.04583	4236.375458	12.34995489	2.86308E-25

Tabla 3.2. Coeficientes de la regresión para estimar tendencia

La gráfica de la tendencia de la serie se presenta en la fig. 3.3., la cual es ascendente conforme pasa el tiempo. Una vez estimada la tendencia, se calcula la serie del promedio móvil para tener explícitamente a la componente Tendencia-Ciclo para posteriormente explicar tanto a la componente Tendencia como a la de Ciclo por separado. Se calcula el promedio móvil (TxC) con pesos constantes y de orden 6, es decir, se toman 13 datos para cada promedio y se puede ver que la serie se suaviza, eliminando la estacionalidad que más adelante se tratará; a esta serie se le conoce como desestacionalizada y también se presentan en la tabla 3.3. junto con la serie original del IVA Real.

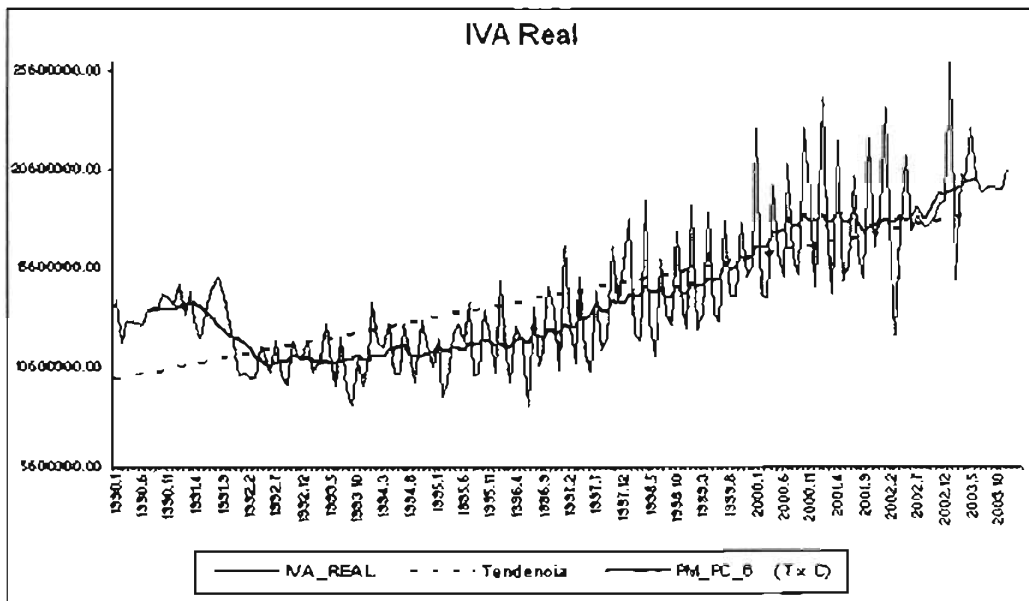


Fig. 3.3. Gráfica de la serie del IVA Real, junto a la componente Tendencia y Tendencia-Ciclo.



El componente ciclo (ver tabla 3.3.) se obtiene como el cociente de la serie de promedios móviles entre la serie de la tendencia y su gráfica se presenta en la fig. 3.4., en donde se puede decir que a principios de 1992 es cuando comienza un nuevo ciclo y conforme transcurre el tiempo, éste se va formando de manera ascendente. Con más temporalidad de la serie se podría dar algún punto máximo para después comenzar a descender, como ocurre en 1991.

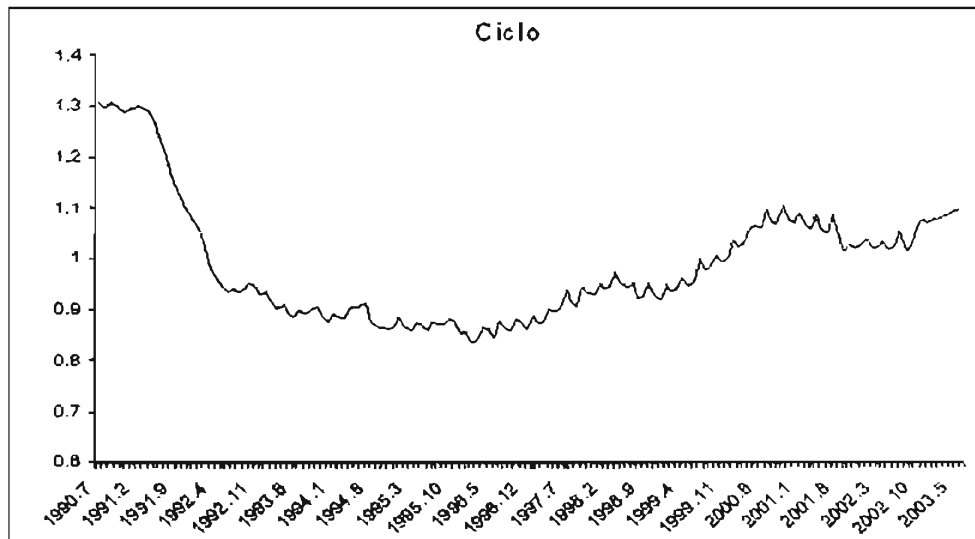


Fig. 3.4. Gráfica del ciclo de la serie del IVA Real de 1990 a 2003

Para realizar el análisis del componente estacional es necesario tener a la serie sin la tendencia y así poder hacer conclusiones apropiadas, pues al tener a la serie nivelada (sin tendencia) es posible comparar el mes de un año con el de cualquier otro año.

Tabla 3.3. Resultados del análisis de Tendencia y Promedios Móviles.

t	Fecha	IVA	INPC	IVA_REAL	Tendencia	PM_PC_6 (T x C)	Ciclo
1.00	1990.1	2092649	15.01	13941641.44	9940680.299		
2.00	1990.2	1796237	15.35	11701905.19	9992999.345		
3.00	1990.3	2005865	15.62	12841183.36	10045318.39		
4.00	1990.4	2032419	15.86	12816121.15	10097637.44		
5.00	1990.5	2060314	16.14	12769190.79	10149956.48		
6.00	1990.6	2094462	16.49	12701091.80	10202275.53		
7.00	1990.7	2274152	16.79	13543760.04	10254594.57	13411479.57	1.307850785
8.00	1990.8	2324410	17.08	13611155.47	10306913.62	13350239.66	1.295270355
9.00	1990.9	2334803	17.32	13479862.39	10359232.67	13556815.04	1.308669809
10.00	1990.10	2506146	17.57	14264039.52	10411551.71	1365541.87	1.302931806
11.00	1990.11	2550674	18.04	14142000.54	10463870.76	13503399.14	1.290478395
12.00	1990.12	2563445	18.60	13778542.98	10516189.8	13543424.36	1.287864199
13.00	1991.1	2815794	19.08	14758739.72	10568508.85	13686864.30	1.295061063
14.00	1991.2	2551794	19.41	13145522.69	10620827.9	13808711.62	1.300153976
15.00	1991.3	2832692	19.69	14387385.02	10673146.94	13852559.01	1.297888906
16.00	1991.4	2577320	19.89	12954632.20	10725465.99	13831388.66	1.289583938
17.00	1991.5	2412395	20.09	12008265.68	10777785.03	13652593.53	1.266734629
18.00	1991.6	2697806	20.30	13289518.60	10830104.08	13340251.21	1.231774978
19.00	1991.7	2983027	20.48	14565811.04	10882423.12	13068156.55	1.200849884
20.00	1991.8	3119677	20.62	15127775.19	10934742.17	12699773.27	1.161414972
21.00	1991.9	2953600	20.83	14181171.56	10987061.22	12458464.85	1.133921492
22.00	1991.10	2782200	21.07	13204647.84	11039380.26	12245836.27	1.109286571
23.00	1991.11	2578143	21.59	11939702.82	11091699.31	12105424.02	1.091394896
24.00	1991.12	2228152	22.10	10081550.37	11144018.35	11976020.11	1.074659044
25.00	1992.1	2304600	22.50	10241312.39	11196337.4	11871425.53	1.060295443
26.00	1992.2	2270073	22.77	9969757.14	11248656.45	11522879.08	1.024378257
27.00	1992.3	2302090	23.00	10008513.13	11300975.49	11102381.62	0.982426839
28.00	1992.4	2697325	23.21	11623213.61	11353294.54	10924244.85	0.962209235
29.00	1992.5	2599728	23.36	11129272.84	11405613.58	10759901.75	0.943386489
30.00	1992.6	2428418	23.52	10326014.92	11457932.63	10695291.87	0.933439933
31.00	1992.7	2823300	23.67	11929788.99	11510251.67	10833833.83	0.941233444
32.00	1992.8	2389400	23.81	10034707.29	11562570.72	10833146.34	0.936915034
33.00	1992.9	2320500	24.02	9661308.22	11614889.77	10912573.93	0.939533147
34.00	1992.10	2870409	24.19	11865393.56	11667208.81	11126160.67	0.9536266
35.00	1992.11	2699800	24.39	11068187.55	11719527.86	11056486.61	0.943424236
36.00	1992.12	2746057	24.74	11099774.29	11771846.9	10936587.02	0.929045978
37.00	1993.1	2976600	25.05	11882595.83	11824185.95	11072872.38	0.936461178
38.00	1993.2	2584160	25.25	10232375.05	11876485	10887421.12	0.916720825
39.00	1993.3	2794800	25.40	11002315.85	11928804.04	10776775.02	0.9034246
40.00	1993.4	3266400	25.55	12785140.69	11981123.09	10875416.77	0.907712632
41.00	1993.5	2753790	25.69	10717450.92	12033442.13	10701327.03	0.889298915
42.00	1993.6	2472900	25.84	9570578.07	12085761.18	10700396.22	0.885372138
43.00	1993.7	3140900	25.96	12097724.59	12138080.22	10913814.84	0.899138467
44.00	1993.8	2484600	26.10	9518922.70	12190399.27	10903400.14	0.894425187
45.00	1993.9	2260400	26.30	8596307.97	12242718.32	11003890.53	0.898811052
46.00	1993.10	2889403	26.40	10943650.89	12295037.36	11137044.61	0.905816248
47.00	1993.11	2546415	26.52	9602226.98	12347356.41	10934973.33	0.885612513
48.00	1993.12	2954320	26.72	11056087.00	12399675.45	10888760.88	0.878148861
49.00	1994.1	3736100	26.93	13874216.42	12451994.5	11134404.47	0.894186427
50.00	1994.2	3179600	27.07	11747204.71	12504313.54	11079430.02	0.886048641
51.00	1994.3	3139237	27.21	11538750.04	12556632.59	11094403.14	0.883549236
52.00	1994.4	3481198	27.34	12733318.97	12608951.64	11427155.23	0.906273223
53.00	1994.5	2790601	27.47	10158214.08	12661270.68	11480360.61	0.906730525
54.00	1994.6	2793100	27.61	10116689.05	12713589.73	11558250.02	0.90912561
55.00	1994.7	3539605	27.73	12763944.69	12765908.77	11631308.55	0.911122644
56.00	1994.8	3171380	27.86	11383056.73	12818227.82	11259869.63	0.878426393
57.00	1994.9	2725500	28.06	9713573.24	12870546.87	11112084.95	0.863373178
58.00	1994.10	3644800	28.21	12922085.26	12922865.91	11167818.09	0.864190511
59.00	1994.11	3299400	28.36	11635320.75	12975184.96	11167533.88	0.860683984
60.00	1994.12	3036410	28.61	10614789.34	13027504	11302468.75	0.86758513

t	Fecha	IVA	INPC	IVA_REAL	Tendencia	PM_PC_6 (T x C)	Ciclo
61.00	1995.1	3563595	29.68	12005847.88	13079823.05	11587533.73	0.885909059
62.00	1995.2	2798696	30.94	9045510.49	13132142.09	11378990.66	0.866499203
63.00	1995.3	3219404	32.76	9826003.85	13184461.14	11298897.44	0.856834217
64.00	1995.4	4336121	35.37	12263280.87	13236780.19	11586246.97	0.875307047
65.00	1995.5	4691300	36.85	12729624.20	13289099.23	11536225.97	0.868096909
66.00	1995.6	4529444	38.02	11912367.43	13341418.28	11430593.48	0.856774988
67.00	1995.7	5362891	38.80	13822533.82	13393737.32	11767428.98	0.878576957
68.00	1995.8	3965031	39.44	10052884.77	13446056.37	11710495.83	0.870924196
69.00	1995.9	4152910	40.26	10315844.79	13498375.42	11763895.72	0.87150455
70.00	1995.10	5536975	41.09	13475117.23	13550694.46	11980823.80	0.884148324
71.00	1995.11	5166313	42.10	12271812.21	13603013.51	11940351.52	0.877772525
72.00	1995.12	4460997	43.47	10262098.37	13655332.55	11622113.28	0.851104375
73.00	1996.1	6752144	45.03	14993650.90	13707651.6	11762116.75	0.858069427
74.00	1996.2	5191740	46.08	11265716.89	13759970.64	11521565.71	0.837324876
75.00	1996.3	4587297	47.10	9739709.10	13812289.69	11624781.40	0.841625947
76.00	1996.4	6125479	48.44	12646068.83	13864608.74	11965154.02	0.86299976
77.00	1996.5	5788842	49.32	11737141.18	13916927.78	11977627.30	0.860651682
78.00	1996.6	4306904	50.12	8592527.11	13969246.83	11831304.37	0.846953634
79.00	1996.7	6981059	50.84	13732412.58	14021565.87	12329518.18	0.87932534
80.00	1996.8	5509404	51.51	10695370.29	14073884.92	12150467.58	0.8633343
81.00	1996.9	5963486	52.34	11394688.69	14126203.97	12110896.45	0.857335522
82.00	1996.10	7810930	52.99	14740688.86	14178523.01	12529949.15	0.883727391
83.00	1996.11	7335726	53.79	13637269.95	14230842.06	12420433.44	0.872782748
84.00	1996.12	5756594	55.51	10369614.06	14283161.1	12311255.97	0.861941966
85.00	1997.1	9531387	56.94	16738877.90	14335480.15	12759593.52	0.890070886
86.00	1997.2	7333418	57.90	12665993.10	14387799.19	12580392.62	0.874379219
87.00	1997.3	6302305	58.62	10751292.23	14440118.24	12680654.69	0.878154491
88.00	1997.4	8998867	59.25	15187394.18	14492437.29	13089538.93	0.903197901
89.00	1997.5	6710201	59.79	11222364.57	14544756.33	13022489.28	0.895339116
90.00	1997.6	6224092	60.32	10317834.02	14597075.38	13189185.91	0.903549894
91.00	1997.7	8775006	60.85	14420915.32	14649394.42	13783511.27	0.940892905
92.00	1997.8	7000203	61.39	11402800.90	14701713.47	13441716.68	0.914295923
93.00	1997.9	7457818	62.15	11998777.17	14754032.52	13374857.02	0.906522132
94.00	1997.10	10469187	62.65	16710183.83	14806351.56	14012869.56	0.94640935
95.00	1997.11	8786376	63.35	13869043.35	14858670.61	13841587.28	0.931549507
96.00	1997.12	10152706	64.24	15804326.17	14910989.65	13829925.23	0.927498814
97.00	1998.1	11877691	65.64	18085843.79	14963308.7	14270967.97	0.953730773
98.00	1998.2	8211804	66.79	12295548.21	15015627.74	14185158.06	0.944692976
99.00	1998.3	7971009	67.57	11796817.49	15067946.79	14285666.40	0.948083146
100.00	1998.4	12989255	68.20	19045455.24	15120265.84	14707118.16	0.972675899
101.00	1998.5	8909798	68.74	12960724.59	15172584.88	14458294.61	0.952922308
102.00	1998.6	7700502	69.56	11070757.83	15224903.93	14355641.48	0.942905226
103.00	1998.7	11272543	70.23	16051389.75	15277222.97	14588563.49	0.95492247
104.00	1998.8	9433913	70.90	13305386.51	15329542.02	14151711.83	0.92316599
105.00	1998.9	9157501	72.05	12709409.20	15381861.07	14229752.73	0.925099549
106.00	1998.10	12773615	73.09	17477650.09	15434180.11	14745306.67	0.955367021
107.00	1998.11	10023023	74.38	13475477.64	15486499.16	14306671.13	0.923815704
108.00	1998.12	8550648	76.19	12534552.71	15538818.2	14296023.25	0.920019983
109.00	1999.1	14711531	78.12	18832312.31	15591137.25	14831631.16	0.951286037
110.00	1999.2	9930174	79.17	12416772.22	15643456.29	14687895.70	0.938916274
111.00	1999.3	10635291	79.90	13310079.89	15695775.34	14759255.33	0.940332988
112.00	1999.4	14817117	80.64	18499018.77	15748094.39	15158088.50	0.962534776
113.00	1999.5	10824322	81.12	13343193.10	15800413.43	14974355.85	0.947719242
114.00	1999.6	10470107	81.66	12822302.23	15852732.48	15156314.05	0.956069502
115.00	1999.7	14822785	82.20	18033660.88	15905051.52	15938123.98	1.002079368
116.00	1999.8	11723206	82.66	14182828.67	15957370.57	15583533.47	0.976572763
117.00	1999.9	11878405	83.46	14233061.79	16009689.62	15716022.09	0.981656888
118.00	1999.10	15028479	83.98	17894240.34	16062008.66	16209787.49	1.00920052
119.00	1999.11	12785301	84.73	15089125.67	16114327.71	16045955.71	0.995757068
120.00	1999.12	13556785	85.58	15840934.25	16166646.75	16180992.08	1.000887341

t	Fecha	IVA	INPC	IVA_REAL	Tendencia	PM_PC_6 (T x C)	Ciclo
121.00	2000.1	19686012	86.73	22698081.81	16218965.8	16803300.27	1.03602785
122.00	2000.2	12444688	87.50	14222635.66	16271284.84	16664368.79	1.024158138
123.00	2000.3	12440203	87.98	14139124.35	16323603.89	16748188.12	1.026010447
124.00	2000.4	17457203	88.48	19729030.04	16375922.94	17398675.39	1.062454645
125.00	2000.5	14538413	88.82	16369205.66	16428241.98	17512891.40	1.066023463
126.00	2000.6	13489407	89.34	15098685.92	16480561.03	17478629.95	1.060560373
127.00	2000.7	18756296	89.69	20912308.63	16532880.07	18122361.63	1.096140633
128.00	2000.8	14634503	90.18	16227551.45	16585199.12	17778830.77	1.071969691
129.00	2000.9	13873803	90.84	15272480.01	16637518.17	17780714.71	1.068711964
130.00	2000.10	20753398	91.47	22689396.21	16689837.21	18395436.26	1.10219387
131.00	2000.11	17877066	92.25	19379048.58	16742156.26	18025000.52	1.076623599
132.00	2000.12	13655005	93.25	14643726.74	16794475.3	17978045.99	1.070473812
133.00	2001.1	22700011	93.77	24209446.17	16846794.35	18379586.83	1.090984222
134.00	2001.2	17084108	93.70	18232180.58	16899113.39	18031231.32	1.066992741
135.00	2001.3	13434574	94.30	14247126.92	16951432.44	17937342.16	1.05816085
136.00	2001.4	20973608	94.77	22130504.51	17003751.49	18473216.47	1.086420046
137.00	2001.5	14166185	94.99	14913365.40	17056070.53	18008815.29	1.055859569
138.00	2001.6	15004665	95.21	15758796.82	17108389.58	17932410.09	1.048164703
139.00	2001.7	19296086	94.97	20318696.77	17160708.62	18640453.75	1.086228673
140.00	2001.8	15651303	95.53	16383686.96	17213027.67	17981821.87	1.044663508
141.00	2001.9	14469603	96.42	15006992.42	17265346.72	17518139.81	1.014641646
142.00	2001.10	21539400	96.85	22238845.99	17317665.76	17779122.79	1.026646606
143.00	2001.11	16189200	97.22	16652180.90	17369984.81	17713572.32	1.019780531
144.00	2001.12	17899355	97.35	18385780.97	17422303.85	17912876.78	1.028157753
145.00	2002.1	23431872	98.25	23848294.37	17474622.9	18105240.14	1.036087603
146.00	2002.2	15363996	98.19	15647231.66	17526941.94	17904055.83	1.02151624
147.00	2002.3	12044697	98.69	12204313.82	17579260.99	18009192.03	1.024456719
148.00	2002.4	17504296	99.23	17639905.63	17631580.04	18246231.84	1.034860847
149.00	2002.5	21157564	99.43	21278348.49	17683899.08	17992501.92	1.017451063
150.00	2002.6	17489824	99.92	17504323.33	17736218.13	18176255.62	1.024810108
151.00	2002.7	18296770	100.20	18259520.58	17788537.17	18759991.59	1.05461126
152.00	2002.8	17806865	100.59	17703300.69	17840856.22	18078848.85	1.013339754
153.00	2002.9	17961688	101.19	17750457.56	17893175.27	18445339.66	1.030858938
154.00	2002.10	18384438	101.64	18088509.98	17945494.31	19039199.80	1.060945966
155.00	2002.11	19405911	102.46	18940357.02	17987813.36	19430629.09	1.079610545
156.00	2002.12	19593929	102.90	19040978.97	18050132.4	19353483.81	1.072207304
157.00	2003.1	26836697	103.32	25974348.63	18102451.45	19503836.77	1.077414118
158.00	2003.2	15534252	103.61	14993438.67	18154770.49	19612876.44	1.080315306
159.00	2003.3	21281351	104.26	20411612.20	18207089.54	19767492.16	1.085703024
160.00	2003.4	20808944	104.44	19924495.64	18259408.59	19906999.58	1.09023244
161.00	2003.5	23660809	104.70	22728486.48	18311727.63	20019083.98	1.093238409
162.00	2003.6	21124596	104.19	20275459.75	18364046.68	20146870.11	1.097082275
163.00	2003.7	20303234	104.34	19458911.82			
164.00	2003.8	20592412	104.65	19677036.27			
165.00	2003.9	20753182	105.28	19713305.15			
166.00	2003.10	20671575	105.66	19564053.91			
167.00	2003.11	20823499	106.54	19545607.20			
168.00	2003.12	22042863	107.00	20601576.69			

La serie del IVA Real sin tendencia se presenta en la fig. 3.5. y se puede observar como la serie ya está nivelada en cero, aunque parece tener problemas de varianza no constante, ya que conforme avanza el tiempo el IVA se aleja de su nivel, pero eso se analizará más adelante. Teniendo ya esta serie, se calcula la componente estacional, es decir el comportamiento que año tras año lleva en promedio la serie mensualmente y se observa que para los años transcurridos de 1993 a 2001 los datos fluctúan de la misma manera, ya que en abril, julio y octubre es cuando toma los valores más altos de recaudación. mientras que en marzo, junio, septiembre y diciembre es cuando se recibe menos IVA por parte de

la gente. Quizá en 1995 exista una ligera desviación respecto a los demás años, pero eso se explica por la crisis que existió en México, lo que obligó que la recaudación sufriera modificaciones ligeramente acentuadas entre mayo y julio de ese año. Cabe mencionar que la serie está desfasada un mes, pues por ejemplo, lo que se recauda en enero es lo que la gente realmente pagó en diciembre, pero que se reporta un mes después.

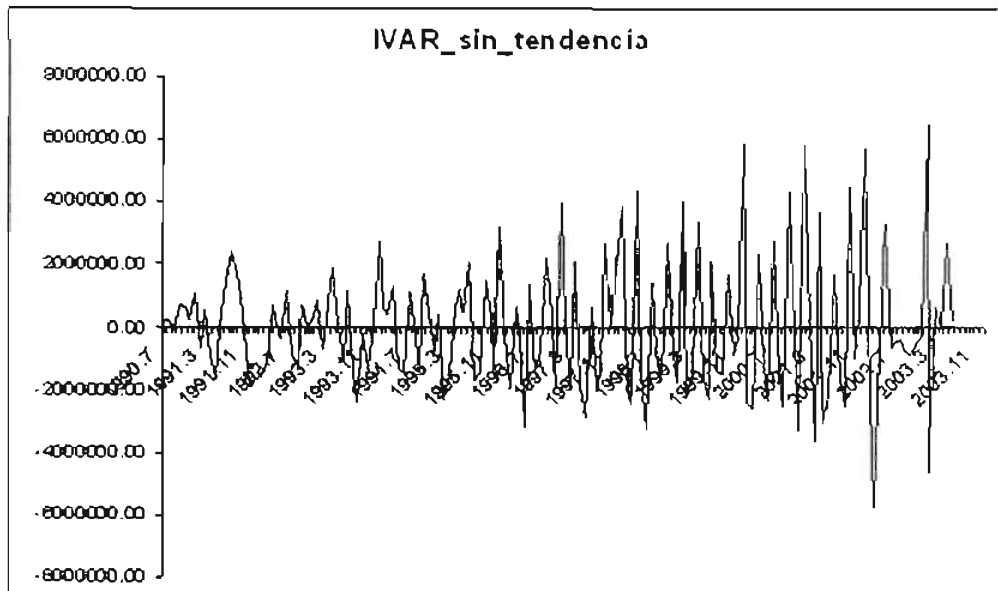


Fig. 3.5. Gráfica del IVA Real sin tendencia.

Por otro lado, mientras que en los años citados anteriormente el comportamiento era el mismo que el componente estacional promedio, para los restantes las fluctuaciones son diferentes, ya que en 1991 y 1992 que fueron los años en donde se dio la intervención al cobrar un 5% más de lo que se venía cobrando, hizo que la serie se desestabilizara y perdiera el comportamiento en que venía desarrollándose. Así también se observa que en 2002 y 2003 la serie cambió ligeramente, en el sentido de que para estos años los puntos de máximo y mínimo recaude no son los mismos o no se muestran tan claros como en los demás años, por ejemplo en mayo se tienen los valores más altos, cuando en todos los años anteriores los máximos se alcanzaban en abril. Las gráficas de la componente estacional comparadas con los años de 1990 a 2003 se presentan en la fig. 3.6. y 3.7.

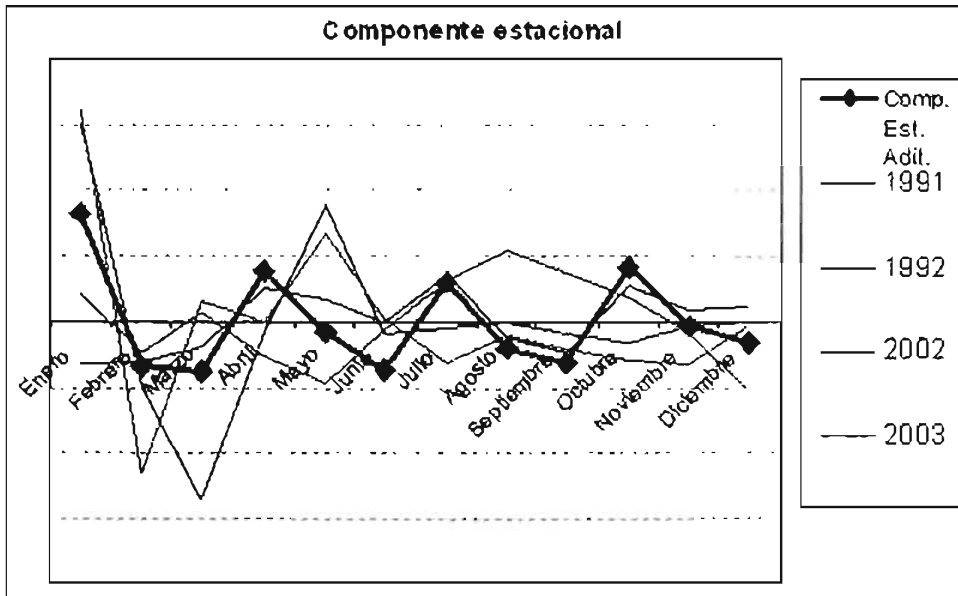


Fig. 3.6. Gráfica del componente estacional para los años 1991, 1992, 2002 y 2003.

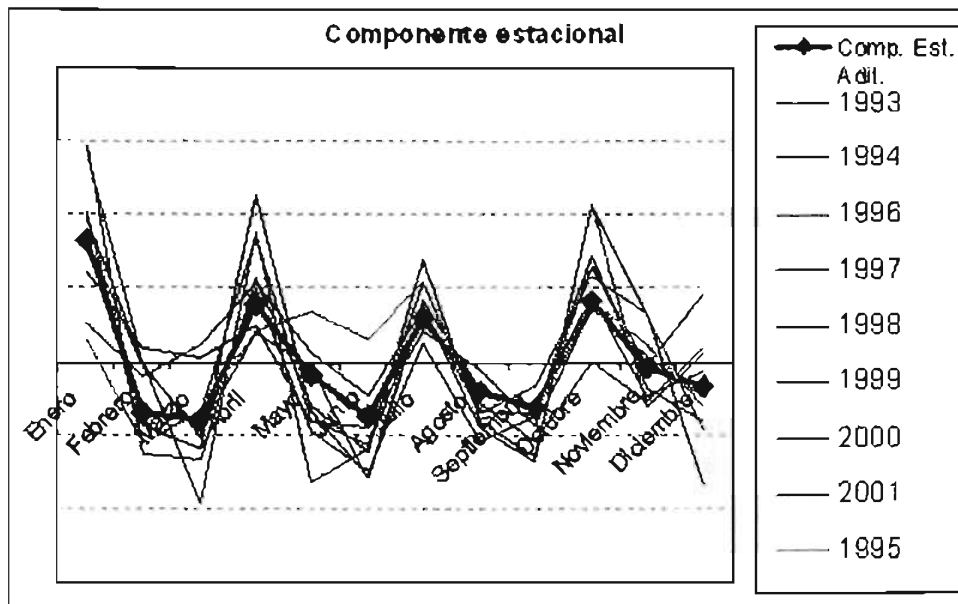


Fig. 3.7. Gráfica del componente estacional para los años de 1993 a 2001.

Por lo presentado anteriormente se puede concluir que la serie del IVA Real presenta tanto tendencia como estacionalidad de período 3 ó 12 (meses), lo cual será removido más

adelante con diferencias estacionales y no estacionales. Es más probable que la serie presente una estacionalidad trimestral, ya que cada tres meses es cuando se observan los puntos máximos y mínimos, aunque no debe descartarse una posible estacionalidad semestral o anual. A continuación se procede a identificar alguna transformación que estabilice a la varianza, pues como se observa en la fig. 3.6. la serie, ya sin tendencia, tiene problemas de varianza no constante.

Aplicando el algoritmo de Guerrero presentado en el capítulo 2.12 el valor de  $\alpha$  obtenido fue de 1.6, por lo que la transformación sugerida para  $Z = IVAR$  fue la siguiente:

$$T(Z) = -\frac{Z^{-0.6} - 1}{0.6}$$

La serie transformada se presenta en la fig. 3.8. en donde se observa que la varianza se mantiene constante, respecto a la serie presentada en la fig. 3.2. en donde conforme avanzaba el tiempo, los picos se iban haciendo más grandes, por lo que se tomará esta serie para continuar el análisis.

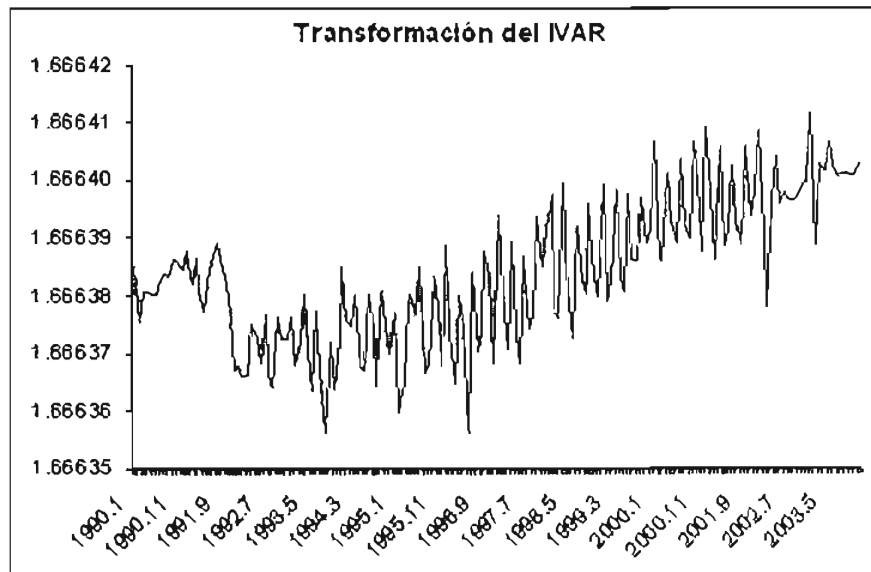


Fig. 3.8. Serie del IVA Real transformada según la metodología de Guerrero.

El siguiente paso es remover esa intervención localizada a finales de 1991, que se logró mediante una intervención explicada en el Apéndice C. De esta forma, los residuales ajustados, es decir ya sin intervención, formarán una nueva serie que tendrá las mismas características de la serie transformada y será ésta la empleada para el análisis de la *fas* y la

*fap* y la estimación de los parámetros para el modelo. En la fig. 3.9. se presenta la gráfica comparativa de las series antes mencionadas.

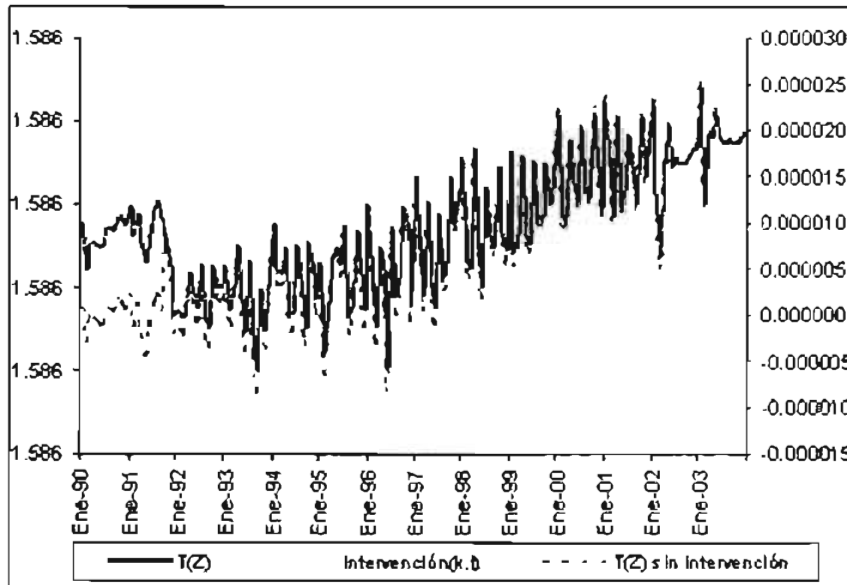


Fig. 3.9. Comparación de la serie transformada, serie sin intervención y la intervención que fue removida.

Se puede observar en la gráfica anterior que el nivel de la serie transformada toma otro rumbo, pues ya está al mismo nivel que el resto de la serie. La estimación de los parámetros para  $I_{k,t}$  fueron  $\omega_0 = 0.00000939$ ,  $\omega_1 = 0.00001006$  y  $\delta = 1.0625$  y el modelo inicial con el que se partió para la estimación de los residuales que mejor se ajusten a la serie transformada, fue la media de los valores de los siguientes dos años en donde se dio el cambio de nivel, es decir, el promedio de los datos de enero de 1992 a diciembre de 1993, que fue de 1.58589 y esto es posible ya que la serie se comporta constante en ese período.

Como puede observarse, la serie sin intervención presenta tendencia y ya no tiene problemas de varianza constante, lo que facilita en gran medida su análisis con los modelos *ARIMA*. El análisis de la función de autocorrelación simple (*fas*) y de la función de autocorrelación parcial (*fap*) se lleva a cabo con mucho cuidado, ya que en el análisis de estacionalidad no fue posible determinar el período estacional, aunque después de muchos intentos y con pruebas y errores encontrados, se determinó que la estacionalidad sería de orden trimestral, por lo que se harán diferencias no estacionales pues la serie presenta tendencia y diferencias estacionales de orden 3, con el objeto de determinar el modelo que estime los parámetros.



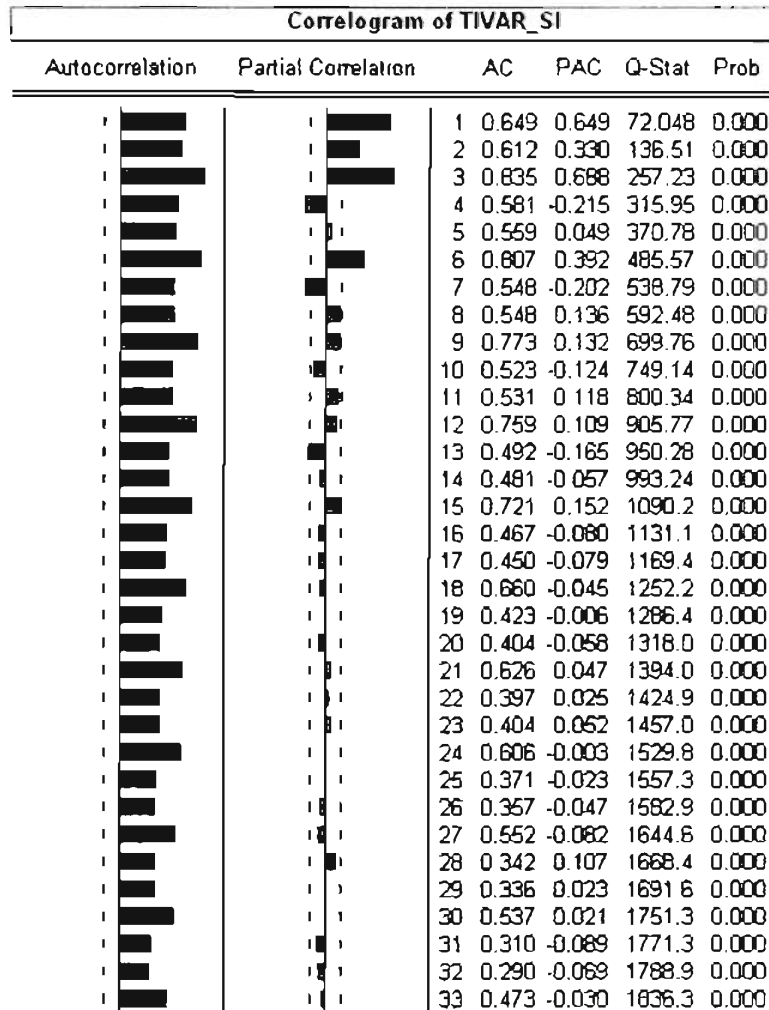


Fig. 3.10. Correlograma de la serie transformada sin intervención.

En la fig. 3.10., se muestra la *fas* de la serie sin intervención, en donde puede apreciarse el decaimiento lineal en la parte no estacional, es decir en la parte regular y el mismo comportamiento descendente en la parte estacional, con picos en los rezagos 3, 6, 9, 12,... , lo que viene a confirmar que la elección de la estacionalidad debe ser trimestral. Por otro lado, la *fap* de la serie tiene un comportamiento descendente en la parte estacional en los rezagos 3, 6, 9, 12 y en la parte regular podría tomarse como un decaimiento súbito o exponencial, lo cual hace suponer el uso de parámetros autorregresivos y de media móvil. Por lo pronto se realizarán la diferencia regular y estacional de orden 3.

Al construir la serie diferenciada regular y estacionalmente, se observa en su correlograma (ver fig. 3.11.) que todavía existe el decaimiento lineal en la parte estacional.

pero eso no es problema ya que esos rezagos altos ayudarán a construir el modelo de mínimos cuadrados que estime los parámetros con los que se modelará a la serie.

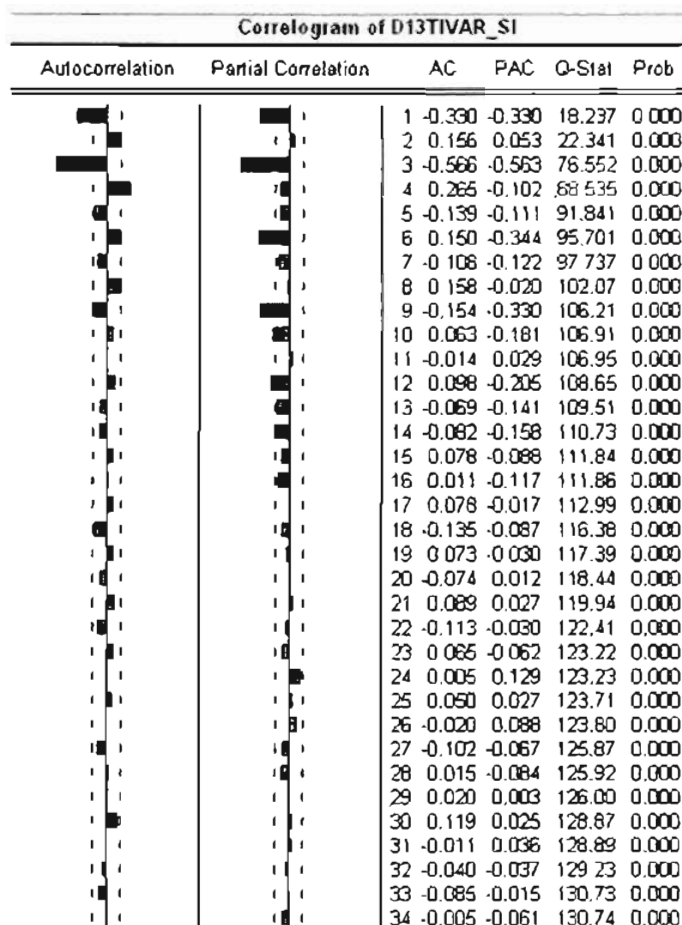


Fig. 3.11. Correlograma de la serie transformada diferenciada regular y estacionalmente.

Con la *fas* de la serie diferenciada se tiene que en los rezagos de orden 3, los coeficientes de autocorrelación simple y parcial son significativamente distintos de cero y que, por un lado, la *fas* decae súbitamente en el rezago 6, que corresponde al segundo rezago estacional y la *fap* lo hace exponencialmente, lo que supone que para la parte estacional se utilicen parámetros de media móvil. Para la parte regular se tiene que en la *fas* y la *fap*, el primer rezago es distinto de cero significativamente y después de éste, ambas decaen ya sea exponencial o súbitamente, eso dependerá de los parámetros estimados, pues es común que gráficamente no pueda detectarse exactamente el tipo de decaimiento y por tanto, se puede suponer el uso de parámetros autorregresivos y de media móvil para la parte regular.

Es posible probar que ya no es necesaria otra diferencia regular o estacional, inspeccionando la gráfica y observar que ya no presenta tendencia, o también analizando la *fas* en el primer período, la cual es de -0.330, que no es considerada como alta y no hace a la serie como un paseo aleatorio, cuyo proceso tiene un coeficiente de autocorrelación alto en el primer rezago, lo que indicaría que no es estacionario. Y a pesar de que la serie diferenciada no presente síntomas de ser estacionaria, es posible ajustar un modelo que explique los rezagos en que las autocorrelaciones son distintas de cero significativamente.

De esta forma, se construyen modelos que incluyen términos autorregresivos y de medias móviles en la parte regular y solamente de medias móviles en la parte estacional, variando los órdenes de 1 a 6. Después de haber intentado con varios modelos y verificar las hipótesis de los residuales de la estimación (los cuales no se presentan por cuestión de espacio en el trabajo), se encontró una combinación de parámetros de medias móviles, que satisfacen los supuestos.

El modelo estimado, la significancia y otros estadísticos que se detallan, se presentan en la tabla 3.4. a continuación:

Dependent Variable: D(TIVAR\_SI,1,3)  
 Method: Least Squares  
 Date: 02/22/05 Time: 00:16  
 Sample(adjusted): 1990:05 2003:12  
 Included observations: 164 after adjusting endpoints  
 Convergence achieved after 18 iterations  
 Backcast: 1989:09 1990:04

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.730409	0.074817	-9.762603	0.0000
MA(2)	-0.258629	0.076105	-3.398341	0.0009
SMA(3)	-0.897415	0.076153	-11.78437	0.0000
SMA(6)	0.165401	0.080232	2.061522	0.0409
R-squared	0.642931			
Adjusted R-squared	0.636236			
Akaike info criterion	-22.46597			
Schwarz criterion	-22.39036			
Durbin-Watson stat	1.970775			
Inverted MA Roots	.99	.86	.64	-.26
	-.32+.55i	-.32-.55i	-.43-.75i	-.43+.75i

Tabla 3.4. Resultados de la estimación del modelo *ARIMA*.

De la serie original, que consistía de 168 observaciones, sólo quedaron 164, pues se hizo una diferencia regular y una estacional de orden 3. Con estos datos, se obtuvo una significancia por debajo del 5% para todos los parámetros, siendo el de orden 6 el que está muy cercano a ya no ser parte de la región de rechazo para la hipótesis de la nulidad del parámetro. Es curioso observar que no fue necesario el uso de parámetros autorregresivos, esto quiere decir que la serie está basada, principalmente, en los errores cometidos en el pasado (de hace 6 meses a lo más) que en sus propios valores. Además de ser significativos los parámetros, se observa que cumplen las condiciones de invertibilidad, ya que  $\theta_1 + \theta_2 < 1$ ,  $\theta_2 - \theta_1 < 1$  y  $-1 < \theta_2 < 1$ .

Dado que no es un modelo de regresión, se desprecia el valor de  $R^2$  y  $R^2$  ajustado, que aunque no está mal, no es posible su análisis para las series de tiempo. De acuerdo al estadístico Durbin-Watson se rechaza la hipótesis nula de que los residuales están correlacionados a un período (y sólo a un período). Por el criterio de Akaike se obtuvo el valor más pequeño de todos los modelos analizados, lo que lo convierte en el mejor modelo por su verosimilitud respecto a la muestra. Además por el criterio de Schwarz el coeficiente no cambia mucho y esto se debe a que la muestra es la misma. Por tanto se tomará este modelo como definitivo, salvo que el análisis de residuales resulta inadecuado.

Lo anterior fue una revisión de los supuestos de los parámetros y del modelo en general, ahora es turno de analizar a los residuales, los cuales tienen su correlograma en la fig. 3.12.

Según el gráfico de la *fas* y la *fap* se observa que por el criterio de Ljung-Box todos los coeficientes de autocorrelación simple están dentro de las bandas de confianza, y esto se precisa con el p-value mayor a 0.05, lo cual indica que a un 95% de confianza, los errores se distribuyen como ruido blanco. Con esto se ha probado el supuesto de independencia y que los errores tienen una media cero.

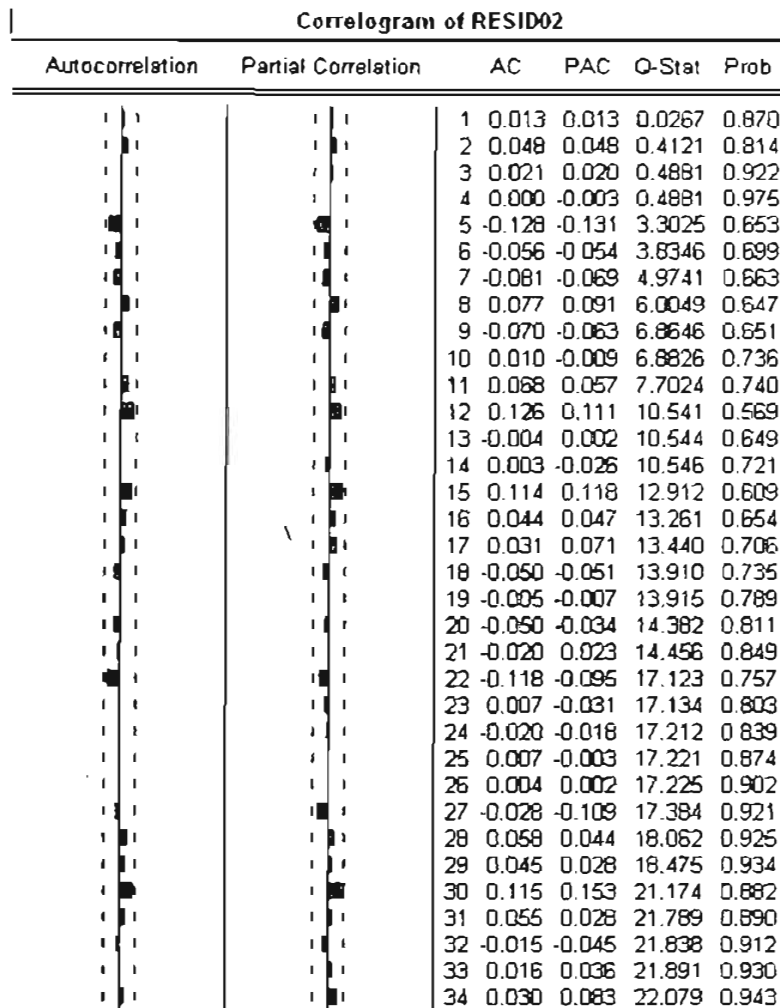


Fig. 3.12. Correlograma de los residuales.

Para analizar el supuesto de varianza constante, se realiza el gráfico de los residuales y en la fig. 3.13. se observa como el comportamiento de éstos viene de forma constante en el tiempo y salvo algunos valores extremos, puede concluirse que los residuales tienen varianza constante.

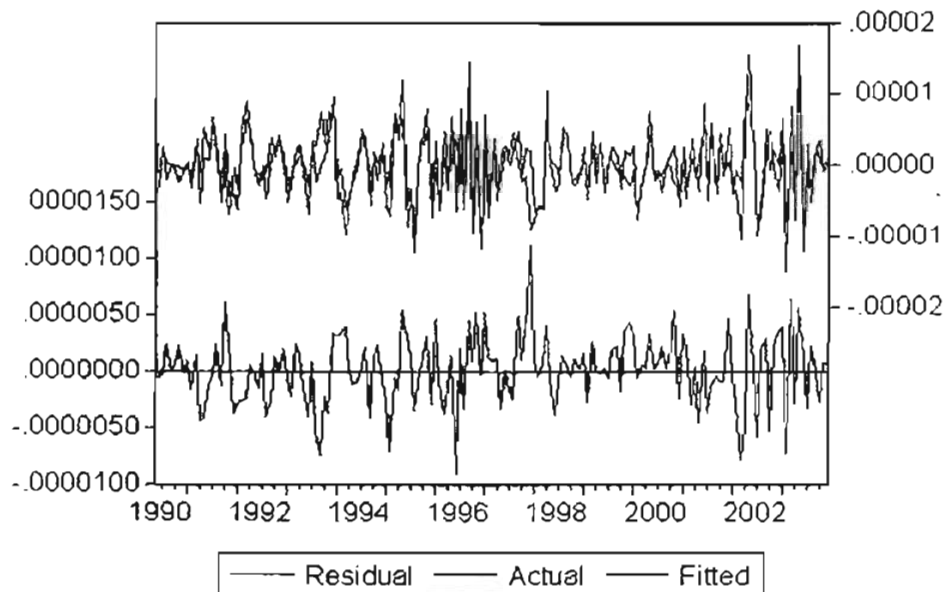


Fig. 3.13. Gráfico de los residuales, serie actual y estimación.

Otro aspecto que se debe revisar es el supuesto de normalidad, y esto se probará con la prueba de Jarque-Bera, la cual se muestra en la fig. 3.14. y confirma, dado su p-value o nivel de significancia mayor a 0.05, que no es posible rechazar la hipótesis nula que la serie se distribuye de manera normal.

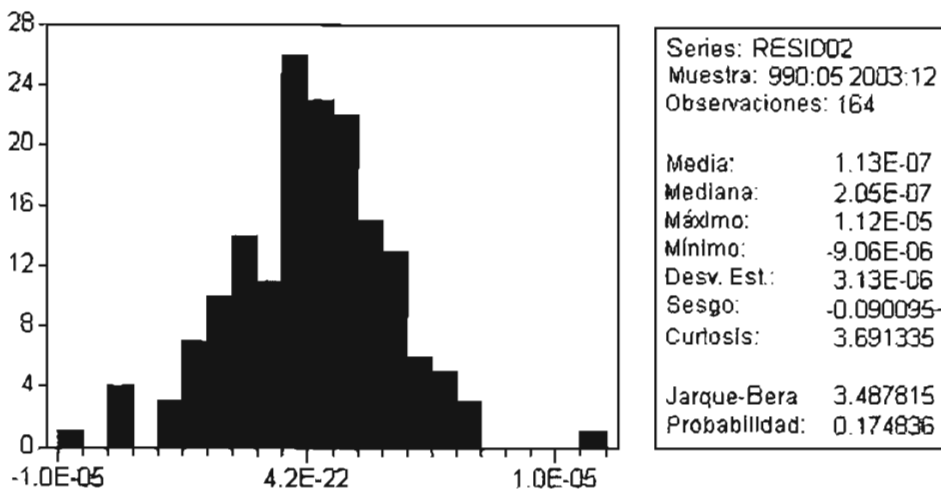


Fig. 3.14. Verificación del supuesto de normalidad de los residuales.

De esta forma, una vez revisados todos los supuestos, se concluye que el modelo  $ARIMA(0, 1, 1) \times (0, 1, 1)_3$  es el adecuado para el pronóstico del IVA y una vez teniendo la estimación, se realizan operaciones inversas al proceso que llevó a tener estos parámetros y poder comparar el pronóstico con lo publicado en los meses de 2004.

Con el modelo propuesto se obtiene la serie de pronóstico a 6 meses, misma que no puede tener un horizonte más largo debido a que el parámetro de orden 6 en media móvil no es tan significativo como los demás, lo que ocasiona que haya errores de pronóstico muy grandes. De tal forma que la serie pronosticada se presenta en la fig. 3.15.

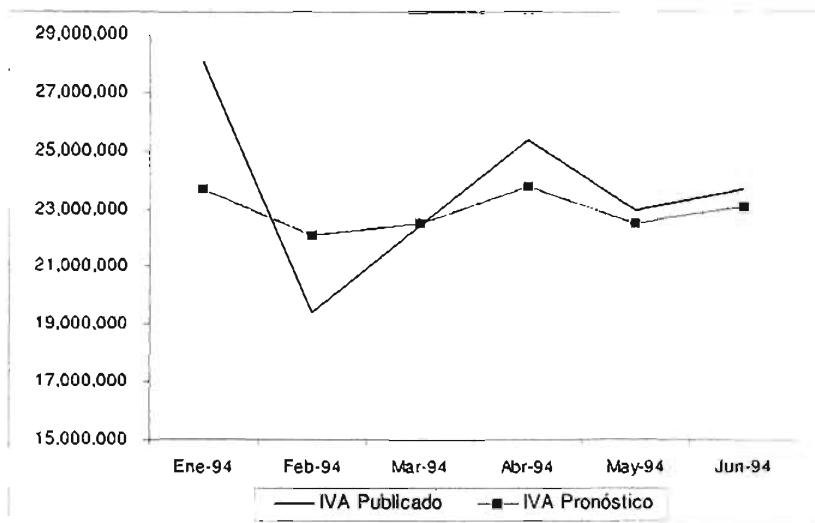


Fig.3.15. Gráfica comparativa de la serie del IVA original y su pronóstico.

Según se observa gráficamente en el primer período es en donde se presenta el mayor error, mismo que se explica por la inestabilidad de la recaudación del IVA a principios de año, y este mismo efecto se observa en las fig. 3.6. y 3.7., en donde el componente estacional para enero es el que tiene mayor dispersión. Por otro lado, se presentan las operaciones inversas (ver Tabla 3.5.) que llevaron a tener la serie del IVA Original Pronosticada, para poder compararla con la publicada por la Secretaría de Hacienda.

Inicialmente, con la serie de pronóstico de los residuales (que representan la serie desintervenida) se obtiene la serie con la intervención, pues hay que recordar que,

$$T(Z_t) = I_{k,t} + N_t$$

Por tal motivo, basta con sumar la serie de los residuales pronosticados con el modelo inicia que generó la intervención. De esta manera se obtuvo la serie transformada, la cual hay que operar inversamente mediante la función:

$$Z_t = [(-0.6 * T(Z_t)) + 1]^{(\frac{1}{0.6})} * C$$

en donde  $C$  es una constante que corrige el hecho de aplicar funciones inversas para tener la serie original (ver Apéndice D). El valor para cada pronóstico se presenta en la Tabla 3.5.

Una vez obtenida la serie sin transformación (Serie Real Pronóstico), se procede a aplicarle el efecto de la inflación, mediante la serie del INPC, aplicando la operación inversa que llevó a quitarle ese efecto. Finalmente se tiene la serie pronóstico del IVA, la cual es comparable a la publicada por Hacienda y tanto, el residuo como los resultados de las operaciones enunciadas anteriormente, se muestran a continuación:

Fecha	Pronóstico IVA Desintervenida	Modelo + Intervención	T(IVAR)	Desv Pronóstico	Var Pronóstico	c
Ene-94	0.00002131	1.585889	1.58590997	0.00000317	0.0000000000100	1.0149
Feb-94	0.00001946	1.585889	1.58590812	0.00000328	0.0000000000108	1.0146
Mar-94	0.00001986	1.585889	1.58590852	0.00000328	0.0000000000108	1.0149
Abr-94	0.00002111	1.585889	1.58590977	0.00000333	0.0000000000111	1.0163
May-94	0.00001986	1.585889	1.58590851	0.00000334	0.0000000000112	1.0154
Jun-94	0.00002040	1.585889	1.58590905	0.00000334	0.0000000000112	1.0158

Fecha	IVA Real Pronóstico	INPC	IVA Pronóstico	IVA Publicado	Residuo
Ene-94	22,015,744.69	107.661	23,702,370.89	28,061,544.00	4,359,173.11
Feb-94	20,391,115.55	108.305	22,084,597.70	19,439,571.00	- 2,645,026.70
Mar-94	20,730,282.63	108.672	22,528,012.74	22,443,279.00	- 84,733.74
Abr-94	21,860,371.63	108.836	23,791,954.06	25,403,607.00	1,611,652.94
May-94	20,735,600.61	108.563	22,511,190.09	22,987,485.00	476,294.91
Jun-94	21,212,082.72	108.737	23,065,382.39	23,710,003.00	644,620.61

Tabla 3.5. Resultados del pronóstico.

Como puede observarse los residuos representan una cantidad despreciable respecto a lo pronosticado y además disminuyen conforme avanza el horizonte de pronóstico. De esta forma, se concluye que el modelo elegido fue correcto, ya que cumple con todos los supuestos expuestos en el capítulo 2 y tiene un buen grado de pronóstico.



Finalmente se presenta el modelo empleado para el pronóstico de la serie del IVA:

$$\nabla \nabla_3 W_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta_3 B^3 - \theta_6 B^6) \varepsilon_t$$

$$\Rightarrow (1 - B)(1 - B^3)W_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} + \theta_1 \theta_3 \varepsilon_{t-4} \\ + \theta_2 \theta_3 \varepsilon_{t-5} - \theta_6 \varepsilon_{t-6} + \theta_1 \theta_6 \varepsilon_{t-7} + \theta_2 \theta_6 \varepsilon_{t-8}$$

$$\Rightarrow W_t = W_{t-1} + W_{t-3} - W_{t-4} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} + \theta_1 \theta_3 \varepsilon_{t-4} + \theta_2 \theta_3 \varepsilon_{t-5} - \theta_6 \varepsilon_{t-6} +$$

$$\therefore W_t = W_{t-1} + W_{t-3} - W_{t-4} + \varepsilon_t - (-0.730409)\varepsilon_{t-1} \\ - (-0.258629)\varepsilon_{t-2} - (-0.897415)\varepsilon_{t-3} + (-0.730409)(-0.897415)\varepsilon_{t-4} \\ + (-0.258629)(-0.897415)\varepsilon_{t-5} - (0.165401)\varepsilon_{t-6} + (-0.730409)(0.165401)\varepsilon_{t-7} \\ + (-0.258629)(0.165401)\varepsilon_{t-8}$$

donde,

$$W_t = -\frac{Z^{-0.6}-1}{0.6}$$

### 3.2. Análisis de la Serie del Agregado Monetario M4.

En la actualidad se ha difundido por los medios de comunicación, que el Banco de México ha implementado la política monetaria de incrementar el denominado corto. Este término representa un recorte en la oferta monetaria, es decir, una disminución del dinero que existe en un país, lo cual implica un impacto económico importante en la economía de los individuos y empresas. Por ello, el propósito de este trabajo es analizar, el comportamiento de la demanda de dinero, en particular se analizará el agregado monetario M4, el cuál es el agregado monetario más completo ya que es definido en función del resto de los agregados.

En la nueva estructura considerada por el Banco de México, se considera un indicador estrecho M1, que incluye los billetes y monedas en poder del público, las cuentas de cheques en moneda nacional y extranjera en poder de residentes del país y a los depósitos en cuenta corriente con interés en moneda nacional. El agregado siguiente es el M2 el cuál mide el esfuerzo de ahorro interno total e incluye al resto de los instrumentos bancarios y no bancarios en poder de los residentes del país. El agregado que sigue es el denominado

M3, el cuál incorpora al agregado M2 otros activos financieros internos que se encuentran en poder de no residentes del país. Por último se tiene al agregado M4, el cual incluye también la captación que realiza la banca mexicana a través de sus agencias en el extranjero de los residentes y de los no residentes. Por ello, el agregado M4 es el agregado más completo en la economía, de ahí el interés de su análisis. En resumen, el M4 se forma de la siguiente manera:

- M1=** Billetes y monedas en poder del público
  - Cuentas de cheques m.n en bancos residentes
  - Cuentas de cheques en moneda extranjera en bancos residentes
  - Depósitos en cuenta corriente m.n. en bancos residentes.
- M2 =** M1 + activos financieros internos en poder de residentes.
- M3 =** M2 + activos financieros internos en poder de no residentes.
- M4 =** M3 + captación de sucursales y agencias de bancos mexicanos en el exterior.

Los datos fueron obtenidos del portal del Banco de México<sup>(3)</sup> y desde enero de 1986 se tienen mensualmente 216 datos hasta llegar a diciembre de 2003. Con estos datos, tratará de pronosticarse los valores del M4 para el 2004 y compararlo con lo publicado en la página antes mencionada para el mismo año. Para comenzar el análisis es indispensable graficar la serie como se muestra en la Fig. 3.16. y explicar los detalles más importantes. Cabe mencionar que los datos obtenidos del BANXICO son datos traídos a pesos corrientes, es decir, afectados por la inflación (INPC), por lo que hay que indexar al M4 para reflejar su valor real. La forma de llevar esto a cabo es dividir la serie original (M4) por el INPC y multiplicarlo por 100 y así tenerlo sobre alguna base. La gráfica de esta serie, que será llamada M4 Real, se presenta también en la Fig. 3.16.

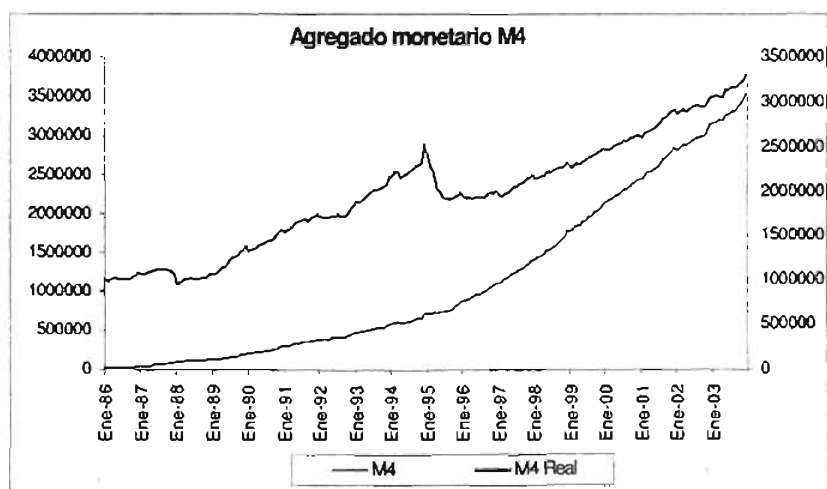


Fig. 3.16. Gráfica de la serie M4 Original y el Real.

<sup>3</sup> Ver referencia 16 de la bibliografía.

Como primer paso se explora la gráfica y se identifican componentes que influyen en el pronóstico de la serie. Inicialmente se observa que la serie M4 Real tiene una tendencia ascendente, así como un cambio de nivel entre los años 1995 y 1996, lo cual indicaría que será necesario hacer algún análisis de intervención. A simple vista no se percibe algún patrón de estacionalidad, pero eso se analizará con más detalle.

Es necesario identificar cada componente por separado, ya que su análisis será de vital importancia en el pronóstico de la serie. El componente de tendencia suele traer consigo a la estacionalidad, es decir, el hecho de tener una serie en la que el nivel medio aumente conforme avanza el tiempo, no permite tener un nivel constante con el cual se comparen los valores año tras año. Es por eso que se identificará mediante promedios móviles y se tratará de eliminar con diferencias regulares más adelante.

La técnica de promedios móviles se hará con pesos constantes (SIM\_PC\_6) y de longitud 6, en donde se supondrá que la serie M4 Real tiene una estacionalidad anual, es decir, que debió pasar un año para que los valores de la serie fueran iguales o al menos tuvieran el mismo comportamiento. Esta suposición se confirmará más adelante con un estudio de la posible estacionalidad.

Por otra parte se calculan los promedios móviles con pesos normales (SIM\_PN\_6), en donde los pesos van cambiando de acuerdo al binomio de Newton, de tal forma que asigne los mismos pesos en las colas y en el dato central otorgue mayor peso. De esta forma se analizaría únicamente los patrones estacionales, en caso de existir tal. Las gráficas de las series de promedios móviles se presentan en la Fig. 3.17.

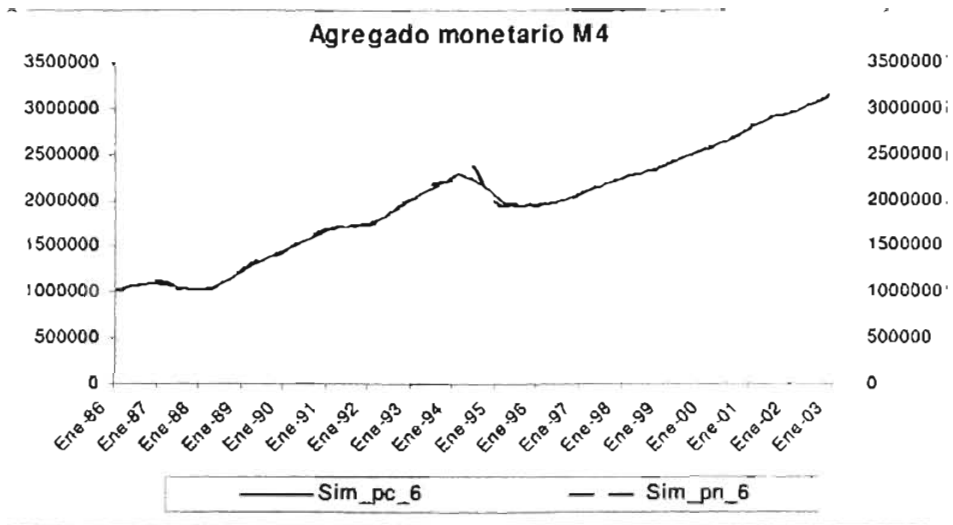


Fig. 3.17. Gráfica de promedios móviles a pesos constantes y normales del M4 Real.

Se observa en la figura anterior que ambas series son prácticamente las mismas, lo que indicaría que no existe estacionalidad en esta serie, ya que ambos promedios móviles modelan la tendencia. Para confirmar el hecho de que no existe patrón estacional se grafican las componentes estacionales por mes en todos los años, como se aprecia en la Fig. 3.18.

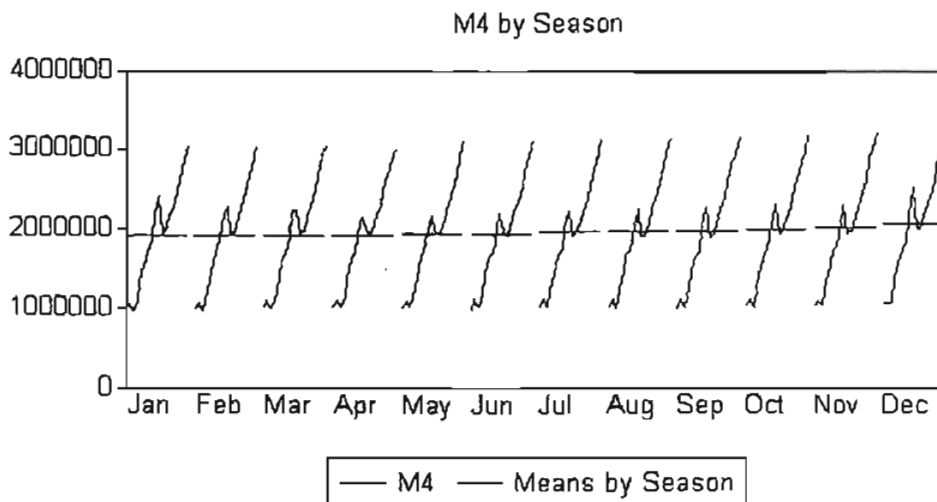


Fig. 3.18. Comportamiento de la serie en todos los años mensualmente.

Es notorio que el patrón estacional no existe para este agregado, ya que mes a mes la media por año es casi la misma, lo que indica que no hay mes que sobresalga de otro por tener mayor o menor demanda de dinero. Esto quizá sea debido a que el M4 contempla demasiados parámetros que al promediarse toma un nivel constante año tras año.

Para dejar bien claro que la serie sí presenta tendencia, se presenta el Boxplot en la Fig. 3.19. en donde es posible observar que conforme avanza el tiempo, las medias van aumentando, lo que es un indicio de que la serie tiene tendencia ascendente.

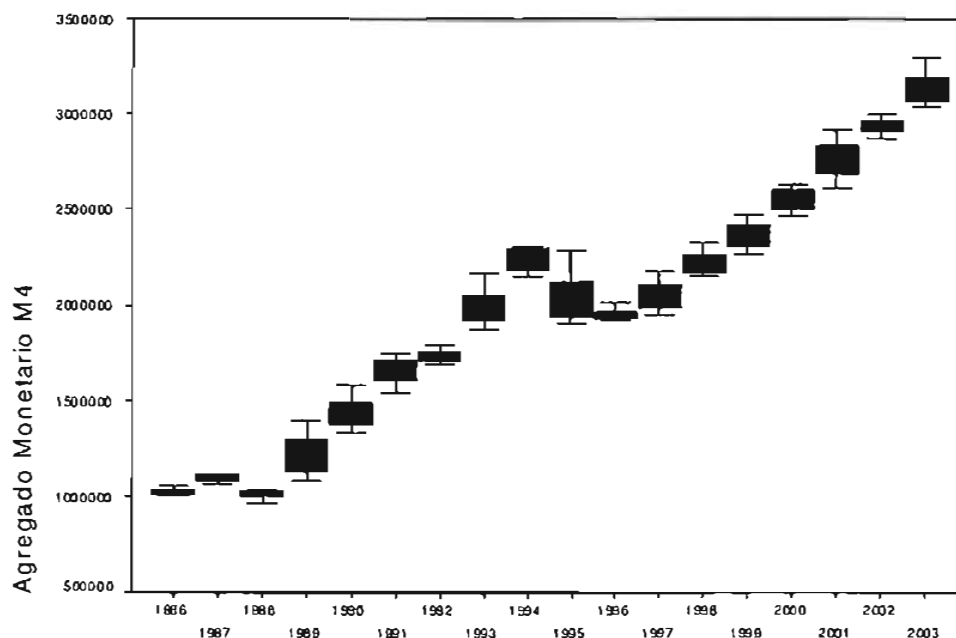


Fig. 3.19. Boxplot de la serie M4 Real.

Una vez analizado tanto la tendencia ascendente como la estacionalidad nula (al menos en estos análisis), se analiza la varianza de la serie para saber si es estable o no, en caso de no serlo se realiza alguna transformación que permita tener varianza constante.

Para estabilizar la varianza aplicamos el método sugerido por Guerrero, formando bloques de 12 datos, en donde se calcula el promedio y desviación estándar para cada uno. El propósito es hallar un valor óptimo  $\lambda$  (una lambda que se explica en la sección 2.12.) que minimice el coeficiente de variación calculado con los estadísticos mencionados anteriormente para cada bloque. El valor óptimo de lambda calculado con el método de Guerrero resultó ser 0.59009432<sup>(4)</sup>.

<sup>4</sup> Lo que indica que la posible transformación de la variable original implica una función de tipo raíz cuadrada.

Por otra parte se calculó otro valor de lambda por medio del método de regresiones, es decir, se encontró la relación entre los logaritmos de las desviaciones estándar y los logaritmos de los promedios encontrados para los mismos bloques de 12 datos formados con el método anterior. La lambda encontrada con este método fue de 0.609564851, la cual es muy parecida a la encontrada con el método de Guerrero. Finalmente se optó por tomar a lambda la del primer método y la gráfica de la serie transformada en la siguiente figura:

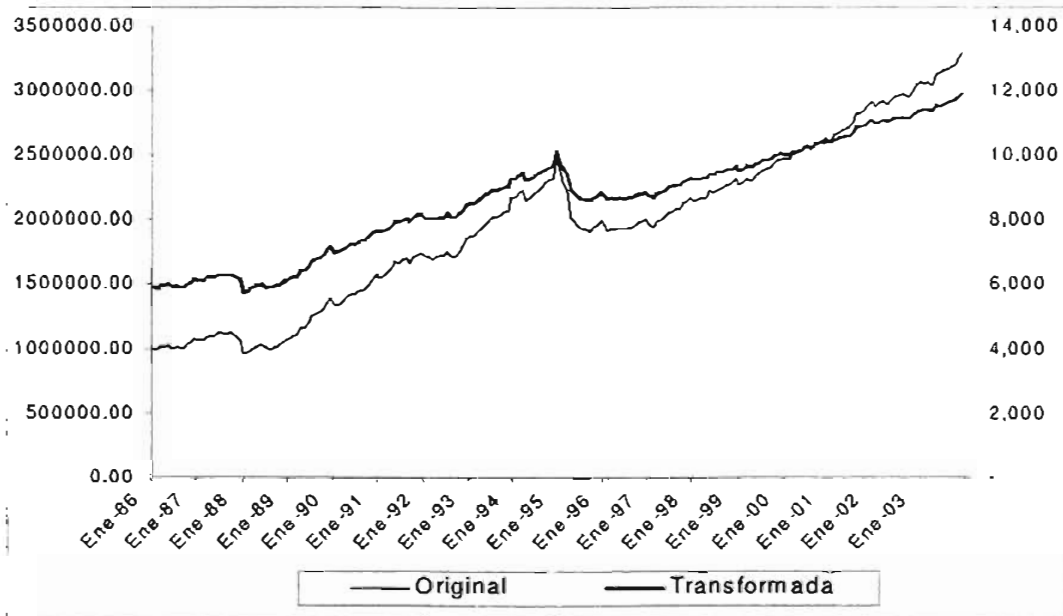


Fig. 3.20. Comparación de la serie original y la transformada por el método de Guerrero.

Al observar la gráfica, puede observarse como no existe un cambio sustancial en el comportamiento de los datos, lo cual hace pensar que no es necesaria una transformación de la varianza. Lo anterior, puede corroborarse con la fig. 3.21. que muestra la serie diferenciada y en donde puede notarse que conforme avanza el tiempo no existe volatilidad alguna en su comportamiento.

Quizá el único cambio significativo sea el que se presenta en diciembre de 1994 y en abril de 1995, pero esto se debe a un cambio de nivel que ocurrió en la serie original, mismo que se estudiará con un análisis de intervención.

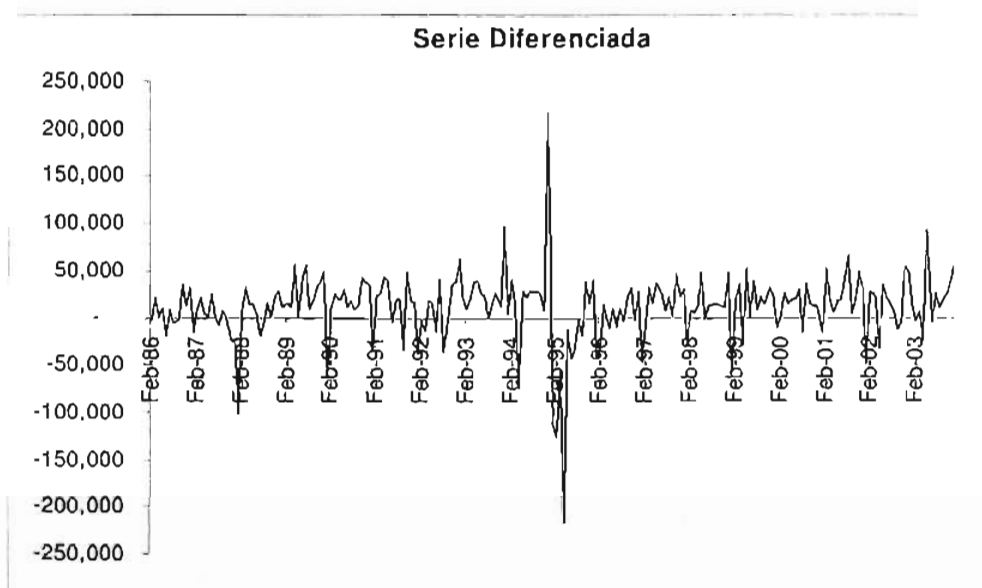


Fig. 3.21. Gráfica de la serie original diferenciada.

Esta intervención podría hacerse tomando como modelo inicial, la predicción que resulte de ajustar una recta a la serie, mediante la técnica de regresión. Esta forma de hacerlo es válida pero existe el problema de perder ajuste en los datos en el período de cambio de nivel, por lo que se realizará la intervención con la serie diferenciada, para sólo tomar en cuenta pulsos y no escalones.

Se sabe que al intervenir la serie, lo que se considera para establecer un criterio de buena o mala intervención son los residuos que ésta genera, ya que es posible ajustar a la serie del agregado M4 mediante:

$$M4 \text{ Real} = \text{Modelo inicial} + \text{Intervención} + \text{Residuales}$$

La intervención a realizar será la misma que se utilizó con la serie del IVA, en donde básicamente el propósito es modelar la intervención como una serie con un escalón ascendente a partir de diciembre de 1994, el cual toma su nivel después de abril de 1995. Los valores de las 3 constantes fueron los siguientes:

$w_0 =$	399762.691
$w_1 =$	425393.665
$\text{delta} =$	1.03222913

Por otro lado, como modelo inicial se tomó la media de los 6 meses siguientes a abril de 1995 de la serie original, ya que es en esos meses en donde la serie permanece casi constante y es posible modelar el escalón que se forma. La serie desintervenida, así como la misma intervención y la serie original, se presentan en la fig. 3.22., en donde es posible observar como el escalón desaparece y la serie parece tomar un ascenso constante.

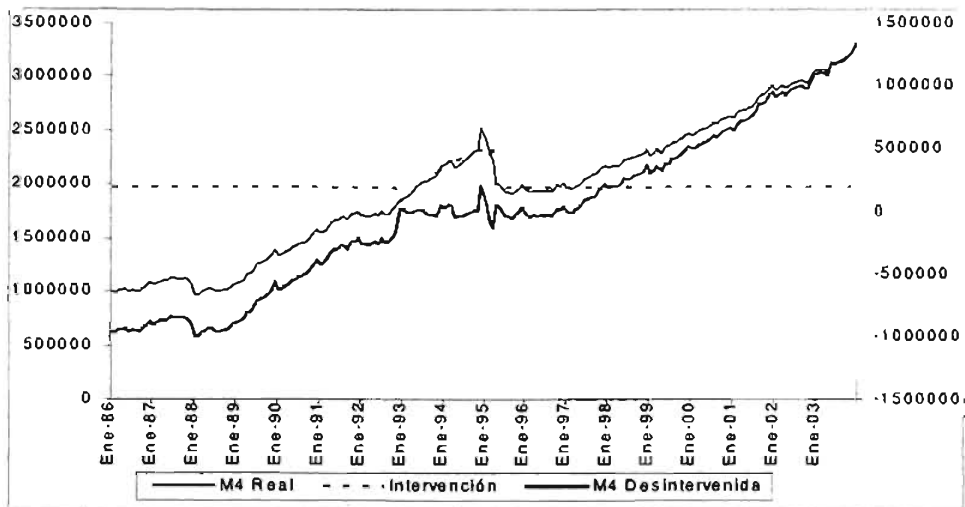


Fig. 3.22. Gráfica comparativa de la serie original y su intervención.

Continuando con la modelación de la serie del agregado M4 Real, se obtiene la función de autocorrelación simple (*fas*) y la función de autocorrelación parcial (*fap*) de la serie desintervenida en M4\_SI, lo que se presenta en la fig. 3.23.



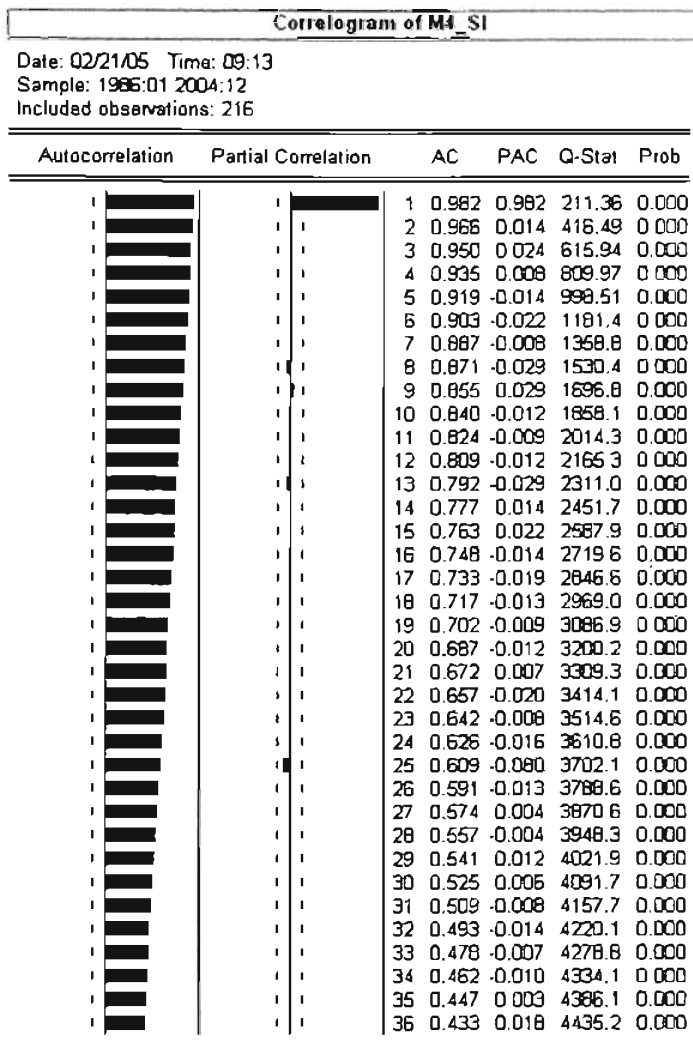


Fig. 3.23. Gráfica de la FAS y FAP de la serie M4 desintervenida

En el gráfico anterior es fácil reconocer que la serie tiene tendencia, ya que la *fas* decrece linealmente, además de que el primer valor es mayor a 0.9, lo que indica que el proceso no es estacionario, es decir, similar a una caminata aleatoria. Se aplica la primera diferencia y la serie D1M4SI se muestra en la fig. 3.24.

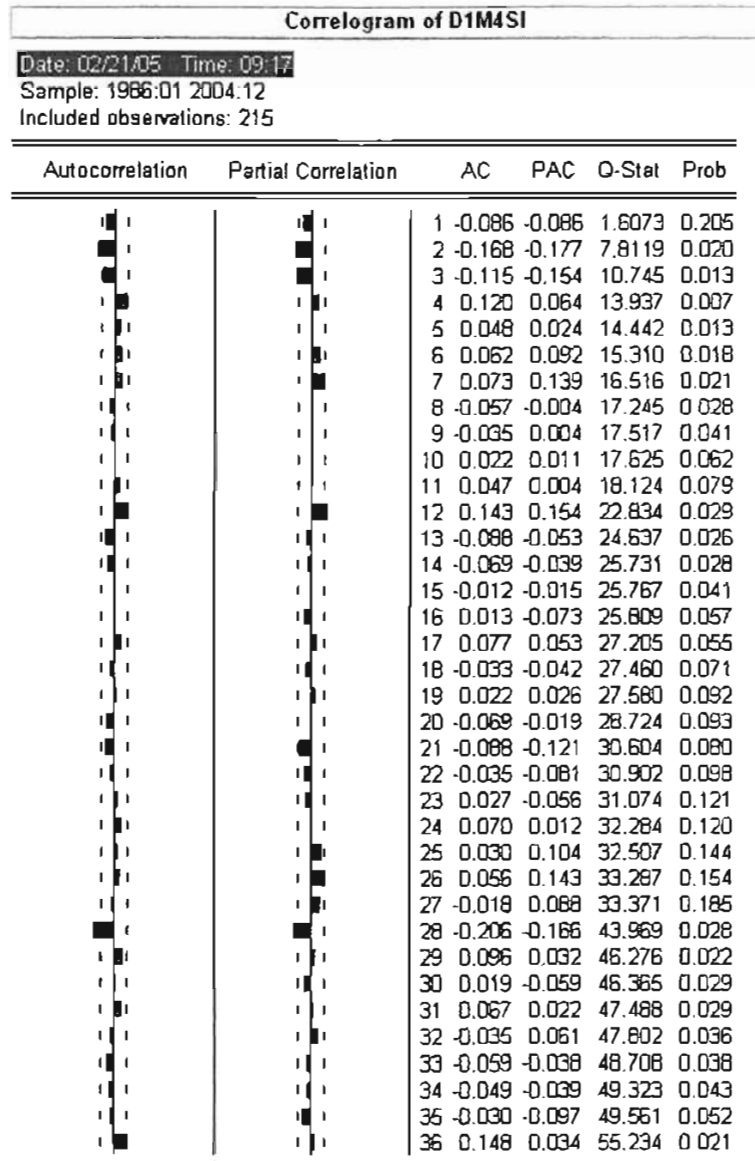


Fig. 3.24. Gráfica de la FAS y FAP de la serie M4 desintervenida diferenciada

En la figura anterior se observa que para la parte estacional, los AC, que en este trabajo es la *fas*, decrecen lentamente, mientras que los PAC, correspondientes a las *fap*, lo hacen súbitamente, lo que estaría indicando un posible modelo AR estacional de período 12, es decir un SAR(12). Por otro lado, en la parte regular se observa una caída lenta a partir del segundo rezago, pues tanto en los AC y PAC el primer rezago es significativamente igual a cero y no sale de las bandas de confianza.

Se debe ser cuidadoso al analizar la parte regular, ya que el decaimiento en los AC es casi súbito, mientras que en los PAC es lento, lo que sugeriría utilizar un modelo de medias móviles, pero al ser distinto de cero hasta el segundo rezago, se toma un MA(2) para modelar la parte regular. Los resultados de este modelo se pueden ver en la Tabla 3.6.

Dependent Variable: D(M4\_SI,1,0)  
 Method: Least Squares  
 Date: 02/21/05 Time: 00:58  
 Sample(adjusted): 1987:02 2003:12  
 Included observations: 203 after adjusting endpoints  
 Convergence achieved after 10 iterations  
 Backcast: 1986:12 1987:01

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11109.49	2605.371	4.264072	0.0000
AR(12)	0.149428	0.071164	2.099784	0.0370
MA(2)	-0.138648	0.070391	-1.969684	0.0503
R-squared	0.045421			
Adjusted R-squared	0.035875			
Akaike info criterion	23.86666			
Schwarz criterion	23.91562			
Durbin-Watson stat	2.226185			
Inverted AR Roots	.85	.74+.43i	.74 -.43i	.43 -.74i
	.43+.74i	.00+.85i	-.00 -.85i	-.43 -.74i
	-.43+.74i	-.74+.43i	-.74 -.43i	-.85
Inverted MA Roots	.37	-.37		

Tabla 3.6. Resultados del modelo ajustado a la serie desintervenida diferenciada.

Por la salida anterior se tiene que los parámetros son significativos al menos a un 94.7% de confianza. Además por el estadístico Durbin-Watson se observa que los residuales no están correlacionados en un rezago. Las raíces invertidas de los parámetros autorregresivos indican que no hay riesgo de seguir teniendo raíces unitarias. Cabe mencionar que este modelo fue el que tuvo mejor AIC, pues fue el menor de entre algunos otros que se probaron, pero que no se presentan, para no extender más este trabajo.

Al tener ya un modelo que explica el comportamiento del agregado M4, es posible pronosticar al menos para 2004 los valores de esta serie, pero su precisión dependerá del comportamiento de sus residuales, pues éstos deben cumplir ser un proceso de ruido blanco, es decir, que tengan media cero y varianza constante, para afirmar que el modelo es adecuado.

En la fig. 3.25 se tiene el correlograma de esta serie diferenciada y con la prueba Ljung-Box, para probar la hipótesis de que los residuos son ruido blanco, se observa que salvo en el rezago 4, cuya probabilidad de no ser ruido blanco es ligeramente mayor al 5%, se puede afirmar que los residuales están todos dentro de la banda de confianza. Por tanto la serie de residuales es ruido blanco.

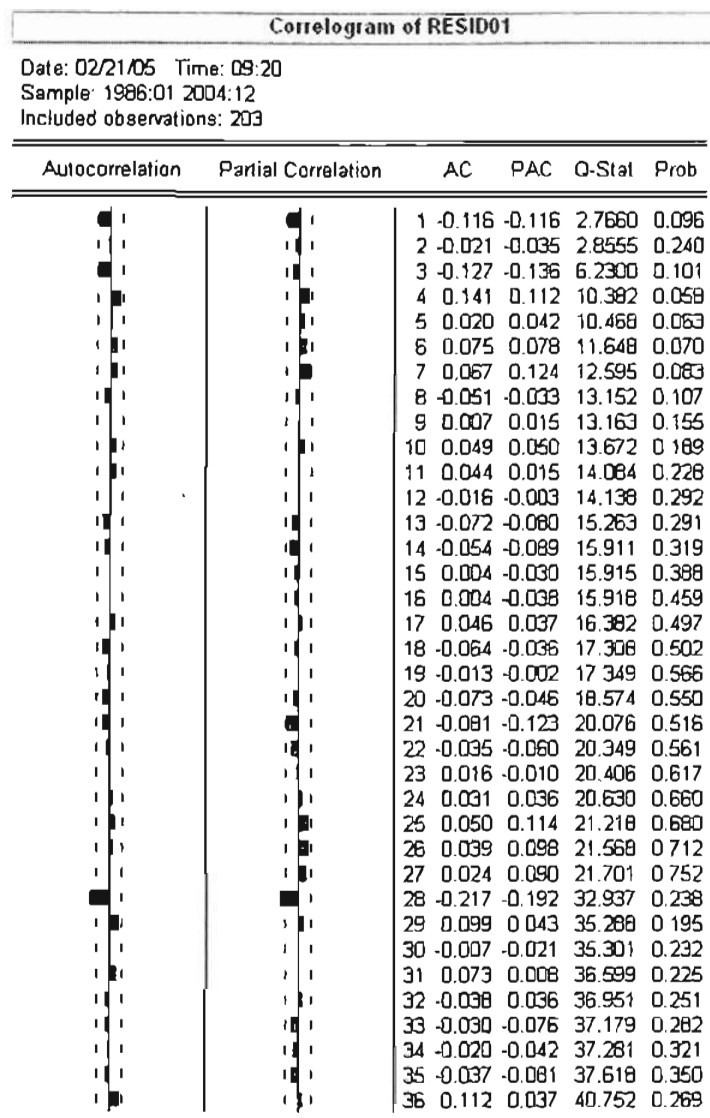


Fig.3.25 Correlograma de los residuales.

Al haber comprobado la aleatoriedad de los residuales, se puede pronosticar para el 2004 la serie del agregado monetario M4. Para esto se deben hacer operaciones inversas al proceso que generó la serie de residuales de la intervención y esto se hace sabiendo que,

$$\text{Residuales} = \text{M4 Real} - \text{Modelo inicial} - \text{Intervención}$$

Por tal motivo, se despeja la serie del M4 Real y se obtienen los valores para el 2004. No es necesario aplicar alguna corrección, como la sugerida por Guerrero, porque no se tomó en cuenta la transformación generada por el valor lambda, ya que se tenía varianza constante desde la serie original, afortunadamente.

Una vez que se obtuvo el pronóstico del M4 Real, se aplica el efecto de la inflación de manera inversa, obteniendo así el pronóstico del M4, el cual se compara con el publicado en el BANXICO. Lo anterior se presenta a continuación en la Tabla 3.7.

Fecha	Modelo + Intervención	Residuales Intervención	M4 Real Pronóstico	M4 Pronóstico	M4 Observado	INPC	Residuos
Ene-04	1968263.50	1340044.18	3,308,307.68	3,561,757.13	3,546,040.30	107.661	15717
Feb-04	1968263.50	1343402.20	3,311,665.70	3,586,699.54	3,552,560.63	108.305	34139
Mar-04	1968263.50	1353852.65	3,322,116.15	3,610,210.07	3,643,001.12	108.672	-32791
Abr-04	1968263.50	1359645.11	3,327,908.60	3,621,962.61	3,629,314.00	108.836	-7351
May-04	1968263.50	1383019.22	3,351,282.72	3,638,253.06	3,628,653.09	108.563	9600
Jun-04	1968263.50	1391564.35	3,359,827.85	3,653,376.01	3,698,941.45	108.737	-45565
Jul-04	1968263.50	1404935.53	3,373,199.03	3,677,529.04	3,707,319.70	109.022	-29791
Ago-04	1968263.50	1415802.71	3,384,066.21	3,712,151.43	3,708,607.37	109.695	3544
Sep-04	1968263.50	1427967.07	3,396,230.57	3,756,298.94	3,778,436.84	110.602	-22138
Oct-04	1968263.50	1441070.25	3,409,333.75	3,796,906.81	3,851,575.29	111.368	-54668
Nov-04	1968263.50	1456328.91	3,424,592.41	3,846,433.70	3,888,885.81	112.318	-42452
Dic-04	1968263.50	1473726.33	3,441,989.83	3,873,959.56	4,007,907.78	112.550	-133948

Tabla 3.7. Tabla de resultados del pronóstico

Observando la tabla resumen anterior y en mi opinión se tiene que el pronóstico en general fue bastante acertado, ya que tiene un margen de error de 3.4% y es debido a que el horizonte de pronóstico es el más alejado del origen, para el caso de Dic04. De hecho la gráfica que se presenta a continuación en la Fig. 3.26 es solamente para el 2004, pues en las series completas, no se distinguía esta diferencia.

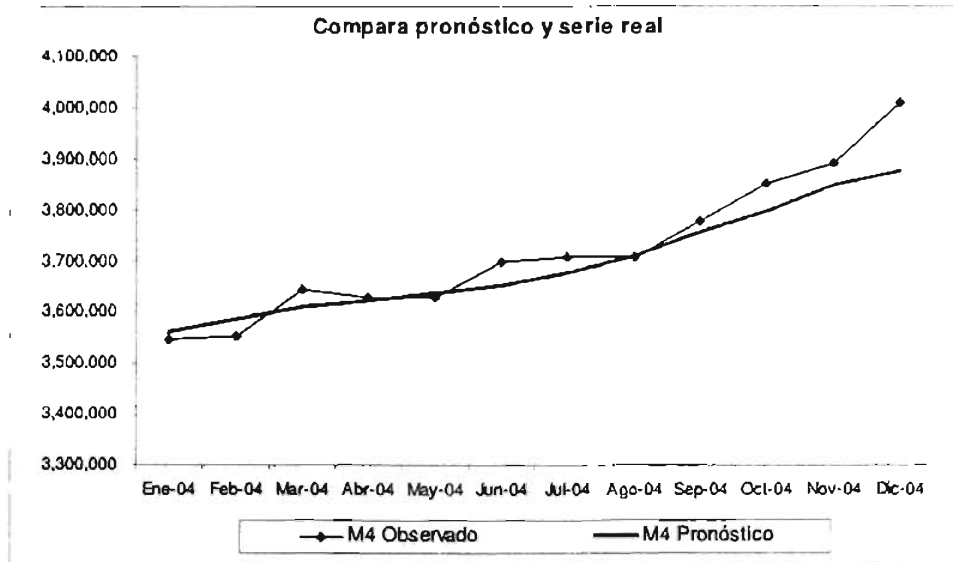


Fig. 3.26. Comparación de los pronósticos y de la serie observada.

Por último se especifica el modelo *ARIMA* empleado para el análisis, el cual se ha validado mediante el comportamiento de los residuales y los errores del pronóstico. El modelo *ARIMA* empleado fue el  $(0, 1, 2) \times (1, 0, 0)_{12}$ , el cual puede representarse como sigue:

$$(1 - \phi_{12}B^{12})(1 - B)Z_t = (1 - \theta_2B^2)\varepsilon_t$$

$$\Rightarrow Z_t = Z_{t-1} + \phi_{12}(Z_{t-12} - Z_{t-13}) + \varepsilon_t - \theta_2\varepsilon_{t-2}$$

$$\therefore Z_t = Z_{t-1} + (0.149428)(Z_{t-12} - Z_{t-13}) + \varepsilon_t - (-0.138648)\varepsilon_{t-2}$$

Con la aplicación del modelo *ARIMA* Multiplicativo a estas dos series se puede constatar la potencia y eficiencia de éstos, utilizando únicamente la historia de las variables de interés. El pronóstico será más preciso dependiendo el horizonte de predicción y de la estacionalidad de la serie.

## CONCLUSIONES.

El análisis de series de tiempo ha sido una herramienta poderosa en diversos procesos de pronóstico en los que el hombre ha estado involucrado. Al ser una serie de tiempo una sucesión de observaciones en distintos momentos de su transcurso, es posible estudiar los diversos componentes que la forman y con esto, concretar el estudio de la variable de interés a través del tiempo, que para su tratamiento se toma en la realidad como una variable discreta con períodos equidistantes, tal como se ha visto en esta tesis.

La suposición básica que está detrás y soporta a este análisis, es que los factores que ocasionan patrones de actividad en el pasado y en el presente, continuarán haciéndolo de la misma forma en el futuro. Por tanto, es preciso identificar y aislar tales factores, lo cual fue atendido en el primer capítulo de este trabajo, cuyo objetivo fue siempre el análisis de variables a base de su propia historia y no su descripción por el comportamiento de otras, para así poder hacer pronósticos más certeros.

Hoy en día existen métodos más efectivos que los descritos en el primer capítulo, pero como en toda situación que se analiza, es mejor conocer las limitaciones de otros métodos para abordar los problemas con métodos más actualizados, los cuales incluyen los desarrollados por Box y Jenkins, mismos que fueron descritos en el capítulo dos.

Antes del estudio de esos modelos y después del análisis de las componentes por separado, fue interesante analizar a las series desde el punto de vista gráfico, pues se pudo observar, por ejemplo, para la serie del IVA, que a medida que avanzaba el tiempo, el nivel de la serie iba aumentando considerablemente, lo cual indicaba una desestabilización de la varianza que se corrigió con una de las transformaciones explicadas en el segundo capítulo. Para la serie del M4 ocurrió que no fue necesario estabilizar la varianza, pues la serie no mostraba tal efecto.

Por otro lado, se observó que las series tenían muy acentuado un efecto de intervención, que afectaba el comportamiento de la misma. En la serie del IVA se observó a finales de 1991 y para la serie del M4 entre 1995 y 1996. Estas intervenciones fueron realizadas por el gobierno mexicano como medidas preventivas en la economía de esos tiempos. Después de analizar estas intervenciones, se logró removerlas y así trabajar con una serie que mantuviera el comportamiento que la serie traía en el pasado.

Y es que sin estos análisis no es posible afrontar un pronóstico sin errores que

seguramente saldrán a la luz de los resultados, ya que la serie por sí misma trae consigo errores y modificaciones que afectan su comportamiento y que los modelos propuestos no son capaces de describir. Después de estos análisis es más fácil analizar la serie con el propósito de pronosticar en un horizonte adecuado.

Un concepto básico en el análisis de la serie es el de estacionariedad, ya que éste permite tener a la variable de interés sin las componentes descritas anteriormente y en muchos casos el factor de estacionalidad hace que la serie carezca de ese comportamiento, pues hace que varíe con una aputa cíclica. Por ejemplo, en el estudio de la serie del IVA se observaba que la serie aparte de tener tendencia incluía una estacionalidad trimestral, las cuales fueron removidas con diferencias regulares y estacionales. Este método de diferencias es el más utilizado para remover este tipo de efectos en las variables, pues elimina patrones originados por el pasado de la serie. Por otro lado, la serie del M4 incluía únicamente tendencia, la cual fue removida con una diferencia regular. De esta forma, ambas series se convirtieron en estacionarias, y si no lo fueron del todo, al menos su estudio se redujo a la explicación de ciertos rezagos.

El análisis anterior pudo realizarse por medio de un correlograma, el cual plasma gráficamente las funciones de autocorrelación simple y parcial, y mediante bandas de confianza determina si el proceso es estacionario o no. Además es fácil identificar el orden que los modelos deben tener para explicar a la serie y así poder pronosticar. En la serie del IVA se observaba un decaimiento lineal en la parte regular y estacional (entendiendo como regular el paso de un período a otro, y estacional como el paso de una estación a otra), lo cual forzaba a aplicar diferencias. En la serie M4, solamente había un decaimiento lineal en la parte regular, pues como anteriormente se dijo, no existió algún patrón estacional que pudiera estudiarse.

Una vez identificadas las diferencias a aplicarse en ambas variables, fue fácil la identificación de los modelos (al menos fue más fácil que con la serie original). Por lo que a la serie del IVA respecta, el correlograma y la forma de decaer de los coeficientes, proponían el uso de parámetros autorregresivos y de media móvil, regular y estacionalmente, ya que los coeficientes fuera de las bandas de confianza así lo proponían; aunque los supuestos del modelo hicieron que al final únicamente se incluyeran a los de media móvil. Se removió a la constante, pues se estaba trabajando con una serie diferenciada. El orden de los parámetros lo proporcionó el número de coeficientes fuera de las bandas de confianza.

Respecto al M4, el correlograma mostró un decaimiento exponencial en la función de autocorrelación parcial y un tanto súbito en la simple, lo que sugirió aplicar parámetros de media móvil, siendo el número de rezagos significativamente distintos de cero, los que daban el orden este parámetro. Lo anterior fue para la parte regular, mientras que para la parte estacional ocurrió lo contrario, es decir, la función de autocorrelación simple, decayó de forma exponencial y la parcial de manera súbita, lo que ayudó a elegir parámetros



autorregresivos.

Una vez identificados los órdenes de los parámetros, se utilizó un paquete estadístico y que, por mínimos cuadrados, estimó el valor de estos coeficientes, lo que a su vez cumplieron los supuestos establecidos para ellos. Con estos coeficientes se realizó la estimación de la serie para períodos comparables a la realidad. Para la serie del IVA se pronosticó de enero a junio de 2004 y para el M4 de enero a diciembre de 2004. La diferencia estribó en el uso de una estacionalidad menor para la primer serie que para la segunda, lo que hace que se acorte el horizonte de pronóstico pues el modelo pierde fuerza conforme avanza el tiempo.

El siguiente paso fue analizar los residuales generados por los datos observados y las predicciones del modelo. Los residuos para el IVA cumplieron todos los supuestos, lo cual hizo suponer que el modelo era el adecuado, lo mismo ocurrió con la serie del M4. Ambos modelos fueron los apropiados incluso con el criterio de selección de Akaike, pues arrojaron menores coeficientes que con otros modelos analizados.

Así, después de comprobar los supuestos de los parámetros y en general de los modelos propuestos, se sigue con el proceso de pronóstico, ya que se deben realizar las operaciones inversas al proceso que llevó a la serie con la que se extrajo el modelo. Para el IVA, los pasos fueron un poco complicados, ya que el modelo fue más robusto que con la serie del M4, además de que para la primera serie fue necesaria una transformación y para la segunda no ocurrió esto.

Es decir, para la serie del IVA primero se le sumó la intervención encontrada para tener a la serie transformada intervenida. Después se aplicó la operación inversa a la transformación, afectada por un factor de corrección, la cual llevó a tener a la serie neta, misma que no es comparable aun con la serie observada, ya que debe tener consigo el efecto de la inflación, mismo que se aplicó y de esta forma ambas series pudieron ser diferenciadas para obtener un error en el pronóstico, el cual fue aceptable. Esto mismo se utilizó para la serie del M4, con la salvedad de que no se aplicó la operación inversa de la transformación.

Al final se obtuvieron pronósticos bastante acertados para ambas variables, tomando en cuenta que la historia de las series era demasiado amplias, como para guardar y utilizar el comportamiento que ambas traían en su pronóstico.

A manera de resumen, se establecen etapas que deben seguirse si se quiere realizar un pronóstico con modelos ARIMA. La primera etapa consiste en tener a una serie con varianza constante y libre de intervenciones, para después decidir qué transformaciones aplicar para convertir a la serie observada en estacionaria y así determinar un modelo ARIMA para ésta, es decir, los órdenes  $p$  y  $q$  de su estructura autorregresiva y de media

móvil y, si el proceso es estacional, los órdenes  $P$  y  $Q$  de la estructura estacional.

Una vez seleccionado provisionalmente un modelo para la serie estacionaria, comienza la segunda etapa, que es la de estimación, donde los parámetros  $AR$  y  $MA$  se estiman por máxima verosimilitud o por mínimos cuadrados, además de obtener sus errores estándares y residuos. Aquí es en donde se revisa que los parámetros cumplan sus supuestos, tales como significancia, invertibilidad, etc.

La tercera etapa es la diagnosis, donde se comprueba que los residuos no tienen una estructura de dependencia y representan un proceso de ruido blanco. Si se acepta que los residuos no contienen información, el proceso termina, en caso contrario, se modifica al modelo incorporando tal información y se repiten las etapas anteriores.

Como la etapa cuarta y final, se tiene el identificar el mejor modelo de los estimados, mediante criterios de selección que están basados en la verosimilitud de los mismos.

Finalmente, es necesario y útil nuevamente recalcar que la metodología utilizada para generar pronósticos en series, principalmente económicas, está basada en fundamentos matemáticos que dan una herramienta poderosa al estadístico que desea realizar pronóstico de variables empleando únicamente la historia de las mismas, pues es común que no se conozcan los efectos que otras variables puedan tener para la variable de interés, lo que hace que no se tomen en cuenta muchos de los supuestos que otros modelos aplicables en estos casos deben cumplir. El pronóstico de series con modelos  $ARIMA$  Multiplicativos tiene también deficiencias y desventajas, aunque como alguna vez Box afirmó: "Todos los modelos son incorrectos, pero alguno de ellos son útiles".

# APÉNDICES

## APÉNDICE A.

*Prop.* Un proceso que sigue una tendencia polinomial de orden  $h$ , es un proceso integrado de orden  $h$ .

*Demostración.*

Es decir, se tiene que demostrar que  $\nabla^h Z_t$  es un proceso estacionario, o sea, que  $\nabla^h u_t$  cumple con las características de estacionariedad, pues  $\nabla^h Z_t$  está compuesto básicamente por la diferencia de procesos estacionarios  $\{\nabla^h u_t\}$ .

Aplicando sucesivamente el operador diferencia, se tiene:

$$\nabla u_t = u_t - u_{t-1}$$

$$\begin{aligned} \nabla^2 u_t &= \nabla \nabla u_t \\ &= \nabla(u_t - u_{t-1}) \\ &= (u_t - u_{t-1}) - (u_{t-1} - u_{t-2}) \\ &= u_t - 2u_{t-1} + u_{t-2} \\ &= \sum_{i=0}^2 \binom{2}{i} (-1)^i u_{t-i} \end{aligned}$$

$$\begin{aligned} \nabla^3 u_t &= \nabla \nabla^2 u_t \\ &= \nabla(u_t - 2u_{t-1} + u_{t-2}) \\ &= (u_t - 2u_{t-1} + u_{t-2}) - (u_{t-1} - 2u_{t-2} + u_{t-3}) \\ &= u_t - 3u_{t-1} + 3u_{t-2} - u_{t-3} \\ &= \sum_{i=0}^3 \binom{3}{i} (-1)^i u_{t-i} \end{aligned}$$

$$\begin{aligned} \nabla^h u_t &= \nabla \nabla^{h-1} u_t \\ &= \nabla \left( \sum_{i=0}^{h-1} \binom{h-1}{i} (-1)^i u_{t-i} \right) \\ &= \sum_{i=0}^{h-1} \binom{h-1}{i} (-1)^i u_{t-i} - \sum_{i=0}^{h-1} \binom{h-1}{i} (-1)^i u_{t-i-1} \\ &= \sum_{i=0}^{h-1} \binom{h-1}{i} (-1)^i u_{t-i} + \sum_{i=1}^h \binom{h-1}{i-1} (-1)^i u_{t-i} \\ &= \left\{ \sum_{i=0}^{h-1} \left[ \binom{h-1}{i} + \binom{h-1}{i-1} \right] (-1)^i u_{t-i} \right\} + u_t + (-1)^h u_{t-h} \end{aligned}$$

Para poder simplificar esta expresión, se utiliza el siguiente teorema:

*Teorema 1. (Teorema de Pascal). Sean  $n, m$  números enteros y no negativos, entonces:*

$$\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$$

Además, sabiendo que  $\binom{h}{0} = 1$  y  $\binom{h}{h} = 1$ , se puede reducir la  $h$ -ésima diferencia de  $u_t$  como sigue:

$$\nabla^h u_t = \sum_{i=0}^h \binom{h}{i} (-1)^i u_{t-i}$$

Nótese que del siguiente teorema,

*Teorema 2. (Teorema del binomio). Sean  $a$  y  $b$  números reales y  $n$  entero, entonces:*

$$(a - b)^n = \sum_{i=0}^n \binom{n}{i} (-1)^i a^{n-i} b^i$$

se puede concluir que  $\nabla^h$  es en realidad un binomio elevado a la  $h$ -ésima potencia. Aplicando el operador esperanza, se tiene:

$$\begin{aligned} E(\nabla^h u_t) &= \sum_{i=0}^h \binom{h}{i} (-1)^i E(u_{t-i}) \\ &= \sum_{i=0}^h \binom{h}{i} (-1)^i \mu \\ &= \mu \sum_{i=0}^h \binom{h}{i} (-1)^i \\ &= \mu(1 - 1)^h = 0 \end{aligned}$$

y de igual manera, se calcula la autocovarianza del proceso,

$$\begin{aligned} Cov(\nabla^h u_t, \nabla^h u_{t+k}) &= E(\nabla^h u_t, \nabla^h u_{t+k}) \\ &= E\left[ \left( \sum_{i=0}^h \binom{h}{i} (-1)^i u_{t-i} \right) \left( \sum_{j=0}^h \binom{h}{j} (-1)^j u_{t+k-j} \right) \right] \\ &= \sum_{i=0}^h \sum_{j=0}^h \binom{h}{i} \binom{h}{j} (-1)^{i+j} E(u_{t-i} u_{t+k-j}) \\ &= \sum_{i=0}^h \sum_{j=0}^h \binom{h}{i} \binom{h}{j} (-1)^{i+j} \gamma_{i-j+k} \end{aligned}$$

la cual depende solamente del rezago y no del tiempo; en particular cuando  $k = 0$ , se

obtiene  $Var(\nabla^h u_t) = \sum_{i=0}^h \sum_{j=0}^h \binom{h}{i} \binom{h}{j} (-1)^{i+j} \gamma_{i-j}$ , que es constante, pues tampoco depende del tiempo  $t$ , ni de  $k$ .

Con esto queda demostrado que la  $h$ -ésima diferencia de un proceso, que sigue una tendencia polinomial de orden  $h$ , es estacionaria, además de que la estacionariedad de un proceso se conserva a pesar del uso del operador  $\nabla$ .

Puede seguir aplicándose el operador diferencia  $\nabla$  a un proceso que ya estacionario, y éste seguirá siéndolo, sin embargo, el hecho de sobrediferenciar la serie, ocasiona la pérdida de observaciones innecesariamente, pues cada vez que se aplica el operador  $\nabla$  se pierde una observación.

## APÉNDICE B.

Si la raíz  $G_i^{-1}$ ,  $i = 1, \dots, k$ , es de multiplicidad  $m_i$ , la solución está dada por,

$$x_t = \sum_{i=1}^N G_i^t \sum_{j=1}^{m_i-1} t^j A_{ij}$$

en donde  $N$  el número de raíces diferente de la ecuación auxiliar. Por tanto, se tendría que:  $\sum_{i=1}^N m_i = k$  y  $\prod_{i=1}^N (1 - G_i B)^{m_i} = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_k B^k$

Para demostrar lo anterior serán de utilidad las demostraciones de los siguientes lemas:

**LEMA 1.** Sea  $Z_t = At^j$ ,  $m$  un número no negativo,  $A$  una constante y  $j$  un entero no negativo menor que  $m$ . Entonces  $(1 - B)^m Z_t = 0$ .

*Demostración (Por inducción sobre  $m$ ).*

Para  $m = 1$ , se tiene que como  $j$  es un entero no negativo menor que  $m$ , entonces  $j$  sólo puede ser cero, por tanto  $Z_t = At^0 = A$ . Sustituyendo  $Z_t$ , se tiene que  $(1 - B)Z_t = (1 - B)A = A - A = 0$ .

Supóngase que es válido para  $m = k - 1$ , es decir,  $(1 - B)^{k-1} Z_t = (1 - B)^{k-1} At^j = 0$ , donde  $j < k - 1$ . Se demostrará que también es válido para  $m = k$ , con  $Z_t = At^j$  y  $j < k$ .

$$\begin{aligned} (1 - B)^k Z_t &= (1 - B)^{k-1} (1 - B) At^j \\ &= (1 - B)^{k-1} [At^j - BA^j] \\ &= (1 - B)^{k-1} [At^j - A(t-1)^j] \\ &= (1 - B)^{k-1} (-A)[(t-1)^j - t^j] \end{aligned}$$

desarrollando  $(t-1)^j$  (con una de las múltiples versiones del teorema del binomio), el término anterior puede expresarse como  $(1 - B)^{k-1} (-A) \left\{ \left[ \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} t^i \right] - t^j \right\}$ , y cuando  $i = j$ , se tiene que el último sumando se elimina con  $t^j$ , por lo que la expresión anterior puede reducirse a:

$$(1 - B)^{k-1} \left\{ (-A) \sum_{i=0}^{j-1} \binom{j}{i} (-1)^{j-i} t^i \right\} = - \left\{ \sum_{i=0}^{j-1} (1 - B)^{k-1} A \binom{j}{i} (-1)^{j-i} t^i \right\}$$

En esta última expresión se observa que en cada uno de los sumandos puede aplicarse la hipótesis de inducción, ya que se cumple que  $i < j-1 < j < k-1$ , donde  $i = 0, 1, 2, \dots, j-1$ . Así para cada  $i$  fija en la expresión final, se tiene que  $A\binom{j}{i}(-1)^{j-i}$  es constante. Por lo cual  $(1-B)^{k-1}A\binom{j}{i}(-1)^{j-i}t^j = 0$ , para cada  $i = 0, 1, 2, \dots, j-1$  y entonces se cumple para  $m = k$ . Por lo tanto  $(1-B)^m Z_t = 0$ .

**LEMA 2.** Sea  $Z_t = G't^j$ ,  $m$  un entero no negativo,  $G \neq 0$  y  $j$  un entero no negativo menor que  $m$ . Entonces  $(1-GB)^m Z_t = 0$ .

*Demostración.*

Nótese que,

$$\begin{aligned} (1-GB)Z_t &= (1-GB)G't^j \\ &= G't^j - BG^{t+1}t^j \\ &= G't^j - G'(t-1)^j \\ &= G'[t^j - (t-1)^j] \\ &= G'(t^j - Bt^j) \\ &= G'(1-B)t^j \end{aligned}$$

Si se repite esta operación para  $m = 2$ .

$$\begin{aligned} (1-GB)^2 Z_t &= (1-GB)^2 G't^j \\ &= (1-GB)(1-GB)G't^j \\ &= (1-GB)G'(1-B)t^j \\ &= (1-B)[(1-GB)G't^j] \\ &= G'(1-B)^2 t^j \end{aligned}$$

y así sucesivamente hasta obtener la expresión para  $m$ .

$$(1-GB)^m Z_t = G'(1-B)^m t^j$$

Entonces, por el **LEMA 1**, si  $A = 1$  entonces  $(1-B)^m t^j = 0$ , por lo tanto

$$(1-GB)^m Z_t = 0$$

Ahora, con la ayuda de estos dos lemas, se va a comprobar que (2.9) es solución de la ecuación en diferencia de orden  $k$ . Sustituyendo en (2.6) se tiene que:

$$\begin{aligned}
& \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} - \phi_1 \sum_{i=1}^N G_i^{t-1} \sum_{j=0}^{m_i-1} (t-1)^j A_{ij} - \phi_2 \sum_{i=1}^N G_i^{t-2} \sum_{j=0}^{m_i-1} (t-2)^j A_{ij} - \dots - \\
& = \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} - \phi_1 B \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} - \phi_2 B^2 \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} - \dots - \phi_k B^k \sum_{i=1}^N \\
& = \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_k B^k) = \sum_{i=1}^N G_i^t \sum_{j=0}^{m_i-1} t^j A_{ij} (1 - G_i B)^{m_i} (1 \\
& = 0
\end{aligned}$$

pues para cada  $i$  y  $j$  fijas,  $A_{ij} t^j G_i^t (1 - G_i B)^{m_i} = 0$ , por el **LEMA 2**.



## APÉNDICE C.

Una intervención se entiende como la ocurrencia de un evento exógeno al comportamiento histórico de la variable en estudio y esto provoca que uno o más residuales sean mayores a lo esperado. Si dichos residuales pueden explicarse por la ocurrencia de esos eventos ajenos a la serie, convendría reducir la influencia de estos fenómenos con la técnica de intervención de la serie.

La anterior situación propone que la influencia de estos residuales sea analizada para entender:

1. Su contribución a la medición de la variable que se estudia,
2. O su influencia en el modelo propuesto y su consecuente afectación en la estimación de los parámetros propuestos.

La técnica de intervención fue diseñada por Box y Tiao en 1975, quienes distinguieron tres tipos de efectos anormales en las series:

a) Aquellos que se dejan ver como una extrema elevación o extrema caída momentánea del nivel, pero que dejan su estructura estocástica intacta. Por ejemplo, una observación atípica.

b) Los que ejercen una influencia sostenida sobre el comportamiento de la serie (no momentáneamente), pero que dejan su estructura estocástica intacta. Por ejemplo, un cambio de nivel.

c) Los efectos que independientemente de influir en la serie, sí alteran la parte estocástica de la serie. Por ejemplo, un cambio de varianza.

Bajo la metodología de Box y Tiao, el análisis de intervención puede verse como una extensión de análisis de los procesos ARIMA. El modelo de intervención para la serie transformada (o no transformada) que contiene los efectos exógenos es el siguiente:

$$T(Z_t) = I_{k,t} + N_t$$

donde  $N_t$  está modelado por:

$$\Phi(B)(1-B)^d N_t = \Theta(B)\varepsilon_t$$

que es un modelo *ARIMA* que lo convierte en un proceso estacionario e invertible y la función  $I_{k,t}$  es una función que permite representar los efectos de intervención ocurrida al momento  $k$ . Para el caso de la serie del IVA transformada, se utilizará la siguiente función de Intervención:

$$(1 - \delta B) \nabla I_{k,t} = (\omega_0 - \omega_1 B) P_{k,t}$$

donde:

$$I_{k,t} = \begin{cases} 0 & \text{si } t < k \\ \omega_0 & \text{si } t = k \\ \frac{[\omega_0 - \omega_1 - (\omega_0 \delta - \omega_1) \delta^{t-k}]}{(1-\delta)} & \text{si } t > k \end{cases}$$

y  $P_{k,t}$  se conoce como la función *pulso*, la cual se define como sigue:

$$P_{k,t} = \begin{cases} 1 & \text{si } t = k \\ 0 & \text{si } t \neq k \end{cases}$$

Para un mejor entendimiento de esta técnica, puede consultarse el libro (1) de la bibliografía a partir de la pag. 235.

## APÉNDICE D.

En la mayoría de las series analizadas, será necesario el pronóstico para la serie original  $Z_t$  y no para la serie transformada  $T(Z_t)$ , pues ésta última no tiene demasiada interpretación. Al tratar de realizar la operación inversa que convierta a  $T(Z_t)$  en  $Z_t$ , las propiedades óptimas del pronóstico  $\hat{T}(Z_t)(h)$ , siendo  $h$  el horizonte de pronóstico, no se conservan necesariamente. Si la transformación no fue lineal, que es lo más común en los casos reales, es posible realizar el cálculo de un factor  $C$  que permite corregir aproximadamente el sesgo que introduce la aplicación de la transformación inversa.

El factor  $C$  para los casos en que la transformación realizada fue del tipo Box-Cox es el siguiente (ver pag. 224 del libro (1) de la bibliografía),

$$C_{t,\lambda}(h) = \begin{cases} \left\{ \frac{1}{2} + \frac{\sqrt{[1-2\lambda(\lambda-1)[1+\lambda\hat{T}(Z_t)(h)]^{-2}\widehat{Var}[e_t(h)]}}{2} \right\}^{\frac{1}{\lambda}} & \text{si } \lambda \leq 1 \text{ y } \lambda \neq 0 \\ \exp\left\{\frac{\widehat{Var}[e_t(h)]}{2}\right\} & \text{si } \lambda = 0 \end{cases}$$

donde  $\lambda$  es el valor obtenido con el Método de Guerrero y  $\widehat{Var}[e_t(h)]$  es el pronóstico de la varianza de los residuales.

# BIBLIOGRAFÍA

- 1.- GUERRERO, Víctor M. *Análisis Estadístico de Series de Tiempo Económicas*. México: Colección CBI. 1ª Edición. 1991.
- 2.- HAMILTON, James D. *Time Series Analysis*. Ed. Princeton University Press. 2a Ed. 1994.
- 3.- CHATFIELD, Chris. *The Analysis of Times Series: An Introduction*. Ed. Chapman & Hall/CRC. 6a Edición. 2004.
- 4.- HANKE, John E. / REITSCH, Arthur G. *Pronósticos en los Negocios*. Ed. Prentice Hall. 5ª Edición.
- 5.- QUENOUILLE, M.H. *The Joint Distribution of Serial Correlations*. *Annals of Mathematical Statistics*. 1949.
- 6.- GUJARATI, Damodar N. *Econometría*. Ed. Mc Graw Hill. 2ª Edición. 1990.
- 7.- SHUMWAY, Robert. *Times Series Analysis and its Applications*. Ed. Springer. 1ª Edición. 2000.
- 8.- BROCKWELL, Peter J. / DAVIS, Richard A. *Time Series: Theory and Methods*. Ed. Springer. 3ª Edición. 1990.
- 9.- YAFFEE, Robert A. *An Introduction to Time Series Analysis and Forecasting: with Applications of SAS and SPSS*. Ed. AP. 1ª Edición. 2000.

- 10.- FULLER, Wayne A. *Introduction to Statistical Time Series*. Wiley Series in Probability and Statistics. 2a Edición. 1996.
- 11.- PULIDO, Antonio San Román. *Predicción Económica y Empresarial*. Ed. Pirámide. 3ª Edición. 1989.
- 12.- TAYLOR, S. J. *Modelling Financial Times Series*. Ed. Chichester U.K. 1ª Edición. 1986.
- 13.- BROCKWELL, Peter J./ DAVIS, Richard A. *Introduction to Time Series and Forecasting*. Ed. Hardcover. 1996.
- 14.- JOHNSTON, J. *Métodos de Econometría*. Ed. Vicens Vives. 2001.
- 15.- <http://www.metu.edu.tr/~eruygur/econ302/probset/tables.pdf>
- 16.- [www.banxico.org.mx](http://www.banxico.org.mx)
- 17.- [www.sat.gob.mx](http://www.sat.gob.mx)
- 18.- <http://www.oilnergy.com/1gnymex.htm>
- 19.- <http://www-personal.buseco.monash.edu.au/~hyndman/TSDL/>
- 20.- [halweb.uc3m.es/esp/docencia/serieslic/stcapx.pdf](http://halweb.uc3m.es/esp/docencia/serieslic/stcapx.pdf)
- 21.- <http://es.wikipedia.org/wiki/IVA>