



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

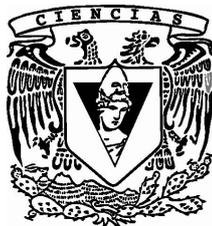
**LA DINÁMICA DE LA INFLACIÓN EN MÉXICO:
UNA APROXIMACIÓN ECONOMÉTRICA**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE:
ACTUARIO**

P R E S E N T A:

RICARDO CÉSAR RAMÍREZ GARCÍA



M. EN C. MICHAEL ROJAS ROMERO

2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

1.-Datos del alumno

Ramírez
García
Ricardo César
5556196490
Universidad Nacional Autónoma de México
Facultad de ciencias
Actuaría
098308482

2.-Datos del tutor

M en C
Michel
Rojas
Romero

3.-Datos del sinodal 1

Mat
Margarita Elvira
Chávez
Cano

4.-Datos del sinodal 2

Dr
Alberto
Contreras
Cristán

5.-Datos del sinodal 3

M en C
Inocencio Rafael
Madrid
Ríos

5.-Datos del sinodal 4

M en C
César Eduardo
Sousa
Mondragón

AGRADECIMIENTOS.

Quiero agradecer a Dios, por permitirme llegar a esta etapa de mi vida.

A mis abuelitos por ser un gran ejemplo y por apoyarme incondicionalmente.

A Mi mamá y mi hermana por estar siempre conmigo, y apoyarme en todo momento.

A la UNAM por permitirme terminar una carrera, y haber disfrutado de esas maravillosas instalaciones durante mi carrera.

A toda mi familia, que ha estado siempre conmigo.

A mis amigos por impulsarme y alentarme para salir adelante.

A mi tutor ; el Maestro Michel por dedicarme tiempo para salir adelante con mis tesis.

A mis sinodales; por sus valiosas observaciones y comentarios que aportaron a mi tesis.

ÍNDICE

INTRODUCCIÓN.....	2
CAPÍTULO I. MODELOS ECONOMETRICOS DE SERIES DE TIEMPO ESTACIONARIAS	3
1.1. DEFINICIÓN DE CONCEPTOS BÁSICOS.....	3
1.1.1. SERIE DE TIEMPO	3
1.1.2. PROCESOS ESTOCÁSTICOS.....	4
1.1.3. PROCESOS LINEALES GENERALES	5
1.1.3.1. Estacionariedad de un proceso lineal general	8
1.1.3.2. Invertibilidad de un proceso lineal general	9
1.2. MODELOS ARIMA	10
1.2.1. MODELOS AUTORREGRESIVOS DE ORDEN p , $AR(p)$	10
1.2.1.1. Estacionariedad e invertibilidad de un $AR(p)$	11
1.2.1.2. Función de autocorrelación de un $AR(p)$	13
1.2.1.3. Función de autocorrelación parcial de un $AR(p)$	15
1.2.2. PROCESOS DE PROMEDIOS MÓVILES DE ORDEN q , $MA(q)$	16
1.2.2.1. Estacionariedad e invertibilidad para un $MA(q)$	17
1.2.2.2. Función de autocorrelación para un proceso $MA(q)$	18
1.2.2.3. Función de autocorrelación parcial de un $MA(q)$	20
1.2.3. PROCESOS AUTORREGRESIVOS Y DE PROMEDIOS MÓVILES MIXTOS $ARMA(p, q)$	20
1.2.3.1. Estacionariedad e invertibilidad de un $ARMA(p, q)$	20
1.2.3.2. Función de autocorrelación de un $ARMA(p, q)$	21
1.2.3.3. Función de autocorrelación parcial de un $ARMA(p, q)$	22
1.2.4. MODELOS ARIMA	23
1.2.4.1. Procesos integrados de orden d	24
1.2.4.2. Prueba de Dickey-Fuller	26
1.2.4.3. Prueba de Dickey-Fuller Aumentada	27
1.2.4.4. Estimación de parámetros	27
Método de máxima verosimilitud	28
Estimación no lineal de parámetros	30
1.2.4.5. Función de pronósticos para modelos estacionarios.....	35
1.2.4.6. Función de pronósticos para modelos no estacionarios.....	40
1.3. METODOLOGÍA DE BOX-JENKINS	42
CAPÍTULO II. UN MODELO ECONOMETRICO DEL ÍNDICE DE INFLACIÓN SUBYACENTE DE MÉXICO (1990-2008).....	45
RESUMEN Y CONCLUSIONES	56
ANEXOS	57
REFERENCIAS BIBLIOGRÁFICAS	62

INTRODUCCIÓN

Objetivo de la tesis

El objetivo de esta tesis es estudiar la dinámica de la inflación en México durante el periodo de enero de 2000 a abril de 2009. Este objetivo se aborda mediante la especificación, estimación y diagnóstico de un modelo econométrico para la variable de serie de tiempo índice de inflación subyacente de México en el periodo 1990/01-2009/04 (enero de 1990 a abril de 2009). Este periodo corresponde al comportamiento relativamente más estable de la serie índice de la inflación subyacente y al cual se pudo ajustar un modelo ARIMA.

Cabe mencionar que se entiende dinámica en el sentido de los modelos ARIMA con ecuaciones con diferencias estocásticas con coeficientes constantes. Las trayectorias convergentes o no de los modelos indican la dinámica de la serie de tiempo.

Para definir el índice de inflación subyacente se dan las definiciones y conceptos siguientes.

De acuerdo a la definición que da el Banco de México, la inflación es la tasa de crecimiento promedio de un periodo a otro de los precios de **una canasta de bienes y servicios**¹.

La estabilidad de los precios es uno de los objetivos que persigue todo Banco Central a través de la **política monetaria**, refiriéndose dicho objetivo a la tasa de variación promedio de los precios o tasa de inflación, ya que sus repercusiones pueden afectar gravemente el bienestar de la sociedad. La preocupación por este fenómeno surge por las consecuencias que éste produce en la redistribución de la renta, la riqueza entre diversos grupos sociales e individuos, en la tasa de crecimiento de la economía, y en la estabilidad política y económica de un país.

La institución encargada de calcular y publicar la inflación en nuestro país es el Banco de México; sin embargo, este no realiza una labor de predicción de la inflación, simplemente se limita a estimarla con base en expectativas de los analistas.

Los instrumentos para medir y darle seguimiento a la inflación son los índices de precios. En México se tiene el Índice Nacional de Precios al Consumidor (INPC) y el Índice de la Inflación Subyacente, ambos elaborados y publicados por el Banco de México de manera quincenal y mensual.

¹ Para ver el significado de los conceptos que están en negritas de la parte del objetivo de tesis véase el anexo 2.

El Índice Nacional de Precios al Consumidor (INPC) es un indicador económico diseñado específicamente para medir el cambio promedio de los precios en el tiempo, mediante una canasta ponderada de bienes y servicios representativa del consumo de las familias urbanas de México.

Los 315 **conceptos genéricos** que integran la canasta de bienes y servicios del INPC se clasifican o agrupan en subconjuntos que responden a necesidades particulares de análisis, entre las clasificaciones más conocidas están la de por **objeto del gasto**, la que se refiere al sector de origen de los bienes y servicios y la de durabilidad de los bienes e inflación subyacente.

La inflación subyacente, en particular, es una clasificación útil para tomar decisiones de política económica y para la elaboración de predicciones sobre la inflación, debido a que los bienes y servicios que integran la canasta con la que se mide dicha inflación no están sujetos a decisiones de carácter administrativo, marcada estacionalidad o alta volatilidad.

El Índice de Inflación Subyacente busca capturar la tendencia del movimiento en los precios. Debido a ello, se excluyen del INPC los genéricos que presentan alta volatilidad en sus precios: productos agropecuarios, bienes y servicios administrados y concertados, educación privada.

Organización de la tesis

Para estudiar la dinámica de la Inflación en México, esta tesis se divide en dos capítulos, conclusiones y tres anexos. En el primer capítulo se presentan los conceptos básicos de la teoría ARIMA de series de tiempo estacionarias. En el capítulo dos presento un modelo econométrico ARIMA para la serie índice de la inflación subyacente. La tesis termina con un resumen y algunas conclusiones.

El software que se utilizó en esta tesis es el software econométrico EViews, versión 5.1.

Capítulo I. Modelos econométricos de series de tiempo estacionarias

1.1. Definición de conceptos básicos.

1.1.1. Serie de tiempo

El análisis de series de tiempo, se puede dividir en dos tipos de métodos: los de dominio de frecuencia y los de dominio de tiempo. En el primer tipo de métodos, una serie de tiempo Z_t se expresa como la suma de senos y cosenos, que varían en forma independiente con amplitudes aleatorias. El segundo método se basa en modelar directamente las relaciones de una serie y su pasado. Esta tesis utiliza el segundo método.

La mayoría de los conceptos que se exponen en esta sección se basan en Box- Jenkins(1990) .

Definición 1. Una serie de tiempo se define como un conjunto de observaciones generadas sucesivamente a través del tiempo.

Si los valores futuros de una serie de tiempo están determinados exactamente por alguna función matemática, entonces se dice que la serie de tiempo es determinista.

Si los valores futuros pueden ser descritos sólo en términos de una función de probabilidad, la serie de tiempo es no determinista o estocástica. En este trabajo las series de tiempo son estocásticas y se describen dentro del contexto de procesos estocásticos, para ello es necesario definir un proceso estocástico.

1.1.2. Procesos Estocásticos

Definición 2. Un proceso estocástico¹ $\{Z(t), t \in T\}$ es una colección de variables aleatorias indexadas a un tiempo t . Esto es, para cada $t \in T$, $Z(t)$ es una variable aleatoria. El índice t es interpretado como tiempo. El conjunto T es llamado el conjunto índice del proceso. Por notación se usará Z_t en vez de $Z(t)$.

Cuando T es un conjunto contable² se dice que el proceso estocástico es discreto en el tiempo. Cuando T es un intervalo de la recta real, el proceso estocástico es un proceso continuo en el tiempo.

¹ Ross, Sheldon. M., 2002

² Por conjunto contable se entiende un conjunto finito o infinito numerable

Con base en la definición anterior, y dado que nuestros datos son mensuales, la serie IIS (Índice de la inflación subyacente) se puede ver como una sucesión de observaciones generadas por un proceso estocástico discreto

Definición 3. Los procesos estocásticos cuyos momentos de primer y segundo orden (media y varianza)³ son constantes, y su función de autocovarianza solo depende de la diferencia en el tiempo, se les denomina procesos estacionarios de segundo orden ó procesos estacionarios en covarianza.⁴

Los procesos descritos en esta tesis son estacionarios en covarianza.

Un importante proceso estacionario en covarianza es el llamado proceso de ruido blanco.

Definición 4. Un proceso ruido blanco es una colección de variables aleatorias $\varepsilon_t, \varepsilon_{t-1}, \dots$ no autocorrelacionadas, con media cero y varianza constante, es decir:

$$E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2. \quad (1.1)$$

Ya que las variables aleatorias ε_t no están autocorrelacionadas, la función de autocovarianza de un ruido blanco esta dada por

$$\begin{aligned} \gamma_k &= E[(\varepsilon_t - \mu_t)(\varepsilon_{t+k} - \mu_{t+k})] \\ &= E[(\varepsilon_t - 0)(\varepsilon_{t+k} - 0)] \\ &= E(\varepsilon_t \varepsilon_{t+k}). \end{aligned}$$

Ahora, por definición la varianza del ruido blanco esta dada por

$$\text{Var}(\varepsilon_t) = E(\varepsilon_t^2) - [E(\varepsilon_t)]^2,$$

y asumiendo (1.1) se tiene que

$$\sigma_\varepsilon^2 = E(\varepsilon_t^2).$$

Por lo tanto la función de autocovarianza esta dada por

$$\gamma_k = E(\varepsilon_t \varepsilon_{t-k}) = \begin{cases} \sigma_\varepsilon^2 & k = 0 \\ 0 & k \neq 0 \end{cases}, \quad (1.2)$$

La función de autocorrelación del ruido blanco esta dada por

³ Lo conceptos de varianza, covarianza, autocovarianza, función de covarianza y autocorrelación se introducen en el anexo 2.

⁴ También existen los procesos estacionarios estrictamente, pero para fines de esta tesis nos basta la definición de procesos estacionarios en covarianza.

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

1.1.3. Procesos lineales generales

En el análisis de series de tiempo, existen dos representaciones útiles para representar estas series. Una es escribir el proceso Z_t como una combinación lineal de una sucesión de variables aleatorias no correlacionadas,

$$Z_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \mu + \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}, \quad (1.3)$$

donde μ es un parámetro que determina el nivel del proceso, $\psi_0 = 1$ y la sucesión $\{\varepsilon_t\}$ es un proceso de ruido blanco.

Utilizando el operador rezago⁵, podemos escribir (1.3) como:

$$\tilde{Z}_t = \psi(B)\varepsilon_t.$$

donde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots \quad (1.4)$$

$\psi(B)$ es un operador lineal que transforma ε_t en Z_t y se denomina función de transferencia del filtro.

Una suma infinita de variables aleatorias se define como el límite en media cuadrática de una suma parcial finita⁶. Entonces, Z_t en (1.3) se define como

$$E \left[\left(\tilde{Z}_t - \sum_{j=0}^n \psi_j \varepsilon_{t-j} \right)^2 \right] \rightarrow 0 \quad \text{cuando} \quad n \rightarrow \infty,$$

donde $\tilde{Z}_t = Z_t - \mu$.

La otra forma útil es escribir el proceso Z_t en una representación autorregresiva de orden infinito, en la cual la variable Z en el periodo t depende de sus propios valores, observados en periodos anteriores a t y

⁵ El operador rezago B se define de la siguiente manera: $BZ_t = Z_{t-1}$; $B^m Z_t = Z_{t-m}$

⁶ Ver Wei, William, W. S. pag. 23.

ponderados de acuerdo con los coeficientes $\pi_1, \pi_2, \pi_3, \dots$, más un choque aleatorio ε_t , esto es,

$$\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + \varepsilon_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + \varepsilon_t \quad (1.5)$$

donde $\pi_1 = 1$.

La relación entre las ponderaciones ψ y π de los modelos dados en (1.3) y (1.5) se puede expresar haciendo uso del operador rezago B . La ecuación (1.3) se puede escribir como

$$\tilde{Z}_t = \left(1 + \sum_{j=1}^{\infty} \psi_j B^j \right) \varepsilon_t.$$

Haciendo uso del operador lineal dado en (1.4) se tiene que

$$\tilde{Z}_t = \psi(B) \varepsilon_t, \quad (1.6)$$

donde

$$\psi(B) = 1 + \sum_{j=1}^{\infty} \psi_j B^j = \sum_{j=0}^{\infty} \psi_j B^j$$

con $\psi_0 = 1$.

Por otro lado (1.5) se puede escribir como

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j \right) \tilde{Z}_t = \varepsilon_t$$

ó

$$\pi(B) \tilde{Z}_t = \varepsilon_t \quad (1.7)$$

donde $\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$

Después de aplicar el operador $\psi(B)$ en ambos lados de (1.7) se obtiene

$$\psi(B) \pi(B) \tilde{Z}_t = \psi(B) \varepsilon_t = \tilde{Z}_t$$

A partir de la expresión anterior, Box-Jenkins⁷ concluyen que

$$\psi(B)\pi(B) = I,$$

es decir

$$\pi(B) = \psi^{-1}(B) \quad (1.8)$$

La relación dada en la ecuación (1.8) sirve para encontrar las π una vez que se conocen las ψ y viceversa.

A continuación se obtienen las condiciones que debe satisfacer el proceso lineal general dado en (1.3) para que sea estacionario, es decir se analiza bajo que condiciones la media y la varianza del proceso son constantes, y la covarianza sólo depende de la diferencia en el tiempo.

1.1.3.1. Estacionariedad de un proceso lineal general

Supongamos que la media de Z_t es μ_t , es decir

$$E(Z_t) = \mu_t$$

entonces de (1.3) se tiene

$$\mu_t = \mu + E(\varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \dots)$$

y esta esperanza existe sólo si

$$\sum_{i=0}^{\infty} |\psi_i| < \infty \quad (1.9)$$

Si esto se cumple, entonces

$$\begin{aligned} E(Z_t) &= \mu + E\left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right) \\ &= \mu + E(\varepsilon_t) + \psi_1 E(\varepsilon_{t-1}) + \psi_2 E(\varepsilon_{t-2}) + \dots \\ &= \mu \end{aligned} \quad (1.10)$$

De la definición de varianza de Z_t , y sabiendo de (1.10) que $E(Z_t) = \mu$, se tiene

$$\text{Var}(Z_t) = E(Z_t - E(Z_t))^2 = E(Z_t - \mu)^2 = E\left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right)^2$$

⁷ Ver Box-Jenkins, 1990, pag. 49.

$$\begin{aligned}
&= \psi_0^2 E(\varepsilon_t)^2 + \psi_1^2 E(\varepsilon_{t-1})^2 + \dots + \sum_{i \neq j} \psi_i \psi_j \varepsilon_i \varepsilon_j \\
&= \psi_0^2 \sigma_\varepsilon^2 + \psi_1^2 \sigma_\varepsilon^2 + \psi_2^2 \sigma_\varepsilon^2 + \dots \\
&= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i^2
\end{aligned} \tag{1.11}$$

Para la covarianza se tiene

$$\begin{aligned}
\gamma_k &= E[(Z_t - E(Z_t))(Z_{t+k} - E(Z_{t+k}))] \\
&= E[(Z_t - \mu)(Z_{t+k} - \mu)] \\
&= E[(\psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots)(\psi_0 \varepsilon_{t+k} + \psi_1 \varepsilon_{t+k-1} + \psi_2 \varepsilon_{t+k-2} + \dots)] \\
&= E\left(\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t+k-j}\right) \\
&= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}
\end{aligned} \tag{1.12}$$

$$\rho_k = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+k}}{\sum_{i=0}^{\infty} \psi_i^2} \tag{1.13}$$

Ahora, la varianza, y la autocovarianza en (1.11) y (1.12) no dependen del tiempo, pero para que tengan sentido, se debe cumplir que

$$\sum_{i=0}^{\infty} \psi_i^2 < \infty, \quad \sum_{i=0}^{\infty} \psi_i \psi_{i+k} < \infty,$$

y esto ocurre si se cumple⁸ la condición dada en (1.9), por lo tanto para que el proceso (1.3) sea estacionario se tiene que cumplir que

$$\sum_{j=0}^{\infty} |\psi_j| < \infty.$$

1.1.3.2. Invertibilidad de un proceso lineal general

El concepto de invertibilidad es independiente del concepto de estacionariedad, y se puede aplicar a series que no sean estacionarias. Tiene que ver con la posibilidad de expresar una serie de tiempo como el proceso dado en (1.5), en términos de sus observaciones pasadas y que la serie se reversible a su media, es decir que varié en torno a su media.

⁸ Ver Guerreo, 2003, pag 30.

El proceso de lineal general bajo ciertas condiciones se puede escribir como un proceso autorregresivo infinito⁹. Sin embargo la condición de estacionariedad del proceso lineal general, no garantiza que una vez que el proceso se ha escrito en términos de sus valores pasados, este no se aleje de la media. La condición que se debe de cumplir para que el proceso sea invertible es

$$\sum_{j=0}^{\infty} |\pi_j| < \infty \quad (1.14)$$

Si se cumple esta condición, entonces la serie

$$\pi(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots) = \sum_{j=0}^{\infty} \pi_j B^j \quad \text{con } \pi_0 = 1$$

converge para $|\pi_i| \leq 1$.

En resumen, se tiene que el proceso (1.3) es estacionario si $\sum_{j=0}^{\infty} |\psi_j| < \infty$ y es

invertible si $\sum_{j=0}^{\infty} |\pi_j| < \infty$.

Como nuestro objetivo es encontrar el mecanismo generador de los datos del índice la inflación subyacente, a continuación se presentan los modelos que pueden generar nuestra serie.

1.2. Modelos ARIMA

La función de autocorrelación y de autocorrelación parcial son muy importantes a la hora de ajustar un modelo, ya que la forma de estas funciones nos permite identificar un posible modelo, y el orden probable del modelo. Otra condición importante a la hora de identificar un modelo es ver si este es invertible y estacionario, y esto se verifica con cierta propiedad que tiene que cumplir las raíces unitarias. Por ello estos conceptos se abordan para cada uno de los modelos que se presentan.

1.2.1. Modelos Autorregresivos de orden p , $AR(p)$

⁹ Ver Koopmans, 1974, pag. 254.

En éste modelo el valor actual del proceso se expresa como una función de los valores que tomó la serie en un número finito de realizaciones anteriores más un choque aleatorio ε_t .

$$\text{Si } \tilde{Z}_{t-i} = Z_{t-i} - \mu,$$

son las desviaciones con respecto a μ , entonces

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t, \quad (2.1)$$

y se denomina proceso autorregresivo de orden P . Si se define el operador autorregresivo como:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (2.2)$$

entonces el proceso autorregresivo puede escribirse abreviadamente como:

$$\phi(B)\tilde{Z}_t = \varepsilon_t. \quad (2.3)$$

Este modelo contiene $p+2$ parámetros desconocidos $\mu, \phi_1, \phi_2, \dots, \phi_p, \sigma_\varepsilon^2$

Como lo procesos que nos interesan, deben de ser estacionarios e invertibles, a continuación se ve que condiciones debe cumplir un proceso $AR(p)$ para que cumpla con estas propiedades.

1.2.1.1. Estacionariedad e invertibilidad de un $AR(p)$

El proceso autorregresivo es un caso particular del proceso lineal general. Para un proceso autorregresivo de orden p sabemos que

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t \quad (2.4)$$

$$\tilde{Z}_{t-1} = \phi_1 \tilde{Z}_{t-2} + \phi_2 \tilde{Z}_{t-3} + \dots + \phi_p \tilde{Z}_{t-p-1} + \varepsilon_{t-1} \quad (2.5)$$

$$\tilde{Z}_{t-2} = \phi_1 \tilde{Z}_{t-3} + \phi_2 \tilde{Z}_{t-4} + \dots + \phi_p \tilde{Z}_{t-p-2} + \varepsilon_{t-2}, \quad (2.6)$$

...

...

Sustituyendo (2.5), en el lado derecho de (2.4) eliminamos \tilde{Z}_{t-1} , de igual forma podemos sustituir (2.6) para eliminar \tilde{Z}_{t-2} , continuando con este proceso obtenemos una serie infinita de ε_t 's en el lado derecho de (2.4). Simbólicamente se tiene que

$$\phi(B)\tilde{Z}_t = \varepsilon_t, \quad (2.7)$$

que es equivalente a

$$\tilde{Z}_t = \psi(B)\varepsilon_t, \quad (2.8)$$

con $\psi(B) = \phi^{-1}(B)$.

Como $\phi(B)$ es un polinomio en términos de B , por el teorema fundamental del álgebra¹⁰, se puede expresar como

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B),$$

donde $G_1^{-1}, \dots, G_p^{-1}$ son las raíces de $\phi(B) = 0$.

Como queremos expresar el proceso $AR(p)$ en la forma del proceso lineal general, tomamos $\phi^{-1}(B)$, y lo expresamos en fracciones parciales, entonces

$$\tilde{Z}_t = \phi^{-1}(B)\varepsilon_t = \sum_{i=1}^p \frac{K_i}{1 - G_i B} \varepsilon_t.$$

A partir de la ecuación anterior, Box-Jenkins¹¹ concluyen que si la serie es convergente si $|B| \leq 1$, es decir que si las $\psi_j = \sum_{i=1}^p K_i G_i^j$ son absolutamente sumables, entonces las $|G_i| < 1$ para $i \in \{1, 2, \dots, p\}$. Y que si esto ocurre, entonces el proceso $AR(p)$, representa un proceso estacionario.

Como las raíces de la ecuación característica $\phi(B) = 0$ son de la forma G_i^{-1} , entonces éstas deben estar fuera del círculo unitario

A la hora de ajustar el modelo, esta condición es muy importante en la etapa de diagnóstico de la metodología de Box-Jenkins.

Por ejemplo, para el caso del proceso $AR(1)$, de (2.1) se tiene que

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \varepsilon_t. \quad (2.10)$$

Haciendo uso del operador rezago de (2.10) se tiene que

$$(1 - \phi_1 B)\tilde{Z}_t = \varepsilon_t$$

y entonces la ecuación característica es

¹⁰ Ver el anexo 2.

¹¹ Ver Box-Jenkins, 1990, pag. 55.

$$1 - \phi_1 B = 0$$

cuya raíz es $B = \frac{1}{\phi_1}$. Por lo tanto para que el proceso sea estacionario se tiene que cumplir que $\left| \frac{1}{\phi_1} \right| > 1$, de donde se sigue esta condición es equivalente a que $|\phi_1| < 1$.

Como la serie

$$\pi(B) = \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

es finita, no hay restricciones para los parámetros de un modelo autorregresivo para que se asegure su invertibilidad.

1.2.1.2. Función de autocorrelación de un $AR(p)$

De (2.1) tenemos que la expresión de un $AR(p)$ es

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

Multiplicando la ecuación anterior por \tilde{Z}_{t-k} se tiene que:

$$\tilde{Z}_{t-k} \tilde{Z}_t = \phi_1 \tilde{Z}_{t-k} \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-k} \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-k} \tilde{Z}_{t-p} + \tilde{Z}_{t-k} \varepsilon_t. \quad (2.11)$$

Tomando valores esperados en (2.11) de ambos lados de la igualdad se tiene

$$E(\tilde{Z}_{t-k} \tilde{Z}_t) = E(\phi_1 \tilde{Z}_{t-k} \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-k} \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-k} \tilde{Z}_{t-p} + \tilde{Z}_{t-k} \varepsilon),$$

teniendo en cuenta que $E(\varepsilon_t \varepsilon_{t-k}) = 0$ y que $E(\tilde{Z}_t) = 0$ se tiene que¹²

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad \text{si } k > 0. \quad (2.12)$$

Dividiendo (2.12) por γ_0 se tiene que

$$\frac{\gamma_k}{\gamma_0} = \phi_1 \frac{\gamma_{k-1}}{\gamma_0} + \phi_2 \frac{\gamma_{k-2}}{\gamma_0} + \dots + \phi_p \frac{\gamma_{k-p}}{\gamma_0}, \quad \text{si } k > 0, \quad (2.13)$$

¹² $E(\tilde{Z}_t) = E(Z_t - \mu) = E(Z_t) - \mu = \mu - \mu = 0$, ya que μ es la media del proceso.

de donde

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad \text{si } k > 0. \quad (2.14)$$

Si se sustituyen los valores $1, \dots, p$ en (2.14) se obtiene un conjunto de ecuaciones lineales para $\phi_1, \phi_2, \dots, \phi_p$ en términos de $\rho_1, \rho_2, \dots, \rho_p$, que son las ecuaciones de Yule-Walker:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (2.15)$$

A partir de éste conjunto de ecuaciones podemos obtener las estimaciones de los parámetros reemplazando las correlaciones teóricas por las correspondientes estimaciones obtenidas a partir de los datos muestrales. Si se define

$$\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}, \quad \boldsymbol{\rho}_p = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix}, \quad \mathbf{P}_p = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix}$$

la solución de (3.15) para los parámetros $\boldsymbol{\phi}$ en términos de las autocorrelaciones se puede escribir como

$$\boldsymbol{\phi} = \mathbf{P}_p^{-1} \boldsymbol{\rho}_p$$

1.2.1.3. Función de autocorrelación parcial de un $AR(p)$

La función de autocorrelación parcial (FACP), es una correlación condicional, que cuantifica la dependencia lineal entre \tilde{Z}_t y \tilde{Z}_{t+k} , sin tener en cuenta $\tilde{Z}_{t+1}, \tilde{Z}_{t+2}, \dots, \tilde{Z}_{t+k-1}$, es decir la FACP esta dada por la siguiente correlación

$$\phi_{kk} = \text{corr}(\tilde{Z}_t, \tilde{Z}_{t+k} \mid \tilde{Z}_{t+1}, \tilde{Z}_{t+2}, \dots, \tilde{Z}_{t+k-1})$$

Lo que se quiere calcular es la contribución de \tilde{Z}_{t+k} , para explicar linealmente a \tilde{Z}_t , y se mide mediante ϕ_{kk} .

Los parámetros $\phi_{11}, \phi_{22}, \dots, \phi_{pp}$, pueden obtenerse a partir de las ecuaciones de Yule-Walker (2.15), como

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \\ &\dots \\ \phi_{pp} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-4} & \rho_3 \\ & & & \dots & & \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & \rho_1 & \rho_p \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-3} & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-4} & \rho_{p-3} \\ & & & \dots & & \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & \rho_1 & 1 \end{vmatrix}}. \end{aligned}$$

Para un proceso $AR(p)$ la función de autocorrelación parcial es nula para valores de k superiores a p . La función de autocorrelación parcial puede estimarse ajustando a los datos procesos autorregresivos de orden sucesivo, mediante el método de mínimos cuadrados, y seleccionando la estimación del último valor obtenido $\hat{\phi}_{kk}, k \in N$.

Si los valores de los parámetros no están demasiado próximos a la frontera de estacionariedad, pueden utilizarse las estimaciones aproximadas que suministran las ecuaciones de Yule-Walker una vez reemplazados las ρ_j por r_j , resultando entonces:

$$\begin{pmatrix} 1 & r_1 & \dots & r_{k-1} \\ r_1 & 1 & \dots & r_{k-2} \\ \dots & \dots & \dots & \dots \\ r_{k-1} & r_{k-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \hat{\phi}_{k1} \\ \hat{\phi}_{k2} \\ \dots \\ \hat{\phi}_{kk} \end{pmatrix} = \begin{pmatrix} r_{k1} \\ r_{k2} \\ \dots \\ r_{kk} \end{pmatrix}.$$

1.2.2. Procesos de Promedios Móviles de orden q , $MA(q)$

En éste proceso se hace depender linealmente a \tilde{Z}_t de un número finito de realizaciones previas de la variable aleatoria ε_t . Así la expresión para un proceso $MA(q)$ está dada por

$$\tilde{Z}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}. \quad (2.16)$$

Si definimos el operador de promedios móviles de orden q como

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (2.17)$$

Podemos escribir el proceso $MA(q)$ como

$$\tilde{Z}_t = \theta_q(B) \varepsilon_t, \quad (2.18)$$

el cual contiene $q+2$ parámetros desconocidos.

$$\mu, \theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2.$$

1.2.2.1. Estacionariedad e invertibilidad para un $MA(q)$

A continuación se deducen las condiciones para que un proceso $MA(q)$ sea invertible. De (2.16) tenemos que un proceso $MA(q)$, esta dado por

$$\begin{aligned} \tilde{Z}_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}. \\ &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t \\ &= \theta(B) \varepsilon_t \end{aligned}$$

De la última igualdad se puede obtener

$$\varepsilon_t = \theta^{-1}(B) \tilde{Z}_t.$$

Como $\theta(B)$ es un polinomio en términos de B , por el teorema fundamental del algebra lo podemos expresar como

$$\theta^{-1}(B) = \prod_{i=1}^q (1 - H_i B)$$

Expandiendo este producto en fracciones parciales se tiene

$$\pi(B) = \theta^{-1}(B) = \sum_{i=1}^q \left(\frac{M_i}{1 - H_i B} \right) \tilde{Z}_t.$$

Expresando $\frac{1}{1 - H_i B}$ como $\sum_{j=0}^{\infty} H_i^j B^j$ y sustituyendo en la ecuación anterior se tiene

$$\pi(B) = \sum_{i=1}^q M_i \left(\sum_{j=0}^{\infty} H_i^j B^j \right) \tilde{Z}_t$$

Box-Jenkins¹³ mencionan que esta serie converge si los coeficiente $\pi_j = -\sum_{i=1}^q M_i H_i^j$ son absolutamente sumables, es decir si $|H_i| < 1$, para $i = 1, 2, \dots, q$. Si esto se cumple entonces el proceso es invertible.

Como las raíces de la ecuación característica $\theta(B) = 0$, son de la forma H_i^{-1} , entonces las raíces de la ecuación característica

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q = 0$$

deben de estar fuera del círculo unitario.

Para el caso de un $MA(1)$, se tiene que

$$\tilde{Z}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} = (1 - \theta_1 B) \varepsilon_t$$

por lo tanto la ecuación característica está dada por $1 - \theta_1 B = 0$, que tiene como raíz $B = \frac{1}{\theta_1}$, por lo tanto para que el proceso sea invertible se debe cumplir que $|\theta_1| < 1$.

Como la serie

$$\psi(B) = \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

es finita, no se necesita ninguna restricción sobre los parámetros del proceso $MA(q)$, para que este sea finito.

¹³ Ver Box-Jenkins, 1990, pag. 70.

1.2.2.2. Función de autocorrelación para un proceso $MA(q)$

La autocovarianza de orden k para un proceso $MA(q)$ está dada por la expresión

$$\begin{aligned}
 \gamma_k &= E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-k} - \theta_1 \varepsilon_{t-k-1} - \theta_2 \varepsilon_{t-k-2} - \dots - \theta_q \varepsilon_{t-k-q})] \\
 &= E[\varepsilon_t \varepsilon_{t-k} - \theta_1 \varepsilon_t \varepsilon_{t-k-1} - \theta_2 \varepsilon_t \varepsilon_{t-k-2} - \dots - \theta_q \varepsilon_t \varepsilon_{t-k-q} \\
 &\quad - \theta_1 \varepsilon_{t-1} \varepsilon_{t-k} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-k-1} + \theta_1 \theta_2 \varepsilon_{t-1} \varepsilon_{t-k-2} + \dots + \theta_1 \theta_q \varepsilon_{t-1} \varepsilon_{t-k-q} \\
 &\quad - \theta_2 \varepsilon_{t-2} \varepsilon_{t-k} + \theta_2 \theta_1 \varepsilon_{t-2} \varepsilon_{t-k-1} + \theta_2^2 \varepsilon_{t-2} \varepsilon_{t-k-2} + \dots + \theta_2 \theta_q \varepsilon_{t-2} \varepsilon_{t-k-q} \\
 &\quad \vdots \\
 &\quad - \theta_q \varepsilon_{t-q} \varepsilon_{t-k} + \theta_q \theta_1 \varepsilon_{t-q} \varepsilon_{t-k-1} + \theta_q \theta_2 \varepsilon_{t-q} \varepsilon_{t-k-2} + \dots + \theta_q^2 \varepsilon_{t-q} \varepsilon_{t-k-q}]
 \end{aligned}$$

Entonces como ε_t es un ruido blanco, tomando esperanzas y haciendo $k = 1, 2, 3, 4, \dots, k, k+1$, se tiene que

$$\gamma_k = \begin{cases} \sigma_\varepsilon^2 (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q) & k = 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (2.19)$$

En particular cuando $k = 0$, para la varianza se tiene que

$$\gamma_0 = \sigma_\varepsilon^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \quad (2.20)$$

por lo tanto la función de autocorrelación es

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & k = 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (2.21)$$

La función de autocorrelación para un proceso $MA(q)$ es nula para retardos superiores a q . Si las correlaciones $\rho_1, \rho_2, \dots, \rho_k$ son conocidas, las q ecuaciones anteriores pueden resolverse para los parámetros $\theta_1, \theta_2, \dots, \theta_k$. Sin embargo, a diferencia de las ecuaciones de Yule-Walker para un proceso $MA(q)$ que eran lineales, estas ecuaciones son no lineales, excepto en el caso

particular de que $q = 1$; por lo tanto, se resuelven de manera iterativa, mediante un método apropiado.

1.2.2.3. Función de autocorrelación parcial de un $MA(q)$

Un proceso $MA(q)$ (el cual es equivalente a un proceso $AR(\infty)$), tiene todas sus autocorrelaciones parciales distintas de cero, aunque la FACP muestre convergencia a cero.

1.2.3. Procesos autorregresivos y de promedios móviles mixtos

$ARMA(p, q)$

Cuando deseamos adquirir una mayor flexibilidad en el ajuste de series de tiempo puede resultar conveniente incluir tanto términos autorregresivos como de promedios móviles, así un modelo mixto con estos dos procesos tiene la siguiente forma:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (2.22)$$

La ecuación (2.22) también se puede escribir como

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

que en notación más resumida es

$$\phi(B) \tilde{Z}_t = \theta(B) \varepsilon_t,$$

donde $\phi(B)$ y $\theta(B)$ son polinomios de grado p y q en B .

En este modelo existen $p + q + 2$ parámetros desconocidos a estimar. Este modelo es un $ARMA(p, q)$

1.2.3.1. Estacionariedad e invertibilidad de un $ARMA(p, q)$

El proceso $ARMA(p, q)$ se puede ver de dos maneras:

1.- Como un proceso autorregresivo de orden p

$$\phi(B) \tilde{Z}_t = \varepsilon_t$$

donde ε_t sigue un proceso de promedios móviles de orden q

$$\varepsilon_t = \theta(B)a_t$$

2.-Como un proceso de promedios móviles de orden q

$$\tilde{Z}_t = \theta(B)b_t$$

donde b_t sigue un proceso autorregresivo de orden p

$$\phi(B)b_t = a_t,$$

De lo anterior se tiene que

$$\phi(B)\tilde{Z}_t = \theta(B)\phi(B)b_t = \theta(B)a_t.$$

Box-Jenkins mencionan que los términos del proceso de promedios móviles de la derecha de la ecuación anterior, no afectan el argumento de la sección 1.1.3.1., que establece las condiciones para la estacionariedad para un proceso autorregresivo.

Entonces, $\phi(B)\tilde{Z}_t = \theta(B)a_t$ va a definir un proceso estacionario si la ecuación característica $\phi(B) = 0$ tiene todas sus raíces fuera del círculo unitario. Similarmente, las raíces de $\theta(B) = 0$ deben estar fuera del círculo unitario para que el proceso sea invertible.

El proceso estacionario e invertible $ARMA(p, q)$, tiene una representación de promedios móviles infinitos

$$\tilde{Z}_t = \psi(B)a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j},$$

donde $\psi(B) = \theta^{-1}(B)\phi(B)$, y una representación autorregresiva de orden infinito

$$\pi(B)\tilde{Z}_t = \tilde{Z}_t - \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} = a_t,$$

donde $\pi(B) = \theta^{-1}(B)\phi(B)$, donde los coeficientes π_i y ψ_j son absolutamente sumables.

1.2.3.2. Función de autocorrelación de un $ARMA(p, q)$

De (2.22) tenemos que un proceso $ARMA(p, q)$ se representa por

$$\tilde{Z}_t = \phi_1 \tilde{Z}_t + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q},$$

multiplicando la ecuación anterior por \tilde{Z}_{t-k} se tiene que

$$\tilde{Z}_t \tilde{Z}_{t-k} = \phi_1 \tilde{Z}_{t-1} \tilde{Z}_{t-k} + \dots + \phi_p \tilde{Z}_{t-p} \tilde{Z}_{t-k} + \varepsilon_t \tilde{Z}_{t-k} - \theta_1 \varepsilon_{t-1} \tilde{Z}_{t-k} - \dots - \theta_q \varepsilon_{t-q} \tilde{Z}_{t-k}, \quad (2.23)$$

tomando esperanzas en (2.23), sabiendo que ε_t es un ruido blanco y $E(\tilde{Z}_t) = 0$, se tiene

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + \gamma_{z\varepsilon}(k) - \theta_1 \gamma_{z\varepsilon}(k-1) - \dots - \theta_q \gamma_{z\varepsilon}(k-q), \quad (2.24)$$

donde $\gamma_{z\varepsilon}(k)$ es la función de autocovarianza entre Z_t y ε_t , y se define como $\gamma_{z\varepsilon}(k) = E(\tilde{Z}_{t-k} \varepsilon_t)$. Como \tilde{Z}_{t-k} depende sólo de los choques aleatorios ocurridos hasta $t-k$, si expresamos \tilde{Z}_{t-k} por medio de un proceso de media móvil infinito, se tiene que

$$\tilde{Z}_{t-k} = \psi(B) \varepsilon_{t-k} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-k-j},$$

entonces se sigue que

$$\gamma_{z\varepsilon}(k) = \begin{cases} 0 & k > 0 \\ \psi_{-k} \sigma_\varepsilon^2 & k \leq 0. \end{cases}$$

Por lo tanto (2.24) se puede expresar como

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} - \sigma_\varepsilon^2 (\theta_k \psi_0 + \theta_{k+1} \psi_1 + \dots + \theta_q \psi_{q-k}), \quad (2.25)$$

con la convención de que $\theta_0 = -1$

de (2.25) se deduce que

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}, \quad k \geq q+1$$

y por lo tanto:

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k \geq q+1.$$

1.2.3.3. Función de autocorrelación parcial de un $ARMA(p, q)$

Un proceso $ARMA(p, q)$ tendrá asociada una FACP que no desaparece después de un número finito de retrasos, ya que el proceso $ARMA(p, q)$ se puede expresar como un proceso autorregresivo infinito.

Hasta ahora hemos obtenido los conceptos teóricos para saber cuando un serie de tiempo es estacionaria e invertible, y como identificar un modelo para este tipo de series, sin embargo en la práctica la mayoría de las series de tiempo son no estacionarios, y este también es el caso de la serie de tiempo Inflación subyacente. A continuación se presentan un tipo de series no estacionarios, pero que se pueden transformar en series estacionarias, mediante el operador diferencia.

1.2.4. Modelos ARIMA

Las series de tiempo que no son estacionarias, no oscilan en torno a un valor central medio. Aunque el nivel absoluto en torno al cual oscila la serie sea cambiante con el tiempo, en algunas ocasiones se observa que al tomar la serie en primeras diferencias, el comportamiento tiende a ser más regular, y este es el caso de la serie de tiempo Inflación subyacente.

Los modelos autorregresivos e integrados de promedios móviles (ARIMA), pueden ser vistos como una generalización de los modelos ARMA vistos en la sección anterior. Lo que se hace al aplicar el operador diferencia ∇^d es eliminar una posible tendencia polinomial de orden d , presente en la serie que se este analizando. Si el proceso original $\{Z_t\}$ no es estacionario porque presenta una tendencia polinomial no determinista, es posible construir el proceso estacionario

Este comportamiento se puede modelar mediante el operador autorregresivo generalizado en el cual se pretende separar en el operador autorregresivo la influencia de las raíces unitarias, definiéndose como $\{W_t\}$, en donde

$$W_t = \nabla^d Z_t, \quad \forall t$$

y para esta nueva serie ya es posible obtener un modelo ARMA $\phi(B)W_t = \theta(B)\varepsilon_t$, lo cuál es equivalente a considerar un modelo ARIMA

$$\phi(B)\nabla^d \tilde{Z}_t = \theta(B)\varepsilon_t \quad d \geq 1, \quad (2.26)$$

para $\{\tilde{Z}_t\}$, donde ε_t , es un ruido blanco.

1.2.4.1. Procesos integrados de orden d

Definición 5. *Un proceso $\{Z_t\}$ se hace estacionario después de diferenciarlo d -veces, se dice que el proceso es integrado de orden d , y se denota por $I(d)$.*

En la práctica para saber si un proceso es integrado de orden d , se usa la prueba de hipótesis de Dickey-Fuller

Las series de tiempo pueden presentar dos tipos de tendencia, determinista y estocástica.

Una serie de tiempo que presenta el primer tipo de tendencia puede ser aproximada por un proceso del tipo:

$$Z_t = \alpha + \beta_t + \varepsilon_t, \quad (2.27)$$

donde t captura la tendencia de la serie y ε_t es un ruido blanco.

En la práctica sin embargo, muchas series no caen dentro de la categoría anterior, un ejemplo es la caminata aleatoria, la cual presenta tendencia estocástica.

Definición 6. Una serie de tiempo Z_t es una caminata aleatoria si satisface:

$$Z_t = Z_{t-1} + \varepsilon_t, \quad (2.28)$$

donde Z_0 es un número real que denota el valor de inicio del proceso y ε_t es un ruido blanco.

Si consideramos un proceso un proceso $AR(1)$

$$Z_t = \rho Z_{t-1} + \varepsilon_t \quad (2.29)$$

tenemos tres casos interesantes para los valores del coeficiente ρ

1. $|\rho| < 1$,
2. $\rho = 1$,
3. $|\rho| > 1$.

Cuando $|\rho| < 1$ tenemos un proceso $AR(1)$ estacionario, esto se sigue de las condiciones de estacionariedad dadas en la sección 1.1.3.1.

Cuando $\rho = 1$, tenemos una caminata aleatoria, como la definida en (2.28), que es un caso especial de un $AR(1)$. Se dice que este proceso tiene una raíz unitaria, ya que la raíz del polinomio característico:

$$1 - B = 0,$$

es uno. De la definición 5, de proceso integrado, tenemos que la caminata aleatoria es un proceso integrado de orden 1, y se denota por $I(1)$, ya que se vuelve un proceso estacionario después de diferenciar la serie una vez. En el caso de la caminata aleatoria se tiene que

$$Z_t - Z_{t-1} = \nabla Z_t = \varepsilon_t, \quad (2.30)$$

es un proceso estacionario, ya que un ruido blanco lo es.

Cuando $|\rho| > 1$ se tiene que

$$Z_t - Z_{t-1} = \nabla Z_t = (\rho - 1)Z_{t-1} + \varepsilon_t. \quad (2.31)$$

de donde se puede ver que Z_t no es estacionaria, ya que no sólo depende del proceso de ruido blanco, sino también del proceso Z_{t-1} que no es estacionario, ya que $\rho - 1 > 0$. Por lo tanto un proceso $AR(1)$ con un coeficiente 1 es $I(1)$, pero el mismo proceso con un coeficiente de 1.01 no es, ya que aunque apliquemos el operador diferencia no se vuelve estacionario.

Las pruebas de hipótesis sobre raíces unitarias, se usan para saber con un cierto nivel de significancia si una serie presenta una raíz unitaria.

Dadas n observaciones Z_1, Z_2, \dots, Z_n , el estimador máximo verosímil de ρ es el estimador de mínimos cuadrados¹⁴.

$$\hat{\rho} = \left(\sum_{t=1}^n Z_{t-1}^2 \right)^{-1} \sum_{t=1}^n Z_t Z_{t-1}. \quad (2.32)$$

1.2.4.2. Prueba de Dickey-Fuller

Esta sección y la que se presenta a continuación están basadas en Enders (1994).

La prueba de Dickey-Fuller consiste en correr una regresión, y determinar si ρ es estadísticamente igual a uno, para saber si la serie tiene una raíz unitaria.

Los valores críticos de la distribución ρ cuando vale uno, fueron obtenidos por D. A. Dickey mediante simulaciones de Montecarlo y publicados en Fuller (1976), esta distribución es conocida como la distribución de Dickey-Fuller. Más recientemente McKinnon a agregado más datos a esta distribución.

Los valores críticos en las tablas de Dickey-Fuller pertenecen a tres tipos de modelo diferentes:

$$\nabla Z_t = (\rho - 1)Z_{t-1} + \varepsilon_t, \quad (2.33)$$

$$\nabla Z_t = \mu + (\rho - 1)Z_{t-1} + \varepsilon_t, \quad (2.34)$$

$$\nabla Z_t = \mu + \beta t + (\rho - 1)Z_{t-1} + \varepsilon_t, \quad (2.35)$$

¹⁴ Una exposición mas detallada de esta parte se encuentra en Fuller, 1976.

Haciendo $\alpha = \rho - 1$ en los tres modelos, la hipótesis nula y la hipótesis alternativa son:

$$H_0 : \alpha = 0,$$

$$H_a : \alpha < 1.$$

y el estadístico de prueba es

$$t_\alpha = \frac{\hat{\rho} - 1}{SE(\hat{\rho})}. \quad (2.36)$$

La diferencia entre las tres regresiones, se debe a la presencia de los elementos determinísticos μ y t . El primer modelo es una caminata aleatoria, el segundo modelo agrega un intercepto o coeficiente de deriva, y el tercero combina los anteriores mas una tendencia lineal.

1.2.4.3. Prueba de Dickey-Fuller Aumentada

Otro tipo de modelos más generales usan otro tipo de estadísticos de prueba que tienen la misma distribución que los estadísticos de las prueba de Dickey-Fuller.

Cuando Z_t sigue un proceso $AR(p)$, Los valores críticos en las tablas de Dickey-Fuller pertenecen a tres tipos de modelos diferentes:

$$\nabla Z_t = (\rho - 1)Z_{t-1} + \sum_{i=2}^p \beta_i \nabla Z_{t-i} + \varepsilon_t, \quad (2.37)$$

$$\nabla Z_t = \mu + (\rho - 1)Z_{t-1} + \sum_{i=2}^p \beta_i \nabla Z_{t-i} + \varepsilon_t, \quad (2.38)$$

$$\nabla Z_t = \mu + \beta t + (\rho - 1)Z_{t-1} + \sum_{i=2}^p \beta_i \nabla Z_{t-i} + \varepsilon_t, \quad (2.39)$$

En coeficiente ρ se usa para hacer la prueba de la raíz unitaria, ya que

$$T(\hat{\rho} - 1) \quad y \quad \frac{(\hat{\rho} - 1)}{SE(\hat{\rho})}.$$

tienen la misma distribución que las tablas de Dickey-Fuller cuando $t \rightarrow \infty$. Por ello se usa el mismo estadístico de prueba y la misma distribución que para la prueba de Dickey-Fuller¹⁵.

¹⁵ Una exposición mas detallada de estos estadísticos se encuentra en Banerjee 1993.

1.2.4.4. Estimación de parámetros

Los conceptos que a continuación se exponen para la parte de exposición de parámetros, están basados en Guerrero (2003) y en Wei (1990).

Una vez que se ha identificado un posible modelo $ARIMA(p, d, q)$ para nuestros datos, tenemos que hacer uso eficiente de los datos disponibles, para obtener los mejores estimadores de nuestro modelo. La estimación de los parámetros de un modelo $ARIMA$ se hace usualmente usando mínimos cuadrados no lineales ($MNLS$). A continuación se presenta el método de máxima verosimilitud y el método de mínimos cuadrados no lineales

Supongamos que después de diferenciar d veces la serie Z_t , el modelo que tenemos es el siguiente:

$$\phi(B)\nabla^d(Z_t) = \theta_0 + \theta(B)\varepsilon_t. \quad (2.40)$$

Entonces los parámetros que tenemos que estimar son:

$$\phi_1, \dots, \phi_p, \theta_0, \dots, \theta_q.$$

Método de máxima verosimilitud

Guerrero, Guzmán¹⁶ menciona que si se parte del supuesto de que $\{\varepsilon_t\}$ es un ruido blanco, con distribución normal, media cero y varianza σ_ε^2 , entonces la función de densidad conjunta de los errores aleatorios esta dada por

$$f(\varepsilon_{d+p+1}, \varepsilon_{d+p+2}, \dots, \varepsilon_N) = (2\pi)^{-(N-d-p)/2} \sigma_\varepsilon^{-N+d+p} \exp\left\{-\sum_{t=d+p+1}^N \frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}\right\}, \quad (2.41)$$

También menciona que si se considera la variable

$$W_t = \nabla^d T(Z_t), \quad (2.42)$$

entonces la expresión (2.41) implica que

$$\varepsilon_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.43)$$

y esto permite obtener la función de densidad conjunta de $W_{d+p+1}, W_{d+p+2}, \dots, W_N$ a partir de (2.41) como

$$f(W_{d+p+1}, W_{d+p+2}, \dots, W_N) = f(\varepsilon_{d+p+1}, \varepsilon_{d+p+2}, \dots, \varepsilon_N) \prod_{t=d+p+1}^N \left| \frac{dW_t}{d\varepsilon_t} \right|$$

¹⁶ Ver Guerrero, Guzmán, 2003, pag. 135.

$$= (2\pi)^{(N-d-p)/2} \sigma_\varepsilon^{-N+d+p} \exp\left\{-\sum_{t=d+p+1}^N (W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_0 + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q})^2 / 2\sigma_\varepsilon^2\right\}$$

(2.44)

Esta función nos permite calcular probabilidades de la distribución normal multivariada una vez que se conocen los parámetros $\boldsymbol{\varphi} = (\phi_1, \dots, \phi_p)'$, θ_0 , $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$, pero lo único que conocemos es $\mathbf{W} = (W_{d+p+1}, W_{d+p+2}, \dots, W_N)'$ y lo que se desconoce es $\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta}$ y σ_ε^2 , por ello se dice que (1.42) es una función de verosimilitud de los parámetros, que depende del vector de observaciones \mathbf{W} , esto es

$$L(\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta}, \sigma_\varepsilon^2 | \mathbf{W}) = (2\pi\sigma_\varepsilon^2)^{-n/2} \exp\left\{-S(\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta}) / 2\sigma_\varepsilon^2\right\}$$

(2.45)

donde

$$\begin{aligned} S(\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta}) &= \sum_{t=1}^n \varepsilon_t^2 \\ &= \sum_{t=1}^n (W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q})^2 \end{aligned}$$

(2.46)

Box - Jenkins proponen que los parámetros sean escogidos de acuerdo al método de máxima verosimilitud (*MMV*). Si se quieren encontrar los estimadores por (*MMV*), la función de verosimilitud dada en (2.44) debe maximizarse con respecto a los parámetros, para obtener los estimadores máximo verosímiles (*EMV*) de los coeficientes del modelo dado en (2.40).

Los estimadores que se obtienen mediante este método tienen muchas propiedades estadísticas importantes. También la función de máxima verosimilitud (de un modelo *ARIMA* correcto) del cuál se obtienen los estimadores de máxima verosimilitud (*EMV*) refleja toda la información de los parámetros contenida en los datos.

Para maximizar $L(\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta}, \sigma_\varepsilon^2 | \mathbf{W})$, una manera de hacerlo es encontrar los valores $\hat{\boldsymbol{\varphi}}, \hat{\theta}_0, \hat{\boldsymbol{\theta}}$ que minimicen $S(\boldsymbol{\varphi}, \theta_0, \boldsymbol{\theta})$ y posteriormente determinar el estimador de σ_ε^2 .

Una vez que se conocen los valores $\hat{\boldsymbol{\varphi}}, \hat{\theta}_0, \hat{\boldsymbol{\theta}}$ que minimizan $S(\hat{\boldsymbol{\varphi}}, \hat{\theta}_0, \hat{\boldsymbol{\theta}})$, entonces se procede a maximizar la función de verosimilitud con respecto a σ_ε^2 .

o, equivalentemente, se maximiza al logaritmo de la función de verosimilitud, esto con el fin de encontrar el estimador de σ_ε^2

$$\begin{aligned} \ell(\sigma_\varepsilon^2 | \mathbf{W}; \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}}) &= \ln[L(\sigma_\varepsilon^2 | \mathbf{W}; \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}})] \\ &= -\frac{n}{2}[\ln(2\pi) - \ln(\sigma_\varepsilon^2)] - \frac{1}{2\sigma_\varepsilon^2} \mathbf{S}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}}). \end{aligned} \quad (2.47)$$

Para lograr dicha maximización se considera la ecuación máximo verosímil

$$\left. \frac{\partial \ell}{\partial \sigma_\varepsilon^2} \right|_{\sigma_\varepsilon^2 = \hat{\sigma}_\varepsilon^2} = -\frac{n}{2\hat{\sigma}_\varepsilon^2} + \frac{\mathbf{S}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}})}{2\hat{\sigma}_\varepsilon^2} = 0, \quad (2.48)$$

que produce

$$\hat{\sigma}_\varepsilon^2 = \frac{\mathbf{S}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}})}{n}. \quad (2.49)$$

que produce un máximo de $\ell(\sigma_\varepsilon^2 | \mathbf{W}; \boldsymbol{\phi}, \boldsymbol{\theta}_0, \boldsymbol{\theta})$; de esta manera, podría usarse (2.49) como estimador de $\hat{\sigma}_\varepsilon^2$, pero en la práctica prefiere utilizarse al estimador insesgado (1.4.41) que se denotará igualmente por $\hat{\sigma}_\varepsilon^2$, y que considera la corrección por grados de libertad usados para estimar los parámetros del modelo

$$\hat{\sigma}_\varepsilon^2 = \frac{\mathbf{S}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}_0, \hat{\boldsymbol{\theta}})}{n - d - p - q - 1}. \quad (2.50)$$

Estimación no lineal de parámetros

Muchas veces se encuentran dificultades al tratar de minimizar $\mathbf{S}(\boldsymbol{\phi}, \boldsymbol{\theta}_0, \boldsymbol{\theta})$, ya que se encuentran ecuaciones no lineales de los parámetros en el caso de los modelos $MA(q)$.

Sabemos que un modelo $ARMA(1,1)$, esta dado por

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (2.51)$$

expresando (2.51) en términos de ε_t se tiene

$$\begin{aligned} \varepsilon_t &= \tilde{Z}_t - \phi_1 Z_{t-1} + \theta_1 \varepsilon_{t-1} \\ &= \tilde{Z}_t - \phi_1 Z_{t-1} + \theta_1 (\tilde{Z}_{t-1} - \phi_1 Z_{t-2} + \theta_1 \varepsilon_{t-2}) \\ &= \tilde{Z}_t - (\phi_1 - \theta_1) \tilde{Z}_{t-1} + \phi_1 \theta_1 \tilde{Z}_{t-2} + \theta_1^2 \varepsilon_{t-2} \\ &\quad \vdots \end{aligned} \quad (2.52)$$

Dicha ecuación no es lineal en términos de los parámetros. Por lo tanto para estimar los parámetros de un proceso $ARMA(p, q)$, se usan mínimos cuadrados no lineales. Para el caso de los modelos $AR(p)$ que no presentan el problema de las ecuaciones no lineales como las dadas en (2.52), la estimación de parámetros por $MCNL$ se reduce al método de estimación de parámetros de mínimos cuadrados ordinarios.

El procedimiento de estimación de parámetros mediante $MCNL$ involucra una técnica de búsqueda iterativa.

Uno de los métodos más usados de $MCNL$ para estimar los modelos $ARIMA$ es el método de Marquardt.

Aquí se expresan las ideas básicas asociadas con dicho método. Este método combina dos métodos numéricos para la solución de ecuaciones lineales: el método de Newton y el método del gradiente.

La ventaja práctica del método de Newton es que este tiende a converger rápidamente a los estimadores de mínimos cuadrados, si es que hay convergencia, la desventaja es que puede no converger. La ventaja práctica del método del gradiente es que en teoría va a converger a los estimadores de mínimos cuadrados, aunque esta puede ser muy lenta.

El método de Marquardt combina lo mejor de estos dos métodos, por lo tanto este método en la mayoría de los casos, cuando se tienen unos valores iniciales adecuados, no sólo converge a los estimadores de mínimos cuadrados sino que lo hace de una manera relativamente rápida. Una forma de obtener valores iniciales adecuados son los estimadores que se obtienen de las ecuaciones Yule-Walker para $\phi_1, \dots, \phi_p, \theta_0, \dots, \theta_q$.¹⁷

A continuación se numeran las etapas básicas de este método:

1. En esta etapa se escogen los $p+q=k$ valores iniciales $\beta^* = (\beta_{1,0}, \beta_{2,0}, \dots, \beta_{k,0})$ del vector de coeficientes $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$, que van a ser estimados.
2. Se calculan la suma de los cuadrados de los residuales SCR_0 asociados con los valores iniciales dados en el paso 1, haciendo uso de (2.44).

Se expande ε_t en series de Taylor alrededor de los valores iniciales $\beta_0 = (\beta_{1,0}, \beta_{2,0}, \dots, \beta_{k,0})$ y sólo se toman los términos lineales, con esto de se tiene

$$[\varepsilon_t] = [\varepsilon_{t,0}] - \sum_{i=1}^k (\beta_i - \beta_{k,0}) X_{t,i}$$

¹⁷ Ver Guerrero, 2003, pag. 130.

donde $[\varepsilon_{t,0}] = [\varepsilon_t | \mathbf{W}, \boldsymbol{\beta}_0]$

$$y \quad X_{t,i} = -\frac{\partial[\varepsilon_t]}{\partial\beta_i} \Big|_{\beta=\beta_0}$$

3. En este paso calculan las derivadas $X_{t,i}$ que se necesitan para el método de Gauss-Newton, esto se hace mediante métodos numéricos.

4. Usando las derivadas obtenidas en el paso 3, se forma un sistema de ecuaciones que son aproximaciones lineales a la relación no lineal entre los residuales ε_t y $\boldsymbol{\varphi}, \boldsymbol{\theta}$ elementos de $\boldsymbol{\beta}_0$. El sistema que se obtiene es el siguiente

$$[\varepsilon_0] = \mathbf{X}(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0) + [\varepsilon_t]$$

5. Se resuelve el sistema de ecuaciones lineales obtenido en el paso anterior. Las correcciones $\mathbf{h} = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_0$, que minimizan $\mathbf{S}(\boldsymbol{\beta}) = \mathbf{S}(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [\varepsilon][\varepsilon]'$, son obtenidas mediante mínimos cuadrados lineales, de donde se obtienen los nuevos estimadores $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_0 + \mathbf{h}$.

Como los estimadores nuevos $\boldsymbol{\beta}_1$ fueron obtenidos solamente de aproximaciones lineales de las ecuaciones no lineales, no necesariamente tienen una SCR menor que SCR_0 .

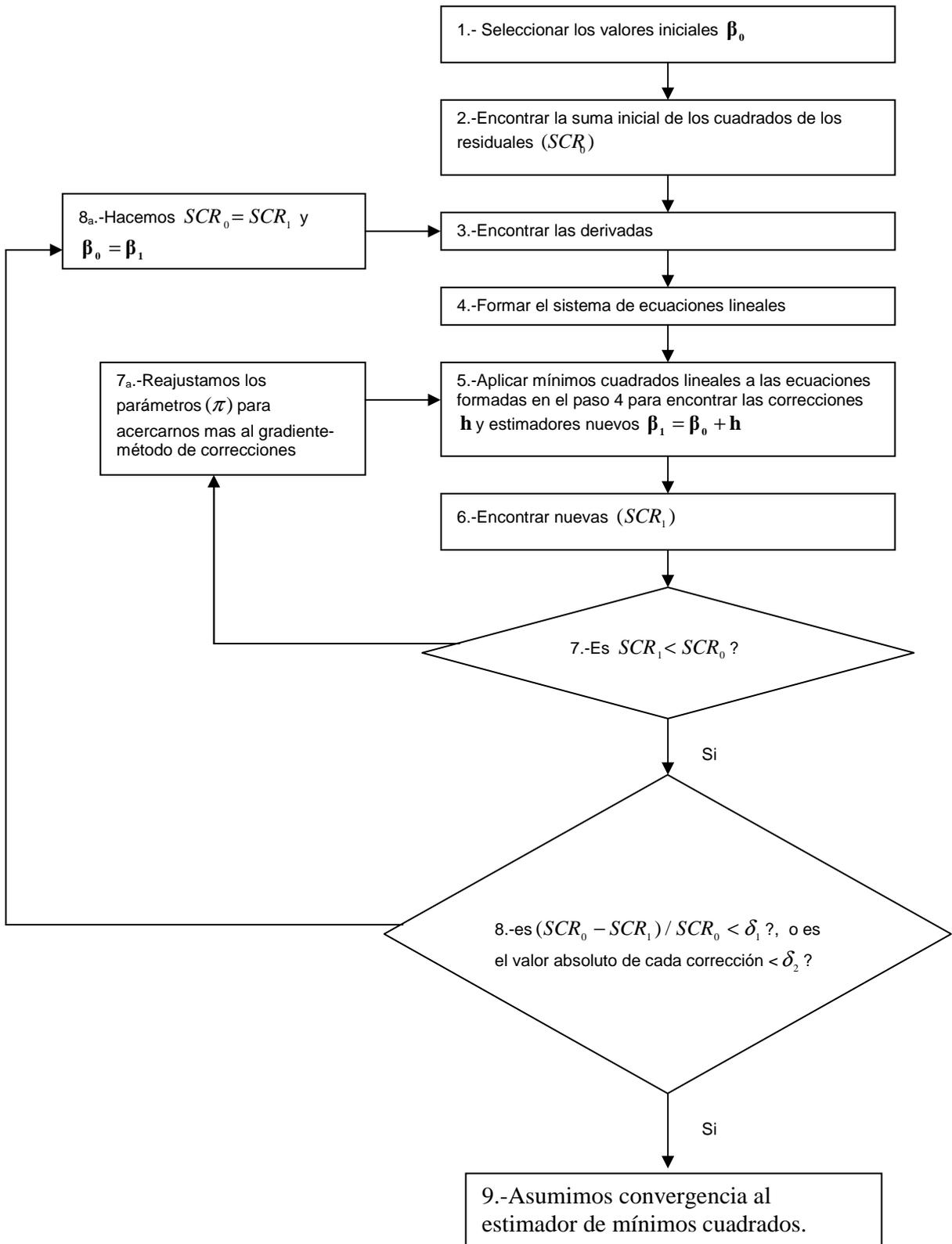
6. Insertar los nuevos estimadores en el modelo.

7. Ver si SCR es menor con los estimadores $\boldsymbol{\beta}_1$ que con los estimadores previos $\boldsymbol{\beta}_0$. Es decir comparar SCR_1 con SCR_0 . Si $SCR_1 < SCR_0$. Se hace una prueba para ver si el método a logrado converger a la mínima SCR. La reducción relativa en SCR puede ser comparada con un parámetro de convergencia δ_1 , o los valores absolutos de las correcciones en \mathbf{h} pueden ser probadas nuevamente mediante algún parámetro de convergencia δ_2 . Si la reducción relativa de SCR es mas pequeña que δ_1 , o alternativamente, si el valor absoluto de las correcciones en \mathbf{h} son menores que δ_2 , entonces se asume que ha habido convergencia, esto es, se asume que los estimadores en $\boldsymbol{\beta}_1$ son los estimadores de mínimos cuadrados. Si se excede los parámetros δ_1 o δ_2 entonces regresamos al paso 3, para encontrar nuevas derivadas y después regresamos al paso 8^a haciendo SCR_0 igual a la nueva SCR_1 y $\boldsymbol{\beta}_0$ igual a los nuevos estimadores $\boldsymbol{\beta}_1$. Es posible que en el paso 7, no obtengamos una SCR menor, entonces usamos el método del gradiente. Si en el paso 7, $SCR_1 > SCR_0$, entonces un parámetro π es incrementado en una cantidad predeterminada y las ecuaciones lineales, que depende de π , son modificadas y nuevas correcciones son obtenidas. A medida que π aumenta, las correcciones se acercan a las correcciones del método del gradiente, esto

significa que los valores absolutos de las correcciones van a ser cada vez más pequeñas, pero son más probables de producir una SCR más reducida.

En la página siguiente se muestra un esquema de los pasos del método de Marquardt.

Algoritmo de Marquardt



1.2.4.5. Función de pronósticos para modelos estacionarios

La parte de pronósticos que a continuación se expone, esta basada en Wei (1970) y Guerrero (2003).

Supongamos que tenemos una serie original $\{Z_t\}$ con n observaciones, que no es estacionaria, pero después de aplicar d veces el operador diferencia a $\{Z_t\}$, obtenemos la serie $\{W_t\}$ estacionaria y con media cero. Supongamos que tenemos un modelo *ARMA* para $\{W_t\}$, donde

$$\phi(B)W_t = \theta(B)\varepsilon_t. \quad (2.53)$$

Como el modelo es estacionario, lo podemos representar en términos de ε_t 's, entonces

$$\begin{aligned} W_t &= \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \dots \\ &= \psi(B)\varepsilon_t, \end{aligned}$$

donde

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j = \frac{\theta(B)}{\phi(B)},$$

con $\psi_0 = 1$.

Para $t = n+l$ tenemos

$$W_{n+l} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n+l-j}, \quad (2.54)$$

que también se puede expresar como

$$\begin{aligned} W_{t+l} &= (\varepsilon_{t+l} + \psi_1\varepsilon_{t+l-1} + \dots + \psi_{l-1}\varepsilon_{t+1}) + (\psi_l\varepsilon_t + \psi_{l+1}\varepsilon_{t-1} + \dots) \\ W_{t+l} &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \sum_{j=l}^{\infty} \psi_j \varepsilon_{t+l-j}, \end{aligned} \quad (2.55)$$

con $\psi_0 = 1$

donde la primera suma corresponde a la información desconocida al tiempo t y la segunda suma corresponde a la información conocida al tiempo t .

Definiremos a

$\tilde{W}_t(l)$ como el pronóstico de la observación W_{t+l} a partir del origen t .

$\hat{W}_t(l)$ como el pronóstico óptimo de la observación W_{t+l} a partir del origen t .

El criterio que se utiliza para determinar cuál es pronóstico óptimo, es el error cuadrático medio mínimo, es decir $\hat{W}_t(l)$ debe de satisfacer la condición siguiente

$$E_t[W_{t+l} - \hat{W}_t(l)] = \min_{\hat{W}_t(l)} E_t[W_{t+l} - \hat{W}_t(l)]^2, \quad (2.56)$$

donde E_t es la esperanza condicional, dada toda la información hasta el momento t , es decir

$$E_t[W_{t+l} - \hat{W}_t(l)]^2 = E[(W_{t+l} - \tilde{W}_t(l))^2 | Z_t, Z_{t-1}, \dots].$$

Como el pronóstico $\tilde{W}_t(l)$ de W_{t+l} es de la forma

$$\begin{aligned} \tilde{W}_t(l) &= \psi^*_l \varepsilon_t + \psi^*_{l+1} \varepsilon_{t-1} + \psi^*_{l+2} \varepsilon_{t-2} + \dots \\ &= \sum_{j=l}^{\infty} \psi^*_j \varepsilon_{t+l-j} \end{aligned} \quad (2.57)$$

se deben determinar los coeficientes ψ^*_j para que $\tilde{W}_t(l)$ sea mínimo. De (2.56) y (2.57) se tiene que el error cuadrático medio para los pronósticos es

$$\begin{aligned} E_t[W_{t+l} - \tilde{W}_t(l)]^2 &= E_t \left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \sum_{j=l}^{\infty} \psi_j \varepsilon_{t+l-j} - \sum_{j=l}^{\infty} \psi^*_j \varepsilon_{t+l-j} \right]^2 \\ &= E_t \left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \sum_{j=l}^{\infty} (\psi_j - \psi^*_j) \varepsilon_{t+l-j} \right]^2 \\ &= E_t \left[\left(\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \right)^2 + 2 \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \sum_{j=l}^{\infty} (\psi_j - \psi^*_j) \varepsilon_{t+l-j} + \left(\sum_{j=l}^{\infty} (\psi_j - \psi^*_j) \varepsilon_{t+l-j} \right)^2 \right] \\ &= E_t \left[\left(\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \right)^2 + 2 E_t \left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \sum_{j=l}^{\infty} (\psi_j - \psi^*_j) \varepsilon_{t+l-j} \right] + E_t \left[\sum_{j=l}^{\infty} (\psi_j - \psi^*_j) \varepsilon_{t+l-j} \right]^2 \right] \end{aligned}$$

como ε_t es un ruido blanco se tiene que $E(\varepsilon_{t+i} \varepsilon_{t+j}) = 0$ para $i \neq j$, y entonces

$$\begin{aligned} &= E_t \left(\sum_{j=0}^{l-1} \psi_j^2 \varepsilon_{t+l-j}^2 + \sum_{i \neq j} \psi_i \psi_j \varepsilon_i \varepsilon_j \right) + E_t \left(\sum_{j=l}^{\infty} (\psi_j - \psi^*_j)^2 \varepsilon_{t+l-j}^2 + \sum_{i \neq j} \psi_i \psi_j \varepsilon_i \varepsilon_j \right) \\ &= E_t \left(\sum_{j=0}^{l-1} \psi_j^2 \varepsilon_{t+l-j}^2 \right) + E_t \left(\sum_{i \neq j} \psi_i \psi_j \varepsilon_i \varepsilon_j \right) + E_t \left(\sum_{j=l}^{\infty} (\psi_j - \psi^*_j)^2 \varepsilon_{t+l-j}^2 \right) + E_t \left(\sum_{i \neq j} \psi_i \psi_j \varepsilon_i \varepsilon_j \right) \end{aligned}$$

$$\therefore E_t[W_{t+l} - \tilde{W}_t(l)]^2 = \sum_{j=0}^{l-1} \psi_j^2 \sigma_\varepsilon^2 + \sum_{j=0}^{\infty} [\psi_j - \psi_j^*]^2 \sigma_\varepsilon^2.$$

La ecuación anterior se minimiza si $\psi_{l-j}^* = \psi_{l-j}$, para $l, l+1, l+2, l+3, \dots$. Por lo tanto

$$\hat{W}_t(l) = \psi_l \varepsilon_t + \psi_{l+1} \varepsilon_{t-1} + \psi_{l+2} \varepsilon_{t-2} + \dots \quad (2.58)$$

Ahora de (2.58) y sabiendo que

$$E[\varepsilon_{n+j} | Z_n, Z_{n-1}, \dots] = \begin{cases} 0 & j > 0 \\ \varepsilon_{n+j} & j \leq 0 \end{cases} \quad (2.59)$$

se tiene

$$E[W_{n+l} | Z_n, Z_{n-1}, \dots] = \psi_l \varepsilon_n + \psi_{l+1} \varepsilon_{n-1} + \psi_{l+2} \varepsilon_{n-2} + \dots$$

Por lo tanto el error en media cuadrático mínimo está dado por

$$\hat{W}_t(l) = E_t(W_{t+l}) \quad (2.60)$$

El error del pronóstico (2.56) con origen en t , está dado por

$$e_t(l) = W_{t+l} - \hat{W}_t(l),$$

que de (2.55) y (2.58) se tiene

$$e_t(l) = \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \quad (2.61)$$

Específicamente, el error que se obtiene al pronosticar un periodo hacia adelante, se obtiene al hacer $l=1$ en la ecuación anterior, es decir

$$e_t(l) = W_{t+l} - \hat{W}_t(l) = \varepsilon_{t+1}. \quad (2.62)$$

Dado que ε_t es un ruido blanco se tiene

$$E_t[e_t(l)] = 0 \quad \text{y} \quad \text{Var}_t[e_t(l)] = \sum_{j=0}^{l-1} \psi_j^2 \sigma_\varepsilon^2 \quad (2.63)$$

Como se desea obtener pronósticos de la serie $\{W_t\}$, si conoce el modelo $\phi(B)W_t = \theta(B)\varepsilon_t$, entonces

$$\begin{aligned}
\hat{W}_t(l) &= E_t(W_{t+l}) \\
&= E_t(\phi_1 W_{t+l-1} + \dots + \phi_p W_{t+l-p} + \varepsilon_{t+l} - \theta_1 \varepsilon_{t+l-1} - \dots - \theta_q \varepsilon_{t+l-q}) \\
&= \phi_1 E_t(W_{t+l-1}) + \dots + \phi_p E_t(W_{t+l-p}) + E_t(\varepsilon_{t+l}) - \theta_1 E_t(\varepsilon_{t+l-1}) - \dots - \theta_q E_t(\varepsilon_{t+l-q})
\end{aligned}$$

donde

$$E_t(W_{t+l-j}) = \begin{cases} W_{t+l-j} & \text{si } j \geq l \\ \hat{W}_t(h-j) & \text{si } j < l \end{cases} \quad (2.64)$$

de donde usando (1.56) y (1.59) se tiene

$$E_t(\varepsilon_{t+l-j}) = \begin{cases} W_{t+l-j} - \hat{W}_{t+l-j-1}(1) & \text{si } j \geq h \\ 0 & \text{si } j < h \end{cases} \quad (2.65)$$

que es la expresión que se usa para calcular los pronósticos de forma recursiva.

Como ejemplo¹⁸ supongamos que el comportamiento de W_t lo describe el siguiente proceso

$$(1 - 0.6B)W_t = (1 + 0.2B)\varepsilon_t$$

o equivalentemente

$$W_t = 0.6W_{t-1} + \varepsilon_t + 0.2\varepsilon_{t-1}$$

Entonces los pronósticos $\hat{W}_t(h)$ se obtienen como

$$\begin{aligned}
\hat{W}_t &= E_t(W_{t+1}) \\
&= 0.6 E_t(W_t) + E_t(\varepsilon_{t+1}) + 0.2 E_t(\varepsilon_t) \\
&= 0.6 E_t(W_t) + 0.2 [W_t + \hat{W}_{t-1}(1)] \\
\hat{W}_t(2) &= 0.6 E_t(W_{t+1}) + E_t(\varepsilon_{t+2}) + 0.2 E_t(\varepsilon_{t+1}) \\
&= 0.6 \hat{W}_t(1)
\end{aligned}$$

En general se tiene que

$$\hat{W}_t(h) = 0.6 \hat{W}_t(h-1) \quad \text{para } h \geq 2$$

Es decir, la generación de pronósticos se vuelve un proceso recursivo. Para hacer los cálculos se supone que $\hat{W}_0(1) = W_1$, es decir, se hace $\varepsilon_1 = 0$.

¹⁸ Los ejemplos para esta parte de pronósticos, fueron tomados de Guerrero Guzmán, 2003.

Así para obtener los pronósticos de nuestro ejemplo para $\hat{W}_t(h)$ para $h \geq 2$, se tienen que calcular los siguientes valores

$$\begin{aligned}\hat{W}_1(1) &= 0.6W_1 \\ \hat{W}_2(1) &= 0.8W_2 - 0.2\hat{W}_1(1) \\ &\dots \\ \hat{W}_t(1) &= 0.8W_t - 0.2\hat{W}_{t-1}(1)\end{aligned}$$

Una vez que se tienen las igualdades anteriores, se utiliza una computadora para encontrar los valores de los pronósticos.

1.2.4.6. Función de pronósticos para modelos no estacionarios

En la práctica la mayoría de las series no son estacionarias, sin embargo hay una relación entre estas y las series estacionarias.

Si \tilde{Z}_t es una serie no estacionaria, pero es integrada de orden d , entonces

$$W_t = \nabla^d T(Z_t)$$

es una serie estacionaria, donde T es alguna transformación.

Para simplificar los cálculos supóngase que $d = 1$. Como W_t es estacionaria, entonces los pronósticos óptimos se obtienen con lo visto en la sección anterior y los pronósticos óptimos de la serie $\{T(Z_t)\}$ se obtienen de la relación

$$E_t(W_{t+h}) = E_t[T(Z_{t+h}) - E_t[T(Z_{t+h-1})]]$$

y entonces $\hat{T}(Z_t)(h) = E_t[T(Z_{t+h-1})]$ está dado por

$$\hat{T}(Z_t)(h) = \begin{cases} T(Z_t) + \hat{W}_t(1) & \text{si } h = 1 \\ T(Z_t)(h-1) + \hat{W}_t(h) & \text{si } h \geq 2 \end{cases}$$

En caso de que el grado de diferenciación sea mayor que $d = 1$, los cambios requeridos en la ecuación anterior son análogos.

Si también decimos que el pronóstico óptimo de $T(Z_{t+h})$ sea su esperanza condicional al tiempo t , es decir, si el modelo para $T(Z_t)$ se puede escribir como

$$\varphi(B)T(Z_t) = \theta(B)\varepsilon_t \quad \text{con} \quad \varphi(B) = \nabla^d \phi(B), \quad (2.66)$$

entonces

$$\begin{aligned}
\hat{T}(Z_t)(l) &= E_t[T(Z_{t+l})] \\
&= E_t[\varphi_1 T(Z_{t+l-1}) + \dots + \varphi_{p+d} T(Z_{t+l-p-d}) \\
&\quad + \varepsilon_{t+l} - \theta_1 \varepsilon_{t+l-1} - \dots - \theta_q \varepsilon_{t+l-q}] \\
&= \varphi_1 E_t[T(Z_{t+l-1})] + \dots + \varphi_{p+d} E_t[T(Z_{t+l-p-d})] \\
&\quad + E_t(\varepsilon_{t+l}) - \theta_1 E_t(\varepsilon_{t+l-1}) - \dots - \theta_q E_t(\varepsilon_{t+l-q}),
\end{aligned}$$

donde

$$E_t[T(Z_{t+l-j})] = \begin{cases} T(Z_{t+l-j}) & \text{si } j \geq l \\ \hat{T}(Z_t)(l-j) & \text{si } j < l \end{cases}, \quad (2.67)$$

que es la expresión equivalente a la (2.64), mientras que la expresión equivalente a (2.65) es

$$E_t(\varepsilon_{t+l-j}) = \begin{cases} T(Z_{t+l-j}) - \hat{T}(Z_{t+l-j-1})(l) & \text{si } j \geq l \\ 0 & \text{si } j < l \end{cases} \quad (2.68)$$

En resumen si se tiene un modelo estacionario se utiliza (2.64) para calcular los pronósticos y se tiene un modelo integrado de orden d , se utiliza (2.67) para calcular los pronósticos.

Intervalos de predicción

En la parte de las aplicaciones prácticas, es importante obtener límites de probabilidad para los pronósticos que se realicen; dichos límites pueden calcularse con la expresión (2.63), la cuál aunada al supuesto de que $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ para toda t , implica que $e_t(l) \{Z_t, Z_{t-1}, \dots\} \sim N(0, \text{Var}[e_t(l)])$ y que

$$[T(Z_{t+l}) - \hat{T}(Z_t)(l)] \{Z_t, Z_{t-1}, \dots\} \sim N\left(0, \text{Var}[e_t(l)]\right).$$

Por lo tanto los límites $100(1-\alpha)\%$ de probabilidad para $T(Z_{t+l})$, condicionados en el conocimiento de las observaciones Z_t, Z_{t-1}, \dots , son

$$\hat{T}(Z_t)(l) \pm Z_{\frac{\alpha}{2}} \left(\sum_{j=0}^{l-1} \psi_j^2 \right)^{\frac{1}{2}} \sigma_\varepsilon, \quad (2.69)$$

en donde $Z_{\frac{\alpha}{2}}$ es el punto porcentual, de modo que $P\left(Z > Z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$ si $Z \sim N(0,1)$.

El intervalo de predicción definido por (2.69) se estima al sustituir a $\psi_1, \psi_2, \dots, \psi_{l-1}$ y σ_ε por sus respectivas estimaciones y la interpretación asociada con el intervalo resultante es que el valor observado de $T(Z_{t+l})$ será cubierto

por el intervalo con una probabilidad de $100(1 - \alpha)\%$. Esto es válido para una l en particular, pero no necesariamente para todas las l que se consideren simultáneamente, es decir la serie pronosticada no tiene porque encontrarse dentro de la banda determinada al variar l

1.3. Metodología de Box-Jenkins

La metodología de Box-Jenkins sirve para identificar, estimar, diagnosticar y utilizar modelos *ARIMA* para series de tiempo estacionarias. Una característica muy importante que deben de tener las series de tiempo para ser analizadas con esta metodología es la estacionariedad en diferencias. Como ya se vio esta se logra aplicando el operador diferencia a la serie, si esta tiene una raíz unitaria.

Los pasos básicos de la metodología de Box-Jenkins, son los siguientes:

1.- Diferenciar para alcanzar la estacionariedad.

Empíricamente para saber si una serie es estacionaria se puede observar su correlograma, si este decae rápidamente a cero en forma exponencial u oscilatoria entonces la serie es estacionaria, de lo contrario se usa el operador diferencia para hacer la serie estacionaria. También se usa la prueba de Dickey-Fuller para saber si una serie es estacionaria, o si tiene una raíz unitaria y se puede hacer estacionaria aplicando el operador diferencia.

2.- Examinar el correlograma para identificar los órdenes adecuados.

A continuación se muestra una tabla con las propiedades de las funciones de autocorrelacion (FAC) y autocorrelacion parcial (FACP), para los modelos *AR*, *MA* y *ARMA*, las cuales nos va a permitir identificar el modelo.

	AR(p)	Ma(q)	ARMA(p,q)
FAC	Decae a cero	Se hace cero después del rezago q	Decae a cero después del rezago q
PACF	Se hace cero después del rezago p	Decae a cero	Decae a cero del rezago p

Cuadro 1. Comportamientos de las FAC y FACP de los modelos econométricos.

3.- Este paso consiste en la estimación del modelo identificado en el paso anterior.

4.- En este paso se verifica si el modelo es estadísticamente adecuado.

Cuando el modelo es inadecuado se regresa al paso 2. Una de las pruebas estadísticas más importantes para ver si un modelo ARIMA es adecuado tiene que ver con la independencia de los choques aleatorios. En la práctica no podemos observar las ε_t , sin embargo tenemos los estimadores de los residuales $\hat{\varepsilon}_t$, que fueron calculados en el paso 3, y sobre estos se hacen pruebas de hipótesis para ver si son independientes.

Una herramienta analítica básica para el diagnóstico es la función de autocorrelación de los residuales estimados $\hat{\varepsilon}_t$. Los coeficientes de autocorrelación se obtienen de la siguiente manera

$$r_k(\hat{\varepsilon}_t) = \frac{\sum_{t=1}^{n-k} (\hat{\varepsilon}_t - \bar{\varepsilon})(\hat{\varepsilon}_{t+k} - \bar{\varepsilon})}{\sum_{t=1}^n (\hat{\varepsilon}_t - \bar{\varepsilon})^2}$$

Si el modelo estimado es adecuado, entonces los choques aleatorios $\{\varepsilon_t\}$ no deben de estar correlacionados, ya que esa es una suposición que se usó en todos los modelos.

Por lo tanto, la función de autocorrelación de los residuales de unos modelos adecuados, debe de tener coeficientes que sean estadísticamente cero.

Una prueba de hipótesis que se usa para saber si los residuales estimados son independientes, es la de Ljung-Box.

La hipótesis nula es:

$$H_0: \rho_1(\varepsilon) = \rho_2(\varepsilon) = \dots = \rho_k(\varepsilon) = 0$$

donde el estadístico de prueba es

$$Q = n(n+2) \sum_{k=1}^k (n-k)r_k(\hat{\varepsilon}_t)$$

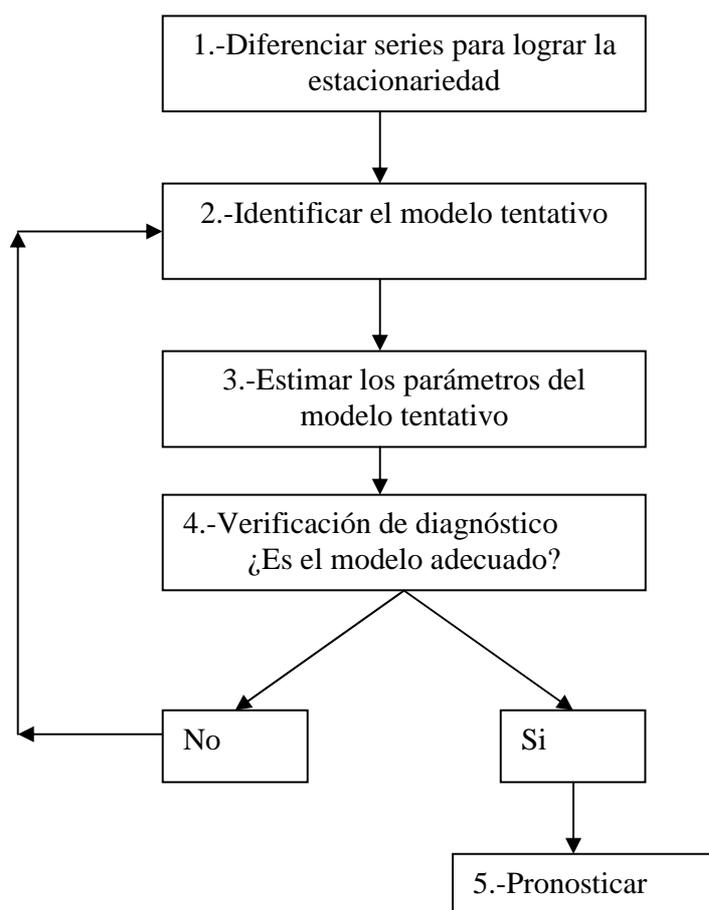
con n el número de observaciones usadas para estimar el modelo.

El estadístico Q se distribuye aproximadamente como una Chi-cuadrada con $(k-m)$ grados de libertad, donde m es el número de parámetros estimados en un modelo ARIMA.

5.- Este paso consiste en pronosticar.

A continuación se presentan estos pasos en forma esquemática en la siguiente figura

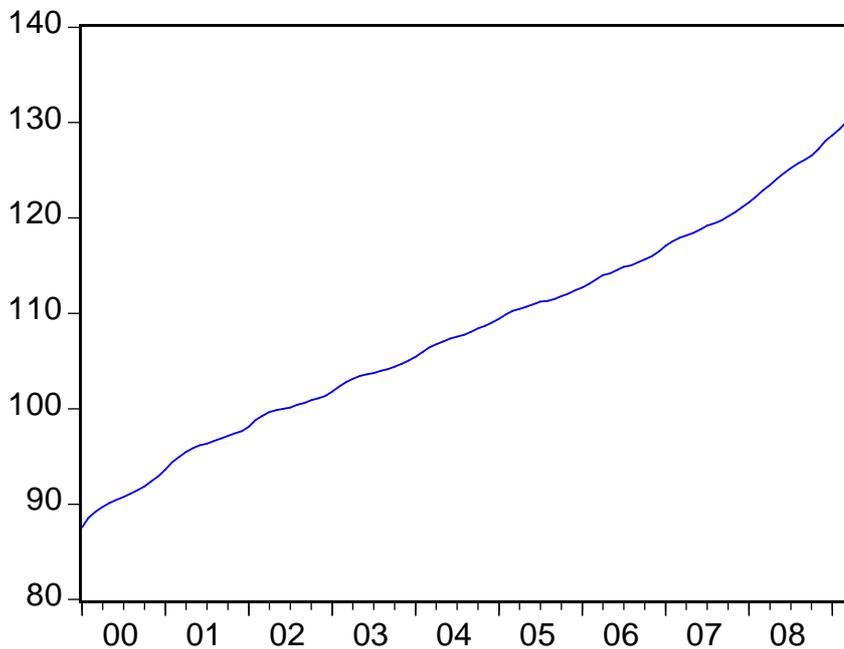
Metodología de Box-Jenkins



Capítulo II. Un modelo econométrico del índice de inflación subyacente de México (1990-2008)

En este capítulo se especifica un modelo econométrico del índice de inflación subyacente (IIS) de México tomando como base para ello la teoría ARIMA y la metodología de Box-Jenkins expuestas en el capítulo I. Según esta metodología, para identificar un modelo econométrico del IIS, se requiere como punto inicial que la serie sea estacionaria en diferencia.

La gráfica de los datos en niveles del IIS se presenta en la gráfica 1. Si trazamos una línea horizontal imaginaria paralela al eje del tiempo, la cual podemos identificar con la media, observamos que la serie IIS se aleja de esta línea, por lo que podemos decir que la serie no tiene media constante. Por otro lado una vez que hemos trazado una línea horizontal que identifiquemos con la media, si trazamos imaginariamente una banda centrada en la media no podemos encerrar la serie en esta banda, y, debido a esto, no podemos decir que la serie tiene varianza constante. De acuerdo con la **definición 3** de estacionariedad dada en el capítulo I, podemos conjeturar que la serie IIS no es estacionaria, ya que la media y la varianza de la serie no son constantes y dependen del tiempo.



Gráfica 1. Índice de la Inflación Subyacente en niveles

El análisis gráfico anterior, muestra que el nivel de la serie IIS no es estacionario. Sin embargo, esto no puede ser concluyente pues la conclusión es heurística. Para corroborar estadísticamente esta conclusión, se aplica una

prueba de raíz unitaria. La prueba de Dickey-Fuller Aumentada, como se estableció en la sección 1.2.4. del capítulo I, es una de estas pruebas y consiste en probar la hipótesis nula, de que la serie IIS tiene raíz unitaria, es decir, $H_0 : \rho = 1$ contra la alternativa de que la serie IIS no tiene raíz unitaria, es decir $H_a : \rho < 1$. Esta prueba se realiza al nivel de significancia convencional de 5%. Los valores críticos para no aceptar o no rechazar la existencia de raíz unitaria son los de MacKinnon. La hipótesis nula no se acepta si el valor absoluto del estadístico de Dickey-Fuller Aumentado (ADF) es mayor al valor absoluto del valor crítico de Mackinnon.

En el cuadro 2 y 3, se muestran los resultados de las pruebas de raíz unitaria de Dikey-Fuller Aumentada y de Phillips-Perron. Como se puede ver en estos cuadros, ambas pruebas no rechazan la existencia de una raíz unitaria en la serie IIS a un nivel de significancia de 5%. Se puede decir entonces que la serie IIS no estacionaria en su nivel.

Null Hypothesis: IIS has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 4 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.305057	0.9984
Test critical values: 1% level	-4.046072	
5% level	-3.452358	
10% level	-3.151673	

*MacKinnon (1996) one-sided p-values.

Cuadro 2. Prueba ADF para la serie IIS.

Null Hypothesis: IIS has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 4 (Fixed using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	0.110038	0.9971
Test critical values: 1% level	-4.042819	
5% level	-3.450807	
10% level	-3.150766	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.027180
HAC corrected variance (Bartlett kernel)	0.071934

Cuadro 3. Prueba P-P para la serie IIS.

A fin de investigar si para la serie IIS transformada, las pruebas ADF y de Phillips-Perron rechazan la presencia de una raíz unitaria, se aplica una transformación logarítmica a la serie IIS y se le aplican las pruebas de raíz unitaria. Los resultados de estas pruebas se presentan en los cuadros 4 y 5. Como se puede observar, la serie transformada log(IIS) también tiene una raíz unitaria y es también no estacionaria.

Null Hypothesis: LOG(IIS) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 4 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.153665	0.5100
Test critical values:		
1% level	-4.046072	
5% level	-3.452358	
10% level	-3.151673	

*MacKinnon (1996) one-sided p-values.

Cuadro 4. Prueba de dickey-Fuller Aumentada para la serie log(IIS).

Null Hypothesis: LOG(IIS) has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 4 (Fixed using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-3.055583	0.1223
Test critical values:		
1% level	-4.042819	
5% level	-3.450807	
10% level	-3.150766	

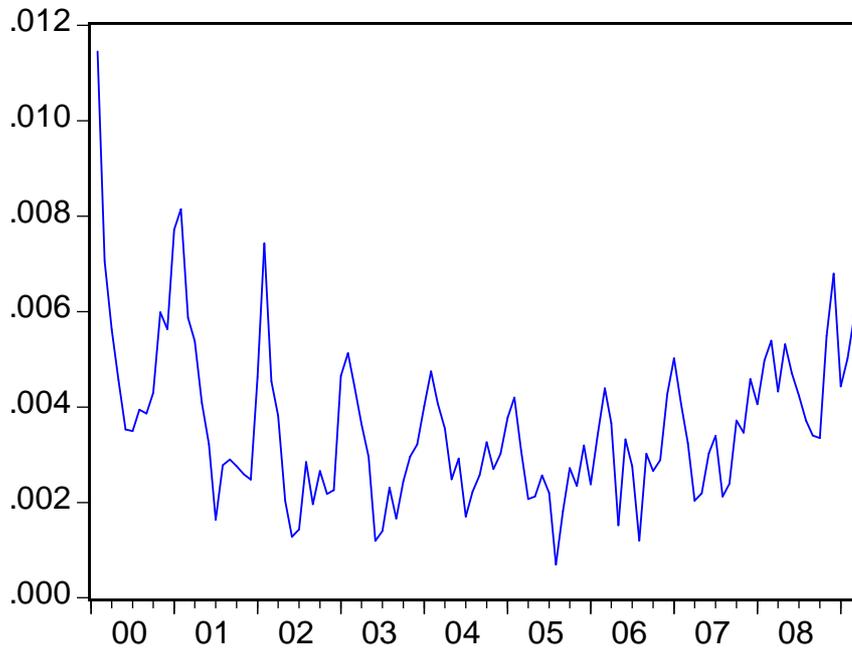
*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	2.37E-06
HAC corrected variance (Bartlett kernel)	6.64E-06

Cuadro 5. Prueba de Phillips-Perron para la serie log(IIS).

Dado que la metodología de Box-Jenkins (1970) para modelar series estacionarias, requiere que la serie a modelar sea estacionaria en diferencia, ahora se investiga si la primera diferencia de la serie log(IIS) tiene esta propiedad. Los cuadros 6 y 7 muestran los resultados de las pruebas ADF y Phillips-Perron de raíz unitaria para la primera diferencia de log(IIS). Como se puede ver en el cuadro, estas pruebas rechazan la hipótesis nula de existencia de raíz unitaria para la primera diferencia de la serie log(IIS). Es decir, la primera diferencia de la serie log(IIS) es una serie estacionaria. Según lo expuesto en la sección 1.2.4.1, del capítulo I, la serie log(IIS) es una serie

integrada de orden uno (I(1)). La grafica de esta serie se muestra en la gráfica 2.



Grafica 2. Primera diferencia de la serie Log(IIS).

Null Hypothesis: D(LOG(IIS)) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 4 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.867285	0.0168
Test critical values: 1% level	-4.046925	
5% level	-3.452764	
10% level	-3.151911	

*MacKinnon (1996) one-sided p-values.

Cuadro 6. Prueba ADF para la primera diferencia de la serie log(IIS).

Null Hypothesis: D(LOG(IIS)) has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 4 (Fixed using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-5.978390	0.0000
Test critical values: 1% level	-4.043609	
5% level	-3.451184	
10% level	-3.150986	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	1.00E-06
HAC corrected variance (Bartlett kernel)	1.27E-06

Cuadro 7. Prueba P-P para la primera diferencia de la serie log(IIS).

Siguiendo la metodología Box-Jenkins, una vez que tenemos una serie estacionaria en diferencia como lo es $\log(IIS)$, la siguiente etapa consiste en *identificar* el modelo ARIMA para esta serie estacionaria en diferencia. En este proceso de identificación se utilizan las funciones de autocorrelacion (FAC) y autocorrelacion parcial (FACP) presentadas en el cuadro 1.

Las FAC y FACP de la serie primera diferencia de $\log(IIS)$ se muestran en cuadro 8. Podemos observar que la FAC decae a cero, mientras que la FACP se hace cero después del rezago 11. De acuerdo con el cuadro 1, el comportamiento de las FAC y FACP muestran que un posible modelo para el $\log(IIS)$ es un modelo AR.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.645	0.645	47.470	0.000
		2	0.395	-0.037	65.404	0.000
		3	0.239	-0.002	72.036	0.000
		4	0.069	-0.130	72.593	0.000
		5	-0.012	0.002	72.611	0.000
		6	-0.014	0.047	72.635	0.000
		7	-0.031	-0.026	72.750	0.000
		8	0.051	0.139	73.064	0.000
		9	0.223	0.234	79.188	0.000
		10	0.306	0.081	90.783	0.000
		11	0.451	0.277	116.32	0.000
		12	0.521	0.146	150.74	0.000
		13	0.366	-0.134	167.87	0.000
		14	0.200	-0.087	173.04	0.000
		15	0.040	-0.120	173.25	0.000
		16	-0.112	-0.090	174.90	0.000
		17	-0.194	-0.106	179.93	0.000
		18	-0.191	-0.039	184.82	0.000
		19	-0.209	-0.120	190.77	0.000
		20	-0.133	-0.070	193.21	0.000
		21	-0.038	-0.108	193.41	0.000
		22	0.076	0.033	194.23	0.000
		23	0.253	0.162	203.35	0.000
		24	0.337	0.143	219.70	0.000
		25	0.184	-0.117	224.62	0.000
		26	0.067	0.067	225.29	0.000
		27	-0.095	-0.093	226.64	0.000
		28	-0.200	0.078	232.71	0.000
		29	-0.230	0.025	240.82	0.000
		30	-0.248	0.004	250.31	0.000
		31	-0.263	-0.054	261.17	0.000
		32	-0.143	0.062	264.43	0.000
		33	-0.072	-0.126	265.27	0.000
		34	0.010	-0.035	265.29	0.000
		35	0.147	-0.051	268.84	0.000
		36	0.168	-0.034	273.54	0.000

Cuadro 8. Función de autocorrelación y autocorrelación parcial del IIS en primera diferencia.

En el mismo cuadro 7, la Función de Autocorrelación Parcial empírica (Partial Correlation) de los rezagos 1, 9 y 11, son estadísticamente diferentes de cero. Por esta razón, el proceso de identificación comienza a partir del siguiente modelo

$$\begin{aligned}
 W_t = & \phi_1 W_{t-1} + \phi_2 W_{t-2} + \phi_3 W_{t-3} + \phi_4 W_{t-4} + \phi_5 W_{t-5} + \phi_6 W_{t-6} + \phi_7 W_{t-7} + \phi_8 W_{t-8} + \phi_9 W_{t-9} \\
 & + \phi_{10} W_{t-10} + \phi_{11} W_{t-11} + \varepsilon_t
 \end{aligned}
 \tag{3.1}$$

Identificado el modelo 3.1 para el IIS, la siguiente etapa en la metodología Box-Jenkins consiste en estimar los parámetros del modelo (3.1). Dado que el modelo identificado es un AR, el criterio de estimación de parámetros que se utiliza es el de mínimos cuadrados ordinarios, como se señalo la sección

1.2.2.4 del capítulo II. Los resultados de la estimación se muestran en el cuadro 9.

Dependent Variable: D(LOG(IIS))
 Method: Least Squares
 Date: 05/30/09 Time: 13:39
 Sample (adjusted): 2001M01 2009M04
 Included observations: 100 after adjustments
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003239	0.000417	7.763175	0.0000
AR(1)	0.674865	0.098416	6.857252	0.0000
AR(2)	-0.211061	0.121619	-1.735424	0.0862
AR(3)	0.179811	0.119682	1.502407	0.1366
AR(4)	-0.100941	0.122096	-0.826737	0.4106
AR(5)	-0.040376	0.124663	-0.323880	0.7468
AR(6)	0.101233	0.124350	0.814093	0.4178
AR(7)	-0.145775	0.125219	-1.164158	0.2475
AR(8)	-0.063367	0.125898	-0.503320	0.6160
AR(9)	0.198374	0.125566	1.579846	0.1177
AR(10)	-0.167830	0.125739	-1.334753	0.1854
AR(11)	0.353830	0.090403	3.913896	0.0002
R-squared	0.648045	Mean dependent var		0.003406
Adjusted R-squared	0.604050	S.D. dependent var		0.001434
S.E. of regression	0.000902	Akaike info criterion		-11.07140
Sum squared resid	7.16E-05	Schwarz criterion		-10.75878
Log likelihood	565.5698	F-statistic		14.73017
Durbin-Watson stat	2.173459	Prob(F-statistic)		0.000000
Inverted AR Roots	.95	.84+.47i	.84-.47i	.41-.80i
	.41+.80i	-.04-.85i	-.04+.85i	-.50+.78i
	-.50-.78i	-.85-.28i	-.85+.28i	

Cuadro 9. Estimación de parámetros del modelo (3.1)

Los parámetros estimados del modelo (3.1) corresponden a la segunda columna del recuadro superior del cuadro 8.

Con el fin de que el modelo identificado en (3.1) y estimado para la primera diferencia del log(IIS) esté correctamente especificado, es decir, para que se pueda utilizar para pronosticar a la serie log(IIS), la siguiente etapa en la metodología de Box-Jenkins consiste en hacer un diagnóstico del modelo (3.1). Tal diagnóstico, como se dijo en la metodología de B-J consiste en que el modelo debe cumplir con algunos supuestos estadísticos para que sea un modelo correctamente especificado. Estos supuestos son los siguientes. (1) Los coeficientes estimados del modelo (3.1) deber ser estadísticamente significativos (distintos de cero). (2) El modelo (3.1) debe cumplir con la condición de estacionariedad e invertibilidad. (3) los residuos estimados del

modelo (3.1) no deben estar autocorrelacionados. Los residuos del modelo (3.1) estimado deben tener una distribución normal.

La prueba de significancia estadística de los parámetros estimados del modelo (3.1) se hace utilizando el estadístico t de Student. En el cuadro 9, los valores de este estadístico corresponden a los valores de la cuarta columna del recuadro superior del cuadro 9. Como se puede observar en el cuadro, al nivel de significancia de 5%, los coeficientes estimados estadísticamente no significativos, son los que corresponden a los parámetros $\phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, \phi_8, \phi_9$ y ϕ_{10} del modelo (3.1). De esta manera, el supuesto de significancia estadística de todos los coeficientes estimados no se cumple para el modelo (3.1). En la metodología recursiva de Box-Jenkins este resultado lleva a eliminar del modelo estimado los parámetros estadísticamente no significativos. Hecho esto, el modelo inicial (3.1) reparametrizado resultante es el siguiente:

$$W_t = \phi_1 W_{t-1} + \phi_{12} W_{t-12} + \phi_{13} W_{t-13} + \varepsilon_t \quad (3.2)$$

Nuevamente, se estiman los parámetros de este modelo (3.2). Como el modelo propuesto es un AR, nuevamente la estimación de parámetros se lleva a cabo mediante el método de mínimos cuadrados ordinarios. Los resultados de la estimación se muestran en el cuadro 10.

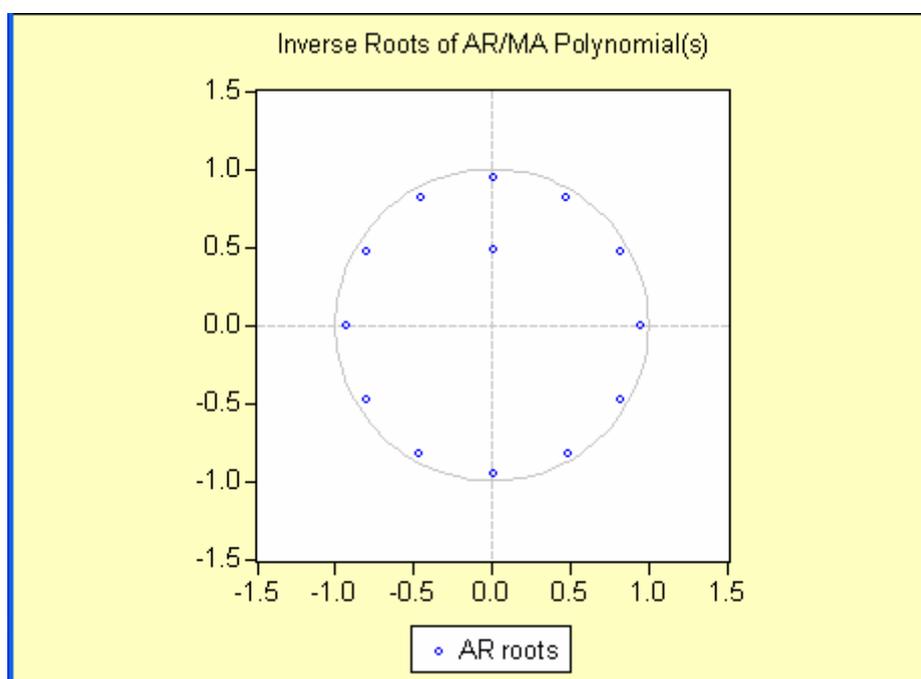
Como se puede ver en el cuadro 10, ahora todos los coeficientes estimados son estadísticamente significativos al nivel de 5% ya que los p-valores del estadístico t de Student son menores que el nivel de significancia de 5%. Se cumple entonces el supuesto estadístico estimaciones estadísticamente significativas.

Otro supuesto estadístico que debe cumplir el modelo (3.2) es el de estacionariedad e invertibilidad. De acuerdo a la sección 1.2.1.1 del capítulo II, todos los modelos AR finitos son invertibles, y como nuestro modelo es finito, entonces lo único que tenemos que validar es que el modelo sea estacionario, es decir que las raíces del modelo estén dentro del círculo unitario del plano complejo. Esta propiedad de estacionariedad se cumple porque, según los resultados del recuadro inferior del cuadro 10, ningunas de las partes reales de las raíces características del polinomio característico asociado al modelo (3.2), es uno. Esto también lo podemos observar en el cuadro 11. Por lo tanto, el modelo (3.2) cumple con el supuesto estadístico de estacionariedad. Así, el modelo (3.2) es estacionario e invertible.

Dependent Variable: D(LOG(IIS))
 Method: Least Squares
 Date: 05/30/09 Time: 13:40
 Sample (adjusted): 2001M03 2009M04
 Included observations: 98 after adjustments
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003182	0.000468	6.803744	0.0000
AR(1)	0.565300	0.084575	6.684035	0.0000
AR(12)	0.492148	0.086616	5.681975	0.0000
AR(13)	-0.242067	0.089636	-2.700558	0.0082
R-squared	0.587333	Mean dependent var		0.003313
Adjusted R-squared	0.574162	S.D. dependent var		0.001291
S.E. of regression	0.000842	Akaike info criterion		-11.28049
Sum squared resid	6.67E-05	Schwarz criterion		-11.17498
Log likelihood	556.7440	F-statistic		44.59545
Durbin-Watson stat	1.980776	Prob(F-statistic)		0.000000
Inverted AR Roots	.96	.83-.47i	.83+.47i	.49
	.48+.81i	.48-.81i	.00+.94i	.00-.94i
	-.47+.81i	-.47-.81i	-.81-.47i	-.81+.47i
	-.94			

Cuadro 10. Estimación de parámetros y estacionariedad del modelo (3.2)



Cuadro 11. Estacionariedad del modelo (3.2).

Ya que el modelo (3.2) tiene coeficientes estadísticamente significativos y es estacionario e invertible, para que sea un modelo correctamente especificado, debe tener residuos no autocorrelacionados. La prueba estadística que se usa para probar la hipótesis nula de residuos autocorrelacionados es la prueba de

Ljung-Box presentada en la sección 1.2.5 del capítulo II. Los resultados de esta prueba se presentan en el cuadro 12.

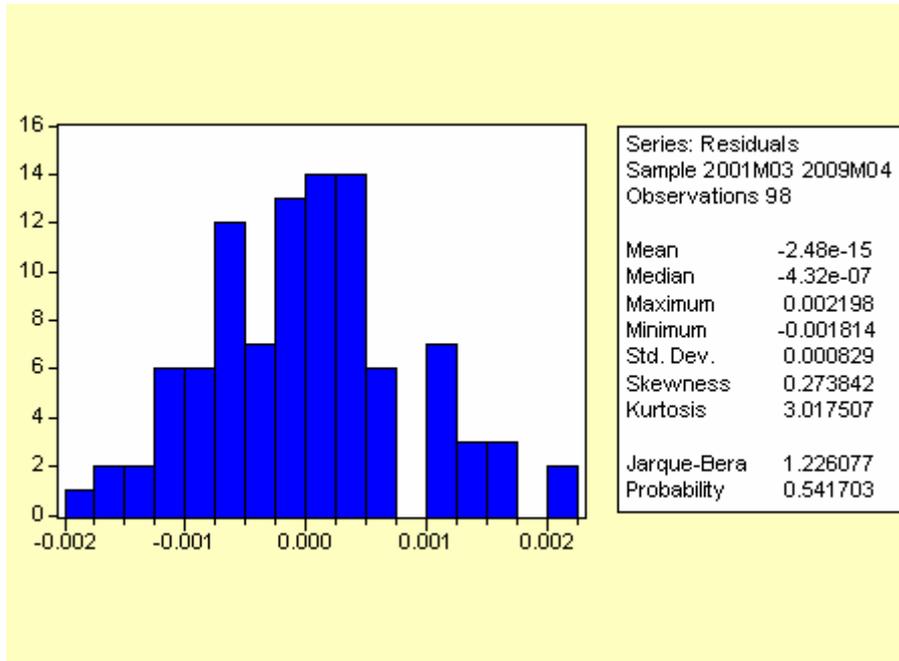
Como se puede ver en el cuadro 12, según el estadístico Q de Ljung-Box de significancia conjunta, no se puede aceptar la hipótesis nula de autocorrelación de residuos. Es decir, los residuos del modelo (3.2) no están autocorrelacionados.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.007	0.007	0.0043	
		2	-0.067	-0.067	0.4679	
		3	0.073	0.074	1.0182	
		4	0.093	0.088	1.9258	0.165
		5	0.077	0.087	2.5515	0.279
		6	0.168	0.179	5.5643	0.135
		7	-0.013	-0.013	5.5833	0.233
		8	-0.058	-0.056	5.9449	0.312
		9	0.230	0.199	11.794	0.067
		10	0.102	0.066	12.958	0.073
		11	0.063	0.086	13.402	0.099
		12	-0.080	-0.117	14.128	0.118
		13	0.111	0.094	15.541	0.114
		14	-0.010	-0.065	15.554	0.159
		15	0.057	-0.011	15.931	0.194
		16	-0.004	-0.046	15.933	0.253
		17	-0.007	-0.001	15.939	0.317
		18	0.010	-0.012	15.950	0.385
		19	-0.035	-0.107	16.100	0.446
		20	0.030	-0.005	16.215	0.509
		21	-0.075	-0.060	16.925	0.528
		22	0.056	0.044	17.328	0.568
		23	0.086	0.114	18.298	0.568
		24	0.111	0.124	19.938	0.525
		25	-0.023	0.062	20.010	0.582
		26	0.078	0.083	20.832	0.591
		27	-0.032	-0.027	20.975	0.640
		28	-0.043	-0.062	21.232	0.680
		29	0.046	-0.011	21.537	0.714
		30	-0.096	-0.131	22.856	0.693
		31	-0.110	-0.162	24.623	0.648
		32	0.070	0.006	25.357	0.660
		33	0.006	-0.107	25.362	0.707
		34	-0.089	-0.070	26.574	0.693
		35	0.073	0.029	27.407	0.698
		36	-0.008	0.053	27.416	0.741

Cuadro 12. Prueba para ver que los residuales del modelo (3.2) son un ruido blanco.

Gráficamente este resultado también se observa en el cuadro 12 en las gráficas de las funciones de autocorrelación (FAC) y de autocorrelación parcial (FACP).

Por ultimo se verifica que los residuos del modelo (3.2) tengan un adistribución normal. Para ello se hace la prueba de normalidad de Jarque-Bera¹. Dado que el p_valor de esta prueba es mayor al 5%, se acepta la hipótesis nula de que los residuales tienen una distribución normal. Estos resultados se muestran en el cuadro 13.



Cuadro 13. Prueba de normalidad de los residuos estimados del modelo (3.2).

En resumen, el modelo (3.2) está correctamente especificado porque satisface los supuestos estadísticos exigidos; tiene parámetros estadísticamente significativos, es estacionario e invertible y tiene residuos no autocorrelacionados y con distribución normal. Por lo tanto, el modelo econométrico correctamente especificado que describe la dinámica de la serie de datos de tiempo IIS es el siguiente

$$\hat{W}_t = 0.5653Z_{t-1} + .492148Z_{t-12} - 0.242067Z_{t-13} + \varepsilon_t \quad (3.3)$$

En la medida en que el modelo (3.3) es un modelo correctamente especificado, este modelo se puede utilizar para pronosticar el valor futuro de la variable de serie de tiempo IIS, como sugiere la metodología aplicada aquí de Box-Jenkins. Aplicando la teoría de pronóstico expuesta en el primer capítulo, se pronostica el valor del IIS para el mes de mayo y junio de 2009. El resultado de este pronóstico se presenta en el cuadro 14. En la columna de variación porcentual se muestra el incremento porcentual que tendría IIS para el mes de junio: el valor del IIS aumentaría en un .427169%.

¹ Ver el anexo 2.

Periodo	Pronostico IIS	Variación porcentual
2009/05	131.2127	
2009/06	131.7732	0.427169

Cuadro 14. Pronóstico.

Para ver la capacidad de pronóstico del modelo, se quito el último dato de nuestras observaciones de la serie inflación subyacente, es decir se quito el dato del IIS correspondiente al mes de abril de 2009. Entonces con los datos del mes de enero de 2000 a marzo de 2009 se pronostico el dato de abril de 2009. Los resultados obtenidos para el dato pronosticado se muestran en el cuadro 15.

Mes	Valor Pronosticado	Valor real
Abr-09	130.7123	130.625

Cuadro 15. Comparación de valor

Resumen y conclusiones

En esta tesis se estudio la dinámica de la inflación en México especificando un modelo econométrico basado en la teoría ARIMA de series de tiempo estacionarias y en la metodología Box-Jenkins. El indicador de inflación utilizado fue el índice de la inflación subyacente (IIS) durante el periodo de enero de 1990 a abril de 2009, Este periodo corresponde al comportamiento relativamente más estable de la serie índice de la inflación subyacente y al cual se pudo ajustar un modelo ARIMA.

Con el fin de especificar un modelo econométrico para la serie IIS se utilizo la metodología recursiva de Box-Jenkins siguiendo las siguientes etapas: (a) se obtuvo una serie estacionaria en diferencias, (2) se identificó un modelo para la serie IID, (3) se estimaron los parámetros del modelo identificado en (2), (4) se diagnosticó estadísticamente el modelo estimado y (5) se usó el modelo que cumplió con los supuestos estadísticos, es decir, el modelo correctamente especificado, para hacer un pronóstico del IIS.

El modelo correctamente especificado para la variable de serie de tiempo IIS en el periodo estudiado fue el siguiente:

$$\hat{W}_t = 0.5653Z_{t-1} + .492148Z_{t-12} - 0.242067Z_{t-13} + \varepsilon_t$$

De acuerdo a los valores que se obtuvieron del pronóstico para el mes de mayo y junio, podríamos decir que el índice de la inflación subyacente mostrara una tendencia creciente. También se concluye que la variación porcentual para el mes de junio será del .427169%.

El alcance de este modelo es de corto plazo es decir para pronosticar uno o dos meses hacia delante. Si se quisieran hacer pronósticos para periodos de tiempo más largos, se tendría que cambiar la periodicidad de la información de los datos con los que se cuenta.

Para periodos donde la variabilidad de la serie del índice de la inflación subyacente es mayor, (por ejemplo el periodo enero de 1990 a abril de 2009), esta serie se podría modelar con otro tipo de modelos econométricos mas generales, los modelos ARCH. Los modelos que se presentaron en esta tesis (modelos (ARIMA(p,d,q))) son un caso particular de estos modelos.

Anexos

Anexo 1

Los datos utilizados en el estudio de la dinámica de la inflación en México se obtuvieron de la página del INEGI en la siguiente ruta:

<http://dgcnesyp.inegi.org.mx/cgi-win/bdieintsi.exe/NIVL1000800220#ARBOL>

La definición de inflación subyacente que se está usando para estos datos es la anterior. La unidad de Medida: Base segunda quincena de junio 2002 = 100. La Inflación subyacente se obtiene eliminando del INPC los bienes y servicios cuyos precios son más volátiles, estos se agrupan en las siguientes categorías: agropecuarias; educación y administrados y concertados.

Periodo	IIS	Periodo	IIS	Periodo	IIS
2000/01	87.56241	2003/05	103.4582	2006/09	115.357
2000/02	88.57075	2003/06	103.5821	2006/10	115.6642
2000/03	89.19724	2003/07	103.7274	2006/11	115.9984
2000/04	89.70001	2003/08	103.9671	2006/12	116.4951
2000/05	90.1104	2003/09	104.1393	2007/01	117.0811
2000/06	90.42869	2003/10	104.3941	2007/02	117.5575
2000/07	90.74474	2003/11	104.7035	2007/03	117.9389
2000/08	91.10334	2003/12	105.0404	2007/04	118.1793
2000/09	91.45578	2004/01	105.4641	2007/05	118.4389
2000/10	91.84965	2004/02	105.9657	2007/06	118.7976
2000/11	92.40085	2004/03	106.3976	2007/07	119.2009
2000/12	92.92265	2004/04	106.7754	2007/08	119.4539
2001/01	93.64237	2004/05	107.0407	2007/09	119.7399
2001/02	94.40772	2004/06	107.3531	2007/10	120.1856
2001/03	94.96339	2004/07	107.536	2007/11	120.6027
2001/04	95.4754	2004/08	107.7754	2007/12	121.1568
2001/05	95.86675	2004/09	108.0529	2008/01	121.6489
2001/06	96.1758	2004/10	108.4058	2008/02	122.2558
2001/07	96.33313	2004/11	108.6987	2008/03	122.9163
2001/08	96.60102	2004/12	109.027	2008/04	123.4487
2001/09	96.88152	2005/01	109.4394	2008/05	124.1062
2001/10	97.14884	2005/02	109.8997	2008/06	124.6914
2001/11	97.40051	2005/03	110.2385	2008/07	125.2206
2001/12	97.64186	2005/04	110.467	2008/08	125.6867
2002/01	98.09456	2005/05	110.7015	2008/09	126.1149
2002/02	98.82619	2005/06	110.9859	2008/10	126.5377
2002/03	99.27588	2005/07	111.2301	2008/11	127.2334
2002/04	99.65512	2005/08	111.3077	2008/12	128.1009
2002/05	99.85688	2005/09	111.5103	2009/01	128.6696
2002/06	99.98468	2005/10	111.8138	2009/02	129.3159
2002/07	100.1282	2005/11	112.0759	2009/03	130.0869
2002/08	100.4134	2005/12	112.434	2009/04	130.625
2002/09	100.6109	2006/01	112.7015		
2002/10	100.8791	2006/02	113.088		
2002/11	101.0986	2006/03	113.5855		
2002/12	101.3274	2006/04	114.0017		
2003/01	101.7995	2006/05	114.1747		
2003/02	102.323	2006/06	114.5542		
2003/03	102.7779	2006/07	114.8712		
2003/04	103.1527	2006/08	115.0094		

Cuadro 1a Datos usados para el análisis econométrico del IIS.

ANEXO 2

Teorema fundamental del algebra.

Todo polinomio complejo $p(z)$ de grado mayor o igual que n , tiene una factorización

$$p(z) = c(z - z_1)^{m_1} \dots (z - z_k)^{m_k}$$

donde las z_j 's son distintas y $m_j \geq 1$. Esta factorización es única, a excepción de la permutación de los factores. (véase Gamelin, Theodore W, 2000)

Variable aleatoria.

Dado un espacio de probabilidad $(\Omega, \mathfrak{S}, P)$, se dice que una función $X : \Omega \rightarrow R$ es una variable aleatoria real si $[X \leq x] \in \mathfrak{S}$ para cualquier $x \in R$.

Varianza

Sea X una variable aleatoria de esperanza finita. Se define la varianza de X , $Var(X)$, mediante la relación:

$$Var(X) = E[(X - E(X))^2]$$

Covarianza

Sean X y Y dos variables aleatorias de varianza finita. Se define la covarianza de X y Y , $Cov(X, Y)$, mediante la relación:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

Función de autocovarianza y Función de autocorrelación

Para un proceso estacionario la covarianza entre Z_t y Z_{t+k} separada por k periodos en el tiempo es la misma para toda t . La autocovarianza al rezago k , se define por

$$\gamma_k = Cov[Z_t, Z_{t+k}] = E[(Z_t - \mu)(Z_{t+k} - \mu)]$$

Similarmente, la autocorrelación al periodo k esta dada por

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{\sqrt{E[(Z_t - \mu)^2][E[(Z_{t+k} - \mu)^2]}} = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{\sigma_z^2}$$

ya que para un proceso estacionario, la varianza $\sigma_z^2 = \gamma_0$ es la misma para el tiempo k y para el tiempo $k + 1$.

La autocorrelación al periodo k , es decir, la correlación entre Z_t y Z_{t+k} , esta dada por

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Lo que implica que $\rho_0 = 1$.

La gráfica de γ_k contra el periodo k es llamada función de autocovarianza $\{\gamma_k\}$ del proceso estocástico.

Similarmente, la gráfica de los coeficientes ρ_k como función del periodo k es llamada función de autocorrelación $\{\rho_k\}$ del proceso.

Prueba de Jarque-Bera

La prueba de Jarque-Bera es una prueba de normalidad, donde la hipótesis nula y la hipótesis alternativa son las siguientes:

$H_0 : X_t$ esta normalmente distribuída.

$H_a : X_t$ no esta normalmente distribuída.

El estadístico de prueba para es:

$$Jarque - Bera = \frac{N - k}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right),$$

donde S es la simetría de la distribución alrededor de la media (skewnness en ingles), K la curtosis (el tamaño del pico de la distribución), N el número de observaciones de la serie y k representa el número estimado de coeficientes para crear las serie.

Una distribución normal tiene una simetría alrededor de la media igual a cero y una curtosis de 3.

Bajo la hipótesis nula de una distribución normal, el estadístico de Jarque-Bera se distribuye como una Ji-cuadrada con dos grados de libertad. La Probabilidad reportada en la prueba de jarque.Bera que muestra E-views es la probabilidad de que le estadístico de Jarque-Bera exceda (en valor absoluto) el valor observado bajo la hipótesis nula.

Anexo 3

Glosario de términos técnicos

Las siguientes definiciones fueron tomadas de la página del Banco de México.

Canasta de bienes y servicios de los índices de precios

Es una muestra amplia de los bienes y servicios representativos de la producción nacional en el periodo base en el caso del INPP y de bienes y servicios representativa de los satisfactores adquiridos por los hogares en el periodo base en el caso del INPC.

Demanda

Cantidad de bienes y servicios que los agentes económicos desean y pueden comprar a un precio dado en un periodo determinado. En teoría la oferta y la demanda son los dos componentes básicos que fijan el precio de los bienes y servicios.

Deseo de cualquier persona por adquirir un bien o servicio económico.

Economía.

Rama de las ciencias sociales que trata de la producción, distribución y consumo de los bienes y servicios.

Política monetaria.

Son las acciones que instrumenta el Banco de México con la finalidad de promover la estabilidad del poder adquisitivo de la moneda del país.

Producto genérico.

Unidad primaria de ponderación de la canasta del INPC, compuesta por un conjunto de productos específicos con características similares. Por ejemplo, todas las marcas y tipos de galletas forman el concepto genérico galletas.

Base de comparación de un índice de precios.

Es el punto de referencia en el tiempo a partir del cual se efectúan las comparaciones de los cambios de precios. El concepto también se conoce como periodo de referencia, o simplemente “base”.

Referencias bibliográficas

- [1] Banxico, "El índice Nacional de Precios al Consumidor: características y Actualización de su Base al año 2000", Banco de México, junio 2002.
- [2] Banerjee, Anindya, Dolado, Juan J, COINTEGRATION, ERROR CORRECTION, AND THE ECONOMETRIC ANALYSIS OF NON-STATIONARY DATA, Oxford University Press, 1993, Great Britain.
- [2] Box, George, P.E., TIME SERIES ANALYSIS forecasting and Control, Prentice-Hall, Inc, 1994, United States of America.
- [3] Brockwell Peter J. and Davis Richard A, *Time series and Methods, second edition*, Springer, United States, 1991.
- [4] Diggle Peter J., Time Series A Bioestatistical Introduction, Oxford University Press Inc., New York, 1990.
- [5] Enders Walter, *Applied Econometric Time Series*, 2da edición, Wiley & Sons, United States, 1994.
- [6] Fuller, Wayne A, Introduction to statistical time series, John Wiley & Sons, United States, 1976.
- [7] Gamelin, Theodore W, COMPLEX ANALYSIS, Springer-Verlag, New York inc, New York USA, 2000.
- [8] Guerrero Guzmán, Victor Manuel, *Análisis estadístico de series de tiempo económicas*, Thomson editores, Mexico 2003.
- [9] Hatanaka, Michio, Time-Series-Based Econometrics Unit roots and Co-Integrations, OXFORD UNIVERSITY PRESS, United States, 1996.
- [10] Koopmans, Lambert H, THE SPECTRAL ANALYSIS OF TIME SERIES, ACADEMIC PRESS, United States of America, 1995.
- [11] Ross, Sheldon, M, Introduction to Probability Models, Eighth Edition, ACADEMIC PRESS, 2003, United States of America.
- [12] Wei William W.S., *Time Series Analysis univariate and multivariate methods*, Addison-Wesley, United States, 1990.