



Casa abierta al tiempo

UNIVERSIDAD AUTONOMA METROPOLITANA

IZTAPALAPA

DIVISION DE CIENCIAS BASICAS
E INGENIERIA

**METODOS PARA DETERMINAR EL
ORDEN DE UN MODELO AUTORE-
GRESIVO Y DE PROMEDIOS MO-
VILES PARA SERIES DE TIEMPO**

U. A. M. IZTAPALAPA BIBLIOTECA

T E S I S

Que para obtener el Grado de
MAESTRIA EN MATEMATICAS
p r e s e n t a
Alejandro Islas Camargo

México, D. F.

Diciembre de 1989

093895

*Deseo expresar mi agradecimiento al
Dr. Victor M. Guerrero por su apoyo
y dirección de este trabajo.*

*Así mismo, al Maestro Alejandro
Alegría por sus valiosos
comentarios y su particular interés
en la lectura de este trabajo.*

CONTENIDO

INTRODUCCIÓN

1.- MÉTODOS BASADOS EN LA TEORIA DE UNA REALIZACIÓN DE UN PROCESO ESTOCÁSTICO

- 1.1 Algunos Conceptos Básicos
- 1.2 El Método de Box-Jenkins
- 1.3 Arreglos "R" y "S"
- 1.4 Función de Autocorrelación Parcial y de Autocorrelación Parcial Generalizada
- 1.5 El Método de la Esquina
- 1.6 Regresiones Iteradas y Función de Autocorrelación Muestral Extendida
- 1.7 Función de Autocorrelación Inversa y Función de Autocorrelación Parcial Inversa

2.- MÉTODOS BASADOS EN LA TEORIA DE PRUEBAS DE HIPÓTESIS ESTADÍSTICAS

- 2.1 Modelaje $ARMA(n,n-1)$ de Pandit-Wu
- 2.2 Criterio de Información de Akaike (AIC)
- 2.3 Métodos Bayesianos.

3.- MÉTODOS BASADOS EN EL ERROR DE PREDICCIÓN UN PASO HACIA ADELANTE

- 3.1 Criterio del Error de Predicción
- 3.2 Criterio de Validación Cruzada
- 3.3 Criterio para una Función de Autocorrelación Transferida

4.- EJEMPLOS

Índice de notación y abreviaturas

v.a (vv.aa.)	significa	variable(s) aleatoria(s)
i.i.d	"	independientes e idénticamente distribuidas.
ssi	"	si y sólo si.
■	"	que la demostración ha terminado
$\xrightarrow{p}$	"	convergencia en probabilidad
$\xrightarrow{D}$	"	convergencia en distribución
$\doteq$	"	equivalencia asintótica
•	"	indefinido
$N(\mu, \sigma^2)$	denota	distribución Normal con media μ y varianza σ^2
$\chi^2_{(n)}$	"	distribución ji cuadrada con n grados de libertad

I. INTRODUCCION.

Los modelos autorregresivos y de promedios móviles de orden (p,q) , usualmente abreviados modelos ARMA (p,q) , son con frecuencia utilizados para describir y predecir observaciones en series de tiempo. Estos modelos comúnmente se representan mediante

$$\Phi_p(B) Z_t = C + \theta_q(B) a_t \quad (1)$$

en donde

$$\Phi_p(B) = U_d(B) \phi_{p-d}(B) = 1 - \sum_{i=1}^p \phi_i B^i, \quad U_d(B) = 1 - \sum_{i=1}^d U_i B^i$$

$$\phi_{p-d}(B) = 1 - \sum_{i=1}^{p-d} \phi_i B^i \quad \text{y} \quad \theta_q(B) = 1 - \sum_{i=1}^q \theta_i B^i$$

son polinomios de retraso de orden $p, d, p-d$ y q respectivamente. B es el operador de retraso tal que $B^k Z_t = Z_{t-k}$ para $k = 0, 1, \dots$ y toda t , C , al igual que las ϕ 's son constantes y $\{a_t: t \in Z\}$ es una sucesión de v.a.i.i.d con distribución $N(0, \sigma_a^2)$. En la práctica, comúnmente se requiere que todas las raíces de $U_d(B)$ estén sobre el círculo unitario y que las de $\phi_{p-d}(B)$ y $\theta_q(B)$ estén fuera del círculo unitario, además de que $\Phi_p(B)$ y $\theta_q(B)$ no tengan factores comunes. Las Z_t 's representan la d -ésima diferencia o cualquier otra transformación apropiada para volver estacionaria alguna serie

de tiempo originalmente no estacionaria; cuando $U_d(B) = (1-B)^d$ el modelo (1) se conoce como un modelo ARIMA (p-d,d,q).

En los últimos años, el problema de la determinación o especificación del orden de un modelo ARMA, ha atraído considerablemente la atención en los trabajos de series de tiempo, y en áreas de investigación como son la teoría de control, la econometría y la estadística, las cuales están estrechamente relacionadas con el análisis de series de tiempo.

Especificar el orden de un modelo ARMA significa determinar el orden (p,q) de éste. Si estos valores se conocen, entonces los parámetros $\Phi = (\phi_1, \phi_2, \dots, \phi_p)$, $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ y σ_a^2 pueden ser estimados. Esta última parte no será considerada en este trabajo; no obstante, algunos de los métodos que aquí se presentan requieren de la estimación de los parámetros de diferentes modelos, tales parámetros estimados pueden entonces ser comparados con el fin de hacer una posible discriminación entre diferentes valores de p y q. En raras ocasiones se conoce el orden real de un modelo y determinarlo es una parte difícil y delicada del análisis de series de tiempo, ya que la selección de (p,q) se basa en sólo una muestra de observaciones del proceso generador de datos. Por este motivo el modelo seleccionado debe verse como una simplificación del modelo asociado con el proceso real.

Varios métodos han sido propuestos y estudiados, la mayoría de los cuales siguen la técnica de Box-Jenkins para construir modelos de series de tiempo. Una cantidad muy extensa de estos métodos se basa en inferencias construidas a partir de patrones de la función de autocorrelación (FAC) muestral y de

la función de autocorrelación parcial (FACP) muestral de la serie. Las correspondientes características de las funciones FAC y FACP teóricas para varios tipos de modelos ARMA (p, q) han sido usadas como patrones de referencia (ver Box y Jenkins 1976, cap3 o, a un nivel más elemental, los caps. 4 y 5 de Guerrero, 1983). Desafortunadamente, esta técnica de reconocimiento de patrones confía en el buen juicio heurístico del analista de la serie. Entonces, elegir el orden apropiado de un modelo mediante una inferencia que parte de patrones de estas dos funciones, con frecuencia resulta ser muy difícil. De aquí que hayan surgido varios métodos alternativos.

Hasta el momento no existe una solución completamente satisfactoria para el problema de la determinación del orden de un modelo de series de tiempo, a partir de la observación de datos muestrales del proceso real, no obstante se han propuesto un gran número de procedimientos para elegir la estructura más adecuada de un modelo. En este trabajo se estudiarán las características más importantes y más útiles de algunos de estos métodos. El objetivo es pues, presentar las técnicas más importantes que actualmente se utilizan en la determinación del orden de un modelo ARMA (p, q). Se tiene conocimiento de algunos trabajos publicados con objetivos similares al presente, la diferencia con el que aquí se presenta es que aquellos están restringidos a modelos de regresión lineal y/o enfocados a un rango más limitado de procedimientos para determinar el orden. Sin embargo, el trabajo de Gooijer, Abraham, Gould y Robinson. (1985) es muy parecido al presente, con la diferencia de que en aquél no se entra en detalle para ninguno de los métodos que cita y aquí se intenta cubrir esa deficiencia.

Una forma de clasificar a los diferentes métodos es

asignándoles alguna propiedad, ya sea ésta de carácter "objetivo" o "subjetivo". En un método subjetivo existe siempre un juicio de apreciación tomado por el analista, éste puede ser la elección de antemano de un nivel de significación o el examen visual de gráficas y tablas para observar el comportamiento o patrón característico de una cierta estadística en particular. Mientras que un método objetivo se caracteriza por el hecho de que no es necesaria la intervención del elemento humano para realizar juicios de apreciación en la especificación del modelo. Esta forma de clasificar los métodos no es necesariamente la más adecuada, pues el hecho de que un método sea clasificado objetivamente no elimina por completo la posible emisión de juicios de apreciación en éste. Sin embargo, esto proporciona un camino para tratar de agrupar los diferentes métodos que aquí se presentan, de acuerdo con alguna característica relevante de los mismos.

La clase de métodos subjetivos puede ser ampliamente dividida en las dos subclases siguientes: a) aquellos métodos que están basados en la teoría de una realización de un proceso estocástico y b) los que hacen uso de la teoría de pruebas de hipótesis estadísticas; los primeros, a excepción del método de la Función de Autocorrelación Inversa y de la Función de Autocorrelación Parcial Inversa, serán tratados con detalle en las secciones 1.2, 1.3, 1.4, 1.5 y 1.6. La razón por la cual este método no se tratará con detalle, es por que sólo presenta ventajas en la determinación del orden de un modelo autorregresivo, mientras que los segundos se presentarán a manera de información en el capítulo 2.

Dentro de los métodos clasificados como objetivos, se presentarán también de manera informativa, aquellos que están basados en el error de predicción un paso hacia adelante, estos

se presentan en el capítulo 3, por último en el capítulo 4 se presentarán aplicaciones de los métodos tratados en las subsecciones de la 1.2, a la 1.6.

Es conveniente mencionar que en este trabajo no se consideran modelos para series de tiempo con efectos periódicos (asociados comúnmente con la estacionalidad) del tipo del modelo multiplicativo estacional sugerido por Box y Jenkins.

CAPÍTULO 1

MÉTODOS BASADOS EN LA TEORIA DE UNA REALIZACIÓN DE UN PROCESO ESTOCASTICO

1.1 ALGUNOS CONCEPTOS BASICOS

En casi todas las ramas de la producción y del conocimiento, se encuentran series de datos que corresponden a fechas sucesivas:

- Ejemplos
- i) Tipo de cambio peso/dólar cada día
 - ii) Venta de algún producto cada semana
 - iii) Índice de inflación de los precios cada mes
 - iv) Producto interno bruto de México cada año

Una serie de este tipo, se llamará serie de tiempo y se denotará su valor al tiempo t por Z_t . El caso en que Z_t sea una función (determinista) de t es de poco interés (y relativamente escaso en la práctica). Es mucho más cercano a la realidad y por lo tanto mucho más eficaz, escoger un modelo en donde cada Z_t sea una variable aleatoria. La sucesión de las Z_t se llama en este caso un proceso. Puesto que los procesos en los que se estará interesado son procesos que se

desenvuelven en el tiempo, sujetos a variaciones que pueden ser descritas por leyes probabilísticas, dichos procesos caen dentro del estudio de los procesos estocásticos.

Para describir lo que es una serie de tiempo dentro del contexto de procesos estocásticos, es necesario primero definir a los procesos estocásticos.

Un proceso estocástico es una familia $Z = \{Z(t); t \in T\}$ de variables aleatorias. Es decir, para cada $t \in T$, en donde T es el conjunto de parámetros o conjunto de índices, $Z(t)$ es una v.a. definida sobre algún espacio de probabilidad $(E, \mathcal{F}, P)$; $Z(t) : E \longrightarrow \mathbb{R}$. Cuando se desea enfatizar que el proceso depende de los eventos elementales $w \in E$, se escribe

$$Z(t) = Z(t, w)$$

Así pues, el proceso se puede ver como una colección de funciones reales definidas sobre $T \times E$, o sea,

$$Z(.,.): T \times E \longrightarrow \mathbb{R}$$

tal que

i) para cada $t \in T$ fijo

$$Z(t,.): E \longrightarrow \mathbb{R} \text{ es una v.a}$$

ii) para cada $w \in E$ fijo

$Z(.,w): T \longrightarrow \mathbb{R}$ es una función de $t \in T$, llamada una realización o trayectoria del proceso $\{Z\}$.

El espacio de estados del proceso es el conjunto S de todos los valores posibles que pueden tomar las variables $Z(t)$

para $t \in T$. Cada punto en S se llama un estado del proceso. Si S es finito o numerable se dirá que el proceso es discreto; si S es un intervalo, se dirá que el proceso es continuo.

El proceso también se clasifica de acuerdo a la naturaleza del conjunto de parámetros T . Si T es numerable, por ejemplo, $T = \{0, \pm 1, \pm 2, \dots\}$, $T = \{0, 1, 2, \dots\}$, etc., se dice que Z es un proceso con parámetro (o en tiempo) discreto. Si T es un intervalo, $T = [a, b]$; $T = [0, \infty)$; $T = (-\infty, \infty)$, etc., $\{Z\}$ es un proceso con parámetro (o en tiempo) continuo.

Con base en lo anterior, se concibe entonces a una serie de tiempo como a la sucesión de observaciones generadas por un proceso estocástico cuyo conjunto de índices se toma en relación al tiempo. Así como existen procesos estocásticos discretos y continuos, existen también series de tiempo discretas y continuas. En este trabajo sólo se considerarán series de tiempo discretas. Asimismo y con el riesgo de crear cierta confusión, en general no se hará distinción entre una variable aleatoria Z y su valor observado, que se denotará también por Z . De esta manera, cuando se tengan N valores sucesivos de una serie de tiempo, se escribirá $Z_1, Z_2, \dots, Z_N$ para denotar a las observaciones hechas en los tiempos $1, 2, 3, \dots, N$.

Puesto que los primeros momentos de las variables aleatorias resumen en buena medida a su distribución, se ha sugerido estudiar a los momentos de primero y segundo orden, es decir

1) La media de $\{Z_t\}$

$$E[Z_t] = \mu_t \quad (1.1.1)$$

ii) La varianza de $\{Z_t\}$

$$\gamma_0 = E [(Z_t - \mu_t)^2] \quad (1.1.2)$$

iii) La covarianza de $\{Z_t\}$

$$\gamma_m = \text{cov} (Z_t, Z_{t+m}) = E [(Z_t - \mu_t)(Z_{t+m} - \mu_{t+m})] \quad (1.1.3)$$

adviértase que es posible hablar de autocovarianza, lo cual sirve para indicar que la covarianza es entre observaciones de la misma variable, aunque retrasada. Sin embargo, para evitar la influencia de las unidades de medida, es preferible trabajar con las autocorrelaciones $\rho_0, \rho_1, \dots$ definidas a través de

$$\rho_m = \frac{E [(Z_t - \mu_t)(Z_{t+m} - \mu_{t+m})]}{E [Z_t - \mu_t]^2} \quad (1.1.4)$$

las cuales generan a la FAC, que es la sucesión $\{\rho_m\}$.

1.1.1 Procesos estacionarios

Como se observó, una serie de tiempo puede pensarse como una realización de un proceso estocástico. En la mayoría de las series de tiempo observadas es imposible observar otra realización. Esto es, no se puede repetir el experimento nuevamente, por ejemplo es imposible regresar en el tiempo y obtener otra sucesión de observaciones sobre la Inflación Interna en México en el periodo 1970-1980.

Por lo tanto, en cada instante del tiempo t , generalmente se tiene sólo una observación Z_t de $\{Z_t\}$. Si se deseara por

ejemplo saber alguna cosa sobre μ_t o γ_t^2 , resulta imposible obtener esta información con apenas una observación Z_t a menos que se tenga alguna suposición adicional sobre el tipo de proceso analizado.

Una primera suposición que normalmente se hace es la de *Estacionariedad*; intuitivamente, un proceso estacionario mostrará el mismo comportamiento en términos generales sin importar el momento en que se observe. Esto es, si se graficara un cierto número de observaciones contiguas de un proceso, la gráfica que se obtendría sería bastante similar a la que se obtuviera al graficar el mismo número de observaciones contiguas, pero m periodos hacia adelante o hacia atrás de los considerados inicialmente.

Formalmente existen dos formas de estacionariedad, la *Estacionariedad estricta* y la *Estacionariedad de segundo orden*, a continuación se define cada una de ellas

Se dirá que un proceso es estrictamente estacionario si la función de densidad para un conjunto arbitrario de variables $Z_t, Z_{t+1}, \dots, Z_{t+m}$ es invariante respecto a desplazamientos en el tiempo, esto es, si se cumple que

$$f(Z_t, Z_{t+1}, \dots, Z_{t+k}) = f(Z_{t+m}, Z_{t+m+1}, \dots, Z_{t+m+k})$$

para toda t, m y k . Esto significa en particular que si los primeros momentos de f existen, se sigue que

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

que

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

y que

$$E [(Z_t - \mu)(Z_{t+m} - \mu)] = E [(Z_{t+k} - \mu)(Z_{t+k+m} - \mu)] = \gamma_m$$

para toda t, m y k . Nótese que en estas expresiones no aparece t , lo cual implica que ni el nivel de la serie, ni su variabilidad dependen del tiempo, y en lo que respecta a la covarianza, no existe dependencia del tiempo, pero sí de la separación (m) que hay entre las variables.

Se dirá que un proceso $\{Z_t\}$ es estacionario de segundo orden si se cumplen las siguientes condiciones

- i) $E [Z_t] = \mu(t) = \mu$, constante para todo t
- ii) $\text{Var} [Z_t] = E [(Z_t - \mu)^2] = \gamma_0$, constante para todo t
- iii) $\text{Cov} [Z_t, Z_{t+m}] = E [(Z_t - \mu)(Z_{t+m} - \mu)] = \gamma_m$.

Por consiguiente, la estacionariedad estricta implica la estacionariedad de segundo orden y, aunque la implicación inversa no siempre se cumple, sí se satisface cuando el proceso tiene como distribución conjunta de todas las variables que lo integran a la Normal multivariada.

Es claro que no todas las series de tiempo pueden considerarse realizaciones de un proceso estacionario, pero mediante una transformación apropiada de estas observaciones (por ejemplo la diferenciación) la serie se puede transformar a una serie estacionaria y utilizar la teoría de procesos estocásticos a este tipo de series.

Finalmente, siendo γ_m la autocovarianza de un proceso

estacionario $\{Z_t\}$, las siguientes propiedades se satisfacen:

- i) $\gamma_0 \geq 0$
- ii) $\gamma_{-m} = \gamma_m$
- iii) $|\gamma_m| \leq \gamma_0$ (1.1.5)
- iv) γ_m es positiva definida, en el sentido de que

$$\sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k \gamma_{j-k} \geq 0$$

La FAC de $\{Z_t\}$ definida por (1.1.4) satisface propiedades análogas a las i) - iv).

La función de autocorrelación, sirve para especificar, junto con μ y σ_a^2 , a un proceso estocástico estacionario; por otro lado, desde un punto de vista práctico de series de tiempo, debe recordarse que se cuenta sólo con una realización finita del proceso, a partir de la cual se hace necesario obtener una estimación de $\{\rho_m\}$. Para lograr dicha estimación, se supondrá que el proceso estacionario en estudio posee ciertas propiedades *Ergódicas*, que permiten la equivalencia entre valores esperados y promedios muestrales obtenidos de una realización suficientemente larga. De esta propiedad se sigue que la media del proceso puede estimarse como la media muestral de la serie observada.

$$\bar{Z} = \frac{\sum_{t=1}^N Z_t}{N} \quad (1.1.1')$$

siempre y cuando $\mu < \infty$ y además se cumpla que

$$\lim_{N \rightarrow \infty} \frac{\sum_{m=1}^N \rho_m}{N} = 0. \quad \text{Asimismo, si se cumple entre}$$

otras condiciones, que $\gamma_m < \infty$, se puede utilizar a

$$\hat{\gamma}_k = \frac{\sum_{t=1}^{N-m} (Z_t - \bar{Z})(Z_{t+m} - \bar{Z})}{N} ; m = 0, 1, \dots \quad (1.1.3')$$

como estimador de γ_m ; por consiguiente, como estimador de ρ_m se utilizará la autocorrelación muestral, que viene dada por

$$\hat{\rho}_m = \frac{\hat{\gamma}_m}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{N-m} (Z_t - \bar{Z})(Z_{t+m} - \bar{Z})}{\sum_{t=1}^N (Z_t - \bar{Z})^2} ; m = 0, 1, \dots \quad (1.1.4')$$

Aunque en la práctica comúnmente no es posible verificar, en base a los datos muestrales, si el proceso es ergódico, si conviene decir que es requisito la estacionariedad para que el proceso pueda ser ergódico.

Es bien sabido que los procesos estacionarios o no estacionarios pueden escribirse ya sea bajo la forma de *Modelos Autorregresivos*

$$\phi_p(B) Z_t = a_t \quad (1.1.6)$$

o bajo la forma de *modelos de Promedios Móviles*

$$Z_t = \theta_q(B) a_t \quad (1.1.7)$$

con $\phi_p(B)$, $\theta_q(B)$, Z_t y a_t definidos como en (1).

La representación (1.1.6) del proceso $\{Z_t\}$ es comúnmente conocida como modelo autorregresivo de orden p y es abreviada $AR(p)$, mientras que la representación (1.1.7) se conoce comúnmente como modelo de promedios móviles de orden q y se abrevia $MA(q)$. Una generalización natural de estos dos modelos son los *Modelos Autorregresivos y de Promedios Móviles* $ARMA(p,q)$ con la representación dada en (1).

1.2 EL MÉTODO DE BOX-JENKINS.

La metodología propuesta por Box y Jenkins (1976) para determinar los valores de p y q de un modelo $ARMA$, consiste básicamente en el reconocimiento de patrones de las FAC teóricas, a partir del comportamiento observado en las FAC muestrales.

1.2.1 Función de autocorrelación de un modelo $AR(p)$

Un proceso $AR(p)$ será estacionario si y sólo si las raíces de la ecuación característica

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0 \quad (1.2.1)$$

se encuentran fuera del círculo unitario, esto es, si en (1)

$U_d(B) = 1$ y por lo tanto (1.1.6) toma la forma

$$\phi_p(B) Z_t = a_t \quad (1.1.6')$$

o

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

Ahora se tiene que

$$\begin{aligned} \gamma_m &= E[Z_t, Z_{t+m}] = \sum_{i=1}^p \phi_i E[Z_t, Z_{t+m-i}] + E[Z_t a_t] \\ &= \sum_{i=1}^p \phi_i \gamma_{k-i} \quad ; \quad \text{si } k > 0 \end{aligned} \quad (1.2.2)$$

Luego entonces, si $\gamma_0 < 0$, se pueden dividir las autocovarianzas $\gamma_1, \gamma_2, \dots, \gamma_p$ entre γ_0 y obtener las llamadas ecuaciones de Yule-Walker

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (1.2.3)$$

las demás autocorrelaciones se pueden obtener de la relación

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} \quad ; \quad m \geq p+1 \quad (1.2.4)$$

que es en realidad otra manera de escribir la ecuación en diferencias

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad ; \quad m \geq p+1$$

cuyas condiciones iniciales están dadas por (1.2.3).

Como el proceso (1.1.6') es estacionario, la FAC de dicho proceso tiende a cero conforme m tiende a infinito, el decaimiento de la FAC depende estrictamente del tipo de raíces que tenga la ecuación característica (1.2.1).

1.2.2 Función de autocorrelación de un modelo MA(q).

El modelo (1.1.7) tiene como media y varianza a

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

y como función de autocovarianza a

$$\gamma_m = E[Z_t Z_{t+m}] = \begin{cases} (-\theta_m + \theta_1 \theta_{m+1} + \dots + \theta_{q-m} \theta_q) \sigma_a^2 & ; \text{si } m = 1, 2, \dots, q \\ 0 & ; \text{si } m \geq q+1 \end{cases}$$

por lo tanto su FAC viene a ser

$$\rho_m = \begin{cases} \frac{-\theta_m + \theta_1 \theta_{m+1} + \dots + \theta_{q-m} \theta_q}{\theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & ; \text{si } m = 1, 2, \dots, q \\ 0 & ; \text{si } m \geq q+1 \end{cases} \quad (1.2.5)$$

lo cual indica que el proceso MA(q) tiene memoria limitada a q periodos.

1.2.3 Función de autocorrelación de un modelo ARMA(p,q)

Para que el proceso autorregresivo y de promedios móviles ARMA(p,q) representado por (1) sea estacionario se requiere que todas las raíces de $\phi_p(x)$ estén fuera del círculo unitario, esto es, $U_d(B) = 1$ y por lo tanto (1) toma la forma

$$\phi_p(B)Z_t = \theta_q(B)a_t \quad (1')$$

o

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

con $\phi_p(B)$, $\theta_q(B)$, a_t y Z_t definidas como en (1).

De (1') se sigue que para $m \geq 0$

$$\gamma_m = E[Z_t Z_{t-m}] = \sum_{i=1}^p \phi_i E[Z_{t-i} Z_{t-m}] - \sum_{j=1}^q \theta_j E[a_{t-j} Z_{t-m}] \quad (1.2.6)$$

obsérvese que Z_t está afectada por los choques aleatorios a_{t-m} , a_{t-m-1} , ..., pero es independiente de a_{t-m+1} , a_{t-m+2} , ..., de donde se sigue que

$$E[a_{t-i} Z_{t-m}] = 0 \quad ; \text{ si } m > i$$

y por lo tanto, (1.2.6) da por resultado

$$\gamma_m = \phi_1 \gamma_{m-1} + \phi_2 \gamma_{m-2} + \dots + \phi_p \gamma_{m-p} \quad ; \text{ para } m > q \quad (1.2.7)$$

mientras que para $m \leq q$, γ_m involucrará a los parámetros $\theta_m, \theta_{m+1}, \dots, \theta_q$. En particular, la varianza está dada por la ecuación

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2 - \theta_1 E[a_t Z_t] - \dots - \theta_q [a_{t-q} Z_t] \quad (1.2.8)$$

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

En lo que respecta a las autocorrelaciones para procesos ARMA(p,q), éstas podrán obtenerse para retrasos mayores a q , de la relación

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} ; \text{para } m > q \quad (1.2.9)$$

y para retrasos menores o iguales a q , la autocorrelación ρ_m involucrará a los parámetros $\theta_m, \theta_{m+1}, \dots, \theta_q$. Lo cual significa que si $q < p$, la FAC estará determinada tan solo por la ecuación en diferencias.

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad ; \text{para } m > p \quad (1.2.10)$$

sujeta a las p condiciones iniciales $\rho_1, \rho_2, \dots, \rho_p$; por otro lado, si $p \leq q$, habrá q autocorrelaciones $\rho_1, \rho_2, \dots, \rho_q$ que no sigan el comportamiento general de la FAC, pero las

autocorrelaciones se comportarán de acuerdo a la ecuación en diferencias

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad ; \text{ para } m > q \quad (1.2.11)$$

y las p condiciones iniciales $\rho_{q-p+1}, \rho_{q-p+2}, \dots, \rho_q$.

Se sabe que la representación de los procesos definidos en (1'), con las condiciones impuestas sobre $\phi_p(B)$ y $\theta_q(B)$, aseguran que a_t es la innovación de $\{Z_t\}$ en el tiempo t . Se dirá que esta representación es MINIMAL, si $\phi(B)$ y $\theta(B)$ no tienen raíces comunes, se sabe también que tal representación minimal existe, es única y es obtenida a partir de cualquier representación del tipo (1') simplemente al eliminar las posibles raíces comunes de $\phi_p(B)$ y $\theta_q(B)$. Las ideas expuestas anteriormente dan la pauta para una primera caracterización de los procesos ARMA(p,q), la cual se llamará "representación minimal ARMA(p,q)". A continuación se da una definición que posteriormente ayudará a la demostración del teorema (1.1), el cual formaliza estas ideas.

Considérese una sucesión de números reales $\{S_m ; m \in \mathbb{Z}\}$ la cual satisface una ecuación en diferencias lineal de orden r para el rango R , esto es, existen $\phi_1, \phi_2, \dots, \phi_r ; \phi_r \neq 0$ tales que

$$S_m - \phi_1 S_{m-1} - \dots - \phi_r S_{m-r} \begin{cases} = 0 & ; \text{ si } m \geq R \\ \neq 0 & ; \text{ si } m = R-1 \end{cases}$$

tal sucesión también satisface otras ecuaciones en diferencias para m mayores que algún rango (posiblemente diferente de R). Entre todas estas ecuaciones, una tiene un orden minimal r_0 . El rango asociado con este orden se denotará por R_0 y la sucesión $\{S_m ; m \in \mathbb{Z}\}$ se dirá que satisface una ecuación en diferencias de orden minimal r_0 para el rango mínimo R_0 .

Usando la definición anterior, se puede dar la primera caracterización de una representación minimal ARMA(p, q) (ver Beguin y Gourieroux (1980) para más detalles).

Teorema 1.2.1 El proceso $\{Z_t\}$ admite una representación minimal ARMA(p, q) si y sólo si la sucesión de autocorrelaciones $\{\rho_m, m \in \mathbb{Z}\}$ satisface una ecuación en diferencias de orden mínimo p , para el rango mínimo $q+1$.

Demostración Ver apéndice 1.

En lo que resta de este trabajo, cuando se dé la representación (1') de los modelos ARMA, se estará pensando en una representación minimal, a menos que se especifique lo contrario.

Del estudio de la función de autocorrelación se puede deducir un criterio para reconocer el proceso MA e identificar su orden : Si $\rho_m = 0$ para $m > q$. Entonces el proceso $\{Z_t\}$ tiene una representación MA(q).

Se trata entonces de encontrar otro criterio de este tipo que permita reconocer los procesos AR. Este criterio se obtendrá a partir de las ecuaciones de Yule-Walker (1.2.3), las cuales pueden reescribirse como

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{bmatrix} = B(n,0) \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{bmatrix}$$

en donde $B(n,m)$ es una matriz de orden $n \times n$ definida por

$$B(n,m) = \begin{bmatrix} \rho_m & \rho_{m-1} & \cdots & \rho_{m-n+1} \\ \rho_{m+1} & \rho_m & \cdots & \rho_{m-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m+n-1} & \rho_{m+n-2} & \cdots & \rho_m \end{bmatrix}$$

Ahora se define la autocorrelación parcial como

$$\phi_{nn} = \frac{|A(n,0)|}{|B(n,0)|} \quad (1.2.12)$$

en donde $A(n,m)$ es una matriz compuesta por las primeras $n-1$ columnas de la matriz $B(n,m)$ y con la n -ésima columna dada por $v = (\rho_{m+1}, \rho_{m+2}, \dots, \rho_{m+n})$. Esto es, si se supone que el proceso es un $ARMA(n,0)$, entonces ϕ_{nn} es la solución de las primeras n ecuaciones de Yule-Walker obtenidas mediante la regla de Cramér, para el último coeficiente autorregresivo.

Si no se conoce n , que es lo que ocurre en la práctica, habrá que hacer el cálculo para $n = 1, 2, \dots$, lo que da la

siguiente relación

$$\phi_{nn} = \begin{cases} \rho_1 & ; \text{ si } n = 1 \\ \frac{|A(n,0)|}{|B(n,0)|} & ; \text{ si } n > 1 \end{cases} \quad (1.2.13)$$

que es conocida como la FACP. Obsérvese que si $\{Z_t\}$ es un AR(p) entonces:

- i) si $n < p$ ϕ_{nn} no será cero en general
- ii) si $n = p$ $\phi_{nn} = \phi_n$ no es cero
- iii) si $n > p$ $\phi_{nn} = 0$

Se puede ver que en el caso de un proceso MA o de un proceso ARMA, la FACP no tiene ninguna razón para ser cero cuando los retrasos sobrepasan un cierto valor. Aunque se observa que en ambos casos, ésta decae en forma exponencial.

De esta manera se ha llegado a un criterio para reconocer los procesos AR e identificar su orden:

Si $\phi_{nn} = 0$ para $n > p$ entonces el proceso $\{Z_t\}$ es un AR(p).

En el caso en donde ni ρ_m ni ϕ_{nn} se vuelven cero para m o n suficientemente grande, habrá que buscar otro tipo de modelo, por ejemplo un ARMA.

En la práctica las correlaciones ρ_m son estimadas, por la relación (1.1.4'). Bartlett (1946) dedujo, en el caso

Gaussiano, la varianza de este estimador. En el caso en el que el cálculo se hace para un MA(q) y para $m > q$ esta varianza vale aproximadamente

$$\text{Var}(\hat{\rho}_m) = \frac{1}{N} \left(1 + 2 \sum_{j=1}^q \rho_j \right) \quad (1.2.14)$$

y puede ser estimada por

$$\widehat{\text{Var}}(\hat{\rho}_m) = \frac{1}{N} \left(1 + 2 \sum_{j=1}^q \hat{\rho}_j \right) \quad (1.2.14')$$

Igualmente, para un AR(p) y para $m > p$, Quenouille (1942) obtuvo la varianza de $\hat{\phi}_{nn}$ y encontró que, aproximadamente

$$\text{Var}(\hat{\phi}_{nn}) = \frac{1}{N} \quad (1.2.15)$$

Además, para N suficientemente grande, $\hat{\rho}_m$ y $\hat{\phi}_{nn}$ tendrán una distribución Gaussiana. Se sabe que el valor absoluto de una v.a Gaussiana tiene una probabilidad de aproximadamente 5% de sobrepasar dos veces su desviación estándar.

Frente a una serie de tiempo, se calculan $\hat{\rho}_m$, $\hat{\phi}_{nn}$, $\widehat{\text{Var}}(\hat{\rho}_m)$, y $\widehat{\text{Var}}(\hat{\phi}_{nn})$, se grafican las $\hat{\rho}_m$, indicando las ordenadas $2\sqrt{\widehat{\text{Var}}(\hat{\rho}_m)}$ y $-2\sqrt{\widehat{\text{Var}}(\hat{\rho}_m)}$. Se podrá considerar que cualquier $\hat{\rho}_m$ comprendida entre esos dos límites, corresponde a un valor cero de ρ_m . Análogamente para la FACP.

se establece que ϕ_{nn} es distinta de cero si el valor calculado

de $\hat{\phi}_{nn}$ se encuentra fuera del intervalo definido por

$$\pm 2\sqrt{\widehat{\text{Var}}(\hat{\phi}_{nn})} \text{ para } n > p.$$

En conclusión, para llevar acabo la etapa de identificación es conveniente recordar las características de la FAC y de la FACP, dichas características se resumen en el cuadro (1.2.1) para una serie estacionaria $\{Z_t\}$ (ver Guerrero, 1983).

CUADRO 1.2.1

COMPORTAMIENTO DE LAS FAC Y FACP PARA PROCESOS AR, MA Y ARMA

PROCESO	FAC	FACP
AR(p)	CONVERGENCIA A CERO, CON COMPORTAMIENTO DICTADO POR LA ECUACION $\phi(B)\rho_k = 0$, PARA $k \geq p$	SOLAMENTE LAS PRIMERAS P AUTOCORRELACIONES PARCIALES SON DISTINTAS DE CERO
MA(q)	SOLO LAS PRIMERAS q AUTOCORRELACIONES SON DISTINTAS DE CERO	SUCESION INFINITA CONVERGENTE A CERO
ARMA(p,q)	COMPORTAMIENTO IRREGULAR DE LAS PRIMERAS q AUTOCORRELACIONES Y DESPUES CONVERGENCIA A CERO DE ACUERDO CON $\phi(B)\rho_k = 0$, PARA $k > q$	SUCESION INFINITA CONVERGENTE A CERO

En resumen, la técnica de Box-Jenkins para identificar

modelos de series de tiempo, se basa en inferencias construidas a partir de patrones de la FAC Muestral (FACM) y de la FACP Muestral de la serie. Esto es, la técnica de Box-Jenkins usa el hecho de que si $\{Z_t\}$ es un $\text{ARMA}(p,0)$, entonces $\phi_{nn} \neq 0$ para $n = p$ y $\phi_{nn} = 0$, para $n > p$, mientras que si $\{Z_t\}$ es un $\text{ARMA}(0,q)$, entonces $\rho_m = 0$ para $m > q$. Dicha técnica emplea la inspección visual de las gráficas de las FACM y FACP. Sin embargo, cuando p y q son ambos diferentes de cero, la FACP y la FAC no poseen gráficas que determinen de manera única los valores de p y q por una simple inspección, en este caso las correspondientes características de las funciones teóricas para varios tipos de modelos $\text{ARMA}(p,q)$ han sido usadas como patrones de referencia (ver por ejemplo Box-Jenkins 1976, cap 3 y Guerrero 1983, cap 4). Desafortunadamente esta técnica de reconocimiento de patrones confía en el buen juicio heurístico del analista de la serie. Entonces, elegir el orden apropiado de un modelo mediante una inferencia que parte de patrones de estas dos funciones, con frecuencia resulta ser muy difícil. De aquí que hayan surgido varios métodos alternativos.

1.3 ARREGLOS "R" Y "S".

Los arreglos R y S tienen su origen en el trabajo de Gray, Houston y Morgan (1978), en el cual se estudia con detalle un nuevo estimador espectral llamado el estimador espectral G. Estos autores ponen de manifiesto que el problema de seleccionar la transformación apropiada G_n , para la estimación espectral, es equivalente al problema de

seleccionar p y q al modelar un proceso $\text{ARMA}(p,q)$. Más aún, dichos autores sugieren y demuestran implícitamente que los arreglos "R" y "S" usados para calcular el estimador espectral G , pueden ser utilizados para determinar los valores de p y q en un modelo $\text{ARMA}(p,q)$. En esta sección se desarrollarán estas ideas con detalle, mostrando cómo sirven para conformar un nuevo criterio que permita la determinación única de p y q , siempre y cuando la FAC teórica sea conocida. También se mostrará cómo, en el caso particular de los modelos no estacionarios, se puede usar este método para identificar y remover la no-estacionariedad debida a las raíces (reales y complejas) que están sobre el círculo unitario.

En la siguiente subsección se presentan algunas definiciones y teoremas necesarios para el desarrollo de las ideas antes mencionadas.

1.3.1 Definiciones y teoremas.

Las siguientes definiciones y teoremas conforman la base para el desarrollo formal de las ideas antes mencionadas. Se pretende pues, en este punto, presentar los conceptos fundamentales que se estarán manejando en el resto de esta sección.

Definición 1.3.1 Sean m, n y h enteros, con $n, h > 0$ y f una función con valores reales que satisface $f_{\mathbf{m}}(mh)$. También

sean

$$H_n[f_m] = \begin{vmatrix} f_m & f_{m+1} & \dots & f_{m+n-1} \\ f_{m+1} & f_{m+2} & \dots & f_{m+n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m+n-1} & f_{m+n} & \dots & f_{m+2n-2} \end{vmatrix}$$

$$H_0[f_m] = 1$$

y

$$H_{n+1}[1; f_m] = \begin{vmatrix} 1 & 1 & \dots & 1 \\ f_m & f_{m+1} & \dots & f_{m+n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m+n-1} & f_{m+n} & \dots & f_{m+2n-1} \end{vmatrix} \quad (1.3.1)$$

Entonces, para $n \in \mathbb{N}$ se define

$$R_n(f_m) = \frac{H_n[f_m]}{H_n[1; f_m]} \quad (1.3.2)$$

y

$$S_n(f_m) = \frac{H_{n+1}[1; f_m]}{H_n[f_m]} \quad (1.3.3)$$

Pye y Atchison (1973) probaron que $R_n[f_m]$ y $S_n[f_m]$ pueden ser calculados recursivamente mediante las siguientes relaciones: hágase

$$S_0(f_m) = 1 \quad \text{para } m \in \mathbb{Z}$$

y

$$R_1(f_m) = f_m \text{ para } m \in \mathbb{Z}$$

entonces se tiene

$$R_{n+1}(f_m) = R_n(f_{m+1}) \left[\frac{S_n(f_{m+1})}{S_n(f_m)} - 1 \right] \quad (1.3.4)$$

y

$$S_n(f_m) = S_{n-1}(f_{m+1}) \left[\frac{R_n(f_{m+1})}{R_n(f_m)} - 1 \right] \quad (1.3.5)$$

para $n \in \mathbb{N}$ y $m \in \mathbb{Z}$.

En (1.3.4) y (1.3.5) se supone que $S_n(f_m) \neq 0$ y $R_n(f_m) \neq 0$. Si $S_n(f_{m+1}) = S_n(f_m) = 0$ se dirá que $R_{n+1}(f_m)$ no está definido. Sin embargo, si $S_n(f_m) = 0$ y $S_n(f_{m+1}) \neq 0$, se define $R_{n+1}(f_m) = \pm \infty$. Se puede dar una definición análoga para $S_n(f_m)$.

Definición 1.3.2 Se dirá que una función f es un elemento de $L(n, \Delta)$ sobre un conjunto de enteros $I = \{m_0, m_0+1, \dots\}$ si existe un conjunto de constantes ϕ_i 's y n es el entero positivo mas pequeño tal que f es solución de la ecuación en diferencias

$$Y_m - \phi_1 Y_{m-1} - \dots - \phi_n Y_{m-n} = 0 \text{ para } m \in I$$

El teorema que se muestra a continuación es la base para el desarrollo de los resultados subsecuentes.

Teorema 1.3.1 Sea $n > 0$ y supóngase que $S_n(f_m)$ y $R_n(f_m)$ en (1.3.2) y (1.3.3) están definidos, con $S_n(f_m) \neq 0$. Entonces $f_m \in L(n, \Delta)$ para $m > m_0 + n$ si y sólo si $S_n(f_m)$, vista como función de m , es constante para $m > m_0$.

Demostración Ver Gray, Houston y Morgan (1978)

1.3.2 Modelos ARMA (p,q) caso estacionario.

En esta subsección se demostrará cómo se puede extender el teorema (1.3.1), de tal forma que se establezca un nuevo criterio para determinar de manera única el valor de p y q , en los modelos ARMA, cuando la FAC teórica es conocida.

Los teoremas (1.3.2) y (1.3.3) (que a continuación se enuncian) muestran que el comportamiento constante de los arreglos S , para ciertos retrasos m de la FAC de un ARMA(p,q) y para un cierto valor n , permite determinar los valores de p y q .

Teorema 1.3.2 Sea $\{Z_t\}$ un proceso estacionario ARMA (p,q) con FAC $\{\rho_m\}$. Supóngase que $S_n(\rho_m)$ y $R_n(\rho_m)$ están definidas, además $p > 0$ y $S_n(\rho_m) \neq 0$. Entonces, para algún entero m_0 y alguna constante $C_1 \neq 0$, se tiene

$$S_n(\rho_m) = C_1 \quad \text{para } m \geq m_0$$

y

$$S_n(\rho_{m_0-1}) \neq C_1 \quad (1.3.6)$$

si y sólo si $n = p$ y $m_0 = q - p + 1$, más aún

$$C_1 = (-1)^p \left[1 - \sum_{i=1}^p \phi_i \right] \quad (1.3.7)$$

Demostración Como $\{Z_t\}$ sigue un modelo ARMA (p, q) , entonces $\{\rho(m)\} \in L(p, \Delta)$ para $m > q + 1 = q - p + 1 + p$, pero no para $m \geq q$. Por lo tanto, la primera parte de este teorema se sigue del teorema (1.3.1). Por otro lado como

$$\rho_m = \sum_{i=1}^n \phi_i \rho_{m-i} \quad \text{para } m \geq q+1$$

al aplicar operaciones elementales en las columnas del siguiente determinante (como sumar las columnas $2, 3, \dots, n$ a la primera, previamente multiplicadas por -1) se tiene que:

$$H_{n+1}[1; \rho_m] = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \rho_m & \rho_{m+1} & \dots & \rho_{m+n} \\ \vdots & \vdots & & \vdots \\ \rho_{m+n-1} & \rho_{m+n} & \dots & \rho_{m+2n-1} \end{vmatrix}$$

$$= (-1)^n \begin{vmatrix} (1 - \phi_1 - \dots - \phi_n) & 1 & \dots & 1 \\ 0 & \rho_{m+1} & \dots & \rho_{m+n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \rho_{m+n} & \dots & \rho_{m+2n-1} \end{vmatrix}$$

$$= (-1)^n (1 - \sum_{i=1}^n \phi_i) H_n[f_m]$$

Entonces

$$C_1 = S_n(\rho_m) = \frac{H_{n+1}[1; \rho_m]}{H_n[\rho_m]}$$

$$= \frac{(-1)^n (1 - \sum_{i=1}^n \phi_i) H_n[\rho_m]}{H_n[\rho_m]}$$

$$= (-1)^n (1 - \sum_{i=1}^n \phi_i) \quad \blacksquare$$

Es conveniente mencionar que Gray, Kelley y McIntire (1978) presentan y demuestran el siguiente teorema, pero una parte de la demostración que ellos dan tiene un error, el cual se comete básicamente en el momento en que ellos aseguran que la condición (1.3.12) se contradice, lo cual no es cierto, pues

en ningún momento se exhibe un entero más pequeño que n tal que (1.3.12) se cumpla. A continuación se enuncia dicho teorema y se presenta su demostración siguiendo la idea original de estos autores, pero corrigiendo el error antes mencionado.

Teorema 1.3.3 Bajo las condiciones del teorema (1.3.2) se tiene que

$$\begin{aligned} S_n(\rho_m) &= C_2 \quad \text{para } m \leq m_1 \\ S_n(\rho_{m+1}) &\neq C_2 \end{aligned} \tag{1.3.8}$$

si y sólo si $n = p$ y $m_1 = -q - p$. Más aún

$$C_2 = - \left[\frac{C_1}{\phi_p} \right] \tag{1.3.9}$$

con C_1 dado en (1.3.7).

Demostración $\Leftarrow$ Como $\{ Z_t \}$ es un proceso estacionario ARMA(p, q), $\{\rho_m\} \in L(p,)$ para $m \geq q+1$. Entonces existen constantes $\phi_1, \dots, \phi_p$, tales que

$$\rho_m = \sum_{i=1}^p \phi_i \rho_{m-i} \quad \text{para } m \geq q+1 \tag{1.3.10}$$

más aún

$$S_p(\rho_{-m}) = \frac{\begin{vmatrix} 1 & 1 & \dots & 1 \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+p} \\ \vdots & \vdots & & \vdots \\ \rho_{-m+p-1} & \rho_{-m+p} & \dots & \rho_{-m+2p-1} \end{vmatrix}}{\begin{vmatrix} \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+p-1} \\ \rho_{-m+1} & \rho_{-m+2} & \dots & \rho_{-m+p} \\ \vdots & \vdots & & \vdots \\ \rho_{-m+p-1} & \rho_{-m+p} & \dots & \rho_{-m+2p-2} \end{vmatrix}} \quad (1.3.11)$$

De (1.3.10) se sigue que el denominador de (1.3.11) es distinto de cero. También de (1.3.10) y como $\rho_m = \rho_{-m}$, se tiene que

$$\begin{aligned} \rho_{-m+p} &= -\frac{\phi_{p-1}}{\phi_p} \rho_{-m+p-1} - \dots - \frac{\phi_1}{\phi_p} \rho_{-m+1} + \frac{1}{\phi_p} \rho_{-m} \\ &\vdots \\ \rho_{-m+2p-1} &= -\frac{\phi_{p-1}}{\phi_p} \rho_{-m+2p-2} - \dots - \frac{\phi_1}{\phi_p} \rho_{-m+p} + \frac{1}{\phi_p} \rho_{-m+p-1} \end{aligned}$$

y

$$H_{p+1}[1; \rho_{-m}] = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+p} \\ \vdots & \vdots & & \vdots \\ \rho_{-m+p-1} & \rho_{-m+p} & \dots & \rho_{-m+2p-1} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & \dots & 1 & 1 + \frac{\phi_{p-1}}{\phi_p} + \dots + \frac{\phi_1}{\phi_p} - \frac{1}{\phi_p} \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+p-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_{-m+p-1} & \rho_{-m+p} & \dots & \rho_{-m+2p-2} & 0 \end{vmatrix}$$

$$= (-1)^p \left(1 + \frac{\phi_{p-1}}{\phi_p} + \dots + \frac{\phi_1}{\phi_p} - \frac{1}{\phi_p} \right) H_p [\rho_{-m}]$$

$$= \frac{(-1)^p}{\phi_p} (\phi_p + \phi_{p-1} + \dots + \phi_1 - 1) H_p [\rho_{-m}]$$

$$= - \frac{C_1}{\phi_p} H_p [\rho_{-m}]$$

entonces

$$S_p(\rho_{-m}) = - \frac{C_1}{\phi_p} = C_2 \quad \text{para } -m \leq -p-q$$

obsérvese que $C_1 \neq 0$ pues el proceso es estacionario.

=> Supóngase ahora que n es el entero positivo más pequeño tal que

$$\begin{aligned} S_n(f_{-m}) &= C_2 \quad ; \quad C_2 \neq 0 \\ \text{y} & \\ S_n(f_{n+1}) &\neq C_2 \quad \text{para } -m \leq M \end{aligned} \quad (1.3.12)$$

Entonces por (1.3.11)

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+n} \\ \vdots & \vdots & & \vdots \\ \rho_{-m+n-1} & \rho_{-m+n} & \dots & \rho_{-m+2n-1} \end{vmatrix}$$

$$= (-1)^n C_2 \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_{-m+n-1} & \rho_{-m+n} & \dots & \rho_{-m+2n-2} & 0 \end{vmatrix}$$

por lo que

$$\begin{vmatrix} 1 & 1 & \dots & 1 & 1 - (-1)^n C_2 \\ \rho_{-m} & \rho_{-m+1} & \dots & \rho_{-m+n-1} & \rho_{-m+n} \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_{-m+n-1} & \rho_{-m+n} & \dots & \rho_{-m+2n-2} & \rho_{-m+2n-1} \end{vmatrix} = 0 \quad (1.3.13)$$

Esto es, un determinante como (1.3.13) puede exhibirse para toda m tal que $-m \leq M$. Entonces, usando la simetría de ρ , para cada m existen constantes $b_1(m), b_2(m), \dots, b_n(m)$ tales que

$$b_1(m) \begin{bmatrix} 1 \\ \rho_m \\ \vdots \\ \rho_{m-n+1} \end{bmatrix} + \dots + b_n(m) \begin{bmatrix} 1 \\ \rho_{m-n+1} \\ \vdots \\ \rho_{m+2n+2} \end{bmatrix} + \begin{bmatrix} 1 - (1)^n C_2 \\ \rho_{m-n} \\ \vdots \\ \rho_{m-2n+1} \end{bmatrix} = 0$$

Entonces se obtiene

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \rho_m & \rho_{m-1} & \dots & \rho_{m-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m-n+2} & \rho_{m-n+1} & \dots & \rho_{m-2n+3} \end{bmatrix} \begin{bmatrix} b_1(m) - b_1(m+1) \\ \vdots \\ \vdots \\ b_n(m) - b_n(m+1) \end{bmatrix} = 0 \quad (1.3.14)$$

Luego, el determinante de la matriz $n \times n$ de (1.3.14) es cero o el vector es nulo. Por otra parte, obsérvese que el determinante de la matriz $n \times n$ de (1.3.14) es el numerador de $S_{n-1}(f_{-m})$, pero por (1.3.5) se sabe que el determinante de esta matriz es distinto de cero, de donde se tiene que:

$$b_i(m) = b_i(m+1) \quad \text{para } i=1,2,\dots,n \text{ y } m \geq -M$$

por lo tanto, las constantes $b_i(m)$ no dependen de m y se tiene que

$$\rho_{m-n+1} = \sum_{i=1}^p \phi_i \rho_{m-n+1-i} \quad \text{para } m-n+1 \geq -M-n+1$$

luego entonces $n = p$ y $M = -q - p$ ■

Ahora se mostrará que estos resultados no varían si en lugar de usar $f_m = \rho_m$ en los teoremas anteriores, ésta es remplazada por $f_m = (-1)^m \rho_m$. La razón por la cual se hace este remplazo es porque $\{(-1)^m \rho_m\}$ usualmente produce una mejor identificación de los valores de p y q para datos de baja frecuencia, mientras que $\{\rho_m\}$ produce una mejor identificación para datos de alta frecuencia. En el caso de los modelos ARMA esto es muy razonable, pues se puede ver que la constante C_1 de la relación (1.3.7) es, excepto por un signo, un factor del denominador de la Densidad Espectral de un ARMA(p, q) evaluada en $w = 0$. En donde la Densidad Espectral está dada por

$$f_z(w) = \frac{\sigma^2}{2\pi} \frac{\left[\sum_{j=0}^q \theta_j e^{i w j} \right] \left[\sum_{l=0}^q \theta_l e^{-i w l} \right]}{\left[\sum_{k=0}^p \phi_k e^{i w k} \right] \left[\sum_{k=0}^p \phi_k e^{-i w k} \right]}$$

(ver Fuller 1976, pag 146). Para datos de alta frecuencia el valor de la Densidad Espectral en $w = 0$ es relativamente pequeño y el denominador es relativamente grande, por lo que C_1 es grande. Para datos de baja frecuencia, la situación es inversa y el valor de la Densidad Espectral en $w = 1/2$ la hace más interesante, pues C_3 (definida posteriormente) es una constante análoga a C_1 , pero para la sucesión $\{(-1)^m \rho_m\}$ y es, excepto por un signo, un factor del denominador de la Densidad Espectral evaluada en $w = 1/2$.

Si $\{Z_t\}$ es una serie de tiempo con FAC $\{\rho_m\}$, entonces la serie $\{Y_t\} = \{(-1)^t Z_t\}$ tiene FAC $\{(-1)^m \rho_m\}$. En particular si $\{Z_t\}$ sigue un proceso ARMA (p, q) la sustitución directa ϕ_k por $(-1)^k \phi_k$, θ_k por $(-1)^k \theta_k$ y $a_t = (-1)^t W_t$ muestran que

$\{Y_t\}$ satisface la misma ecuación en diferencias que $\{Z_t\}$. Por lo tanto, las demostraciones de los siguientes teoremas se siguen inmediatamente de los teoremas (1.3.2) y (1.3.3) respectivamente.

Teorema 1.3.4 Supóngase que las condiciones del teorema (1.2.2) se cumplen con ρ_m remplazada por $(-1)^m \rho_m$. Entonces, para algun entero m_0 y alguna constante $C_3 \neq 0$,

$$\begin{aligned} S_n((-1)^m \rho_m) &= C_3 \quad \text{para } m \geq m_0 \\ S_n((-1)^{m-1} \rho_{m-1}) &\neq C_3 \end{aligned} \quad (1.3.15)$$

si y sólo si $n = p - m_0 + 1$, más aún

$$C_3 = (-1)^p \left(1 - \sum_{i=1}^p (-1)^i \phi_i \right) \quad (1.3.16)$$

Teorema 1.3.5 Bajo las condiciones del teorema (1.3.4)

$$S_n((-1)^m \rho_m) = C_4 \quad \text{para } m \leq m_1$$

$$S_n((-1)^{m_1+1} \rho_{m_1+1}) \neq C_4$$

si y sólo si $n = p$ y $m_1 = -q - p$, más aún

$$C_4 = \frac{(-1)^{p+1} C_3}{\phi_p}$$

Corolario 1.3.1 Bajo las condiciones del teorema (1.3.2)

$$R_{n+1}(\rho_m) = 0, \quad \text{para } m \geq m_0 \text{ y } m \leq m_1$$

$$R_{n+1}((-1)^m \rho_m) = 0 \quad \text{para } m \geq m_0 \text{ y } m \leq m_1$$

y

$$R_{n+1}(\rho_{q-p}) \neq 0$$

$$R_{n+1}(\rho_{-q-p}) \neq 0$$

$$R_{n+1}((-1)^{q-p} \rho_{q-p}) \neq 0$$

$$R_{n+1}((-1)^{-q-p+1} \rho_{-q-p}) \neq 0$$

si y sólo si $n = p$, $m_0 = q - p + 1$ y $m_1 = -q - p - 1$.

Demostración De la relación (1.3.4) y de los teoremas (1.3.2), (1.3.4), (1.3.3), y (1.3.5) se sigue que :

$$R_{n+1}(\rho_m) = R_n(\rho_{m+1}) \left[\frac{C_1}{C_1} - 1 \right] = 0$$

$$R_{n+1}((-1)^m \rho_m) = R_n((-1)^m \rho_{m+1}) \left[\frac{C_3}{C_3} - 1 \right] = 0 \quad \text{para } m \geq m_0$$

$$R_{n+1}(\rho_m) = R_n(\rho_{m+1}) \left[\frac{C_2}{C_2} - 1 \right] = 0$$

$$R_{n+1}((-1)^m \rho_m) = R_n((-1)^m \rho_{m+1}) \left[\frac{C_4}{C_4} - 1 \right] = 0 \quad \text{para } m \leq m_1$$

Por último, también de la relación (1.3.4) y de los teoremas (1.3.2), (1.3.4), (1.3.3) y (1.3.5) se sigue que:

$$R_{n+1}(\rho_{q-p}) = R_n(\rho_{q-p+1}) \left[\frac{C_1}{S_n(\rho_{q-p})} - 1 \right] \neq 0$$

$$R_{n+1}((-1)^{q-p} \rho_{q-p}) = R_n((-1)^{q-p+1} \rho_{q-p+1}) \left[\frac{C_3}{S_n((-1)^{q-p} \rho_{q-p})} - 1 \right] \neq 0$$

$$R_{n+1}(\rho_{-q-p}) = R_n(\rho_{-q-p+1}) \left[\frac{S_n(\rho_{-q-p+1})}{C_2} - 1 \right] \neq 0$$

$$R_{n+1}((-1)^{-q-p} \rho_{-q-p}) = R_n((-1)^{-q-p+1} \rho_{-q-p+1})$$

$$\left[\frac{S_n((-1)^{-q-p+1} \rho_{-q-p+1})}{C_4} - 1 \right] \neq 0$$

respectivamente ■

Para mostrar la manera como los resultados anteriores pueden ser utilizados en la identificación del orden de un modelo estacionario ARMA (p,q), se construirán arreglos de números, los cuales se llamarán arreglos "R" y "S". Estos serán calculados mediante el programa que se muestra en el Apéndice 2 y se ordenarán como se muestra en las tablas 1.3.1 y 1.3.2. Además, ya que ρ_n se utiliza en todos los casos, se acortará la notación, de tal forma que ahora:

$$R_n(\rho_m) = R_n(m) \quad Y \quad S_n(\rho_m) = S_n(m).$$

Aunque el corolario (1.3.1) caracteriza tanto a p como a q , se usará sólo en raras ocasiones para determinar el valor de p , ya que en el momento de sustituir la FAC teórica $\{\rho_m\}$ por la FACM $\{\rho_m\}$ definida en la relación (1.1.4'), el comportamiento de las columnas de los arreglos R resulta ser muy inestable. En este caso, se identificará el valor de p , como el de la primera columna de los arreglos S (tabla 1.3.1) que tenga el comportamiento descrito por los teoremas (1.3.2) y (1.3.3). En contraste, se mostrará con algunos ejemplos que el comportamiento de los arreglos R (tabla 1.3.2) es más adecuado para identificar el valor de q .

T A B L A 1.3.1
VALORES $S_n(f_m)$

n m	1	2	. . .	k
-l	$S_1(-l)$	$S_2(-l)$	. . .	$S_k(-l)$
-l+1	$S_1(-l+1)$	$S_2(-l+1)$	. . .	$S_k(-l+1)$
.	.	.	.	.
.	.	.	.	.
-1	$S_1(-1)$	$S_2(-1)$	. . .	$S_k(-1)$
0	$S_1(0)$	$S_2(0)$	. . .	$S_k(0)$
1	$S_1(1)$	$S_2(1)$	. . .	$S_k(1)$
.	.	.	.	.
.	.	.	.	.
j	$S_1(j)$	$S_2(j)$	. . .	$S_k(j)$

093895

T A B L A 1.3.2
VALORES $R_n(f_m)$

$\begin{matrix} n \\ m' \end{matrix}$	1	2	. . .	k+1
-1	$R_1(-1)$	$R_2(-1)$	. . .	$R_{k+1}(-1)$
-1+1	$R_1(-1+1)$	$R_2(-1+1)$	. . .	$R_{k+1}(-1+1)$
.	.	.	.	.
.	.	.	.	.
-1	$R_1(-1)$	$R_2(-1)$	. . .	$R_{k+1}(-1)$
0	$R_1(0)$	$R_2(0)$	. . .	$R_{k+1}(0)$
1	$R_1(1)$	$R_2(1)$	. . .	$R_{k+1}(1)$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
j	$R_1(j)$	$R_2(j)$	. . .	$R_{k+1}(j)$

En términos de los arreglos R y S, de las tablas 1.3.1 y 1.3.2 (con $f_m = \rho_m$), los teoremas (1.3.2) y (1.3.3), junto con el corolario (1.3.1), aseguran que $\{Z_t\}$ es un proceso estacionario ARMA(p,q) si y sólo si los arreglos R y S asociados con éste, tienen el patrón que se muestra en las tablas 1.3.3 y 1.3.4.

En las tablas 1.3.3 y 1.3.4 se tiene que $f_m = \rho_m$, pero los arreglos no difieren en sus patrones si ρ_m es remplazada por $f_m = (-1)^m \rho_m$, sólo obsérvese que en este caso la columna 1 de los arreglos R se compone de los valores dictados por $(-1)^m \rho_m$.

En la tabla 1.3.3 se incluye la columna $p+1$ aunque en ella aparezcan varios elementos indefinidos. Tales elementos estarán indicados por un valor de la forma $[\frac{0}{0} -1]$. Esto es, al momento de sustituir ρ_m por $\hat{\rho}_m$, se espera que la columna que tenga las características de la columna p , vaya seguida de una columna de alta variabilidad, pues en la práctica la cantidad $\frac{0}{0}$ será remplazada por un cociente de números pequeños. Por otra parte, de la relación (1.3.4) se sigue que el comportamiento de ceros, observado en la columna $p+1$ de los arreglos R , se mantendrá en las columnas $p+2, p+3, \dots$

TABLA 1.3.3

$$S_n(\rho_m)$$

COMPORTAMIENTO DE LOS ARREGLOS S PARA UN PROCESO ARMA(p, q)

n m'	1	. . . p	$p+1$
-1	$S_1(-1)$	. . . C_2	$C_2[\frac{0}{0} -1]$
-1+1	$S_1(-1+1)$	. . . C_2	$C_2[\frac{0}{0} -1]$
.	.	.	.
.	.	.	.
-q-p-2	$S_1(-q-p-2)$	. . . C_2	$C_2[\frac{0}{0} -1]$
-q-p-1	$S_1(-q-p-1)$	. . . C_2	$\pm \infty$
-q-p	$S_1(-q-p)$	. . . C_2	$\left. \begin{array}{l} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right\} \begin{array}{l} 2q \text{ termino} \\ \text{no} \\ \text{constantes} \end{array}$
.	.	.	
.	.	.	
.	.	.	
q-p	$S_1(q-p)$	.	$-C_1$
q-p+1	$S_1(q-p+1)$	. . . C_1	$C_1[\frac{0}{0} -1]$
.	.	.	.
.	.	.	.
j	$S_1(j)$	. . . C_1	$C_1[\frac{0}{0} -1]$

T A B L A 1.3.4

$$R_n(\rho_m)$$

COMPORTAMIENTO DE LOS ARREGLOS R PARA UN PROCESO ARMA(p,q)

$\begin{matrix} n \\ m \end{matrix}$	1	2	. . .	p+1
-1	$\rho(-1)$	$R_2(-1)$	. . .	0
-1+1	$\rho(-1+1)$	$R_2(-1+1)$	. . .	0
.	.	.	.	.
.	.	.	.	.
-q-p-2	$\rho(-q-p-2)$	$R_2(-q-p-2)$	. . .	0
-q-p-1	$\rho(-q-p-1)$	$R_2(-q-p-1)$	. . .	0
-q-p	$\rho(-q-p)$	$R_2(-q-p)$	. . .	* 0
.	.	.	.	.
.	.	.	.	.
q-p	$\rho(q-p)$	$R_2(q-p)$	. . .	* 0
q-p+1	$\rho(q-p+1)$	$R_2(q-p+1)$	. . .	0
.	.	.	.	.
.	.	.	.	.
j	$\rho(j)$	$R_2(j)$	. . .	0

1.3.3 El estadístico $D(n,m)$

Hasta ahora, el procedimiento de identificación que se ha propuesto, está basado en el reconocimiento de un patrón en los arreglos R y S. En esta subsección se presenta un estadístico que fué propuesto por Gray, Kelley y McIntire (1978) para medir la concordancia del patrón muestral observado, con el patrón

esperado para un proceso ARMA(p,q).

Al inspeccionar las tablas 1.3.3 y 1.3.4 se produce la motivación para la siguiente definición, en la cual se hace uso de las relaciones (1.3.4) y (1.3.5).

Definición 1.3.3 Para cualesquier par de enteros m y n tales que $n = 1, 2, \dots, N$ y $m = 0, 1, \dots, M$ se define

$$D(n, m) = v/u$$

en donde

$$v = \left| \frac{S_{n+1}(f_{-m-n-1})}{S_n(f_{-n-m})} \right|$$

y

$$u = \left[\frac{S_n(f_{m-n+1}) + S_{n+1}(f_{m-n})}{S_{n-1}(f_{m-n+2}) + S_n(f_{m-n+1})} \right]^2 + \sum_{i=1}^3 \left[\frac{R_{n+1}(f_{-m-n-1})}{R_n(f_{-m-n-i+1})} \right]^2$$

$$+ \left[\frac{R_{n+1}(f_{m-n+1})}{R_n(f_{m-n+1+1})} \right]^2$$

$$\text{con } S_{-1}(f_m) = 0 \quad \text{y} \quad R_0(f_m) = 1$$

De acuerdo a la definición del estadístico D , indudablemente se deja de usar parte de la información, sin embargo, como se

mostrará con algunos ejemplos, dicho estadístico resulta ser suficientemente sensible para reaccionar ante diferentes valores de p y q , y de esta manera ayuda a la identificación del modelo ARMA más adecuado para representar a la serie observada. Por otro lado, debe indicarse que no se conoce la distribución del estadístico D y por ello no es posible realizar pruebas de significación estadística.

Aunque no queda claro a partir de las tablas 1.3.3 y 1.3.4, en las relaciones (1.3.4) y (1.3.5), sí se puede observar fácilmente que el comportamiento de ceros que aparece en la columna $p+1$ de los arreglos R , tiende a continuar en las columnas $p+2$, $p+3$,... debido a la multiplicación por R_n en el cálculo de R_{n+1} . También a partir de estas relaciones, se puede mostrar, que los efectos "infinito y $-C_1$ " que aparecen en la columna $p+1$ de los arreglos S , tienden a continuar en las columnas $p+2$, $p+3$,... Sin embargo, estos efectos se nulifican en $D(n,m)$ por las divisiones que se hacen en el numerador y en el denominador; de aquí que un valor grande de $D(n,m)$ sugiera la concordancia con un modelo ARMA (p,q) . Por lo tanto, la estimación de p y q basada en D , se define mediante la siguiente regla : seleccionar el modelo ARMA (p,q) cuando

$$D(p,q) = \max D(n,m) \quad ; \quad n = 1, \dots, N \quad ; \quad m = 0, 1, \dots, M.$$

Es siempre recomendable la inspección de los arreglos R y S aun cuando el estadístico $D(n,m)$ sea distintivo, ya que el estadístico D es suficiente sólo si se asegura la estacionariedad del proceso. Sin embargo, si el proceso es no-estacionario o estacional, en el sentido de que existen columnas en los arreglos R tengan números "cercaños a cero"

intercalados entre valores no cercanos a cero, el estadístico D no será de utilidad. Por lo tanto, se recomienda alternar el uso del estadístico D con una inspección visual de los arreglos R y S.

1.3.4 Modelos ARMA (p,q) caso no-estacionario

Nótese que hasta ahora se ha centrado la exposición en modelos que son estacionarios, pero en la práctica lo más común es que las series que se analizan sean no estacionarias, ya sea porque exhiben algún tipo de tendencia, por su varianza no constante o porque estén influidas por algún factor de tipo semideterminista como puede ser la estacionalidad. No obstante, si el modelo ARMA (p,q) es no estacionario, podría transformarse a un modelo estacionario mediante la aproximación de un modelo ARIMA (p,d,q). Esto es, se observa la autocorrelación muestral y si ésta no tiene un decaimiento rápido se aplica el operador diferencia en los datos hasta producir en la autocorrelación muestral un decaimiento suficientemente rápido. Por otra parte, si el modelo es un ARMA (p,q), con raíces de su ecuación característica sobre el círculo unitario, la diferenciación en los datos puede no producir estacionariedad. En efecto, es claro que para tales modelos el procedimiento será adecuado sólo cuando la unidad es "casi una solución" de la ecuación característica del polinomio autorregresivo.

En esta subsección se generalizarán los teoremas de las subsecciones anteriores para lograr la identificación de un modelo no estacionario ARMA(p,q). La siguiente definición y el siguiente teorema son útiles para el desarrollo de estos

resultados e indican además en qué sentido se presenta la no-estacionariedad.

Definición 1.3.4 Se dirá que un modelo ARMA (p,q) es puramente no-estacionario si todas las raíces de su ecuación característica están sobre el círculo unitario.

Teorema 1.3.6 Sean $r_1, r_2, \dots, r_p$ las raíces de la ecuación característica $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$ de un modelo ARMA (p,q), y supóngase que los ϕ_i 's son tales que $|r_i| \geq 1$; para $i = 1, 2, \dots, p$. Entonces, $|\phi_p| = 1$ si y sólo si el modelo es puramente no estacionario.

Demostración Ya que

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = \prod_{i=1}^p (1 - r_i z) = \prod_{i=1}^p (z_1 - r_i)$$

se tiene que

$$\phi_p = (-1)^{p+1} \prod_{i=1}^p r_i^{-1}$$

entonces

$$|\phi_p| = \prod_{i=1}^p |r_i^{-1}|$$

Pero como $|r_i| \geq 1$ para toda i , entonces si $|\phi_p| = 1$ se sigue que $r_i = 1$ para toda i , por lo tanto el modelo es

no-estacionario. Por otra parte, si el modelo es puramente no-estacionario, entonces los $r_i = 1$ para toda i y por lo tanto,

$$|\phi_p| = 1$$

A fin de interpretar la FAC cuando el modelo es no-estacionario, se da la siguiente definición

Definición 1.3.5 Se denotará por $\rho_m(\alpha_1, \alpha_2, \dots, \alpha_p)$ la autocorrelación de un modelo ARMA (p, q) con raíces de la ecuación característica α_i , tales que $|\alpha_i| > 1$. Sean $r_1, r_2, \dots, r_p$ las raíces de la ecuación característica de un modelo ARMA (p, q) dado, con j de estas raíces sobre el círculo unitario, las cuales se denotarán por $r_1, r_2, \dots, r_j$. Entonces, se define

$$\rho_m^* = \begin{cases} \rho_m(r_1, r_2, \dots, r_p) & \text{si } j = 0 \\ \lim_{|\alpha_1|, \dots, |\alpha_j| \rightarrow 1} \rho_m(\alpha_1, \alpha_2, \dots, \alpha_j, r_{j+1}, \dots, r_p) & \text{si } j = 1, \dots, p \end{cases}$$

en donde $|\alpha_1| = |\alpha_2| = \dots = |\alpha_j|$

La definición anterior es compatible con la idea de que en la práctica, las raíces cercanas al círculo unitario e igualmente espaciadas, podrían ser clasificadas como unitarias. En lo que resta de esta sección se hará a $\rho_m^* = \rho_m$, de tal

manera que ρ_m está bien definida, tanto para modelos estacionarios como para no-estacionarios.

En este momento se tienen las bases suficientes para probar el siguiente resultado básico, que ayudará a detectar la no estacionariedad en los datos.

Teorema 1.3.7 Sea $\rho_m = \rho_m^*$ y supóngase que $|r_1| \geq 1$. Supóngase también que existe un número l que sea el entero más pequeño tal que $S_l(\rho_m) = C_f$ para $m \in \mathbb{Z}$, en donde C_f es una constante que depende de que $f_m = \rho_m$ o $f_m = (-1)^m \rho_m$. Entonces, el modelo es no-estacionario.

Demostración Sin pérdida de generalidad, supóngase que $f_m = \rho_m$. Entonces, si $S_l(m) = C_l$ con $C_l \neq 0$ para $m \in \mathbb{Z}$, se sigue del teorema 8 de Gray, et al (1978) que ρ_m satisface una ecuación en diferencias de orden l , esto es, existen constantes

$\phi_1, \phi_2, \dots, \phi_l$ tales que

$$\rho_m = \phi_1 \rho_{m-1} + \dots + \phi_l \rho_{m-l} \quad \text{para } m \in \mathbb{Z}$$

Pero, como $S_l(m) \equiv C_f$ con $C_f \neq 0$, se sigue de los teoremas (1.3.2) y (1.3.3) que $\phi_1 = -1$, entonces por el teorema (1.3.6) se sabe que la ecuación en diferencias para ρ_m tiene la ecuación característica de un modelo puramente no-estacionario; por lo tanto, $j \geq 1$ en la definición (1.3.5) y el modelo es no estacionario. Por otra parte, si $C_f = 0$, ya que l es minimal

se sigue que $H_{1+1}(1;m) = 0$. Sin embargo, la demostración del teorema 8 de Gray, et al (1978) implica que $1 - \sum_{i=1}^1 \phi_i = 0$, lo cual ocurre si y sólo si existe una raíz de la ecuación característica sobre el círculo unitario, por lo tanto $j \geq 1$ en la definición (1.3.6) y el modelo es no-estacionario. ■

Las raíces de la ecuación característica "cercanas" al círculo unitario traen consigo un problema de índole práctico, pues un modelo que matemáticamente sería estacionario, en la práctica no presenta el comportamiento esperado de un modelo estacionario, en el sentido de los arreglos S. Por consiguiente, con fines prácticos es conveniente clasificar a dichas raíces como unitarias y por lo tanto, considerara los modelos respectivos como no-estacionarios. Esta idea queda formalizada en el siguiente teorema.

Teorema 1.3.8 Supóngase que un modelo ARMA (p,q) tiene K raíces de su ecuación característica "cercanas" al círculo unitario, de acuerdo con la definición (1.3.5), N de las cuales son distintas. Entonces, para $m \in \mathbb{Z}$, ρ_m^* satisface una ecuación en diferencias de orden menor o igual que N, cuya ecuación característica tiene todas sus raíces distintas y sobre el círculo unitario.

Demostración (Este teorema se demuestra sólo para el caso de un modelo ARMA (3,1), pues la demostración para modelos de orden mayor encierra un tedioso trabajo algebraico, que no aporta nada a una mejor comprensión)

Teorema 1.3.8a Sean r_1 , r_2 y r_3 las raíces de la ecuación característica de un modelo ARMA (3,1), o sea $1 - \phi_1 z - \phi_2 z^2 - \phi_3 z^3 = 0$, en el cual M de las raíces r_1 , r_2 y r_3 , con $1 \leq M \leq 3$, están sobre el círculo unitario. Sea k el entero tal que $1 \leq k \leq M$, es el número de raíces distintas que tienen módulo 1. Entonces, ρ_m^* definida como en la definición (1.3.5), satisface una ecuación en diferencias lineal y homogénea con coeficientes constantes cuyo orden es menor o igual que k , para $m \in \mathbb{Z}$, y cuya ecuación característica tiene todas sus raíces distintas y sobre el círculo unitario.

Demostración Para un modelo ARMA (3,1) se tiene

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \phi_3 \rho_{m-3} \quad \text{con } m \geq 2 \quad (1.3.17)$$

la solución general de esta ecuación en diferencias está dada por

$$\rho_m = \begin{cases} A_1 r_1^{-|m|} + A_2 r_2^{-|m|} + A_3 r_3^{-|m|} & \text{para } r_1 \neq r_2, r_2 \neq r_3, r_1 \neq r_3 \\ (B_1 + B_2 m) r_1^{-|m|} + B_3 r_3^{-|m|} & \text{para } r_1 = r_2 \neq r_3 \\ (C_1 + C_2 m + C_3 m^2) r_1^{-|m|} & \text{para } r_1 = r_2 = r_3 \end{cases} \quad (1.3.18)$$

En donde A_1 , A_2 , A_3 , B_1 , B_2 , B_3 , C_1 , C_2 y C_3 se determinan con base en las condiciones iniciales $\rho_0 = 1$,

$$\rho_1 = \frac{\phi_1 + \phi_2 \phi_3}{1 - \phi_2 - \phi_3(\phi_1 + \phi_3)} + \theta_1 k_1$$

sean

$$A = \phi_1 + \phi_2 \phi_3 + (\phi_2 + \phi_1 \phi_3) (\phi_1 \phi_3)$$

y

$$B = 1 - \phi_2 - \phi_3(\phi_1 - \phi_3)$$

así que

$$k_1 = \frac{(\phi_1 + \phi_2 \phi_3)A - [1 - \phi_3^2 - \phi_2(\phi_2 + \phi_1 \phi_3)]B}{B \{ B [1 - \theta_1(\phi_1 - \theta_1)] - \theta_1 A \}}$$

y

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 + \phi_3 \rho_{-1} = (\phi_1 + \phi_3) \rho_1 + \phi_2 \cdot$$

Ahora, sean $g_1 = r_1^{-1}$, $g_2 = r_2^{-1}$ y $g_3 = r_3^{-1}$. Entonces, las constantes dadas en la relación (1.3.18) son

$$A_1 = \frac{(g_2 - g_3)[\rho_1(g_2 + g_3) - \rho_2 - \rho_0 g_2 g_3]}{D},$$

$$A_2 = \frac{(g_3 - g_1)[\rho_1(g_3 + g_1) - \rho_2 - \rho_0 g_1 g_3]}{D},$$

$$A_3 = \frac{(g_1 - g_2)[\rho_1(g_1 + g_2) - \rho_2 - \rho_0 g_2 g_1]}{D}$$

en donde

$$D = (g_1 - g_2)(g_1 - g_3)(g_3 - g_2)$$

$$B_1 = \frac{g_3^2 + 2 \rho_1 g_1 - 2 g_1 g_3 - \rho_2}{(g_1 - g_3)^2},$$

$$B_2 = \frac{\rho_2 - \rho_1(g_1 + g_3) + g_1 g_3}{g_1(g_1 - g_3)^2},$$

$$B_3 = \frac{\rho_2 - 2 g_1 \rho_1 + g_1^2}{(g_1 - g_3)^2}, \quad C_1 = 1$$

$$C_2 = \frac{4 \rho_1 g_1 - \rho_2 - 3 g_1^2}{2 g_1^2}$$

$$C_3 = \frac{\rho_2 - 2 \rho_1 g_1 + g_1^2}{2 g_1^2}$$

Finalmente, recuérdese que los valores g_1 , g_2 y g_3 están relacionados con los coeficientes ϕ_1 , ϕ_2 y ϕ_3 mediante las ecuaciones

$$\phi_1 = g_1 + g_2 + g_3; \quad \phi_2 = -(g_1 g_2 + g_1 g_3 + g_2 g_3) \quad \text{y} \quad \phi_3 = g_1 g_2 g_3$$

Esencialmente, la demostración de este teorema se basa en expresiones apropiadas de la solución general de la ecuación (1.3.17) en terminos de g_1 , g_2 y g_3 , y tomando el límite de la solución general sobre las raíces especificadas, próximas al círculo unitario. En cada caso, cada una de las raíces próximas al círculo unitario son escogidas arbitrariamente, teniéndose los resultados siguientes.

Caso 1. Todas las raíces son reales, con $M = 1$ y $k = 1$. Considérese $g_1 \rightarrow 1$, mientras que g_2 y g_3 permanecen fijas con $|g_2| < 1$ y $|g_3| < 1$. Entonces, como $\lim_{g_1 \rightarrow 1} \rho_1 = 1$ y

$$\lim_{g_1 \rightarrow 1} \rho_2 = 1 \quad \text{se tiene que} \quad \lim_{g_1 \rightarrow 1} A_1 = 1, \quad \lim_{g_1 \rightarrow 1} A_2 = 0 \quad \text{y}$$

$\lim_{g_1 \rightarrow 1} A_3 = 0$. Así, la FAC del modelo está dada por $\rho_m^* = 1$

para $m \in \mathbb{Z}$; esto es, ρ_m^* es solución de una ecuación en diferencias de primer orden.

Caso 2. Todas las raíces son reales, con $M = 2$ y $k = 1$. Supóngase que g_1 tiene multiplicidad 2 y tómesese el límite cuando $g_1 \rightarrow 1$, en donde g_3 permanece fija con $|g_3| < 1$. Los resultados son similares a los del caso 1

Caso 3. Todas las raíces son reales, con $M = 3$ y $k = 1$. Supóngase que g_1 es de multiplicidad 3. El resultado es similar al caso 1

Caso 4 Dos raíces son complejas y una es real, con $M = 2$ y $k = 2$. Sean $g_1 = x + iy = d \exp(i\theta)$ y $g_2 = x - iy$

$= d \exp(-i\theta)$ en donde $d = \sqrt{x^2 + y^2}$, $i^2 = -1$ y θ es el ángulo medido en radianes, que satisface la relación $\theta = \tan^{-1} \frac{y}{x}$. Supóngase que g_3 está fija con $|g_3| < 1$, entonces

como el $\lim_{d \rightarrow 1} \rho_1 = \cos \theta$ y $\lim_{d \rightarrow 1} \rho_2 = \cos 2\theta$, se tiene

que $\lim_{d \rightarrow 1} A_1 = \lim_{d \rightarrow 1} A_2 = \frac{1}{2}$ y $\lim_{d \rightarrow 1} A_3 = 0$. Así, la

solución general de la ecuación en diferencias se transforma en

$$\rho_m^* = \frac{1}{2} \exp(i\theta|m|) + \frac{1}{2} \exp(-i\theta|m|)$$

$$= \frac{1}{2} [\cos \theta m + i \sin \theta m] + \frac{1}{2} [\cos \theta m - i \sin \theta m]$$

$$= \cos \theta m \quad \text{para } m \in \mathbb{Z}$$

lo cual es solución de la ecuación en diferencias de segundo orden.

Caso 5 Todas las raíces son reales, con $M = 2$ y $k = 2$. En este caso se tiene que $g_1 \rightarrow 1$, $g_2 \rightarrow -1$ con g_3 fija y $|g_3| < 1$. Aquí los límites son calculados con la restricción de que $g_2 = -g_1$. Entonces, como

$$\rho_1^* = \lim_{g_1 \rightarrow 1} \rho_1 = \frac{2g_3}{1+g_3^2} - \frac{2\theta_1(1-g_3^2)^2}{(1+g_3)^2[(1-\theta_1)(g_3-\theta_1)-g_3\theta_1(3-g_3^2)]}$$

$$\text{y } \lim_{g_1 \rightarrow 1} \rho_2 = 1, \text{ se tiene que } \lim_{g_1 \rightarrow 1} A_1 = \frac{1}{2}(\rho_1^* + 1),$$

$$\lim_{g_1 \rightarrow 1} A_2 = \frac{1}{2}(1 - \rho_1^*) \quad \text{y} \quad \lim_{g_1 \rightarrow 1} A_3 = 0. \quad \text{Así, la}$$

solución general de la ecuación en diferencias se transforma en

$$\rho_m^* = \frac{1}{2}[(\rho_1^* + 1) + (-1)^{|m|}(1 - \rho_1^*)] = \begin{cases} 1 & \text{si } m \text{ es par} \\ \rho_1^* & \text{si } m \text{ es impar} \end{cases}$$

Esto es, $\rho_m^* = \rho_{m-2}^*$ para $m \in \mathbb{Z}$, lo cual implica que ρ_m^* satisface una ecuación en diferencias de segundo orden con raíces ± 1 .

Caso 6 Todas las raíces son reales, con $M = 3$ y $k = 2$. Aquí g_1 es de multiplicidad 2, $g_1 \rightarrow 1$, $g_2 \rightarrow -1$ y los límites son tomados con la restricción de que $g_3 = -g_1$.

$$\text{Entonces, como } \lim_{g_1 \rightarrow 1} \rho_1 = \lim_{g_1 \rightarrow 1} \rho_2 = 1 \text{ Se tiene que } \lim_{g_1 \rightarrow 1} B_1$$

$$= 1 \text{ y } \lim_{g_1 \rightarrow 1} B_2 = 0 \text{ este caso es similar al caso 1, esto es,}$$

$\rho_m^\bullet = \rho_{m-1}^\bullet$ para $m \in \mathbb{Z}$. Nótese que en este caso el orden de la ecuación en diferencias resultante es menor que k , mientras que en los casos anteriores fue igual a k .

Caso 7. Dos raíces son complejas y una es real, con $M = 3$ y $k = 3$. Sean $g_1 = d \exp(i\theta)$, $g_2 = d \exp(-i\theta)$ y calcúlese los límites cuando $|g_1| \rightarrow 1$, $|g_2| \rightarrow 1$ y $g_3 \rightarrow 1$, bajo la restricción $|g_1| = |g_2| = |g_3|$, con $d = g_3$. Entonces se tiene

$$\lim_{g_3 \rightarrow 1} \rho_1 = \frac{2\cos(\theta + 1)}{\cos(\theta + 2)} \quad \text{y} \quad \lim_{g_3 \rightarrow 1} \rho_2 = \frac{\cos(2\cos \theta + 1)}{2 + \cos \theta}$$

por lo que

$$\lim_{g_3 \rightarrow 1} A_1 = \lim_{g_3 \rightarrow 1} A_2 = \frac{1}{2} \left(\frac{1}{2 + \cos \theta} \right) \quad \text{y} \quad \lim_{g_3 \rightarrow 1} A_3 = \frac{1 + \cos \theta}{2 + \cos \theta}$$

Por lo tanto, $\rho_m^\bullet$ es solución de una ecuación en diferencias de tercer orden para $m \in \mathbb{Z}$.

Los casos restantes son omitidos pues existe cierta "equivalencia" con los casos antes mencionados. Por ejemplo si $M = 1$, $k = 1$ y $g_1 \rightarrow -1$, los resultados son similares a los del caso 1 con la diferencia de que $\rho_m^\bullet = -\rho_{m-1}^\bullet$. ■

El teorema (1.3.8), así como el (1.3.9) (que a continuación se enuncia), puede ser utilizado para sugerir una

transformación apropiada que logre la estacionariedad del proceso. El último teorema de esta sección hace uso del teorema (1.3.8) para obtener el recíproco del teorema (1.3.7).

Teorema 1.3.9 Supóngase que el teorema (1.3.8) se cumple, entonces, la condición de que $S_1(m)$ sea constante para $1 \leq p$ es necesaria y suficiente para garantizar la no estacionariedad del modelo.

Demostración La suficiencia se sigue inmediatamente del teorema (1.3.7). Por otra parte, sea $f_m = \rho_m$ y supóngase que el modelo es no-estacionario con FAC $\rho_m = \rho_m^*$. Sea N el número de raíces distintas próximas al círculo unitario, en el sentido de la definición (1.3.5). Por el teorema (1.3.8), ρ_m satisface la ecuación característica de un proceso puramente no-estacionario de orden $1 \leq N$, esto es,

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_1 B^1) \rho_m = 0 \quad (1.3.19)$$

y $|\phi_1| = 1$

luego, ya que ρ_m satisface una ecuación como la (1.3.19), si

$$1 - \sum_{i=1}^1 \phi_i \neq 0 \quad \text{la demostración del teorema 1 de Gray, et al (1971 y$$

1978) implica que los teoremas (1.3.2) y (1.3.3) se cumplen, aunque el modelo sea no-estacionario.

De este modo $\phi_1 = -1$, implica que $C_1 = C_2$ o $S_1(m)$ es constante. Por otra parte, si $\phi_1 = 1$, ya que todas las raíces

están sobre el círculo unitario, se tiene que $1 - \sum_{i=1}^1 \phi_i = 0$ ■

La demostración para el caso $f_m = (-1)^m \rho_m$ es análoga a esta demostración.

Nótese que hasta ahora sólo se ha tratado el caso en donde $p > 0$. Esto no implica que el método no sea útil para aplicarse a los procesos puros de promedios móviles $MA(q)$, pues como se verá a continuación, éste es uno de los casos más fáciles de determinar. Sin embargo, en el caso de los procesos $MA(q)$, los arreglos S no tendrán un patrón fijo, pero el valor de q se puede estimar mediante la primera columna de los arreglos R , pues $R_1(\rho_m) = \rho_m$, esto es, a partir de los valores de la FAC. Obsérvese que en este punto, el método coincide con el de Box-Jenkins.

1.4.- FUNCIÓN DE AUTOCORRELACIÓN PARCIAL Y DE AUTOCORRELACIÓN PARCIAL GENERALIZADA.

En esta subsección se estudiará una extensión de la FACP propuesta por Woodward & Gray (1981), la cual será referida como la Función de Autocorrelación Parcial Generalizada (FACPG). Esta generalización de la FACP es útil para examinar la relación que existe entre los arreglos S del método de Gray, Kelley y McIntire (1978) presentado en la sección anterior y el método de Box-Jenkins. También se mostrará cómo la FACPG resulta ser una herramienta útil en el problema de la identificación del orden de un modelo $ARMA(p,q)$, si es usada junto con los arreglos S , además se presenta una

modificación de los arreglos S, argumentando que con dicha modificación el uso práctico de estos arreglos es más fácil.

1.4.1.- Métodos de identificación

Como se observó en la sección 1.2, la FACP es una herramienta básica en la etapa de identificación de un modelo mediante el método de Box-Jenkins, dicha función está definida en (1.2.13).

Por otra parte, si $\{Z_t\}$ es un proceso ARMA(p,q), la FAC de este proceso no satisface las primeras q ecuaciones de Yule-Walker, pero sí las siguientes q+1, q+2, Usando esta observación se define la FACPG como:

$$\phi_{nn}^{(m)} = \begin{cases} \frac{\rho_{m+1}}{\rho_m} & \text{si } n = 1 \\ \frac{|A(n, m)|}{|B(n, m)|} & \text{si } n > 1 \end{cases} \quad (1.4.1)$$

Así, $\phi_{nn}^{(m)}$ es la solución de las m+1, ..., m+n ecuaciones de Yule-Walker mediante la regla de Cramér, para el último coeficiente autorregresivo. Esto es, si se supone que el proceso es un ARMA(n,m), $\phi_{nn}^{(m)}$ es el último coeficiente autorregresivo. En particular, si $\{Z_t\}$ se representa con un modelo ARMA(p,q), entonces $\phi_{pp}^{(q)} = \phi_p$. Nótese que si $\{Z_t\}$ es un

ARMA(p,q), entonces $\phi_{nn}^{(q)} \neq 0$ para $n = p$ y $\phi_{nn}^{(q)} = 0$ para $n > p$. Por otra parte, si $\{Z_t\}$ sigue un modelo ARMA(p,q), pueden usarse las p ecuaciones de Yule-Walker $q+1, \dots, q+p$ para estimar los parámetros autorregresivos $\phi_1, \dots, \phi_p$. Sin embargo, también se pueden usar las p ecuaciones de Yule-Walker $q+i$ hasta $q+i+p-1$ para $i = 1, 2, \dots$ y obtener la misma estimación de los parámetros autorregresivos, siempre y cuando la FAC teórica sea conocida. Por lo tanto, si $\{Z_t\}$ sigue un modelo estacionario ARMA(p,q), entonces $\phi_{pp}^{(q+i)} = \phi_p$ para $i = 0, 1, \dots$. Estas propiedades hacen posible la identificación única de p y q en un proceso mixto, mediante una simple inspección de la FACPG, siempre y cuando la FAC teórica sea conocida. Las ideas antes mencionadas serán formalizadas y resumidas en el teorema (1.4.3).

Para cada $q = 0, 1, 2, \dots$ y $p = 1, 2, \dots$ los valores de la FACPG serán denotados por $\phi_{pp}^{(q)}$. Nótese que $\phi_{pp}^{(0)}$ genera la FACP usual.

En esta subsección se presentan dos métodos para la identificación de p y q en un modelo mixto ARMA(p,q). El primero de ellos se basa en un arreglo en el cual los elementos (n,m) están dados por $\phi_{nn}^{(m)}$ para $m \geq 0$ y $n \geq 1$. Si $\{Z_t\}$ sigue un modelo ARMA(p,q), entonces el arreglo asociado con este modelo tendrá el patrón mostrado en la tabla 1.4.1. El procedimiento a seguir en la identificación de p y q, es buscar una columna p en la cual ocurra un comportamiento constante y un renglón q en el cual los elementos de las columnas m, con $m > p$, sean iguales a cero.

TABLA 1.4.1

ARREGLO DE LA FACPG PARA UN ARMA(p,q)

ORDEN DEL POLINOMIO DE PROMEDIOS MOVILES	ORDEN DEL POLINOMIO AUTORREGRESIVO					
	1	...	p-1	p	p+1	...
0	$\phi_{11}^{(0)}$	...	$\phi_{p-1p-1}^{(0)}$	$\phi_{pp}^{(0)}$	$\phi_{p+1p+1}^{(0)}$	...
1	$\phi_{11}^{(1)}$	...	$\phi_{p-1p-1}^{(1)}$	$\phi_{pp}^{(1)}$	$\phi_{p+1p+1}^{(1)}$	...
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	
q-1	$\phi_{11}^{(q-1)}$	...	$\phi_{p-1p-1}^{(q-1)}$	$\phi_{pp}^{(q-1)}$	$\phi_{p+1p+1}^{(q-1)}$	...
q	$\phi_{11}^{(q)}$	...	$\phi_{p-1p-1}^{(q)}$	ϕ_p	0	...
q+1	$\phi_{11}^{(q+1)}$	...	$\phi_{p-1p-1}^{(q+1)}$	ϕ_p	*	...
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	

El segundo de los métodos que se presentan en esta sección, consiste en una modificación a los arreglos S estudiados en la sección 1.3. Dicha modificación se resume en el siguiente teorema.

Teorema 1.4.1 Sea $\{Z_t\}$ un proceso estacionario ARMA(p,q) con FAC dada por ρ_n . Supóngase que $S_n(\rho_n)$ está definido para $p > 0$, además sea $f_n = \rho_n$ o $f_n = (-1)^n \rho_n$.

Entonces, $S_k(\rho_{-k-m}) = \mp \infty$ y $S_k(\rho_{-k+m+1}) = (-1)^{k-n} S_n(\rho_{-n+m+1})$

para $k > n$, si y sólo si $n = p$ y $m = q$.

Demostración $\Leftarrow$ Sea $f_m = \rho_m$ y supóngase que $n = p$ y $m = q$, esto es, $\{Z_t\}$ es un proceso ARMA (n, m) , entonces

$$\rho_1 = \sum_{j=1}^n \phi_j \rho_{1-j} \quad \text{para } 1 \geq m+1 \quad (1.4.2)$$

Ahora, como

$$S_{n+1}(\rho_{-n-m-1}) = \frac{H_{n+1+1}[1; \rho_{-n-m-1}]}{H_{n+1}[\rho_{-n-m-1}]}$$

$$= \frac{\begin{vmatrix} 1 & 1 & \dots & 1 \\ \rho_{-n-m-1} & \rho_{-n-m-1+1} & \dots & \rho_{-m} \\ \rho_{-n-m-1+1} & \rho_{-n-m-1+2} & \dots & \rho_{-m+1} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{-m-1} & \rho_{-m} & \dots & \rho_{-m+n+1-1} \end{vmatrix}}{\begin{vmatrix} \rho_{-n-m-1} & \rho_{-n-m-1+1} & \dots & \rho_{-m-1} \\ \rho_{-n-m-1+1} & \rho_{-n-m-1+1} & \dots & \rho_{-m} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{-m-1} & \rho_{-m} & \dots & \rho_{-m+n+1-2} \end{vmatrix}}$$

y además como $\rho_1 = \rho_{-1}$ de (1.4.2) se sigue que

$$\rho_{-n-m-1} = \sum_{j=1}^n \phi_j \rho_{-n-m-1+j} \quad \text{con } -n-m-1 \leq -m-1$$

de donde se observa que la primera columna del determinante

$H_{n+1}[\rho_{-n-m-1}]$ con $i > 0$, puede expresarse como una combinación

lineal de las columnas 2, 3, ..., $n+1$; entonces,

$$H_{n+1}[\rho_{-n-m-1}] = 0.$$

Por otra parte, ya que $\{Z_t\}$ es un ARMA (p, q) se tiene que

$$\rho_q = \sum_{j=1}^n \phi_j \rho_{q-j} \quad \text{y por lo tanto} \quad \rho_q = \sum_{j=1}^n \phi_j \rho_{j-k} \quad \text{de donde}$$

se sigue que $H_{n+1+1}[1; \rho_{-n-m-1}] \neq 0$. Entonces

$$S_{n+1}(\rho_{-n-m-1}) = \mp \infty \quad \text{para } i > 0.$$

Ahora, ya que

$$S_{n+1}(\rho_{-n-m-1+1}) = \frac{H_{n+1+1}[1; \rho_{-n-m-1+1}]}{H_{n+1}[\rho_{-n-m-1+1}]}$$

$$= \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ \rho_{-n+m-1+1} & \rho_{-n+m-1+2} & \cdots & \rho_m & \rho_{m+1} \\ \rho_{-n+m-1+2} & \rho_{-n+m-1+3} & \cdots & \rho_{m+1} & \rho_{m+2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_m & \rho_{m+1} & \cdots & \rho_{m+n+1-1} & \rho_{m+n+1} \end{vmatrix}}{\begin{vmatrix} \rho_{-n+m-1+1} & \rho_{-n+m-1+2} & \cdots & \rho_{m-1} & \rho_m \\ \rho_{-n+m-1+2} & \rho_{-n+m-1+3} & \cdots & \rho_m & \rho_{m+1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_m & \rho_{m+1} & \cdots & \rho_{m+n+1-2} & \rho_{m+n+1-1} \end{vmatrix}}$$

entonces, es fácil deducir de la relación (1.4.2) que la columna $n+i+1$ de $H_{n+1+1}[1; \rho_{-n+m-1+1}]$ puede expresarse como una combinación lineal de las $n+i$, $n+i-1$, ..., $i+1$ columnas. Por lo tanto, de manera análoga a la demostración del teorema (1.3.2) se demuestra que:

$$H_{n+1+1}[1; \rho_{-n+m-1+1}] = (-1)^{n+1+2} \left(1 - \sum_{j=1}^n \phi_j\right) H_{n+1}(\rho_{-n+m-1+1})$$

también como $\rho_m \neq \sum_{j=1}^n \phi_j \rho_{m-j}$, se tiene que $H_{n+1}[\rho_{-n+m-1+1}] \neq 0$,

luego entonces

$$S_{n+1}(\rho_{-n+m-1+1}) = \frac{(-1)^{n+1+2} \left(1 - \sum_{j=1}^n \phi_j\right) H_{n+1}[\rho_{-n+m-1+1}]}{H_{n+1}[\rho_{-n+m-1+1}]}$$

$$= (-1)^1 \left[(-1)^n \left(1 - \sum_{j=1}^n \phi_j \right) \right] = (-1)^1 C_1 = (-1)^1 S_n(\rho_{-n+m+1})$$

la última igualdad surge del teorema (1.3.2). Nótese que $k = n + i$. De manera análoga se puede demostrar que $S_{n+1}(\rho_{-n-m-1+1})$ es indefinido.

=>) Ahora supóngase que:

$$S_{n+1}(\rho_{-n-m-1}) = \mp \infty \quad \text{y} \quad S_{n+1}(\rho_{-n+m-1+1}) = (-1)^1 S_{n+1}(\rho_{-n+m+1})$$

además sea $\{Z_t\}$ un proceso ARMA (p, q) para algún p y q con $p > 0$ y $q \geq 0$. Entonces por la demostración de la necesidad, se sigue que $n \leq p$ y por otra parte se sabe que ρ_n satisface una ecuación en diferencias de orden $m + 1$, pero no de orden m , por lo tanto $n = p$ y $m = q$. La demostración para el caso $f_n = (-1)^n \rho_n$ es análoga. ■

Igual que en la sección 1.3, los valores $S_n(\rho_n) = S_n(m)$ se ordenarán como se muestra en la tabla 1.4.2. Nótese que el arreglo S mostrado en esta tabla, es una modificación de los arreglos S estudiados en la sección anterior, pues el elemento $S_n(m)$ de dicha tabla está localizado en el renglón correspondiente al retraso $m - n + 1$, en vez de estar en el renglón correspondiente al retraso m , como previamente se definió, razón por la cual en el resto de esta

sección se denotará a los arreglos S como S° , indicando que se refiere a los arreglos S con la modificación antes mencionada.

T A B L A 1.4.2

VALORES DE $S_n^*(f_m)$

$\begin{matrix} n \\ m \end{matrix}$	1	2	...	k
-l	$S_1(-l)$	$S_2(-l-1)$	...	$S_k(-l-k+1)$
-l+1	$S_1(-l+1)$	$S_2(-l)$	...	$S_k(-l-k+2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
-1	$S_1(-1)$	$S_2(-2)$	...	$S_k(-k)$
0	$S_1(0)$	$S_2(-1)$	...	$S_k(-k+1)$
1	$S_1(1)$	$S_2(0)$	...	$S_k(-k+2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
j	$S_1(j)$	$S_2(j-1)$	...	$S_k(j-k+1)$

Por lo tanto, en términos de los arreglos S° de la tabla 1.4.2 y de los teoremas (1.3.2), (1.3.3) y (1.4.1), se asegura que el proceso $\{Z_t\}$ sigue un modelo estacionario ARMA(p,q) si y sólo si el arreglo S° asociado con éste, tiene el patrón que se muestra en la tabla 1.4.3

TABLA 1.4.3

COMPORTAMIENTO DE LOS ARREGLOS S^* PARA UN PROCESO ARMA(p,q)

$m \cdot n$	1	2	...	p	p+1	p+1
-1	$S_1(-1)$	$S_2(-1-1)$	...	C_2	*	...
-1+1	$S_1(-1+1)$	$S_2(-1)$	...	C_2	*	...
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
-q-2	$S_1(-q-2)$	$S_2(-q-3)$	...	C_2	*	...
-q-1	$S_1(-q-1)$	$S_2(-q-2)$	...	C_2	$\mp \infty$	...
-q	$S_1(-q)$	$S_2(-q-1)$	...	NC	NC	...
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
-1	$S_1(-1)$	$S_2(-2)$	...	NC	NC	...
0	$S_1(0)$	$S_2(-1)$	...	NC	NC	...
1	$S_1(1)$	$S_2(0)$	...	NC	NC	...
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
q-1	$S_1(q-1)$	$S_2(q-2)$	...	NC	NC	...
q	$S_1(q)$	$S_2(q-1)$	...	C_1	$-C_1$	$(-1)^1 C_1$
q+1	$S_1(q+1)$	$S_2(q)$	...	C_1	*	...
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$

Obsérvese como en este nuevo patrón existen más elementos de juicio para la identificación de los valores de p y q, pues además del comportamiento de la columna p, en donde aparecen

las constantes C_1 y C_2 , se tiene el comportamiento del renglón $-q-1$ en donde el efecto $\mp \infty$ se prolonga a las columnas $p+2, \dots$ y el del renglón q , en donde el efecto $(-1)^1 C_1$ se prolonga también a las columnas $p+2, \dots$.

El siguiente teorema establece la relación que existe entre los arreglos S^* y la FACPG.

Teorema 1.4.2 Sea $\{\rho_m\}$ la FAC de una serie de tiempo.

$$i) \text{ Si } f_m = \rho_m \text{ entonces } \phi_{nn}^{(m)} = \frac{-S_n^*(\rho_{-n+m+1})}{S_n^*(\rho_{-n-m})}$$

$$ii) \text{ Si } f_m = (-1)^m \rho_m \text{ entonces } \phi_{nn}^{(m)} = \frac{(-1)^{n+1} S_n^*(f_{-n+m+1})}{S_n^*(f_{-n-m})}$$

Demostración i) Por definición

$$\frac{S_n(\rho_{-n+m+1})}{S_n(\rho_{-n-m})} = \frac{H_{n+1}[1; \rho_{-n+m+1}] H_n[\rho_{-n-m}]}{H_n[\rho_{-n+m+1}] H_{n+1}[1; \rho_{-n-m}]}$$

En donde los elementos (r, c) de $H_{n+1}[1; \rho_{-n-m}]$ y los elementos $(n+3-r, n+2-c)$ de $H_{n+1}[1; \rho_{-n+m+1}]$ están dados por $\rho_{-n-m+r+c-3}$ para $2 \leq r \leq n+1$ y $1 \leq c \leq n+1$ esto es,

$$H_{n+1}[1; \rho_{-n-m}] = \begin{vmatrix} 1 & 1 & & 1 & 1 \\ \rho_{-n-m} & \rho_{-n-m+1} & \cdots & \rho_{-m-1} & \rho_{-m} \\ \rho_{-n-m+1} & \rho_{-n-m+2} & \cdots & \rho_{-m} & \rho_{-m+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_{-m-1} & \rho_{-m} & \cdots & \rho_{-m+n-2} & \rho_{-m+n-1} \end{vmatrix}$$

es un determinante cuyo elemento (r, c) está dado por $\rho_{-n-m+r+c-3}$ y

$$H_{n+1}[1; \rho_{-n+m+1}] = \begin{vmatrix} 1 & 1 & & 1 & 1 \\ \rho_{-n+m+1} & \rho_{-n+m+2} & \cdots & \rho_m & \rho_{m+1} \\ \rho_{-n+m+2} & \rho_{-n+m+3} & \cdots & \rho_{m+1} & \rho_{m+2} \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_m & \rho_{m+1} & \cdots & \rho_{m+n-1} & \rho_{m+n} \end{vmatrix}$$

es un determinante con elemento $(n+3-r, n+2-c)$ dado por $\rho_{-n-m+r+c-3}$. Ahora obsérvese que intercambiando las columnas c y $n+2-c$ para $c = 1, 2, \dots, [\frac{n+1}{2}]$ y los renglones r y $n+3-r$ para $r = 2, 3, \dots, 1 + [\frac{n}{2}]$ del determinante $H_{n+1}[1; \rho_{-n+m+1}]$, en donde $[.]$ denota la función del mayor entero, definida como $[|x|] = n$ si $n \leq x < n+1$; $n \in \mathbb{Z}$, y además como $\rho_m = \rho_{-m}$ se tiene que:

$$H_{n+1}[1; \rho_{-n-m}] = (-1)^n H_{n+1}[1; \rho_{-n+m+1}]$$

Por otra parte

$$H_n[\rho_{-n+m+1}] = \begin{vmatrix} \rho_{-n+m+1} & \rho_{-n+m+2} & \cdots & \rho_{m-1} & \rho_m \\ \rho_{-n+m+2} & \rho_{-n+m+3} & \cdots & \rho_m & \rho_{m+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_m & \rho_{m+1} & \cdots & \rho_{m+n-2} & \rho_{m+n-1} \end{vmatrix}$$

es un determinante con elementos $(n-r+1, c)$. Mientras que

$$|B(n, m)| = \begin{vmatrix} \rho_m & \rho_{m+1} & \cdots & \rho_{m+n-2} & \rho_{m+n-1} \\ \rho_{m-1} & \rho_m & \cdots & \rho_{m+n-3} & \rho_{m+n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ \rho_{m-n+1} & \rho_{m-n+2} & \cdots & \rho_{m-1} & \rho_m \end{vmatrix}$$

es el determinante de la matriz $B(n, m)$ con elementos (r, c) para $1 \leq c \leq n$ y $1 \leq r \leq n$. Es fácil observar que al intercambiar los renglones r y $n-r+1$ de $B(n, m)$ para $r = 1, 2, \dots, [\frac{n}{2}]$, se tiene que:

$$H_{n+1}[\rho_{-n+m+1}] = (-1)^{[\frac{n}{2}]} |B(n, m)|$$

De manera análoga se puede demostrar que:

$$H_n[\rho_{-n-m}] = (-1)^{[\frac{n-1}{2}]} |A(n, m)|$$

entonces,

$$\begin{aligned}
 - \frac{S_n(\rho_{-n+m+1})}{S_n(\rho_{-n+m})} &= \frac{(-1)^{H_{n+1}[1; \rho_{-n+m+1}]} (-1)^{[\frac{n-1}{2}]} |A(n, m)|}{(-1)^{[\frac{n}{2}]} |B(n, m)| (-1)^n H_{n+1}[1; \rho_{-n+m+1}]} \\
 &= \frac{(-1) (-1)^{[\frac{n-1}{2}]} |A(n, m)|}{(-1)^{[\frac{n}{2}]} (-1)^n |B(n, m)|} = \frac{|A(n, m)|}{|B(n, m)|} = \phi_{nn}^{(n)}
 \end{aligned}$$

ii) La demostración es análoga a la de i). ■

Nótese que aunque los elementos de la FACPG pueden ser calculados a partir de la relación (1.4.1), dicho cálculo se facilita si se usa el teorema (1.4.2) y el programa mostrado en el Apéndice 2.

Como último resultado, a continuación se formalizarán las ideas mencionadas al inicio de esta sección, mediante el siguiente teorema.

Teorema 1.4.3 Sea $\{Z_t\}$ un proceso estacionario ARMA(p, q) con $p > 0$. Entonces

i) $\phi_{pp}^{(m)} = \phi_p$ para $m > q$

ii) Supóngase que el sistema de ecuaciones de Yule-Walker de orden n es no-singular. Para alguna constante m , se obtiene

$\phi_{kk}^{(m)} = 0$ con $k \geq n+1$ y $\phi_{nn}^{(m)} \neq 0$, en donde todas estas cantidades están bien definidas si y sólo si $n = p$ y $m = q$.

Demostración i) De los teoremas (1.3.2), (1.3.3), (1.4.1) y (1.4.2) se sigue que:

$$\phi_{pp}^{(m)} = \frac{-S_p^*(\rho_{-p+m+1})}{S_p^*(\rho_{-p-m})} = \frac{-C_1}{C_2} = \phi_p \quad \text{para } m > q$$

$$\text{pues } C_2 = -\frac{C_1}{\phi_p}$$

ii) Ahora, de los teoremas (1.4.1) y (1.4.2) es claro que:

$$\phi_{n+i, n+i}^{(m)} = \frac{-S_{n+i}^*(\rho_{-n+m-i+1})}{S_{n+i}^*(\rho_{-n-m-i})} = \frac{(-1)^i C_1}{\mp \infty} = 0; \quad i = 1, 2, \dots$$

Por último, también de estos teoremas se tiene que:

$$\phi_{pp}^{(q)} = \frac{-S_p^*(\rho_{-p+q+1})}{S_p^*(\rho_{-p-q})} = \frac{-C_1}{C_2} \neq 0$$

■

A manera de resumen de esta sección, conviene señalar que aquí se ha mostrado el uso dela FACPG. para la

identificación del orden de un modelo ARMA (p,q) . También se ha mostrado la relación que existe entre los arreglos S de Gray, Kelley y McIntire (1978) y la FACPG. Esto es, se demostró que la FACPG, al igual que el estadístico D , representa una condensación de la información que contienen los arreglos S . Es recomendable el uso de los arreglos S^* en vez de los arreglos S , pues estos presentan la información de una manera más fácil de interpretar.

1.5.- EL MÉTODO DE LA ESQUINA.

Beguín, Gourieroux y Monfort (1980) introducen el método de la esquina como una herramienta para la identificación del orden (p,q) de un proceso ARMA. En ésta sección se presenta con detalle el método de la esquina. Dicho método se basa en el reconocimiento de algunos patrones de la FACM. No obstante, estos autores también desarrollan un estadístico de prueba para dichos patrones, basándose en teoría asintótica. En esta sección también se mostrará que existen algunas restricciones implícitas en el uso del estadístico de prueba. Estas restricciones no son obvias a simple vista y no son mostradas por Beguín et al (1980). Sin embargo, en el trabajo de Gooljer & Heuts (1981) se muestran dichas restricciones.

1.5.1.- Caracterización de una representación minimal ARMA(p,q)

Resulta difícil el uso práctico de la primera caracterización minimal de un proceso ARMA(p,q) (teorema 1.2.1) en la identificación del orden (p,q) a partir de $\{\rho_n\}$, en particular porque las ϕ_i 's son desconocidas. Por esta razón, se presenta a continuación una condición equivalente a la del teorema (1.2.1) en la cual las ϕ_i 's no aparecen en forma explícita.

Sea $M(m,n)$ una matriz de orden $n \times n$ definida por:

$$M(m,n) = \begin{bmatrix} \rho_m & \rho_{m-1} & \dots & \rho_{m-n+1} \\ \rho_{m+1} & \rho_m & \dots & \rho_{m-n+2} \\ \vdots & \vdots & & \vdots \\ \rho_{m+n-1} & \rho_{m+n-2} & \dots & \rho_m \end{bmatrix} \quad \text{para } m \geq 0 \text{ y } n \geq 1 \quad (1.5.1)$$

y $\Delta(n,m)$ es el determinante de $M(m,n)$. Antes de enunciar el teorema que dará dicha caracterización, se demostrará el siguiente lema.

Lema 1.5.1 Una sucesión $S_1, S_2, \dots, S_h, \dots$ de números complejos satisface una ecuación en diferencias de orden $p \geq 0$, esto es, existen constantes $\phi_1, \phi_2, \dots, \phi_p$ tales que

$$S_h - \phi_1 S_{h-1} - \dots - \phi_p S_{h-p} = 0 \quad \text{para } h = p+1, p+2, \dots$$

en donde p es el entero más pequeño que tiene esta propiedad, si y sólo si los determinantes $\delta(n)$ para $n \geq 1$, definidos por:

$$\delta(n) = \begin{vmatrix} S_n & S_{n-1} & \dots & S_1 \\ S_{n+1} & S_n & \dots & S_2 \\ \vdots & \vdots & & \vdots \\ S_{2n-1} & S_{2n-2} & \dots & S_n \end{vmatrix}$$

son tales que: $\delta(n) = 0$ para $n \geq p+1$
 $\delta(n) \neq 0$ para $n = p$

Demostración Es inmediata de los resultados mostrados en el Apéndice 3.

Teorema 1.5.1 ("la caracterización de la esquina"). El proceso $\{Z_t\}$ tiene una representación minimal ARMA(p, q) si y sólo si

$$\begin{aligned} \Delta(m, n) &= 0 && \text{para toda } m \geq q + 1 \text{ y } n \geq p + 1 \\ \Delta(m, p) &\neq 0 && \text{para toda } m \geq q \\ \Delta(q, n) &\neq 0 && \text{para toda } n \geq p \end{aligned}$$

Demostración Del teorema (1.2.1), si $\{Z_t\}$ tiene una representación minimal ARMA(p, q), su FAC satisface la siguiente ecuación en diferencias

$$\rho_m = \sum_{i=1}^p \phi_i \rho_{m-1} \quad \text{para } m \geq q+1, \text{ pero} \quad (1.5.2)$$

$$\rho_m \neq \sum_{i=1}^p \phi_i \rho_{m-1} \quad \text{para } m = q \text{ con } \phi_p \neq 0$$

Primera etapa: Primeramente se demostrará que $\Delta(m,n) = 0$ para toda $m \geq q+1$ y $n \geq p+1$. Sean $C^1(m,n), \dots, C^{p+1}(m,n)$ las primeras $p+1$ columnas de la matriz $M(m,n)$, entonces de la relación (1.5.2) se sigue que la primera columna de esta matriz puede expresarse como una combinación de las restantes p columnas, esto es

$$C^1(m,n) = \sum_{i=1}^p \phi_i C^{i+1}(m,n)$$

por lo tanto, $\Delta(m,n) = 0$ para toda $m \geq q+1$ y para toda $n \geq p+1$.

Segunda etapa: Ahora se demostrará que $\Delta(n,m) \neq 0$ para toda $m \geq q$. Para esto considérese la sucesión $S_1 = \rho_{q-p+1}, \dots, S_h = \rho_{q-p+h}, \dots$ la cual satisface una ecuación en diferencias de orden mínimo p . Luego, por la condición necesaria del lema (1.5.1) se tiene que $\delta(p) = \Delta(m,n) \neq 0$. Más generalmente, considérese la sucesión $S_1 = \rho_{q-p+k+1}, \dots, S_h = \rho_{q-p+h+k}, \dots$ para $k \geq 0$, que también satisface una ecuación en diferencias de orden mínimo p , entonces

$$\delta(p) = \Delta(q+k,p) \neq 0 \text{ para } k \geq 0$$

Tercera etapa: Por último se demostrará que $\Delta(m,n) \neq 0$ para toda $n \geq p$. Como $\Delta(p,q) \neq 0$, es suficiente demostrar que $\Delta(q,p+k) \neq 0$ para $k > 0$. Supóngase que $\Delta(q-p-k+n,n) \neq 0$ y $\Delta(q,p+k) = 0$, además supóngase que $\Delta(q-1,p+k-1) \neq 0$ y considérese la sucesión $S_1 = \rho_{q-p-k+1}, \dots, S_h = \rho_{q-p-k+h}, \dots$ los correspondientes $\delta(n)$ son iguales a $\delta(n) = \Delta(q-p-k+n,n)$ y son tales que:

$$\delta(n) = 0 \quad \text{para toda } n \geq p+k$$

$$\delta(n) \neq 0 \quad \text{para } n = p+k+1$$

por la condición de suficiencia en el lema (1.5.1), la FAC satisface una ecuación en diferencias de orden $p+k-1$ para el rango q , esto es

$$\rho_h - \phi_1 \rho_{h-1} - \dots - \phi_{p+k-1} \rho_{h-p-k+1} = 0 \quad \text{para toda } h \geq q$$

Nótese que $\phi_{h-p-k+1} \neq 0$, pues si este no fuera el caso, $\{\rho_n\}$ tendría que satisfacer una ecuación en diferencias de orden $p+k-2$ para el rango q y por el lema (1.5.1) se tendría que

$$\Delta(q,p+k-1) = 0$$

Aplicando la condición de suficiencia del lema (1.1.1), el proceso $\{Z_t\}$ tiene una representación ARMA($p+k-1, q-1$) lo cual es imposible ya que $\{Z_t\}$ tiene una representación minimal ARMA(p, q). Luego entonces $\Delta(p+k, q) \neq 0$.

=> Ya que $\Delta(q,p) \neq 0$ y $\Delta(q+i,p+i) = 0$ para $i \geq 1$, la FAC satisface una ecuación en diferencias de orden p para el rango $q+1$, esto es

$$\rho_m = \sum_{i=1}^p \phi_i \rho_{m-i} \quad \text{para toda } m \geq q+1$$

y como $\Delta(q+1, p) \neq 0$, usando el mismo argumento de la tercera etapa acerca de la condición necesaria, el coeficiente $\phi_p \neq 0$. Luego, por la condición suficiente del lema (1.1.1), el proceso $\{Z_t\}$ tiene una representación ARMA(p, q). Sólo resta demostrar que esta representación es mínima. Para esto supóngase que este no es el caso, esto es, existe una representación mínima ARMA(p', q') para $\{Z_t\}$ con $p' < p$ y $q' < q$. Luego entonces, de la condición necesaria del lema (1.5.1), el determinante $\Delta(q, p) = 0$, lo cual es imposible. Por lo tanto, $\{Z_t\}$ tiene una representación mínima ARMA(p, q) ■

El teorema (1.5.1) sugiere el siguiente procedimiento para la identificación de p y q , a partir de las autocorrelaciones ρ_m :

- i) Elegir un entero k suficientemente grande tal que $k > \max\{p, q\}$.
- ii) Construir el arreglo $\Delta(m, n)$ para m y $n = 1, 2, \dots, k$.
- iii) El proceso estacionario $\{Z_t\}$ satisface una representación minimal ARMA(p, q) si y sólo si el arreglo- Δ tiene la estructura. del cuadro 1.5.1

CUADRO 1.5.1

PATRON TEORICO DEL ARREGLO- Δ PARA UN PROCESO ARMA(p,q); q ES EL ORDEN DE PROMEDIOS MOVILES, n ES EL ORDEN AUTORREGRESIVO

$n \backslash n$	1	2	...	p	p+1	...	k
1	$\Delta(1,1)$	$\Delta(1,2)$	...	$\Delta(1,p)$	$\Delta(1,p+1)$	...	$\Delta(1,k)$
2	$\Delta(2,1)$	$\Delta(2,2)$	...	$\Delta(2,p)$	$\Delta(2,p+1)$	...	$\Delta(2,k)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
q	$\Delta(q,1)$	$\Delta(q,2)$	...	$\Delta(q,p)$	$\Delta(q,p+1)$	...	$\Delta(q,k)$
q+1	$\Delta(q+1,1)$	$\Delta(q+1,2)$	...	$\Delta(q+1,p)$	0	...	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
k	$\Delta(k,1)$	$\Delta(k,2)$	...	$\Delta(k,p)$	0	...	0

Observación 1

De la sección (1.3) se sabe que si $\{Z_t\}$ es un proceso con FAC $\{\rho_n\}$, el proceso $\{Y_t\} = \{(-1)^t Z_t\}$ tiene FAC $\{(-1)^n \rho_n\}$. En particular, si $\{Z_t\}$ tiene una representación minimal ARMA(p,q), $\{Y_t\}$ también tiene esta representación. Esto implica que los arreglos- Δ obtenidos a partir de $\{\rho_n\}$ o de $\{(-1)^n \rho_n\}$, tienen el mismo comportamiento y ambos pueden ser usados para la identificación de p y q.

Observación 2

Los arreglos- Δ construidos a partir de la Función de Autocorrelación Inversa (FACI) (ver sección 1.7) tienen un comportamiento similar al de los arreglos- Δ construidos con $\{\rho_n\}$, siempre y cuando sean remplazadas p por q y q por p.

1.5.2.- Algunas propiedades de la caracterización de la esquina.

I) Procesos MA(q)

El proceso $\{Z_t\}$ tiene una representación minimal MA(q) si y sólo si los arreglos- Δ tienen el comportamiento del cuadro 1.5.2.

CUADRO 1.5.2

COMPORTAMIENTO DEL ARREGLO Δ PARA UN PROCESO MA(q)

$m \backslash n$	1	2	3	...	k	...
1	$\Delta(1,1)$	$\Delta(1,2)$	$\Delta(1,3)$	...	$\Delta(1,k)$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	
q	$\Delta(q,1)$	$\Delta(q,2)$	$\Delta(q,3)$	...	$\Delta(q,k)$	...
q+1	0	0	0	...	0	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	
k	0	0	0	...	0	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	

Esto es, q+1 es el primer valor de m tal que $\Delta(m,n) = 0$ para toda $m \geq q+1$, y ya que $\Delta(m,1) = \rho_m$, se obtiene el criterio usual de la FAC para determinar el valor de q.

ii) Procesos AR(p)

El proceso $\{Z_t\}$ tiene una representación minimal AR(p) si

y sólo si los arreglos- Δ tienen el comportamiento del cuadro 1.5.3

CUADRO 1.5.3

COMPORTAMIENTO DEL ARREGLO Δ PARA UN PROCESO AR(p)

n	1	...	p	p+1	...	k	...
1	$\Delta(1,1)$	...	$\Delta(1,p)$	0	...	0	...
2	$\Delta(2,1)$	...	$\Delta(2,p)$	0	...	0	...
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	
k	$\Delta(k,1)$	...	$\Delta(k,p)$	0	...	0	...
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	

Esto es, $p+1$ es el primer valor de n tal que $\Delta(1,n) = 0$ para toda $n \geq p+1$. Puede demostrarse que el criterio de los arreglos- Δ para elegir p , es equivalente al criterio de la FACP, pues de la relación (1.2.12) se sabe que

$$\phi_{nn} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & \rho_1 & \rho_n \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-2} & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-3} & \rho_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & \rho_1 & 1 \end{vmatrix}}$$

y como

$$\Delta(1,n) = \begin{vmatrix} \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_n & \rho_{n-1} & \rho_{n-2} & \dots & \rho_1 \end{vmatrix}$$

al intercambiar desde la primera columna de $\Delta(1,n)$ hasta la columna n , se sigue que

$$\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & \rho_1 & \rho_n \end{vmatrix} = (-1)^{n-1} \Delta(1,n)$$

por lo tanto

$$\phi_{nn} = \frac{(-1)^{n-1} \Delta(1,n)}{\Delta(0,n)}$$

en donde $\Delta(0,n) \neq 0$, luego entonces $\phi_{nn} = 0$ si y sólo si $\Delta(1,n) = 0$, lo cual ocurre para $n \geq p+1$. Esto demuestra que el criterio de los arreglos- Δ para la elección de p , es una generalización del procedimiento usual basado en la FACP.

1.5.3.- Identificación de p y q a partir de los arreglos- Δ estimados (Δ).

1) Inspección visual:

En la práctica, los determinantes $\Delta(m,n)$ son estimados por

los determinantes $\Delta(m,n)$ obtenidos al remplazar $\{\rho_m\}$ por $\{\hat{\rho}_m\}$.

Usualmente, los determinantes $\Delta(m,n)$ no son exactamente iguales a cero para $m \geq q+1$ y $n \geq p+1$, pero al ser $\Delta(m,n)$ estimadores consistentes de $\Delta(m,n)$, se espera que los $\Delta(m,n)$ sean "próximos a cero" a partir de las líneas

$$\begin{array}{ll} m = q+1 & n \geq p+1 \\ y & \\ m \geq q+1 & n = p+1 \end{array}$$

La manera formal para decidir cuándo dichos elementos se pueden considerar estadísticamente distintos de cero, es mediante la siguiente prueba de hipótesis.

ii) *Procedimiento estadístico.*

Beguin et al (1980) consideran que un criterio visual es al menos, suficiente para la identificación de p y q . No obstante, es necesario realizar una prueba de hipótesis estadística para probar de manera formal si los $\Delta(m,n)$ se pueden considerar estadísticamente distintos de cero. Se sabe (ver Fuller 1976, pag 256) que para cualquier M , el vector

$$\sqrt{N} [\hat{\rho}_1 - \rho_1, \hat{\rho}_2 - \rho_2, \dots, \hat{\rho}_M - \rho_M]$$

converge a una distribución Normal multivariada con media cero y matriz de covarianzas G_M de orden $M \times M$, cuyo elemento (i,j) está dado por

$$\sum_{m=-\infty}^{\infty} [\rho_m - \rho_{m-1+j} + \rho_{m+j} \rho_{m-1} - 2\rho_m \rho_j \rho_{m-1} - 2\rho_1 \rho_m \rho_{m-j} + \rho_1 \rho_j \rho_m^2]$$

usando este resultado se deduce la distribución

asintótica de cualquier subconjunto finito de determinantes

$\{ \Delta(m_l, n_l) \text{ para } m_l \geq q+1 \text{ y } n_l \geq p+1 \}$ bajo la hipótesis nula

$$H_0: \Delta(m_l, n_l) = 0 \quad (1.5.3)$$

para $\{(m_l, n_l) \text{ con } 1 \leq l \leq T\} \subset \{(m, n); m \geq 0, n > 0\}$.

Para esto se denotará por $\Delta_1, \dots, \Delta_T$ a los determinantes de estos subconjuntos y por $\Delta_1, \dots, \Delta_T$ a los correspondientes $\Delta(n, m)$. Bajo H_0 , el vector

$$\sqrt{N}[\Delta_1(m_1, n_1), \dots, \Delta_T(m_T, n_T)] \xrightarrow{D} N(\underline{0}, \mathcal{G}) \quad (1.5.4)$$

con $\mathcal{G} = A G_M A'$, en donde G_M es la matriz de covarianzas de $\rho_1, \dots, \rho_M$ de orden $M \times M$, A es una matriz de orden $T \times M$ con elementos a_{ij} , tales que

$$a_{ij} = \frac{\partial \Delta(m_1, n_1)}{\partial \rho_j}$$

es la suma de cofactores de la correspondiente matriz $M(m, n)$. y $M = \max_{1 \leq l \leq T} (m_l + n_l - 1)$ esto es, M es el índice máximo m de las

ρ_m 's que aparecen en los determinantes $\Delta_1, \Delta_2, \dots, \Delta_T$. Las matrices $\hat{A}$ y $\hat{G}_M$ son estimadores de las matrices A y G_M obtenidas al remplazar $\{\rho\}_m$ por $\{\hat{\rho}\}_m$. Bajo H_0 , la forma cuadrática

$$Q_T = N(\hat{\Delta}(m_1, n_1), \dots, \hat{\Delta}(m_T, n_T)) [\hat{A} G_M \hat{A}']^{-1} (\hat{\Delta}(m_1, n_1), \dots, \hat{\Delta}(m_T, n_T))'$$

$$(1.5.4)$$

converge en distribución a una $\chi^2(n)$.

El conjunto $\{Q_T \geq C_\alpha\}$, en donde C_α es el punto porcentual superior α de $\chi^2(n)$, es una región crítica asintótica para probar H_0 a un nivel de significación α .

Mientras que la prueba de hipótesis descrita anteriormente parece ser completamente general para cualquier conjunto de índices $\{(m_i, n_i)\}_{i=1}^T$, existen en realidad algunas restricciones sobre el número de índices permitidos y sobre los índices que pueden ser usados en la distribución asintótica (1.5.4) para obtener una matriz de covarianzas no-singular. A continuación se presentan dichas restricciones, las cuales fueron presentadas por primera vez por Gooljer & Heuts (1981).

1.5.4.- Restricciones sobre la prueba de hipótesis del método de la esquina.

Si la matriz $\mathcal{F}$ es no singular, la estadística de prueba χ^2 tendrá un comportamiento inestable para valores de N suficientemente grandes y propiedades estadísticas no satisfactorias para valores pequeños de N .

Las restricciones que se encuentran en esta prueba de hipótesis se pueden clasificar de la siguiente manera:

- a) *Restricciones sobre el número de índices permitidos.*
- b) *Restricciones sobre la matriz de derivadas parciales.*

Primeramente se estudiarán las restricciones clasificadas como:

a) *RESTRICCIONES SOBRE EL NÚMERO DE ÍNDICES PERMITIDOS.*

Estas restricciones provienen del simple hecho de que $\mathcal{S}$ es singular si $T > M$, pues se sabe que el $\text{rango}(A) \leq \min \{M, T\}$ y $\text{rango}(G_M) \leq M$. También se sabe que el rango de un producto de dos matrices es a lo más, igual al rango de una de estas dos matrices, de donde se sigue que el $\text{rango}(AG_M A') \leq \min \{T, M\}$.

En el método de la esquina existen varias configuraciones del conjunto de índices $\{(m_i, n_i)\}_{i=1}^T$ que pueden ser de interés, dentro de estas están:

i) *Líneas horizontales o verticales*

Para las configuraciones horizontales o verticales se tienen los siguientes conjuntos de índices

$$\left\{ (m_i, n_i) \right\}_{i=1}^T = \left\{ (m_0, n) \right\}_{n=n_0}^{n_0+T-1} \quad \text{y} \quad \left\{ (m, n_0) \right\}_{m=m_0}^{m_0+T-1}$$

respectivamente, para algunos m_0 y n_0 . Entonces $M = m_0 + n_0 + T - 2$ en cada caso y por lo tanto $\mathcal{S}$ será singular si $m_0 + n_0 \leq 1$. Sin embargo, como en estos casos sólo se está probando la hipótesis para índices $m_0 \geq 1$ y $n_0 \geq 1$, se sigue que $M > T$, luego entonces $\mathcal{S}$ es no singular y por lo tanto no existen restricciones en la prueba de hipótesis para este tipo de configuraciones. Dichas configuraciones se presentan en las siguientes tablas.

T A B L A 1.5.4

CONJUNTO DE INDICES EN UNA LINEA HORIZONTAL

$m \quad n$	...	n_0	n_0+1	...	n_0+r-1
$\vdots$		$\vdots$	$\vdots$		$\vdots$
m_0	...	$\Delta(m_0, n_0)$	$\Delta(m_0, n_0+1)$	...	$\Delta(m_0, n_0+r-1)$

T A B L A 1.5.5

CONJUNTO DE INDICES EN UNA LINEA VERTICAL

$m \quad n$	...	n_0
$\vdots$		$\vdots$
m_0	...	$\Delta(m_0, n_0)$
m_0+1	...	$\Delta(m_0+1, n_0)$
$\vdots$		$\vdots$
m_0+r-1	...	$\Delta(m_0+r-1, n_0)$

ii) Rama ortogonal

Para esta configuración se tiene el siguiente conjunto de índices

$$\left\{ (m_1, n_1) \right\}_{l=1}^T = \left\{ (m_0, n) \right\}_{n=n_0}^{n_1} \cup \left\{ (m, n_0) \right\}_{m=m_0+1}^{m_1}$$

con $m_0 < m_1$ y $n_0 \leq n_1$. Entonces $T = (m_1 - m_0) + [n_1 - (n_0 - 1)]$
 $= (m_1 - m_0) + (n_1 - n_0) + 1$ y

$$M = \max\{n_1 + m_0 - 1, m_1 + n_0 - 1\}$$

por lo tanto, $\mathcal{G}$ es singular si $T > M$

si y sólo si $(m_1 - m_0) + (n_1 - n_0) + 1 > n_1 + m_0 - 1$

si y sólo si $m_1 - n_0 + 2 > 2m_0$

y

si y sólo si $(m_1 - m_0) + (n_1 - n_0) + 1 > m_1 + n_0 - 1$

si y sólo si $n_1 - m_0 + 2 > 2n_0$

Para el caso particular en el que las ramas sean de igual longitud, se tiene que

$d = m_1 - m_0 = n_1 - n_0$, entonces $n_1 + m_0 - 1 = m_1 + n_0 - 1$ y $\mathcal{G}$ será singular si $T > M$

si y sólo si $m_1 - m_0 + n_1 - n_0 + 1 > n_1 + m_0 - 1$

si y sólo si $d + 2 > m_0 + n_0$.

La representación de esta configuración se da en la siguiente tabla

T A B L A 1.5.6
CONJUNTO DE INDICES EN UNA RAMA ORTOGONAL

$m \cdot n$	...	$n_0 - 1$	n_0	$n_0 + 1$	...	n_1
$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
$m_0 - 1$	...	$\Delta(m_0 - 1, n_0 - 1)$	$\Delta(m_0 - 1, n_0)$	$\Delta(m_0 - 1, n_0 + 1)$	...	$\Delta(m_0 - 1, n_1)$
m_0	...	$\Delta(m_0, n_0 - 1)$	$\Delta(m_0, n_0)$	$\Delta(m_0, n_0 + 1)$	...	$\Delta(m_0, n_1)$
$m_0 + 1$	...	$\Delta(m_0 + 1, n_0 - 1)$	$\Delta(m_0 + 1, n_0)$			
$m_0 + 2$	...	$\Delta(m_0 + 2, n_0 - 1)$	$\Delta(m_0 + 2, n_0)$			
$\vdots$		$\vdots$	$\vdots$			
m_1	...	$\Delta(m_1, n_0 - 1)$	$\Delta(m_1, n_0)$			

iii) Rectangulares

Para esta configuración, se tiene el siguiente conjunto de índices

$$\left\{ (m, n) \right\}_{1=1}^T = \left\{ (m, n) : m_0 \leq m \leq m_1 \text{ y } n_0 \leq n \leq n_1 \right\}$$

para $m_0 \leq m_1$ y $n_0 \leq n_1$. Entonces $T = (m_1 - m_0 + 1)(n_1 - n_0 + 1)$ y

$M = m_1 + n_1 - 1$. Por lo tanto, $\mathcal{S}$ será singular si

$$(m_1 - m_0 + 1)(n_1 - n_0 + 1) > m_1 + n_1 - 1$$

Por otra parte, si el rectángulo es en realidad un cuadrado con $d = m_1 - m_0 = n_1 - n_0$, entonces $\mathcal{S}$ será singular si

$$(m_1 - m_0 + 1)(n_1 - n_0 + 1) > m_1 + n_1 - 1$$

$$\text{si y sólo si } (m_1 - m_0)(n_1 - n_0) + (m_1 - m_0) + (n_1 - n_0) + 1 > m_1 + n_1 - 1$$

$$\text{si y sólo si } d > \sqrt{m_0 + n_0 - 2}.$$

La siguiente tabla muestra dicha configuración.

TABLA 15.7
CONJUNTO DE INDICES EN LA CONFIGURACION RECTANGULAR

$m \cdot n \dots$	$n_0 - 1$	n_0	$n_0 + 1$	$\dots$	n_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
$m_0 - 1$	$\dots \Delta(m_0 - 1, n_0 - 1)$	$\Delta(m_0 - 1, n_0)$	$\Delta(m_0 - 1, n_0 + 1)$	$\dots \Delta(m_0 - 1, n_1)$	
m_0	$\dots \Delta(m_0, n_0 - 1)$	$\Delta(m_0, n_0)$	$\Delta(m_0, n_0 + 1)$	$\dots \Delta(m_0, n_1)$	
$m_0 + 1$	$\dots \Delta(m_0 + 1, n_0 - 1)$	$\Delta(m_0 + 1, n_0)$	$\Delta(m_0 + 1, n_0 + 1)$	$\dots \Delta(m_0 + 1, n_1)$	
$m_0 + 2$	$\dots \Delta(m_0 + 2, n_0 - 1)$	$\Delta(m_0 + 2, n_0)$	$\Delta(m_0 + 2, n_0 + 1)$	$\dots \Delta(m_0 + 2, n_1)$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
m_1	$\dots \Delta(m_1, n_0 - 1)$	$\Delta(m_1, n_0)$	$\Delta(m_1, n_0 + 1)$	$\dots \Delta(m_1, n_1)$	

b) RESTRICCIONES SOBRE LA MATRIZ DE DERIVADAS PARCIALES (A).

La matriz $\mathcal{S}$ también será singular si la matriz A de derivadas parciales tiene un mal comportamiento. Tales

consideraciones conducen a demostrar que bajo las condiciones del siguiente lema, cada entrada de la matriz A es igual a cero.

Lema 1.5.2 Para cualquier proceso ARMA(p,q) se tiene que $\frac{\partial \Delta(m,n)}{\partial \rho_k} = 0$ para $k > 1$, si $m \geq n + q - p$ y $n \geq p + 2$.

Demostración Obsérvese que la matriz (1.5.1) se puede describir como:

$$M(m,n) = [C^m, C^{m-1}, \dots, C^{m-n+1}]$$

$$\text{en donde } C^i = [\rho_i, \rho_{i+1}, \dots, \rho_{i+m-1}]'$$

Nótese además que $\frac{\partial \Delta(m,n)}{\partial \rho_k}$ es la suma de cofactores de todas las entradas ρ_k en $M(m,n)$. Para demostrar que bajo las condiciones del lema (1.5.2) cada cofactor debe ser igual a cero, es suficiente con mostrar que cualesquier $m-1$ columnas de $M(m,n)$ son linealmente dependientes.

Del teorema (1.5.1) se sigue que

$$C^i = \sum_{l=1}^p \phi_l C^{i-1} \quad \text{para } i = m, m-1, \dots, m-n+p+1 \quad (1.5.6)$$

ahora, sea M^v la matriz que se obtiene al eliminar la columna C^v de $M(m,n)$. Si $v < m-p$ entonces de (1.5.6) se sigue que

$$C^m = \sum_{l=1}^p \phi_l C^{m-1}$$

similarmente si $\nu > m-n+p+1$ de (1.5.6) se sigue que

$$C^{\nu-1} = \sum_{l=1}^p \phi_l C^{\nu-1-l}$$

demostrando con esto que las columnas de M^ν son linealmente dependientes.

Ahora, si $m-p \leq \nu \leq m-n+p+1$, considérense las columnas

$$\left\{ C^l \right\}_{l=m-p-1}^m \quad \text{Se sabe que}$$

$$C^m - \sum_{l=1}^p \phi_l C^{m-l-1} = 0 \quad (1.5.7)$$

supóngase que $\phi_{m-\nu-1} \neq 0$, entonces como también

$$C^{m-1} - \sum_{l=1}^p \phi_l C^{m-l-1} = C^{m-1} - \sum_{\substack{l=1 \\ l \neq m-\nu-1}}^p \phi_l C^{m-l-1} - \phi_{m-\nu-1} C^\nu = 0$$

se tiene que

$$C^\nu = \phi_{m-n-1}^{-1} \left\{ C^{m-1} - \sum_{\substack{l=1 \\ l \neq m-\nu-1}}^p \phi_l C^{m-l-1} \right\}$$

por lo cual

$$0 = C^m - \sum_{l=1}^p \phi_l C^{m-l-1} = C^m - \sum_{\substack{l=1 \\ l \neq m-\nu}}^p \phi_l C^{m-l-1} - \phi_{m-\nu} C^\nu$$

$$\begin{aligned}
&= C^m - \sum_{\substack{l=1 \\ l \neq m-\nu}}^p \phi_l C^{m-1} - \phi_{m-\nu} \left[\phi_{m-\nu-1}^{-1} \left\{ C^{m-1} - \sum_{\substack{l=1 \\ l \neq m-\nu-1}}^p \phi_l C^{m-1-1} \right\} \right] \\
&= C^m - \sum_{\substack{l=1 \\ l \neq m-\nu}}^p \phi_l C^{m-1} - \phi_{m-\nu} \phi_{m-\nu-1}^{-1} C^{m-1} + \phi_{m-\nu} \phi_{m-\nu-1}^{-1} \sum_{\substack{l=1 \\ l \neq m-\nu-1}}^p \phi_l C^{m-1-1} \\
&= C^m - \phi_1 C^{m-1} - \sum_{\substack{l=2 \\ l \neq m-\nu}}^p \phi_l C^{m-1} - \phi_{m-\nu} \phi_{m-\nu-1}^{-1} C^{m-1} + \phi_{m-\nu} \phi_{m-\nu-1}^{-1} \sum_{\substack{l=2 \\ l \neq m-\nu}}^{p+1} \phi_{l-1} C^{m-1} \\
&= C^m - (\phi_1 + \phi_{m-\nu} \phi_{m-\nu-1}^{-1}) C^{m-1} - \sum_{\substack{l=2 \\ l \neq m-\nu}}^p (\phi_l - \phi_{m-\nu} \phi_{m-\nu-1}^{-1} \phi_{l-1}) C^{m-1} \\
&\quad + \phi_{m-\nu} \phi_{m-\nu-1}^{-1} \phi_p C^{m-p-1}
\end{aligned}$$

y esto demuestra que las columnas de M^ν son linealmente dependientes

Si $\phi_{m-\nu-1} = 0$, no habrá contribución de C^ν en (1.5.7) y las columnas de M^ν serán linealmente dependientes ■

Una vez demostrado el lema (1.5.2) es fácil demostrar el siguiente teorema, en cuyas condiciones se observan más claramente las restricciones que hay en la prueba de hipótesis (1.5.3).

Teorema 1.5.2 Con respecto a la prueba de hipótesis (1.5.3), si $m_1 \geq n_1 + q - p$ y $n_1 \geq p + 2$, para $1 \leq l \leq T$, entonces $\mathcal{S}$ es singular.

Demostración Es inmediata del lema (1.5.2).

En términos de las configuraciones consideradas anteriormente, las restricciones del teorema (1.5.2) no tendrán efecto sobre la prueba de hipótesis, al usar la configuración de línea vertical, si $n_0 \leq p + 1$ o línea horizontal si $m_0 \leq q + 1$. Similarmente dichas restricciones no tendrán efecto al usar la configuración ortogonal si $n_0 \leq p + 1$ y $m_0 \leq q + 1$. Por último si se usa la configuración rectangular, tampoco habrá efecto sobre la prueba de hipótesis si $m_1 \leq p + 1$ o si $m_1 \leq q + 1$.

1.5.5.- Equivalencia entre los arreglos R y el método de la esquina.

El siguiente resultado que muestra la equivalencia entre los arreglos R y el método de la esquina es mencionado en el trabajo de Gooijer, Abraham, Gould y Robinson (1985). Sin embargo, la demostración que aquí se presenta no es mencionada en dicho trabajo.

Del corolario (1.3.1) se tiene que $R_{n+1}(\rho_m) = 0$ para $n = p$ y $m \geq q - p + 1$. Por otra parte, de la relación (1.3.4) se sigue que el comportamiento de los ceros observado en la columna $p + 1$ de los arreglos R, se mantendrá en las columnas $p + 2, p + 3, \dots$. Además del teorema (1.5.1) se sabe que

$\Delta(m,n) = 0$ para $m \geq q+1$ y $n \geq p+1$. Ahora, por la relación (1.3.2) es claro que $R_n(\rho_m) = 0$ si y sólo si $H_n[\rho_m] = 0$. Por otra parte, se tiene que

$$\Delta(m,n) = \begin{vmatrix} \rho_m & \rho_{m-1} & \dots & \rho_{m-n+1} \\ \rho_{m+1} & \rho_m & \dots & \rho_{m-n+2} \\ \vdots & \vdots & & \vdots \\ \rho_{m+n-1} & \rho_{m+n-2} & \dots & \rho_m \end{vmatrix}$$

y

$$H_n(f_{m-n+1}) = \begin{vmatrix} \rho_{m-n+1} & \rho_{m-n+2} & \dots & \rho_m \\ \rho_{m-n+2} & \rho_{m-n+3} & \dots & \rho_{m+1} \\ \vdots & \vdots & & \vdots \\ \rho_m & \rho_{m+1} & \dots & \rho_{m+n-1} \end{vmatrix}$$

de donde se observa claramente que, mediante un intercambio de columnas en $\Delta(m,n)$ se obtiene

$$\Delta(m,n) = (-1)^{[|n/2|]} H_n[\rho_{m-n+1}]$$

Con esto se demuestra que el método de la esquina puede considerarse como un caso particular de los arreglos R estudiados en la sección 1.3.

1.6 REGRESIONES ITERADAS Y LA FUNCIÓN DE AUTOCORRELACIÓN MUESTRAL EXTENDIDA (FAME).

En esta sección se estudia con detalle un método útil para la identificación del orden de un modelo de series de tiempo mixto $ARMA(p,q)$, ya sea estacionario o no estacionario, propuesto por Tsay y Tiao (1984). Se desarrollan básicamente las siguientes dos ideas de estos autores.

- i) Proponer estimadores consistentes por mínimos cuadrados para los parámetros autorregresivos.
- ii) Presentar un procedimiento para la especificación del orden de un modelo dentro de la clase de modelos $ARMA(p,q)$ estacionarios y no estacionarios.

El problema de estimar los parámetros autorregresivos de un modelo $ARMA(p,q)$ mediante un ajuste de mínimos cuadrados ordinarios (MCO) ha sido tratado muy ampliamente. Después de que Mann y Wald (1943) probaron que se obtenían estimadores consistentes en el caso autorregresivo ($q = d = 0$) sólo avances marginales se habían logrado hasta Tiao y Tsay (1983). En este trabajo se realiza un estudio completo del problema, probando la consistencia mencionada en los casos $p = 0$ ó $q = 0$.

Para procesos mixtos, aún estacionarios, los estimadores de MCO de los parámetros autorregresivos son sesgados e inconsistentes. La obtención de estimadores de MCO consistentes para procesos $ARMA(p,q)$ en general, se logra en Tsay y Tiao (1984) mediante un proceso de iteración de los ajustes autorregresivos. La idea del proceso iterativo les conduce además, a introducir la Función de Autocorrelación Muestral Extendida (FAME) como una herramienta para la identificación

del orden (p,q) de un modelo ARMA.

La FAME no es más que una generalización de la FAC que en opinión de sus autores, tiene dos grandes ventajas.

- a) Se aplica a series de tiempo no estacionarias, evitando con esto la diferenciación de la serie para volverla estacionaria.
- b) Supera los problemas que se presentan en la identificación de procesos mixtos mediante la FAC y la FACP.

A continuación se procede a presentar con detalle las ideas antes mencionadas, pero antes se recomienda la lectura del Apéndice 4, en el cual se presentan algunas definiciones y resultados básicos para la comprensión del contenido de esta sección.

1.6.1 Regresiones iteradas.

El objetivo principal es encontrar estimadores consistentes por el método de MCO para los parámetros autorregresivos ϕ_1 's. La idea desarrollada por Tsay y Tiao para solventar la inconsistencia en la estimación de los parámetros ϕ_1 's mediante un ajuste $AR(p)$, es la de incluir los residuos de dicho ajuste como nuevo regresor, reestimando el modelo y repitiendo el proceso; es decir, partiendo del modelo

$$Z_t = \sum_{i=1}^p \phi_i Z_{t-i} - \sum_{j=1}^q \theta_j a_{t-j} + a_t \quad (1.6.1)$$

primero se hace un ajuste $AR(p)$ a estos datos, dicha regresión puede escribirse como

$$Z_t = \sum_{i=1}^p \hat{\phi}_{i(p)}^{(0)} Z_{t-i} + e_{p,t}^{(0)} \quad \text{para } t = p+1, \dots, n \quad (1.6.2)$$

en donde el supraíndice (0) indica la autorregresión ordinaria, mientras que el subíndice (p) indica el orden del ajuste AR y

$e_{p,t}^{(0)}$ es el error en dicho ajuste. Se sabe que los estimadores

de MCO $\hat{\phi}_{i(p)}^{(0)}$ son consistentes para los ϕ_i 's, esto es

$$\hat{\phi}_{i(p)}^{(0)} \xrightarrow{P} \phi_i \quad \text{para } i = 1, 2, \dots, p$$

si el proceso en cuestión sigue un modelo puramente $AR(p)$ ($q=0$) o un modelo puramente no estacionario $ARMA(p,q)$ ($\phi(B)=1$). Estos resultados fueron probados por Mann y Wald (1943) para el caso estacionario, mientras que para el caso no estacionario se pueden encontrar en Tiao y Tsay (1983).

Ahora, considérese el caso en que los $\hat{\phi}_{i(p)}^{(0)}$ son sesgados. Ya que los residuales estimados de (1.6.2)

$$e_{p,t}^{(0)} = Z_t - \sum_{i=1}^p \hat{\phi}_{i(p)}^{(0)} Z_{t-i}$$

no son ruido blanco para n suficientemente grande, entonces

los valores retrasados de $\hat{e}_{p,t-j}^{(0)}$ para $j > 0$, contienen alguna información acerca del proceso $\{Z_t\}$. Esto ha motivado la definición de la primera regresión iterada $AR(p)$ como

$$Z_t = \sum_{i=1}^p \hat{\phi}_{1(p)}^{(1)} Z_{t-1} + \beta_{1(p)}^{(1)} \hat{e}_{p,t-1}^{(0)} + e_{p,t}^{(1)} \quad \text{para } t = p+2, \dots, n \quad (1.6.3)$$

en donde el supraíndice (1) indica la primera regresión iterada y $e_{p,t}^{(1)}$ es el error correspondiente a esta primera regresión iterada. Se demostrará posteriormente que los estimadores por MCO $\hat{\phi}_{1(p)}^{(1)}$ de esta regresión son consistentes, esto es

$$\hat{\phi}_{1(p)}^{(1)} \xrightarrow{P} \phi_1 \quad \text{para } i = 1, \dots, p$$

si $1 \geq q$ o $\phi(B) = 1$.

Similarmente, si los $\hat{\phi}_{1(p)}^{(1)}$ son inconsistentes, los valores retrasados de los residuales de (1.6.3)

$$\hat{e}_{p,t}^{(1)} = Z_t - \sum_{i=1}^p \hat{\phi}_{1(p)}^{(1)} Z_{t-1} - \beta_{1(p)}^{(1)} \hat{e}_{p,t-1}^{(0)}$$

también contienen alguna información útil acerca del proceso $\{Z_t\}$. Entonces se define la segunda regresión iterada como

$$Z_t = \sum_{i=1}^p \hat{\phi}_{1(p)}^{(2)} Z_{t-1} + \beta_{1(p)}^{(2)} \hat{e}_{p,t-1}^{(1)} + \beta_{2(p)}^{(2)} \hat{e}_{p,t-2}^{(0)} + e_{p,t}^{(2)}$$

para $t = p+3, \dots, n$ (1.6.4)

en donde $\hat{e}_{p,t-2}^{(0)}$ y $\hat{e}_{p,t-1}^{(1)}$ son los residuales de (1.6.2) y (1.6.3) respectivamente; se demostrará posteriormente que

$$\hat{\phi}_{1(p)}^{(2)} \xrightarrow{P} \phi_1 \quad \text{para } i = 1, \dots, p$$

si $2 \geq q$ o $\phi(B) = 1$.

En la práctica el orden real (p, q) de un proceso ARMA, usualmente es desconocido. Esta consideración conduce a plantear el proceso de regresiones iteradas a partir de un ajuste AR(m), con $m > 0$, con lo que se trata simultáneamente el problema de estimación y el de identificación del modelo. Específicamente, la j -ésima regresión iterada AR(m) de una serie de tiempo $\{Z_t\}$ se define como:

$$Z_t = \sum_{i=1}^m \hat{\phi}_{1(m)}^{(j)} Z_{t-i} + \sum_{i=1}^j \hat{\beta}_{1(m)}^{(j)} \hat{e}_{m,t-1}^{(j-1)} + e_{m,t}^{(j)} \quad (1.6.5)$$

para $t = m+j+1, \dots, n$; $j = 0, \dots$ y $m = 1, 2, \dots$

en donde

$$\hat{e}_{m,t}^{(j)} = Z_t - \sum_{i=1}^m \hat{\phi}_{1(m)}^{(j)} Z_{t-i} + \sum_{i=1}^j \hat{\beta}_{1(m)}^{(j)} \hat{e}_{m,t-1}^{(j-1)} \quad (1.6.6)$$

es el residual de la j -ésima regresión iterada AR(m) y los $\hat{\phi}_{1(m)}^{(j)}$'s, $\hat{\beta}_{1(m)}^{(j)}$'s son los estimadores por el método de MCO de los $\phi_{1(m)}^{(j)}$'s y $\beta_{1(m)}^{(j)}$ respectivamente.

1.6.2 Propiedades de los estimadores de las regresiones iteradas.

En esta subsección se demostrará como los estimadores de la j -ésima iteración $\hat{\phi}_{1(m)}^{(j)}$ pueden ser calculados recursivamente mediante mínimos cuadrados por etapas. Para esto, sea

$$\hat{\phi}_m^{(j)}(B) = 1 - \hat{\phi}_{1(m)}^{(j)} B - \dots - \hat{\phi}_{m(m)}^{(j)} B^m \quad (1.6.7)$$

el estimador del polinomio autorregresivo en la j -ésima regresión iterada AR(m) (1.6.5).

Los residuales de las regresiones iteradas, $\hat{e}_{m,t}^{(j)}$, verifican una relación recurrente de gran trascendencia para las propiedades de los estimadores $\hat{\phi}_{1(m)}^{(j)}$, expresada en el siguiente lema.

Lema 1.6.1 Para cualesquiera enteros positivos m y j se verifica que

$$\hat{e}_{m,t}^{(j)} = \hat{e}_{m+1,t}^{(j-1)}$$

Demostración Obsérvese que

$$\hat{e}_{m,t}^{(0)} = Z_t - \sum_{i=1}^m \hat{\phi}_{i(m)}^{(0)} Z_{t-i}$$

y

$$\hat{e}_{m,t-1}^{(0)} = Z_{t-1} - \sum_{i=1}^m \hat{\phi}_{i(m)}^{(0)} Z_{t-1-i} = \hat{\phi}_m^{(0)}(B) Z_{t-1}$$

luego (1.6.3) puede reescribirse como

$$\begin{aligned}
 Z_t &= \sum_{i=1}^m \Phi_{1(m)}^{(1)} Z_{t-i} + \beta_{1(m)}^{(1)} \hat{e}_{m,t-1}^{(0)} + e_{m,t}^{(1)} \\
 &= \sum_{i=1}^m \Phi_{1(m)}^{(1)} Z_{t-i} + \beta_{1(m)}^{(1)} [\hat{\Phi}_m^{(0)}(B) Z_{t-1}] + e_{m,t}^{(1)} \\
 &= (\Phi_{1(m)}^{(1)} + \beta_{1(m)}^{(1)}) Z_{t-1} + (\Phi_{2(m)}^{(1)} - \beta_{1(m)}^{(1)} \hat{\Phi}_{1(m)}^{(0)}) Z_{t-2} + \dots \\
 &\quad + (\Phi_{m(m)}^{(1)} - \beta_{1(m)}^{(1)} \hat{\Phi}_{m-1(m)}^{(0)}) Z_{t-m} - \beta_{1(m)}^{(1)} \hat{\Phi}_{m(m)}^{(0)} Z_{t-m-1} + e_{m,t}^{(1)}
 \end{aligned}$$

por otra parte , se tiene que

$$Z_t = \sum_{i=1}^{m+1} \Phi_{1(m+1)}^{(0)} Z_{t-i} + e_{m+1,t}^{(0)}$$

de donde se sigue que

$$\hat{e}_{m,t}^{(1)} = \hat{e}_{m+1,t}^{(0)}$$

Repetiendo el mismo argumento se llega a la siguiente igualdad

$$\hat{e}_{m,t}^{(j)} = \hat{e}_{m+1,t}^{(j-1)}$$

Lema 1.6.2 Para cualesquiera enteros positivos m y j se verifica que

$$\hat{\Phi}_{m+j}^{(0)}(B) = \hat{\Phi}_m^{(j)}(B) - \sum_{l=1}^j \hat{\beta}_{1(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B) B^l$$

Demostración Se tiene que

$$Z_t = \sum_{i=1}^{m+j} \hat{\phi}_{i(m+j)}^{(0)} Z_{t-1} + e_{m+j,t}^{(0)}$$

entonces

$$\hat{e}_{m+j,t}^{(0)} = Z_t - \sum_{i=1}^{m+j} \hat{\phi}_{i(m+j)}^{(0)} Z_{t-1}$$

por otra parte, como

$$Z_t = \sum_{i=1}^m \hat{\phi}_{i(m)}^{(j)} Z_{t-1} + \sum_{i=1}^j \hat{\beta}_{i(m)}^{(j)} \hat{e}_{m,t-1}^{(j-1)} + e_{m,t}^{(j)}$$

se sigue del lema (1.6.1) que

$$\hat{e}_{m+j,t}^{(0)} = \hat{e}_{m,t}^{(j)}$$

o

$$\hat{\phi}_{m+j}^{(0)}(B) = \hat{\phi}_m^{(j)}(B) - \sum_{i=1}^j \hat{\beta}_{i(m)}^{(j)} \hat{e}_{m,t-1}^{(j-1)} \quad (1.6.8)$$

además como

$$Z_t = \sum_{i=1}^{m+j-1} \hat{\phi}_{i(m+j-1)}^{(0)} Z_{t-1} + e_{m+j-1,t}^{(0)}$$

entonces

$$\hat{e}_{m+j-1,t-1}^{(0)} = Z_{t-1} - \sum_{i=1}^{m+j-1} \hat{\phi}_{i(m+j-1)}^{(0)} Z_{t-1-1} = \hat{\phi}_{m+j-1}^{(0)} B B^1 Z_t$$

y como

$$\hat{e}_{m,t-1}^{(j-1)} = \hat{e}_{m+j-1,t-1}^{(0)}$$

se sigue que (1.6.8) puede reescribirse como

$$\hat{\Phi}_{m+j}^{(0)}(B) = \hat{\Phi}_m^{(j)}(B) - \sum_{l=1}^j \hat{\beta}_{1(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B) B^l$$

Otra importante consecuencia del lema (1.6.1) es una expresión recurrente para los estimadores $\hat{\Phi}_{1(m)}^{(j)}$, que permite en último extremo calcularlos todos en función de estimadores $\hat{\Phi}_{1(m)}^{(0)}$, es decir, a partir de estimadores obtenidos en ajustes autorregresivos ordinarios.

Lema 1.6.3 Para cualesquiera enteros positivos m y j se verifica que

$$\hat{\Phi}_m^{(j)}(B) = \hat{\Phi}_{m+1}^{(j-1)}(B) + \alpha_m^{(j-1)} \hat{\Phi}_m^{(j-1)}(B) B$$

en donde

$$\alpha_m^{(j-1)} = \frac{-\hat{\Phi}_{m+1(m+1)}^{(j-1)}}{\hat{\Phi}_{m(m)}^{(j-1)}}$$

Demostración Este resultado para $j = 1$ es inmediato del lema (1.6.1) y de (1.6.3). Ahora, para $j > 1$ cambiando m por $m+1$ y j por $j - 1$ en la fórmula del lema (1.6.2), se obtiene

$$\hat{\Phi}_{m+j}^{(0)}(B) = \hat{\Phi}_{m+1}^{(j-1)}(B) - \sum_{l=1}^{j-1} \hat{\beta}_{1(m+1)}^{(j-1)} \hat{\Phi}_{m+j-l}^{(0)}(B) B^l$$

también de la fórmula del lema (1.6.2) y de este último

resultado se observa que

$$\begin{aligned}
 \hat{\Phi}_m^{(j)}(B) &= \hat{\Phi}_{m+j}^{(0)} + \sum_{l=1}^j \hat{\beta}_{l(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B)B^l \\
 &= \hat{\Phi}_{m+1}^{(j-1)}(B) + \sum_{l=1}^j \hat{\beta}_{l(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B)B^l - \sum_{l=1}^{j-1} \hat{\beta}_{l(m+1)}^{(j-1)} \hat{\Phi}_{m+j-l}^{(0)}(B)B^l \\
 &= \hat{\Phi}_{m+1}^{(j-1)}(B) + \hat{\gamma}_{1(m)}^{(j)} \hat{\Phi}_{m+j-1}^{(0)}(B)B + \sum_{l=2}^{j-1} \hat{\gamma}_{l(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B)B^l \\
 &\quad + \hat{\beta}_{j(m)}^{(j)} \hat{\Phi}_m^{(0)}(B)B^j \tag{1.6.9}
 \end{aligned}$$

en donde

$$\hat{\gamma}_{l(m)}^{(j)} = \hat{\beta}_{l(m)}^{(j)} - \hat{\beta}_{l(m+1)}^{(j-1)} \quad \text{para } l = 1, 2, \dots, j-1$$

Ahora, aplicando el lema (1.6.2) a $\hat{\Phi}_{m+j-1}^{(0)}(B)$ se sigue que

$$\hat{\Phi}_{m+j-1}^{(0)}(B) = \hat{\Phi}_m^{(j-1)}(B) - \sum_{l=1}^{j-1} \hat{\beta}_{l(m)}^{(j-1)} \hat{\Phi}_{m+j-l-1}^{(0)}(B)B^l$$

sustituyendo esta relación en (1.6.9) se observa que

$$\begin{aligned}
 \hat{\Phi}_m^{(j)}(B) &= \hat{\Phi}_{m+1}^{(j-1)}(B) + \hat{\gamma}_{1(m)}^{(j)} [\hat{\Phi}_m^{(j-1)}(B) - \sum_{l=1}^{j-1} \hat{\beta}_{l(m)}^{(j-1)} \hat{\Phi}_{m+j-l-1}^{(0)}(B)B^l]B \\
 &\quad + \sum_{l=2}^{j-1} \hat{\gamma}_{l(m)}^{(j)} \hat{\Phi}_{m+j-l}^{(0)}(B)B^l + \hat{\beta}_{j(m)}^{(j)} \hat{\Phi}_m^{(0)}(B)B^j
 \end{aligned}$$

de donde se sigue

$$\begin{aligned}
\hat{\Phi}_m^{(j)}(B) - \hat{\Phi}_{m+1}^{(j-1)}(B) - \hat{\gamma}_{1(m)}^{(j)} \hat{\Phi}_m^{(j-1)}(B)B &= \sum_{l=2}^{j-1} \hat{\gamma}_{1(m)}^{(j)} \hat{\Phi}_{m+j-1}^{(0)}(B)B^l \\
- \sum_{l=1}^{j-1} \hat{\gamma}_{1(m)}^{(j)} \hat{\beta}_{1(m)}^{(j-1)} \hat{\Phi}_{m+j-1-l}^{(0)}(B)B^l B + \hat{\beta}_{j(m)}^{(j)} \hat{\Phi}_m^{(0)} B B^j \\
= \sum_{l=2}^{j-1} (\hat{\gamma}_{1(m)}^{(j)} - \hat{\gamma}_{1(m)}^{(j)} \hat{\beta}_{1-1(m)}^{(j-1)}) \hat{\Phi}_{m+j-1}^{(0)}(B)B^l + (\hat{\beta}_{j(m)}^{(j)} - \hat{\gamma}_{1(m)}^{(j)} \hat{\beta}_{j-1}^{(j-1)}) \\
\hat{\Phi}_m^{(0)}(B)B^j \quad (1.6.10)
\end{aligned}$$

Nótese que el lado izquierdo de la relación (1.6.10) es un polinomio en B de grado $m+1$. Por lo tanto, aquellos coeficientes de B^{m+1} para $i \geq 2$ en el polinomio del lado derecho de (1.6.10) deben ser iguales a cero. Más aún, como para cualquier n dado todos los estimadores por MCO de cuadrados de $\hat{\Phi}_{1(m)}^{(j)}$ son diferentes de cero, se tiene que

$$\hat{\beta}_{j(m)}^{(j)} - \hat{\gamma}_{1(m)}^{(j)} \hat{\beta}_{j-1(m)}^{(j-1)} = 0,$$

$$\hat{\gamma}_{1(m)}^{(j)} - \hat{\gamma}_{1(m)}^{(j)} \hat{\beta}_{1-1(m)}^{(j-1)} = 0$$

entonces

$$\hat{\Phi}_m^{(j)}(B) = \hat{\Phi}_{m+1}^{(j-1)}(B) + \hat{\gamma}_{1(m)}^{(j)} \hat{\Phi}_m^{(j-1)}(B)B \quad (1.6.11)$$

finalmente, igualando los coeficientes de B^{m+1} en (1.6.11) se observa que

$$\hat{\gamma}_{1(m)}^{(j)} \hat{\Phi}_{m(m)}^{(j-1)} = - \hat{\Phi}_{m+1(m+1)}^{(j-1)}$$

con lo cual se termina la demostración

■

Obsérvese que se supone que todos los estimadores $\hat{\phi}_{1(m)}^{(j)}$ son diferentes de cero, lo cual se sigue a partir de la hipótesis de que el proceso no es puramente determinístico, esto implica que todos los estimadores obtenidos en las regresiones iteradas son v.v.a.a continuas diferentes de cero con probabilidad uno, para algún n fijo.

Corolario 16.1 Se verifica que

$$\hat{\phi}_{1(m)}^{(j)} = \hat{\phi}_{1-1(m)}^{(j-1)} - \hat{\phi}_{1(m)}^{(j-1)} \frac{\hat{\phi}_{m+1(m+1)}^{(j-1)}}{\hat{\phi}_{1(m)}^{(j-1)}}$$

para $i = 1, 2, \dots, m$; $m \geq 1$ y $j \geq 1$ en donde $\hat{\phi}_{0(m)}^{(j-1)} = 1$.

Demostración Es inmediata del lema (1.6.3)

1.6.3 Relación con las soluciones de las ecuaciones de momentos.

Supóngase que el proceso $\{Z_t\}$ es estacionario. Se estudiará ahora la relación entre los estimadores de los parámetros autorregresivos de un ajuste $AR(m)$ en su j -ésima iteración y las soluciones de las ecuaciones de momentos muestrales (o las ecuaciones generalizadas de Yule-Walker estimadas). Las ecuaciones normales de la j -ésima regresión

iterada AR(m) se obtienen a partir de la siguiente expresión

$$S = \sum_t (e_{m,t}^{(j)})^2 = \sum_t \left\{ Z_t - \sum_{l=1}^m \phi_{l(m)}^{(j)} Z_{t-l} - \sum_{l=1}^j \beta_{l(m)} \hat{e}_{m,t-l}^{(j-1)} \right\}^2$$

entonces

$$\begin{aligned} \frac{\partial S}{\partial \phi_h} &= -2 \sum_t (Z_{t-h}) \left\{ Z_t - \sum_{l=1}^m \phi_{l(m)}^{(j)} Z_{t-l} - \sum_{l=1}^j \beta_{l(m)} \hat{e}_{m,t-l}^{(j-1)} \right\} \\ &= -2 \sum_t (Z_{t-h}) e_{m,t}^{(j)} = 0 \end{aligned}$$

si y sólo si

$$\sum_t Z_{t-h} e_{m,t}^{(j)} = 0 \quad \text{para } h = 1, 2, \dots, m \quad (1.6.12)$$

y

$$\frac{\partial S}{\partial \beta_{l(m)}} = -2 \sum_t \hat{e}_{m,t-l}^{(j-1)} e_{m,t}^{(j)} = 0$$

si y sólo si

$$\sum_t \hat{e}_{m,t-l}^{(j-1)} e_{m,t}^{(j)} = 0 \quad \text{para } l = 1, 2, \dots, j \quad (1.6.13)$$

en donde $\sum_t$ indica que la suma es tomada desde $m + j + 1$ hasta

n a menos que se indique otra cosa. Luego como

$$\hat{e}_{m,t-l}^{(k)} = \hat{e}_{m+k,t-l}^{(0)} = \hat{\phi}_{m+k}^{(0)} (B) Z_{t-l}$$

se sigue que las relaciones (1.6.12) y (1.6.13) se reducen a

$$\sum_t Z_{t-h} e_{m,t}^{(j)} = 0 \quad \text{para } h = 1, 2, \dots, m+j \quad (1.6.14)$$

Nótese que se pueden dividir las $k+j$ ecuaciones de (1.6.14) en los dos siguientes grupos de ecuaciones

i) Para $h = 1, 2, \dots, j$ de la siguiente relación

$$\sum_t Z_{t-h} e_{m,t}^{(j)} = \sum_t Z_{t-h} \left\{ Z_t - \sum_{i=1}^m \phi_{i(m)}^{(j)} Z_{t-i} - \sum_{i=1}^j \beta_{i(m)}^{(j)} \hat{e}_{m,t-i}^{(j-1)} \right\} = 0$$

se tiene que

$$\sum_t Z_{t-h} \left\{ \sum_{i=1}^j \beta_{i(m)}^{(j)} \hat{e}_{m,t-i}^{(j-1)} \right\} = \sum_t Z_{t-h} \left\{ Z_t - \sum_{i=1}^m \phi_{i(m)}^{(j)} Z_{t-i} \right\}$$

o equivalentemente

$$\underline{H}_j \underline{\beta}_j = \underline{V}_j \quad (1.6.15)$$

en donde

$$\underline{\beta}_j = (\beta_{1(m)}^{(j)}, \dots, \beta_{j(m)}^{(j)})' ; \quad \underline{H}_j = \sum_t R_{j,t} Z'_{j,t}$$

$$\text{y } \underline{V}_j = \sum_t Z_{j,t} W_{j,t} ; \quad \text{con } R_{j,t} = (\hat{e}_{m,t-1}^{(j-1)}, \dots, \hat{e}_{m,t-j}^{(0)})'$$

$$Z_{j,t} = (Z_{t-1}, \dots, Z_{t-j})' \quad \text{y} \quad W_{j,t} = Z_t - \sum_{i=1}^m \phi_{i(m)}^{(j)} Z_{t-i}$$

ii) Para $h = j+1, \dots, k+j$

$$\sum_t Z_{t-h} W_{j,t} = \sum_{l=1}^j \beta_{l(m)}^{(j)} \left\{ Z_{t-h} \hat{e}_{m,t-1}^{(j-1)} \right\} \quad (1.6.16)$$

A continuación se discutirán algunas propiedades asintóticas de las soluciones de (1.6.15) y (1.6.16). Para esto, sea $\underline{A}_j^{(\nu)}$ una matriz de orden $h \times h$, en donde

$$\underline{A}_h^{(\nu)} = (\underline{C}_h^{(\nu)}, \underline{C}_h^{(\nu-1)}, \dots, \underline{C}_h^{(\nu-h+1)}) \quad (1.6.17)$$

y

$$\underline{C}_h^{(\nu)} = (\rho_\nu, \rho_{\nu+1}, \dots, \rho_{n+h-1})'$$

Además considérese la siguiente propiedad asintótica de los residuales estimados.

Lema 1.6.4 Se verifica que

$$\frac{\sum_{m+1}^n \hat{e}_{m,t}^{(0)} Z_t}{\sum_1^n Z_t^2} \xrightarrow{p} \frac{|\underline{A}_{m+1}^{(0)}|}{|\underline{A}_m^{(0)}|}$$

Demostración Como

$$\frac{\sum_{m+1}^n \hat{e}_{m,t}^{(0)} Z_t}{\sum_1^n Z_t^2} = \frac{\sum_{m+1}^n [\hat{\Phi}_m^{(0)}(B) Z_t] Z_t}{\sum_1^n Z_t^2} = \hat{\Phi}_m^{(0)}(B) \hat{\rho}_0$$

Por otro lado, obsérvese que mediante operaciones elementales entre columnas se tiene que

$$\frac{|\hat{A}_{m+1}^{(0)}|}{|\hat{A}_m^{(0)}|} = \Phi_m^{(0)}(B)\rho_0$$

por lo tanto, de las propiedades asintóticas de la FAC se sigue que

$$\hat{\Phi}_m^{(0)}(B)\hat{\rho}_0 \xrightarrow{p} \Phi_m^{(0)}(B)\rho_0$$

Ahora, recordando que $\hat{e}_{m,t}^{(j-1)}$ es el residual de la (j-1)-ésima regresión iterada y por MCO, se tiene que

$$\sum_{t_0}^n Z_{t-h} \hat{e}_{m,t}^{(j-1)} = 0 \quad \text{para } h = 1, 2, \dots, m \quad (1.6.18)$$

en donde $t_0 = m+j+1$. Más aún, como $\hat{e}_{m,t}^{(j-1)} = \hat{e}_{m+j-1,t}^{(0)}$ se tiene también que

$$\sum_{t_0}^n Z_{t-h} \hat{e}_{m,t}^{(j-1)} = 0 \quad \text{para } h = m+1, \dots, m+j-1 \quad (1.6.19)$$

Por lo tanto, la suma

$$\sum_t Z_{t-h} \hat{e}_{m,t}^{(j-1)} = - \sum_{t_0}^{n+j-1} Z_{t-h} \hat{e}_{m,t}^{(j-1)} \quad (1.6.20)$$

es $O_p(1)$ dado que sólo tiene un número finito de sumandos (independientes de n).

Ahora bien, la serie Z_t es estacionaria, luego la suma satisface

$$\sum Z_{t-h} \hat{e}_{m,t-1}^{(j-1)} = O_p(1) \quad \text{para } h = i+1, \dots, m+j \quad (1.6.21)$$

Considérese la relación (1.6.15), por los lemas (1.6.1) y (1.6.4) se sigue que los elementos de H_j pueden ser normalizados para obtener

$$\frac{\left[\sum_t Z_{t-h} \hat{e}_{m,t-h}^{(j-h)} \right]}{\sum_t Z_t^2} = \frac{\sum_t Z_{t-h} \hat{e}_{m+j-h,t-h}^{(0)}}{\sum_t Z_t^2} \xrightarrow{p} \frac{|\Delta_{m+j-h+1}^{(0)}|}{|\Delta_{m+j-h}^{(0)}|}$$

$$\text{para } h = 1, 2, \dots, j \quad (1.6.22)$$

por (1.6.21) y (1.6.22) $\left[\sum_t Z_t^2 \right]^{-1} H_j$ converge en probabilidad a una matriz triangular superior no singular. Escribiendo

$\hat{\beta}_j' = (\hat{\beta}_{1(m)}^{(j)}, \dots, \hat{\beta}_{j(m)}^{(j)})$ se sigue que

$$\hat{\beta}_j = \left\{ \left[\sum_t Z_t^2 \right]^{-1} H_j \right\}^{-1} \left[\sum_t Z_t^2 \right]^{-1} \hat{y}_j \quad (1.6.23)$$

dicha solución existe si $\hat{y}_j$ existe, en donde $\hat{y}_j$ se obtiene

a partir de $\underline{Y}_j$ al reemplazar $\hat{W}_{j,t} = \hat{\Phi}_{\underline{m}}^{(j)}(B)Z_t$. En otras palabras, la existencia de la solución de (1.6.15) sólo depende de la existencia de los $\hat{\Phi}_{1(m)}^{(j)}$'s.

Considérese el sistema de ecuaciones (1.6.16). Si $\hat{\Phi}_{1(m)}^{(j)}$ para $i = 1, 2, \dots, m$ son finitos y satisfacen estas ecuaciones, entonces $\hat{\beta}_j = O_p(1)$, lo cual se sigue de (1.6.23). Así pues, dividiendo por $\sum_t Z_t^2$ el sistema (1.6.16) se transforma en

$$\hat{\rho}_h - \sum_{i=1}^m \hat{\Phi}_{1(m)}^{(j)} \hat{\rho}_{h-1} = O_p(n^{-1}) \quad \text{para } h = j+1, \dots, m+j \quad (1.6.24)$$

ahora, de (1.6.23) y (1.6.24) se sigue que si $|\hat{\Delta}_{\underline{m}}^{(j)}| \neq 0$, entonces asintóticamente la ecuación normal (1.6.14) tiene una solución única. También se puede observar que ignorando el término $O_p(n^{-1})$, las ecuaciones en (1.6.24) son en realidad estimaciones de las ecuaciones generalizadas de Yule-Walker consideradas por Woodward y Gray (1981) (ver sección 1.4). Así, cualquier solución $\{\hat{\Phi}_{1(m)}^{(j)} \text{ con } i=1, 2, \dots, m\}$ de las ecuaciones normales (1.6.14), es asintóticamente equivalente a la solución de las ecuaciones generalizadas de Yule-Walker estimadas. Las ideas antes mencionadas se pueden resumir en el siguiente resultado

Lema 1.6.5 Supóngase que $\{Z_t\}$ sigue un modelo estacionario ARMA. Si $|A_m^{(j)}| \neq 0$, entonces los estimadores de los parámetros autorregresivos de un ajuste AR(m) en la j -ésima iteración, $m \geq 1$ y $j \geq 0$, son asintóticamente equivalentes a la solución de las ecuaciones generalizadas de Yule-Walker

$$\rho_h = \sum_{i=1}^m \phi_i \rho_{h-i} \quad \text{para } j+1 \leq h \leq j+m$$

y además son únicos.

1.6.4 Consistencia de los parámetros autorregresivos estimados mediante un ajuste AR(m) en su j -ésima iteración.

En esta subsección se consideran las propiedades de consistencia de los parámetros autorregresivos estimados mediante un ajuste AR(m) en su j -ésima iteración; $m \geq 1$ y $j \geq 0$. Para procesos estacionarios ARMA los resultados de consistencia serán establecidos a través de las estimaciones de las ecuaciones generalizadas de Yule-Walker. Sin embargo, dicha aproximación no se cumple en el caso no estacionario.

1) Procesos estacionarios ARMA

Para el caso estacionario se tiene que $\phi_p(B) = \phi_p$ en (1), esto es $U(B) = 1$. Ahora, Gersch (1970) demuestra que para un modelo estacionario ARMA(p,q) las soluciones de

$$\hat{\rho}_h - \sum_{i=1}^p \hat{\phi}_i \hat{\rho}_{h-i} = 0 \quad \text{para } h = q+1, \dots, q+p \quad (1.6.25)$$

proporcionan estimadores consistentes de los parámetros autorregresivos con una tasa de convergencia $O_p(n^{-1/2})$, esto es

$$\hat{\phi}_i = \phi_i + O_p(n^{-1/2}) \quad \text{con } i = 1, \dots, p$$

Este resultado, junto con la condición $|\Lambda_p^{(q)}| \neq 0$ asegura que las ecuaciones (1.6.25) convergen en probabilidad a las ecuaciones generalizadas de Yule-Walker

$$\rho_h - \sum_{i=1}^p \phi_i \rho_{h-i} = 0 \quad \text{para } h = q+1, \dots, q+p$$

En consecuencia, la consistencia estaría probada si la solución de las ecuaciones de Yule-Walker fuese única. El siguiente lema dá condiciones sobre j y m que aseguran dicha unicidad y, por ende, la consistencia.

Lema 1.6.6 Si $\{Z_t\}$ sigue un modelo ARMA(p, q), entonces se verifica que

a) $|\Lambda_p^{(j)}| \neq 0$ para $j \geq q$

b) $|\Lambda_m^{(q)}| \neq 0$ para $m \geq p$

Demostración Gersch (1970) demuestra que $|\Delta_p^{(q)}| \neq 0$,
por otra parte, se tiene que

$$\Delta_p^{(j+1)} = \begin{bmatrix} \rho_{j+1} & \rho_j & \cdots & \rho_{j-p+2} \\ \rho_{j+2} & \rho_{j+1} & \cdots & \rho_{j-p+3} \\ \vdots & \vdots & & \vdots \\ \rho_{j+p} & \rho_{j+p-1} & \cdots & \rho_{j+1} \end{bmatrix}$$

y como

$$\rho_{j+1} = \sum_{i=1}^p \phi_i \rho_{j+1-i} \quad \text{para } j \geq q$$

se sigue que

$$|\Delta_p^{(j+1)}| = \begin{vmatrix} \phi_1 \rho_j + \phi_2 \rho_{j-1} + \cdots + \phi_p \rho_{j-p+1} & \rho_j & \cdots & \rho_{j-p+2} \\ \phi_1 \rho_{j+1} + \phi_2 \rho_j + \cdots + \phi_p \rho_{j-p+2} & \rho_{j+1} & \cdots & \rho_{j-p+3} \\ \vdots & \vdots & & \vdots \\ \phi_1 \rho_{j+p-2} + \phi_2 \rho_{j+p-3} + \cdots + \phi_p \rho_{j-1} & \rho_{j+p-1} & \cdots & \rho_{j+1} \end{vmatrix}$$

$$= (-1)^{p-1} \phi_p |\Delta_p^{(j)}|$$

además como

$$\underline{A}_{m+1}^{(q)} = \begin{bmatrix} \rho_q & \rho_{q-1} & \cdots & \rho_{q-m} \\ \rho_{q+1} & \rho_q & \cdots & \rho_{q-m+1} \\ \vdots & \vdots & & \\ \rho_{q+m} & \rho_{q+m-1} & \cdots & \rho_q \end{bmatrix}$$

y de

$$\rho_j = \sum_{i=1}^p \phi_i \rho_{j-i} \quad \text{para } j \geq q+1$$

se sigue que

$$|\underline{A}_{m+1}^{(q)}| = \begin{vmatrix} \rho_q & \rho_{q-1} & \cdots & \rho_{q-m} \\ \phi_1 \rho_q + \phi_2 \rho_{q-1} + \cdots + \phi_p \rho_{q-p+1} & \rho_q & \cdots & \rho_{q-m+1} \\ \vdots & \vdots & & \vdots \\ \phi_1 \rho_{q+m-1} + \phi_2 \rho_{q+m-2} + \cdots + \phi_p \rho_{q+m-p} & \rho_{q+m-1} & \cdots & \rho_q \end{vmatrix}$$

$$= \left\{ \sum_{i=0}^p \phi_i \rho_{q-i} \right\} |\underline{A}_m^{(q)}| \quad \text{con } \phi_0 = 1 \quad \text{y } m \geq p$$

luego entonces, la demostración se sigue directamente de estas dos últimas relaciones, puesto que $\phi_p \neq 0$ y

$$\sum_{i=0}^p \phi_i \rho_{q-i} = \phi_p(B) \rho_q \neq 0$$

Como una consecuencia inmediata del lema (1.6.6) se tiene que para un modelo estacionario ARMA(p,q), la solución de las

ecuaciones

$$\hat{\rho}_h - \sum_{i=1}^m \hat{\phi}_i \hat{\rho}_{h-i} = 0 \quad ; \quad h = j+1, \dots, j+m$$

para

$$i) \quad m \geq p \quad y \quad j = q \quad o$$

$$ii) \quad m = p \quad y \quad j > q$$

también proporcionan estimadores consistentes de los parámetros autorregresivos. Entonces, usando el resultado (1.6.24) y el lema (1.6.5), se demuestra fácilmente el siguiente teorema.

Teorema 1.6.1 Supóngase que $\{Z_t\}$ sigue un proceso estacionario ARMA(p,q). Entonces

$$\hat{\phi}_{1(m)}^{(j)} = \phi_1 + O_p(n^{-1/2}) \quad \text{para } i = 1, 2, \dots, m$$

si

$$i) \quad m \geq p \quad y \quad j = q \quad o$$

$$ii) \quad m = p \quad y \quad j > q.$$

en donde se entiende que $\phi_i = 0$ para $i > p$.

ii) Procesos ARMA no estacionarios

Supóngase ahora que el proceso $\{Z_t\}$ sigue un modelo ARMA no estacionario, esto es $U_d(B) \neq 1$ en (1). Para este caso, se demuestra en Tiao y Tsay (1983) que los estimadores de MCO de un ajuste AR(m) con $m \geq d$ son asintóticamente estimados por

$$\hat{\phi}_m^{(0)}(B) \doteq \hat{U}_d^{(0)}(B) \hat{\phi}_{m-d}^{(0)}(B) \quad (1.6.26)$$

en donde los polinomios estimados $\hat{\phi}_m^{(0)}(B)$, $\hat{U}_d^{(0)}(B)$ y $\hat{\phi}_{m-d}^{(0)}(B)$ son, respectivamente, los estimadores de MCO de las siguientes autorregresiones

$$Z_t = \sum_{i=1}^m \phi_{i(m)}^{(0)} Z_{t-i} + e_{m,t}^{(0)}$$

$$Z_t = \sum_{i=1}^d U_{i(d)}^{(0)} Z_{t-i} + e_{d,t}^{(0)}$$

$$W_t = \sum_{i=1}^m \phi_{i(m-d)}^{(0)} W_{t-i} + f_t^{(0)}$$

en donde $W_t = \phi(B) Z_t$ es un proceso estacionario ARMA(p-d,q) igual a

$$\phi_{p-d}(B) Z_t = \theta_q(B) a_t$$

En ese mismo trabajo se demuestra también que

$$\hat{U}_d^{(0)}(B) = U_d(B) + O_p(n^{-1}) \quad (1.6.27)$$

Ahora, regresando a la j-ésima regresión iterada (1.6.5) se tiene el siguiente lema, útil para encontrar estimadores consistentes del factor no estacionario.

Lema 1.6.7 Se verifica que

$$\hat{\Phi}_m^{(j)}(B) \doteq \hat{U}_d^{(0)}(B) \hat{\Phi}_{m-d}^{(j)} \quad \text{para } j \geq 0$$

Demostración La demostración de este lema se hará por inducción sobre j . Supóngase que el lema se cumple para $j = h$, esto es

$$\hat{\Phi}_m^{(h)}(B) \doteq \hat{U}_d^{(0)}(B) \hat{\Phi}_{m-d}^{(h)}(B) \quad \text{para } m \geq d \quad (1.6.28)$$

entonces, por el lema (1.6.3) se tiene que

$$\hat{\Phi}_m^{(h+1)}(B) = \hat{\Phi}_{m+1}^{(h)}(B) + \hat{\alpha}^{(h)} \hat{\Phi}_m^{(h)}(B)B \quad (1.6.29)$$

en donde

$$\hat{\alpha}^{(h)} = \frac{-\hat{\Phi}_{m+1}^{(h)}}{\hat{\Phi}_m^{(h)}}$$

entonces, por la hipótesis de inducción (1.6.28) se tiene que

$$\hat{\Phi}_{m+1}^{(h)}(B) \doteq \hat{U}_d^{(0)}(B) \hat{\Phi}_{m-d+1}^{(h)}(B) \quad (1.6.30)$$

$$\hat{\Phi}_m^{(h)}(B) \doteq \hat{U}_d^{(0)}(B) \hat{\Phi}_{m-d}^{(h)}(B)$$

de donde claramente se sigue que

$$\hat{\alpha}^{(h)} = \frac{-\hat{\Phi}_{m+1}^{(h)}}{\hat{\Phi}_m^{(h)}} \doteq \frac{-\hat{\Phi}_{m-d+1}^{(h)}}{\hat{\Phi}_{m-d}^{(h)}} \quad (1.6.31)$$

por lo tanto, por (1.6.30) y (1.6.31), (1.6.29) se transforma en

$$\begin{aligned}\hat{\phi}_m^{(h+1)}(B) &= \hat{U}_d^{(0)}(B) [\hat{\phi}_{m-d+1}^{(h)}(B) + \hat{\alpha}^{(h)} \hat{\phi}_{m-d}^{(h)}(B)B] \\ &= \hat{U}_d^{(0)}(B) \hat{\phi}_{m-d}^{(h+1)}(B)\end{aligned}$$

lo cual demuestra que este resultado también se cumple para $j = h+1$. Además, de (1.6.26) se sigue que el lema también se cumple para $h = 0$, con lo cual queda completa la demostración

Corolario (1.6.2) Se verifica que

$$\hat{\phi}_d^{(j)}(B) = U(B) + O_p(n^{-1})$$

Demostración Se sigue inmediatamente de (1.6.27) pues

$$\hat{\phi}_d^{(j)} = \hat{U}_d^{(0)}(B) + O_p(n^{-1})$$

Las implicaciones prácticas de este corolario son muy importantes, pues se tiene directamente a partir de este resultado que para un proceso ARMA(p,q) no estacionario, en cualquier iteración del ajuste AR(d) se obtiene una estimación consistente del factor no estacionario $U_d(B)$, con tasa de convergencia $O_p(n^{-1})$.

Ahora, por el lema (1.6.7) y la relación (1.6.27), se sigue que $\hat{\phi}_m^{(j)}(B)$ será un estimador consistente de $\phi_m(B)$ si y sólo si $\hat{\phi}_{m-d}^{(j)}(B)$ lo es para $\phi_{m-d}(B)$. Consecuentemente,

del teorema (1.6.1) se sigue el siguiente resultado.

Teorema 1.6.2 Supóngase que $\{Z_t\}$ sigue un proceso ARMA(p,q) ya sea estacionario o no estacionario. Entonces

$$\hat{\phi}_{i(m)}^{(j)} = \phi_i + O_p(n^{-1/2}) \quad \text{para } i = 1, 2, \dots, m$$

si

$$i) \quad m \geq p \quad \text{y} \quad j = q \quad \text{o}$$

$$ii) \quad m = p \quad \text{y} \quad j > q.$$

en donde se entiende que $\phi_i = 0$ para $i > p$. También,

$O_p(n^{-1/2})$ se transforma en $O_p(n^{-1})$ si $\phi_{m-d}(B) = 1$, $m = p$ y

$j \geq q$.

Del resultado anterior es fácil observar que el proceso de regresiones iteradas se puede manejar, tanto con procesos ARMA estacionarios como con los no estacionarios. Por otra parte, (ver Quinn (1980) y Tiao y Tsay (1983)) se sabe que el comportamiento asintótico de la FACM de un proceso no estacionario, está determinado por aquellas raíces no estacionarias de alta multiplicidad, lo que hace que las estimaciones de las ecuaciones generalizadas de Yule-Walker no siempre proporcionen estimadores consistentes de los parámetros autorregresivos. De esta forma (ver el lema (1.6.5)) se observa que la equivalencia asintótica que existe entre los estimadores de los parámetros autorregresivos de un ajuste AR(m) en su j-ésima iteración y las soluciones de las ecuaciones generalizadas de Yule-Walker estimadas, no se cumple para procesos no estacionarios, específicamente en los

siguientes casos:

- i) Cuando $\phi(B)$ y $U(B)$ en (1), no son iguales a la identidad, esto es, algunas raíces de $\phi(B)$ están sobre y otras fuera del círculo unitario.
- ii) Cuando $\phi(B)$ tiene todas sus raíces sobre el círculo unitario, pero no todas son distintas.

Estos resultados se demuestran en el corolario (2.6) de Tiao y Tsay (1983).

1.6.5 Función de autocorrelación muestral extendida (FAME).

La conexión que se ha puesto de manifiesto entre el orden (p, q) de un proceso ARMA y la consistencia de los estimadores de los parámetros autorregresivos, es una herramienta idónea para la identificación del modelo. Esto se logra mediante la FAME, que se define de la siguiente manera.

1.- Si $p = 0$, entonces $\{Z_t\}$ sigue un modelo MA(q) y su FACM tiene un punto de corte asintótico para retrasos mayores que q . Esto es,

$$\hat{\rho}_{j(0)} \doteq 0 \quad \text{si } j > q \text{ y } p = 0 \quad (1.6.32)$$

en donde el subíndice (0) indica que se trata de la FACM usual y j es el retraso.

2.- Si $p = 1$, entonces $\{Z_t\}$ sigue un modelo ARMA(1, q). En

este caso, si se puede encontrar un estimador consistente

$\hat{\phi}_1$ de ϕ_1 , entonces

$$W_t = Z_t - \hat{\phi}_1 Z_{t-1}$$

sigue asintóticamente un modelo MA(q) y su FACM tendrá la propiedad (1.6.32). Sin embargo, para obtener un estimador consistente de ϕ_1 , el número de iteraciones necesarias en el proceso de regresiones iteradas depende del valor de q, el cual en la práctica es desconocido. Lo anterior conduce a tomar las siguientes consideraciones:

i) Si $q = 0$, entonces, por el teorema (1.6.2) se tiene que

$$\hat{\phi}_{1(1)}^{(j)} \xrightarrow{P} \phi_1 \quad \text{para toda } j \geq 0$$

y en particular la serie transformada $\{W_{1,t}^{(0)}\}$, en donde

$$W_{1,t}^{(0)} = Z_t - \hat{\phi}_{1(1)}^{(0)} Z_{t-1}$$

se comporta asintóticamente como un ruido blanco, por lo tanto

$$\hat{\rho}_1(W_{1,t}^{(0)}) \rightarrow 0 \quad \text{para } l > 0$$

en donde $\{\hat{\rho}_1(x_t)\}$ es la FACM de x_t en el retraso l.

ii) si $q = 1$, entonces

$$\hat{\phi}_{1(1)}^{(j)} \xrightarrow{P} \phi_1 \quad \text{para } j \geq 1$$

y la serie transformada $\{W_{1,t}^{(1)}\}$, en donde

$$W_{1,t}^{(1)} = Z_t - \hat{\phi}_{1(1)}^{(1)} Z_{t-1},$$

sigue el modelo MA(1)

$$W_{1,t}^{(1)} = (1 - \theta_1 B) a_t$$

para n suficientemente grande. Esto implica que

$$\hat{\rho}_1(W_{1,t}^{(1)}) \doteq 0 \quad \text{para } l > 1$$

iii) Más generalmente, si $j \geq q$, entonces

$$\hat{\phi}_{1(1)}^{(j)} \xrightarrow{p} \phi_1 \quad \text{para } j \geq q$$

y la serie transformada $\{W_{1,t}^{(q)}\}$, en donde

$$W_{1,t}^{(q)} = Z_t - \hat{\phi}_{1(1)}^{(q)} Z_{t-1},$$

sigue un modelo MA(q), por lo tanto

$$\hat{\rho}_1(W_{1,t}^{(q)}) \doteq 0 \quad \text{para } l > q$$

Basados en las ideas anteriores, se define la primer FAME de $\{Z_t\}$ como

$$\hat{\rho}_{j(1)} = \hat{\rho}_j(W_{1,t}^{(j)}) \quad (1.6.33)$$

en donde el subíndice (1) denota la primer FAME. Por lo tanto, es fácil observar que

$$\begin{aligned}\hat{\rho}_{j(1)} &\doteq 0 && \text{para } j > q \text{ y } p = 1 \\ &\neq 0 && \text{para } j = q \text{ y } p = 1\end{aligned}\quad (1.6.34)$$

3.- En general, para cualquier entero no negativo m , se define la m -ésima FAME de $\{Z_t\}$ como

$$\hat{\rho}_{j(m)} = \hat{\rho}_j(W_{m,t}^{(j)}) \quad (1.6.35)$$

en donde

$$W_{m,t}^{(j)} = Z_t - \sum_{i=1}^m \hat{\phi}_{1(m)}^{(j)} Z_{t-i}$$

Entonces, si el modelo real es un ARMA(p, q) se tiene que $W_{p,t}^{(j)}$ sigue asintóticamente un modelo MA(q) para $j \geq q$, por lo tanto

$$\begin{aligned}\hat{\rho}_{j(m)} &\doteq 0 && \text{para } j > q \text{ y } m = p \\ &\neq 0 && \text{para } j = q \text{ y } m = p\end{aligned}\quad (1.6.36)$$

nótese que para $m = 0$, $\hat{\rho}_{j(0)}$ se reduce a la FACM usual.

A continuación se dará una demostración formal del comportamiento de la FAME dictado por la relación (1.6.36).

Para esto, primero se demostrará que $\hat{\rho}_j(W_{m,t}^{(j)})$ es asintóticamente equivalente a $\rho_j(W_{m,t})$, en donde

$$W_{m,t} = \phi_m(B) Z_t \quad \text{para } m = p$$

Considérese el modelo ARMA(p,q) de (1). Si $U_d(B) \neq 1$, dicho polinomio se puede factorizar como

$$U_d(B) = \prod_{i=1}^m U_i(B) \quad (1.6.37)$$

en donde

$$U_i(B) = 1 - \sum_{j=1}^{d_i} U_{j(i)} B^j$$

tal que

$$i) \quad \sum_{i=1}^n d_i = d$$

$$ii) \quad U_i(B) \text{ divide a } U_{i+1}(B) \text{ para } 1 \leq i \leq m-1$$

$$iii) \quad U_i(B) = 1 - U_{1(i)} B - \dots - U_{d_1(i)} B^{d_1(i)} \text{ para } 1 \leq i \leq m \text{ es un polinomio de raíces simples}$$

Dicha factorización siempre es posible de modo único. (Ver las ecuaciones (3.1) y (3.2) de Tiao y Tsay (1983) para más detalles). Entonces, definiendo

$$Z_{1,t} = Z_t \quad (1.6.38)$$

$$Z_{i,t} = U_{i-1}(B) Z_{i-1,t}, \text{ para } i = 2, \dots, m+1$$

de tal forma que el proceso original (1) puede reescribirse como

$$Z_t = \sum_{i=1}^m \sum_{\nu=1}^{d_i} U_{\nu(i)} Z_{i,t-\nu} + \sum_{h=1}^{p-d} \phi_h Z_{m+1,t-h} + e_t \quad (1.6.39)$$

en donde $e_t = \theta(B)a_t$.

De (1.6.39) se puede demostrar (ver el teorema 3.1 de Tiao y Tsay 1983) que

$$\hat{U}_{\nu(i)} = U_{\nu(i)} + O_p(n^{i-m+1}) \text{ para } \nu = 1, \dots, d_i \text{ y } i = 1, \dots, m \quad (1.6.4)$$

en donde $\hat{U}_{\nu(i)}$ denota el estimador por MCO de $U_{\nu(i)}$ en la regresión

$$Z_t = \sum_{i=1}^m \sum_{\nu=1}^{d_i} U_{\nu(i)} Z_{i,t-\nu} + W_t,$$

la cual, por supuesto, es equivalente al ajuste autorregresivo AR(d). Además, de la transformación (1.6.38), los estimadores de MCO del ajuste autorregresivo AR(d) sobre $\{Z_t\}$ están dados por

$$\hat{U}_d(B) = 1 - \sum_{i=1}^m \left\{ \sum_{\nu=1}^{d_i} \left[\sum_{k=1}^{i-1} U_k(B) \right] \hat{U}_{\nu(i)} B^\nu \right\} \quad (1.6.41)$$

Por otro lado, del lema (2.5) y de las ecuaciones (3.7) y (3.8) de Tiao y Tsay (1983) se sigue que

$$\sum_{i=1}^n Z_{i,t} Z_{j,t} = O_p(n^{2n-1-j+2}) \text{ para } i, j = 1, 2, \dots, m \quad (1.6.42)$$

$$\sum_{i=1}^n Z_{i,t} Z_{m+1,t} = O_p(n^{m-1+1}) \text{ para } i = 1, 2, \dots, m$$

Ahora, considérese la j -ésima iteración del ajuste autorregresivo $AR(p)$ de Z_t para $j \geq q$. Entonces por el lema (1.6.7), la transformación (1.6.38) y la ecuación (4.7) de Tiao y Tsay (1983) se tiene que

$$\hat{\phi}_p^{(j)}(B) = \hat{U}_d^{(0)}(B) - U(B) \sum_{h=1}^{p-d} \hat{\phi}_h^{(j)} B^h \quad (1.6.43)$$

Luego, sea $\rho_1(x_t)$ la FACM de x_t en el retraso 1. Entonces se tiene el siguiente resultado.

Lema 1.6.8 Si $\{Z_t\}$ sigue un modelo $ARMA(p,q)$, entonces

$\hat{\rho}_1(\hat{e}_t^{(j)})$ es asintóticamente equivalente a $\rho_1(e_t)$, en donde

$$\hat{e}_t^{(j)} = \hat{\phi}_p^{(j)}(B)Z_t \quad \text{y} \quad e_t = \theta(B)a_t$$

Demostración De (1.6.43), (1.6.41) y (1.6.38) se tiene que

$$\begin{aligned} \hat{e}_t^{(j)} &= \hat{\phi}_p^{(j)}(B)Z_t \\ &= Z_t - \sum_{i=1}^m \sum_{\nu=1}^d \hat{U}_{\nu(1)} Z_{1,t-\nu} - \sum_{h=1}^{p-d} \hat{\phi}_h Z_{m+1,t-h} \end{aligned}$$

sustituyendo (1.6.39) por Z_t se tiene

$$\hat{e}_t^{(j)} = e_t - \sum_{i=1}^m \sum_{\nu=1}^d (\hat{U}_{\nu(1)} - U_{\nu(1)})Z_{1,t-\nu} - \sum_{h=1}^{p-d} (\hat{\phi}_h^{(j)} - \phi_h)Z_{m+1,t-h}.$$

después, de (1.6.42) y (1.6.40) se tiene que

$$\sum_{m+1, t}^n Z_{m+1, t}^2 = O_p(n) \quad \text{y} \quad \sum_t^n e_t^2 = O_p(n),$$

de donde se sigue que

$$\sum_t^n \hat{e}_t^{(j)} \hat{e}_{t-1}^{(j)} = \sum_t^n e_t e_{t-1} [1 + O_p(n^{-1/2})]$$

Por lo tanto el resultado se cumple ■

Entonces, se sigue de este lema que

$$\begin{aligned} \hat{\rho}_l(\hat{e}_t^{(j)}) &\doteq 0 && \text{para } l > q \\ &\neq 0 && \text{para } l = q \end{aligned}$$

Comparando (1.6.32) con (1.6.36), se puede observar que la p -ésima FAME de un proceso ARMA(p, q), tiene exactamente las mismas propiedades de punto de corte asintótico que la FACM de un proceso MA. Además, como cada elemento individual de la FAME es la autocorrelación muestral de una serie de tiempo transformada, sus propiedades muestrales pueden ser obtenidas a partir de las propiedades asintóticas de la FACM.

1.6.6 El problema de sobreajuste.

Para modelos ARMA(p, q) con p conocida, la propiedad de

punto de corte asintótico de la FAME en (1.6.36) hace de ésta una herramienta útil en la identificación del valor de q . si embargo, estas propiedades dependendn críticamente de las

propiedades de consistencia de los estimadores $\hat{\Phi}_{1(m)}^{(j)}$ en un ajuste AR(m) establecidas en el teorema (1.6.2). En la práctica, los valores de p y q son desconocidos, a continuación se estudiará el comportamiento asintótico de estos estimadores y los efectos que tendrán sobre la FAME, cuando demasiados retrasos e iteraciones son usadas. Este problema es conocido como el problema de sobreajuste.

En lo que sigue de esta sección, se dirá que el polinomio estimado

$$\hat{\Phi}_m^{(j)}(B) \xrightarrow{p} G(B)$$

si todos los coeficientes estimados en $\hat{\Phi}_m^{(j)}(B)$ convergen en probabilidad a los correspondientes coeficientes en $G(B)$, en donde $G(B)$ es algún polinomio en B . De la fórmula recursiva del lema (1.6.3) cambiando j por $q+i$ y $m = p+h$ se obtiene

$$\hat{\Phi}_{p+h}^{(q+1)}(B) = \hat{\Phi}_{p+h+1}^{(q+1-1)}(B) + \hat{\alpha}_h^{(1-1)} \hat{\Phi}_{p+h}^{(q+1-1)}(B)B \quad (1.6.44)$$

en donde

$$\hat{\alpha}_h^{(1-1)} = \frac{-\hat{\Phi}_{p+h+1(p+h+1)}^{(q+1-1)}}{\hat{\Phi}_{p+h(p+h)}^{(q+1-1)}} \quad (1.6.45)$$

A continuación se estudiará el comportamiento asintótico de

$\hat{\Phi}_{p+h}^{(q+1)}(B)$ para $i \geq 1$ y $h \geq 1$, suponiendo que $\{Z_t\}$ sigue un

modelo ARMA(p,q). Para motivar esta discusión considérese el caso $i = 1$. Entonces, por el teorema (1.6.2) se tiene que

$$\hat{\Phi}_{p+h}^{(q)}(B) \xrightarrow{p} \Phi(B) \quad \text{para } h \geq 0 \quad (1.6.46)$$

y

$$\hat{\Phi}_{p+1(p+h)}^{(q)} = O_p(n^{-1/2}) \quad \text{para } i = 1, \dots, h; h \geq 1 \quad (1.6.47)$$

Entonces en (1.6.44) tomando $i = 1$ y para $h \geq 1$, se tiene que

$$\hat{\alpha}_h^{(0)} = \frac{-\hat{\Phi}_{p+h+1(p+h+1)}^{(q)}}{\hat{\Phi}_{p+h(p+h)}^{(q)}} \xrightarrow{p} \alpha_h^{(0)} = \frac{-\Phi_{p+h+1(p+h+1)}^{(q)}}{\Phi_{p+h(p+h)}^{(q)}}$$

y

$$\hat{\Phi}_{p+h}^{(q+1)}(B) \xrightarrow{p} \Phi(B) (1 + \alpha_h^{(0)} B) \quad (1.6.48)$$

en donde $\alpha_h^{(0)}$ es una v.a continua con valores reales y por lo tanto diferente de cero con probabilidad uno. Por otro lado, del teorema (1.6.2) o por el hecho de que

$$\hat{\Phi}_{p(p)}^{(q)} \xrightarrow{p} \Phi_p$$

se tiene que para $h = 0$

$$\hat{\alpha}_h^{(0)} \xrightarrow{p} 0$$

y

$$\hat{\Phi}_p^{(q+1)}(B) \xrightarrow{p} \Phi(B) \quad (1.6.49)$$

La implicación de los resultados (1.6.48) y (1.6.49) sobre la FAME es la siguiente: Tomando

$$W_{p+h,t}^{(q+1)} = \hat{\Phi}_{p+h}^{(q+1)}(B) Z_t$$

se tiene que para n suficientemente grande

$$W_{p+h,t}^{(q+1)} = \begin{cases} \theta(B)a_t & \text{si } h = 0 \\ (1 + \alpha_h^{(0)}B) \theta(B)a_t & \text{si } h > 0 \end{cases} \quad (1.6.50)$$

de donde se observa fácilmente que

$$\hat{\rho}_{q+1(p+h)} = \begin{cases} 0 & \text{si } h = 0 \\ C(h,1) & \text{si } h > 0 \end{cases} \quad (1.6.51)$$

en donde $C(h,1)$ es una constante entre -1 y 1 , diferente de cero con probabilidad uno. Comparando (1.6.51) con (1.6.34) se puede observar que mientras $\hat{\rho}_{q+1(p)}$, $i > 0$, posee la propiedad de punto de corte asintótico, similar a la de la FAME de un modelo $MA(q)$, esta propiedad ya no es cierta para $\hat{\rho}_{q+1(p+h)}$ cuando más de p parámetros autorregresivos han sido incluidos en el ajuste autorregresivo.

Ahora se extiende el resultado (1.6.48) al caso general

$i \geq 1$. Primero se define la función polinomial $H_h^{(1)}(B)$ para $i \geq 0$ y $h \geq 1$ como:

$$H_h^{(i)}(B) = H_{h+1}^{(i-1)}(B) + \alpha_h^{(i-1)} H_h^{(i-1)}(B)B, \quad \text{para } i > 0 \quad (1.6.52)$$

y

$$H_h^{(0)}(B) = 1$$

en donde $\alpha_h^{(i-1)}$ es una v.a. continua. Nótese que $H_h^{(1)}(B)$ es un polinomio de grado i , y que

$$H_h^{(1)}(B) = 1 + \alpha_h^{(0)}$$

además, el último término de $H_h^{(i)}(B)$ es igual a

$$\prod_{\nu=1}^i \alpha^{(1-\nu)}$$

El siguiente lema presenta de manera formal la generalización del resultado (1.6.48).

Lema 1.6.9 Supóngase que $\{Z_t\}$ sigue un modelo ARMA(p,q). Entonces para $i \geq 1$ y $h \geq i$

$$\hat{\Phi}_{p+h}^{(q+1)}(B) \xrightarrow{p} \Phi(B) H_h^{(i)}(B) \quad (1.6.53)$$

Demostración La demostración de este lema es por inducción sobre i . De (1.6.48) se sigue que el resultado es cierto para $i = 1$. Supóngase que (1.6.53) se cumple para $i_0 \geq 1$. Entonces de (1.6.44) se tiene que

$$\hat{\Phi}_{p+h}^{(q+1, 0+1)}(B) = \hat{\Phi}_{p+h+1}^{(q+1, 0)}(B) + \hat{\alpha}_h^{(1, 0)} \hat{\Phi}_{p+h}^{(q+1, 0)}(B)B \quad (1.6.54)$$

Luego, por hipótesis de inducción se sigue que

$$\hat{\Phi}_{p+h+1}^{(q+1)_0}(B) \xrightarrow{p} \Phi(B) H_{h+1}^{(1)_0}(B)$$

(1.6.55)

y

$$\hat{\Phi}_{p+h}^{(q+1)_0}(B) \xrightarrow{p} \Phi(B) H_h^{(1)_0}(B)$$

lo cual implica que para $h > i_0$

$$\hat{\Phi}_{p+h(p+h+1)}^{(q+1)_0} \xrightarrow{p} 0$$

(1.6.56)

$$\hat{\Phi}_{p+h+1(p+h+1)}^{(q+1)_0} \longrightarrow 0$$

pues, $\hat{\Phi}_{p+h}^{(q+1)_0}(B)$ converge a un polinomio de grado $p+i_0$. De (1.6.47) se sigue que la tasa de convergencia en (1.6.56) es la misma en ambos casos. Luego, de (1.6.56) se tiene que $\hat{\alpha}_h^{(1)_0}$ converge en probabilidad a una v.a. $\alpha_h^{(1)_0}$, la cual junto con (1.6.54) y (1.6.55) implican que

$$\begin{aligned} \hat{\Phi}_{p+h}^{(q+1)_0+1}(B) &\xrightarrow{p} \Phi(B) H_{h+1}^{(1)_0}(B) + \alpha_h^{(1)_0} \Phi(B) H_h^{(1)_0}(B) B \\ &= \Phi(B) [H_{h+1}^{(1)_0} + \alpha_h^{(1)_0} H_h^{(1)_0}(B) B] \\ &= \Phi(B) H_h^{(1)_0+1} \quad \text{para } h \geq i_0+1 \end{aligned}$$

El comportamiento asintótico de $\hat{\Phi}_{p+h}^{(q+1)}(B)$ para $h < i$

se resume en el siguiente lema

Lema 1.6.10 Supóngase que $\{Z_t\}$ sigue un modelo ARMA(P,q). Entonces para $h \geq 0$ y $l = 0, 1, \dots$

$$\hat{\Phi}_{p+h}^{(q+h+1)}(B) \xrightarrow{P} \Phi(B)H_{h+1}^{(h)}(B)$$

Demostración La demostración de este lema es por inducción sobre l . Del lema (1.6.9) se sigue que el resultado es cierto para $l = 0$. Ahora, para $l = 1$ se tiene que

$$\hat{\Phi}_{p+h}^{(q+h+1)}(B) = \hat{\Phi}_{p+h+1}^{(q+h)}(B) + \hat{\alpha}_h^{(h)} \hat{\Phi}_{p+h}^{(q+h)}(B)B \quad (1.6.57)$$

en donde

$$\hat{\alpha}_h^{(h)} = \frac{-\hat{\Phi}_{p+h+1}^{(q+h)}(p+h+1)}{\hat{\Phi}_{p+h}^{(q+h)}(p+h)}$$

Así, del lema (1.6.9) se sigue que

$$\hat{\Phi}_{p+h+1}^{(q+h)}(B) \xrightarrow{P} \Phi(B)H_{h+1}^{(h)}(B)$$

y

$$\hat{\Phi}_{p+h}^{(q+h)}(B) \xrightarrow{P} \Phi(B)H_h^{(h)}(B)$$

por lo que

$$\hat{\Phi}_{p+h+1(p+h+1)}^{(q+h)} \xrightarrow{p} 0$$

y

$$\hat{\Phi}_{p+h(p+h)}^{(q+h)} \xrightarrow{p} \Phi_p \prod_{\nu=1}^h \alpha_h^{(h-\nu)}$$

pues, $\hat{\Phi}_{p+h+1}^{(q+h)}$ y $\hat{\Phi}_{p+h}^{(q+h)}$ convergen en probabilidad a un polinomio de grado $p+h$. Nótese que $\Phi_p \prod_{\nu=1}^h \alpha_h^{(h-\nu)}$ es el último término del polinomio $\Phi(B)H_h^{(h)}(B)$. Luego entonces, en (1.6.57) se tiene que

$$\hat{\alpha}_h^{(h)} \xrightarrow{p} 0$$

y

$$\hat{\Phi}_{p+h}^{(q+h+1)}(B) \xrightarrow{p} \Phi(B)H_{h+1}^{(h)}(B)$$

por lo tanto, el lema se cumple para $l = 1$.

Ahora supóngase que el lema se cumple para $l = 1, \dots, l_0$.

Entonces para $l = l_0 + 1$ se tiene que

$$\hat{\Phi}_{p+h}^{(q+h+1, 0+1)}(B) = \hat{\Phi}_{p+h+1}^{(q+h+1, 0)}(B) + \hat{\alpha}_h^{(h+1, 0)} \hat{\Phi}_{p+h}^{(q+h+1, 0)}(B)B \quad (1.6.58)$$

en donde

$$\hat{\alpha}_h^{(h+1, 0)} = \frac{-\hat{\Phi}_{p+h+1(p+h+1)}^{(q+h+1, 0)}}{\hat{\Phi}_{p+h(p+h)}^{(q+h+1, 0)}}$$

luego, por hipótesis de inducción se sigue que

$$\hat{\Phi}_{p+h+1}^{(q+h+1)}(B) \xrightarrow{p} \Phi(B) H_{h+1,0}^{(h+1)}(B)$$

y

(1.6.59)

$$\hat{\Phi}_{p+h}^{(q+h+1)}(B) \xrightarrow{p} \Phi(B) H_{h+1,0}^{(h)}(B)$$

Así

$$\hat{\Phi}_{p+h+1(p+h+1)}^{(q+h+1)}(B) \xrightarrow{p} \Phi_p \prod_{\nu=1}^{h+1} \alpha_{h+1,0}^{(h+1-\nu)}$$

y

$$\hat{\Phi}_{p+h(p+h)}^{(q+h+1)}(B) \xrightarrow{p} \Phi_p \prod_{\nu=1}^h \alpha_{h+1,0}^{(h-\nu)}$$

por lo tanto, de (1.6.58) se tiene el siguiente resultado

$$\hat{\alpha}_h^{(h+1)} \xrightarrow{p} \frac{- \Phi_p \prod_{\nu=1}^{h+1} \alpha_{h+1,0}^{(h+1-\nu)}}{\Phi_p \prod_{\nu=1}^h \alpha_{h+1,0}^{(h-\nu)}} = - \alpha_{h+1,0}^{(h)} \quad (1.6.60)$$

pues,

$$\prod_{\nu=1}^{h+1} \alpha_{h+1,0}^{(h+1-\nu)} = \alpha_{h+1,0}^{(h)} \alpha_{h+1,0}^{(h-1)} \dots \alpha_{h+1,0}^{(0)}$$

y

$$\prod_{\nu=1}^h \alpha_{h+1,0}^{(h-\nu)} = \alpha_{h+1,0}^{(h-1)} \dots \alpha_{h+1,0}^{(0)}$$

entonces de (1.6.59) y (1.6.60) se sigue que

$$\begin{aligned}
\hat{\Phi}_{p+h}^{(q+h+1, 0+1)} &\xrightarrow{P} \Phi(B)H_{h+1,0}^{(h+1)}(B) - \hat{\alpha}_{h+1,0}^{(h)} \Phi(B)H_{h+1,0}^{(h)}(B)B \\
&= \Phi(B) \left[H_{h+1,0}^{(h+1)}(B) - \hat{\alpha}_{h+1,0}^{(h)} H_{h+1,0}^{(h)}(B)B \right] \\
&= \Phi(B)H_{h+1,0+1}^{(h)}(B)
\end{aligned}$$

en donde la última igualdad se sigue de la siguiente relación

$$H_{h+1,0}^{(h+1)}(B) = H_{h+1,0+1}^{(h)}(B) + \hat{\alpha}_{h+1,0}^{(h)} H_{h+1,0}^{(h)}(B)B$$

por lo tanto, el resultado se cumple para $l = l_0 + 1$ ■

Los resultados del teorema (1.6.2) y los lemas (1.6.9) y (1.6.10) se resumen en el siguiente teorema.

Teorema 1.6.3 Supóngase que $\{Z_t\}$ sigue un modelo ARMA(p,q). Entonces el polinomio estimado en la j-ésima iteración $\hat{\Phi}_m^{(j)}(B)$ converge en probabilidad a los siguientes polinomios

$$\hat{\Phi}_m^{(j)}(B) \xrightarrow{P} \Phi(B) \begin{cases} H_{m-p}^{(j-p)}(B)a_t & \text{si } j-q \leq m-p \\ H_{(j-q)}^{(m-p)}(B)a_t & \text{si } j-q > m-p \end{cases}$$

Demostración Es inmediata de los lemas (1.6.9) y (1.6.10).

Corolario 1.6.3 Sea

$$W_{m,t}^{(j)} = \hat{\phi}_m^{(j)}(B) Z_t$$

Entonces para n suficientemente grande

$$W_{m,t}^{(j)} = \theta(B) \begin{cases} H_{m-p}^{(j-q)}(B) a_t & \text{si } 0 \leq j-q \leq m-p \\ H_{(j-q)}^{(m-p)}(B) a_t & \text{si } j-q > m-p \geq 0 \end{cases}$$

en donde $H_h^{(i)}$ es un polinomio de grado i .

A partir del corolario (1.6.3) es claro que los efectos de sobreajustar $m-p > 0$ retrasos autorregresivos en el polinomio estimado $\hat{\phi}_m^{(j)}(B)$, en general tienden a incrementar el orden del polinomio de promedios móviles de la serie transformada $\{W_{m,t}^{(j)}\}$. En donde el número de términos adicionales está dado por el $\min \{m-p, j-q\}$. Así, de la definición de la FAME y el corolario (1.6.3), se sigue que para un modelo ARMA(p,q) la FAME tiene el siguiente comportamiento asintótico

$$\hat{\rho}_{j(m)} = \begin{cases} C(m-p, j-q) & \text{si } 0 \leq j-q \leq m-p \\ 0 & \text{si } j-q > m-p \geq 0 \end{cases} \quad (1.6.61)$$

pues, si $j-q > m-p$, entonces $j > m-p+q$, luego $\min \{j-q, m-p\} = m-p$, por lo tanto $\{W_{m,t}^{(j)}\}$ sigue un proceso MA(q-p+m), por lo tanto

$$\hat{\rho}_{j(m)} = 0 \quad \text{si } j-q > m-p \geq 0$$

de manera análoga se demuestra la otra igualdad. Obsérvese que $C(m-p, j-q)$ es una constante diferente de cero o una v.a continua acotada entre -1 y 1.

1.6.8 La identificación de modelos ARMA mediante la FAME.

Las propiedades asintóticas de la FAME mostradas en la relación (1.6.61), pueden ser utilizadas en la práctica para identificar modelos ARMA(p,q). Para este propósito resulta útil arreglar las autocorrelaciones muestrales como se muestra en la siguiente tabla

TABLA 1.6.1
LA FUNCION DE AUTOCORRELACION MUESTRAL EXTENDIDA

MA AR	0	1	2	3	...
0	$\hat{\rho}_{1(0)}$	$\hat{\rho}_{2(0)}$	$\hat{\rho}_{3(0)}$	$\hat{\rho}_{4(0)}$	...
1	$\hat{\rho}_{1(1)}$	$\hat{\rho}_{2(1)}$	$\hat{\rho}_{3(1)}$	$\hat{\rho}_{4(1)}$	...
2	$\hat{\rho}_{1(2)}$	$\hat{\rho}_{2(2)}$	$\hat{\rho}_{3(2)}$	$\hat{\rho}_{4(2)}$	...
3	$\hat{\rho}_{1(3)}$	$\hat{\rho}_{2(3)}$	$\hat{\rho}_{3(3)}$	$\hat{\rho}_{4(3)}$	...
...			...		

En el primer renglón de esta tabla se localiza la FAC, mientras que el segundo renglón muestra la primera FAME, el tercero la segunda FAME y así sucesivamente. Obsérvese que los renglones están numerados del 0,1,2, ... lo cual indica el orden AR, de manera similar se numeran las columnas para indicar el orden de promedios móviles MA.

Para ilustrar el uso de la tabla (1.6.1), supóngase que el modelo real es un ARMA(p,q). De las propiedades de la FAC es bien sabido que $\hat{\rho}_{j(0)} \neq 0$ para $j \geq q+1$. Ahora, en (1.6.61) se puede observar que

$$i) \text{ cuando } m = p, \quad \hat{\rho}_{j(p)} = 0 \quad \text{para } j \geq q+1$$

$$ii) \text{ cuando } m = p+1, \quad \hat{\rho}_{j(p+1)} = 0 \quad \text{para } j > q+2$$

y así sucesivamente.

Para identificar el orden (p,q) de un modelo ARMA mediante la FAME se busca en la tabla (1.6.1) el "punto de corte triangular", es decir el punto (c_1, c_2) tal que las rectas $m = c_1$ y $m-j = c_2$ delimitan un triángulo de ceros, tal como el de la tabla (1.6.2), en donde x denota un valor diferente de cero y 0 es un valor igual a cero. El renglón y columna coordenadas del vértice de este triángulo, corresponden precisamente con el orden AR y MA respectivamente

TABLA 1.6.2

COMPORTAMIENTO ASINTOTICO DE LA FASE EN UN MODELO ARMA(p,q)

AR	...	q-1	q	q+1	q+2	q+3	...
...	:	:	:	:	:	:	:
p-1	x	x	x	x	x	x	...
p	x	0	0	0	0	0	...
p+1	x	x	0	0	0	0	...
p+2	x	x	x	0	0	0	...
p+3	x	x	x	x	0	0	...
...	:	:	:	:	:	:	:

En la práctica, para muestras finitas, los elementos $\hat{\rho}_{j(m)}$ no serán iguales a cero. Por lo tanto, para detectar el orden de un modelo ARMA se utiliza una tabla "simbólica" similar a la tabla (1.6.1) en cuyo lugar (m, j) se coloca un cero si $\hat{\rho}_{m, j+1}$ es asintóticamente nulo (al nivel de significación del 5%), es decir, si $|\hat{\rho}_{m, j+1}| \leq 2\sigma_{m, j+1}$, siendo $\sigma_{m, j+1}^2$ la varianza asintótica de $\hat{\rho}_{m, j+1}$. En caso contrario, se coloca, por ejemplo, una x. A estos efectos, se precisa una estimación de la varianza asintótica de los coeficientes $\hat{\rho}_{m, j+1}$. Tiao y Tsay proponen como estimación a $(N-m-j)^{-1}$, aproximación que resulta de la fórmula de Bartlett, bajo la hipótesis de que $\{W_{m,t}^{(j)}\}$ es ruido blanco, aunque reconocen que posiblemente sea una estimación sesgada a la baja. Este estimador de $\sigma_{a,t}^2$

queda pendiente de posteriores investigaciones.

El problema que presenta este método es la falta de unicidad, pues la teoría de la FAME sólo determina el comportamiento asintótico para $j \geq q$ y $m \geq p$, pero no impide la existencia de puntos de corte para $m < p$.

1.7 FUNCIÓN DE AUTOCORRELACIÓN INVERSA Y FUNCIÓN DE AUTOCORRELACIÓN PARCIAL INVERSA.

Es conveniente recordar que esta sección se presenta de manera informativa por la razón de que el método que aquí se muestra sólo proporciona ventajas en la determinación del orden de un modelo puramente autorregresivo.

Cleveland (1972) y Chatfield (1979) proporcionan argumentos en favor del uso de la Función de Autocorrelación Inversa (FACI) en lugar de la FAC propuesta en el método de Box-Jenkins. El m -ésimo retraso de la autocovarianza inversa, denotado por $\gamma_1(m)$ se define como el coeficiente de B^m en la función generatriz

$$\Gamma_1(B) = \sum_{k=-\infty}^{\infty} \gamma_1(k) B^k,$$

la cual satisface que $\Gamma_1(B)\Gamma(B) = 1$, en donde $\Gamma(B)$ es la

función generatriz de autocovarianzas. La función $\Gamma_1(B)$ se conoce usualmente como la función generatriz de autocovarianzas inversas y

$$\Gamma_1(B)/\gamma_1(0)$$

como la función generatriz de autocorrelaciones inversas. Puede demostrarse que para un proceso ARMA(p,q) como en (1')

$$\Gamma_1(B) = \frac{\phi_p(B)\phi_p(B^{-1})}{\theta_q(B)\theta_q(B^{-1})} \frac{1}{\sigma_a^2} \quad (1.7.1)$$

Esto implica que el m-ésimo retraso de la autocorrelación inversa de un proceso AR(p), es igual a

$$\rho_1(k) = \begin{cases} \frac{\left[-\phi_m + \sum_{j=1}^{p-m} \phi_j \phi_{j+m} \right]}{\left[1 + \sum_{j=1}^p \phi_j^2 \right]} & \text{si } m \leq p \\ 0 & \text{si } m > p \end{cases} \quad (1.7.2)$$

Entonces, en el caso autorregresivo $\rho_1(m)$ tiene un punto de corte para retrasos mayores que p. Esta propiedad de la FACI hace que ésta sea una opción alternativa de la FACP. Sin embargo, la estimación de $\rho_1(m)$ no resulta tan fácil como la de la FACP. Un método para estimar la FACI de una serie de tiempo, es aproximar dicha serie mediante un proceso AR de orden suficientemente grande para estimar sus parámetros usando las ecuaciones de Yule-Walker (1.2.3) y después, mediante la

ecuación (1.7.2), estimar la autocorrelación inversa. Bhansali (1980) deriva las propiedades distribucionales asintóticas de la autocorrelación inversa muestral. En particular se ha demostrado que para un proceso de ruido blanco, la FACI muestral se distribuye asintóticamente como una normal con media cero y varianza $1/n$. De aquí que la significación de la FACI puede ser comparada con los valores $1.96/n^{1/2}$ a un nivel del 5%.

También se ha presentado (ver Chatfield (1979)) como una herramienta de especificación a una función conocida como la Función de Autocorrelación Parcial Inversa (FACPI). Dicha función puede ser definida como la FACP del modelo inverso

$$\theta_q(B)Z_t = \phi_p(B)a_t$$

La FACPI puede ser calculada a partir de las ecuaciones de Yule-Walker (1.2.4) al reemplazar la autocorrelación por la autocorrelación inversa. El comportamiento de la FACPI es similar al de la FAC, sin embargo, su uso es limitado debido a los problemas que presenta su estimación.

Para un estudio más completo de este método se recomiendan los artículos de Bhansali (1980, 1983a).

En Bhansali (1983a) los estimadores de los parámetros autorregresivos de la FACI son utilizados para estimar el orden de un proceso MA mediante un criterio similar al Criterio del error final de predicción propuesto por Bhansali y Dowhum (1977) (ver capítulo 3). También, en este trabajo se deriva la distribución asintótica de dichos estimadores, Abraham y Ledolter (1984), en este artículo se investiga la utilidad de la FACI como herramienta para la especificación del orden de un modelo de series de tiempo. También, se presentan algunos

ejemplos de simulación para procesos autorregresivos, los cuales indican que la FACI es menos eficiente que la FACP. Mientras que Hipel, McLeod y Lemox (1977) dan ejemplos del uso de la FACI en la identificación de modelos de series de tiempo.

CAPÍTULO 2

MÉTODOS BASADOS EN LA TEORÍA DE PRUEBAS DE HIPÓTESIS ESTADÍSTICAS

2.1 MODELAJE ARMA(n,n-1) DE PANDIT-WU

Esta metodología se basa en el hecho de que, si se observara en tiempo discreto un proceso AR(p) que originalmente es un proceso en tiempo continuo, el proceso observado sería un ARMA(p,p-1).

Pandit (1973) y Pandit-Wu (1984) a través de un estudio de dinámica de sistemas físicos, sugieren el siguiente procedimiento para modelar series de tiempo:.

i) Como toda serie estacionaria puede ser representada por un proceso AR(∞), se pueden obtener estimaciones por mínimos cuadrados condicionados o por máxima verosimilitud para los modelos ARMA(2n,2n-1).

ii) Para cada estimación de un nuevo modelo se calcula el cociente

$$F = \frac{A_1 - A_0}{4} / \frac{A_0}{N-4n+4} \sim F(4, N-4n+4) \quad (2.1.1)$$

en donde N es el número de observaciones, A_1 es la suma de cuadrados de los residuales del modelo ARMA(2n,2n-1) y A_0 es la suma de cuadrados de los residuales del modelo ARMA(2n+2,2n+1).

Si el valor de F obtenido mediante este cociente, excede el valor de la distribución $F(4, N-4n-4)$ leído en tablas para un nivel de significación α ; entonces se obtiene una disminución significativa en la suma de cuadrados residuales al pasar del modelo $\text{ARMA}(2n, 2n-1)$ al $\text{ARMA}(2n+2, 2n+1)$, por lo que debe de tomarse como modelo más adecuado al $\text{ARMA}(2n+2, 2n+1)$. En caso contrario, esto es, si el valor de F calculada es menor que el valor de F leído en tablas, el modelo más adecuado será el $\text{ARMA}(2n, 2n-1)$.

iii) Una vez que se ha tomado la decisión de parar en un modelo $\text{ARMA}(2n, 2n-1)$, el criterio F puede ser usado para probar otros modelos no probados antes de esta etapa, como por ejemplo los modelos $\text{ARMA}(2n-1, 2n-2)$ y $\text{ARMA}(2n-1, m)$ con $m < 2n-1$. Para llevar a cabo estas pruebas puede hacerse uso de los intervalos de confianza de los parámetros como guía. Si los parámetros ϕ_{2n} y θ_{2n-1} son pequeños y sus intervalos de confianza contienen al cero, se ajusta el modelo de orden impar $\text{ARMA}(2n-1, 2n-2)$ y se vuelve a aplicar el criterio F a este modelo para determinar si es el más adecuado; de la misma forma se puede verificar si un modelo con un número menor de parámetros de promedios móviles es el más adecuado.

Nótese que el criterio F está justificado sólo cuando uno de los dos modelos el $\text{ARMA}(2n, 2n-1)$ ó el $\text{ARMA}(2n+2, 2n+1)$ es el adecuado. Por ejemplo, supóngase que el modelo $\text{ARMA}(6, 5)$ es el más adecuado para describir el comportamiento de una serie de tiempo, entonces el uso del criterio F está estrictamente justificado para probar los modelos $\text{ARMA}(4, 3)$ contra el $\text{ARMA}(6, 5)$ y el $\text{ARMA}(6, 5)$ contra el $\text{ARMA}(8, 7)$, pero no está justificado para probar el modelo $\text{ARMA}(2, 1)$ contra el $\text{ARMA}(4, 3)$ ya que ambos modelos son inadecuados. En un caso como

este último, la reducción de la suma de cuadrados residuales es frecuentemente tan grande que se dudaría del criterio F, por lo que se recomienda no parar el procedimiento. Como una medida de precaución, la decisión de parar el procedimiento en un determinado modelo debe tomarse después de que las autocorrelaciones de los residuales sean lo suficientemente pequeñas, por ejemplo, que se encuentren dentro de una banda de $\pm 2/(N)^{1/2}$.

Obsérvese que el criterio F empleado para determinar el orden de un modelo ARMA(2n, 2n-1), es análogo al caso de un problema de regresión lineal con r parámetros, en el que se desea probar si Δ de estos son estadísticamente iguales a cero a través del criterio

$$F = \frac{A_1 - A_0}{\Delta} / \frac{A_0}{N-r} \sim F(\Delta, N-r)$$

en donde $F(\Delta, N-r)$ denota la distribución F con Δ y $N-r$ grados de libertad y A_0 , A_1 se definen igual que en (2.1.1). Por lo que en (2.1.1) se tiene que:

$$r = (2n+2) + (2n+1) + 1 = 4n + 4 \quad y$$

$$\Delta = (4n+4) - [4(n-1) + 4] = 4.$$

2.2 CRITERIO DE INFORMACIÓN DE AKAIKE (AIC).

El bien conocido concepto de información de Fisher se ha convertido a través del principio de máxima verosimilitud y de

estadísticas auxiliares, en la base de algunos métodos para determinar el orden de un modelo de series de tiempo. Esto mismo no se puede decir de otros conceptos de información que tienen una fuerte correspondencia con la interpretación probabilística de la entropía termodinámica. El uso del concepto de entropía es relativamente nuevo en el área del análisis de series de tiempo; antes de entrar en detalle en dicho concepto se tratará la relación estadística que guardan n observaciones independientes de una variable aleatoria Y.

Supóngase que Y es una variable aleatoria absolutamente continua caracterizada por la función de densidad de probabilidad $f(Y|\bar{\beta})$, la cual está determinada por el vector de parámetros $\bar{\beta}=(\beta_1, \beta_2, \dots, \beta_k)'$; $\bar{\beta} \in \mathbb{R}^k$. Supóngase además que existe un vector de valores teóricos $\bar{\beta}^*$ de $\bar{\beta}$ y denótese la densidad teórica $f(Y)$ por $f(Y|\bar{\beta}^*)$. La bondad de ajuste de $f(Y|\bar{\beta}^*)$ con respecto a $f(Y|\bar{\beta})$ puede ser medida a través de la entropía definida por:

$$B(\bar{\beta}^*, \bar{\beta}) = \int_{\mathbb{R}} f(Y|\bar{\beta}^*) \log f(Y|\bar{\beta}) d_Y - \int_{\mathbb{R}} f(Y|\bar{\beta}^*) \log f(Y|\bar{\beta}^*) d_Y \quad (2.2.1)$$

(ver Akaike, 1978b, para una justificación formal), en donde el primer término de (2.2.1) determina la bondad de ajuste de $f(Y|\bar{\beta})$ con respecto a $f(Y|\bar{\beta}^*)$, mientras que el segundo término es una constante para $f(Y|\bar{\beta}^*)$ dada.

En vez de maximizar el criterio de entropía (2.2.1), el siguiente criterio de información puede ser minimizado;

$$I(\bar{\beta}^*, \bar{\beta}) = -B(\bar{\beta}^*, \bar{\beta}) = \int_{\mathbb{R}} \{\log f(Y|\bar{\beta}^*) - \log f(Y|\bar{\beta})\} f(Y|\bar{\beta}^*) d_Y \quad (2.2.2)$$

Esta última expresión es conocida como la medida de información de Kullback-Leibler (ver Kullback, 1959), dicha medida es mayor que cero a menos que $f(Y|\bar{\beta}^*) = f(Y|\bar{\beta})$.

Si se supone que $\bar{\beta}^*$ es suficientemente cercano a $\bar{\beta}$ se puede derivar una regla de decisión a partir de (2.2.2), para esto sea $\bar{\beta} = \bar{\beta}^* + \Delta\bar{\beta}$, en donde $\Delta(\bar{\beta}) = (\Delta\beta_1, \Delta\beta_2, \dots, \Delta\beta_k)'$ es un vector de orden $k \times 1$ con norma "pequeña", entonces $I(\bar{\beta}^*, \bar{\beta})$ tiene la siguiente expansión en serie de Taylor

$$I(\bar{\beta}^*, \bar{\beta}^* + \Delta\bar{\beta}) \approx \int_{\mathbb{R}} \left\{ -\sum_i (\Delta\bar{\beta})_i \frac{\partial \log f(Y|\bar{\beta}^*)}{\partial \bar{\beta}_i^*} - \frac{1}{2} \sum_i \sum_j (\Delta\bar{\beta})_i (\Delta\bar{\beta})_j \frac{\partial^2 \log f(Y|\bar{\beta}^*)}{\partial \bar{\beta}_i^* \partial \bar{\beta}_j^*} \right\} f(Y|\bar{\beta}^*) d_Y$$

si $f(Y|\bar{\beta}^*)$ es una función regular, el primer término del lado derecho de esta relación se anula, por lo que $I(\bar{\beta}^*, \bar{\beta}^* + \Delta\bar{\beta}) = 1/2 \|\bar{\beta}\|_I^2$, en donde $\|\bar{\beta}\|_I^2 = \Delta\bar{\beta}' I(\bar{\beta}^*) \Delta\bar{\beta}$, I es la matriz de información de Fisher, y $\|\cdot\|_I^2$ es la norma Euclídeana.

Ahora, restrínjase $\bar{\beta}$ a un subespacio Θ_s de dimensión S con $S = 1, 2, \dots, K-1$, mientras que el vector de parámetros teóricos $\bar{\beta}^*$ está en un subespacio de dimensión K , con $K > S$.

Denótese por $\bar{\beta}_s^*$ la proyección de $\bar{\beta}^*$ sobre Θ_s en el sentido de la norma $||\cdot||_1^2$, se puede demostrar que

$$2l(\bar{\beta}^*, \bar{\beta}_s) \approx ||\bar{\beta}_s^* - \bar{\beta}^*||_1^2 + ||\bar{\beta}_s - \bar{\beta}^*||_1^2 \quad (2.2.3)$$

en donde $\bar{\beta}_s \in \Theta_s$ y es cercano a $\bar{\beta}_s^*$.

Reemplazando $\bar{\beta}_s$ en (2.2.3) por el vector de variables aleatorias $\hat{\beta}_s$, que es la restricción del estimador de máxima verosimilitud de $\bar{\beta}^*$ en Θ_s y por el hecho de que $n||\bar{\beta}_s^* - \hat{\beta}_s||_1^2$ se distribuye asintóticamente como una χ^2 con s grados de libertad, se puede obtener que para n grande (ver Huber, 1967)

$$2E[l(\bar{\beta}^*, \hat{\beta}_s)] \approx ||\bar{\beta}_s^* - \bar{\beta}^*||_1^2 + S/n \quad (2.2.4)$$

La ecuación (2.2.4) produce una extensión en la medida tal que $\hat{\beta}_s$ se desvía del vector de parámetros teóricos $\bar{\beta}^*$; esta ecuación implica que la desviación esperada de $\hat{\beta}_s$ con respecto a $\bar{\beta}^*$ tiene dos componentes, una de estas representa el error cometido al seleccionar un espacio de parámetros aproximados para $\bar{\beta}_s^*$ y la otra es el error debido a la estimación del vector de parámetros especificado.

En la práctica, para muestras finitas es imposible minimizar directamente la ecuación (2.2.4), ya que el primer término de ésta es desconocido y necesita ser estimado. Akaike (1973) demuestra que bajo ciertas condiciones de regularidad, el cociente de verosimilitud estadística:

$$LR(Y) = -2 \sum_{i=1}^n \log [f(y_i | \hat{\beta}_s) / f(y_i | \hat{\beta})],$$

en donde $\hat{\beta}$ es un estimador de máxima verosimilitud de $\bar{\beta}$, se distribuye como una variable χ^2 no centrada con K-S grados de libertad y parámetro de no centralidad $||\bar{\beta}_s^* - \bar{\beta}^*||_1^2$, tal que

$$n^{-1}\{LR(Y+2S-K)\} \quad (2.2.5)$$

es un estimador insesgado de (2.2.4). En la práctica (2.2.5) puede ser simplificada eliminando términos comunes de todos los modelos ya que sólo se está interesado en la elección de un mínimo. Por lo tanto, sólo se necesita minimizar la relación

$$-2\text{Ln}(Y; \hat{\beta}_s) + 2s \quad (2.2.6)$$

sobre $s = 1, 2, \dots, k-1$. Este procedimiento es conocido como el Criterio de Información de Akaike (AIC), en donde el primer término de (2.2.6) proporciona la penalización de un mal ajuste, mientras que el segundo término de esta relación proporciona la penalización por seleccionar un orden alto del modelo y la no confiabilidad de éste.

Akaike (1973, 1974) fué el primero en reconocer la importancia de la entropía para determinar el orden de un modelo ARMA(p,q). Cuando una serie de tiempo es generada por un proceso ARMA(p,q) normalmente distribuido, el valor máximo aproximado del log de la función de verosimilitud Ln es igual a:

$$\hat{L}_n = -1/2 n \log \hat{\sigma}_a^2 - 1/2 n \quad (2.2.7)$$

ver Box y Jenkins (1976, cap 5). Ignorando el segundo término de la ecuación (2.2.7), que es independiente de los parámetros y del orden (p,q) del proceso, se obtiene la siguiente expresión equivalente para (2.2.6):

$$AIC(p,q) = n \log \hat{\sigma}_a^2 - 1/2 n \quad (2.2.8)$$

El procedimiento de Akaike sugiere la elección del orden (p,q) como aquél para el cual $AIC(p,q)$ alcanza su mínimo, con $p,q = 0,1,2, \dots, L$, en donde L es el límite superior previamente asignado, el cual no necesariamente tiene que ser igual a p o q .

Una razón por la cual algunos autores no consideran el criterio AIC para determinar el orden de un modelo para series de tiempo, es por el hecho de que la estimación mediante dicho criterio no es consistente y porque asintóticamente sobreestima el orden teórico del modelo. Shibata (1976) demuestra este resultado para los modelos $AR(p)$, dicho resultado también se cumple para el criterio del EFP (ver capítulo 3).

Por supuesto, la inconsistencia del criterio AIC para los modelos AR implica inconsistencia en los modelos ARMA. Sin embargo, las demostraciones de estos dos casos son muy diferentes. Si (p^*, q^*) es el orden teórico del modelo $ARMA(p,q)$ y $p^* < P$, $q^* < Q$, en donde P y Q son los valores máximos considerados entonces bajo condiciones muy generales,

$$\hat{L}_n = -1/2 n \log \hat{\sigma}_a^2 - 1/2 n \quad (2.2.7)$$

ver Box y Jenkins (1976, cap 5). Ignorando el segundo término de la ecuación (2.2.7), que es independiente de los parámetros y del orden (p,q) del proceso, se obtiene la siguiente expresión equivalente para (2.2.6):

$$AIC(p,q) = n \log \hat{\sigma}_a^2 - 1/2 n \quad (2.2.8)$$

El procedimiento de Akaike sugiere la elección del orden (p,q) como aquél para el cual $AIC(p,q)$ alcanza su mínimo, con $p,q = 0,1,2, \dots, L$, en donde L es el límite superior previamente asignado, el cual no necesariamente tiene que ser igual a p o q .

Una razón por la cual algunos autores no consideran el criterio AIC para determinar el orden de un modelo para series de tiempo, es por el hecho de que la estimación mediante dicho criterio no es consistente y porque asintóticamente sobreestima el orden teórico del modelo. Shibata (1976) demuestra este resultado para los modelos $AR(p)$, dicho resultado también se cumple para el criterio del EFP (ver capítulo 3).

Por supuesto, la inconsistencia del criterio AIC para los modelos AR implica inconsistencia en los modelos ARMA. Sin embargo, las demostraciones de estos dos casos son muy diferentes. Si (p^*, q^*) es el orden teórico del modelo $ARMA(p,q)$ y $p^* < P$, $q^* < Q$, en donde P y Q son los valores máximos considerados entonces bajo condiciones muy generales,

Hannan (1982b, teorema 2) demuestra que el orden obtenido mediante el criterio AIC seguramente es sobreestimado. Hannan (1982b, pag 462) también demuestra este resultado para el caso de modelos multivariados ARMA(p,q).

Akaike (1978b), Jones(1974) y Tong(1977), presentan algunas aplicaciones del criterio AIC para determinar el orden de un modelo AR, mientras que Kitagawa (1977) y Ozaki (1977) presentan aplicaciones del criterio AIC para determinar el orden de un modelo ARMA. Neftci (1982) aplica el criterio AIC para especificar el orden del modelo de algunas series de tiempo económicas, por último Kitagawa(1981), Kitagawa y Akaike (1978) y Kozin y Nakajima (1980) presentan una discusión sobre el uso del criterio AIC para modelar series de tiempo no estacionarias.

2.3 MÉTODOS BAYESIANOS.

Entre los métodos para determinar el orden de un modelo ARMA(p,q) que se han presentado hasta el momento, no figuran los métodos Bayesianos. Dichos métodos a diferencia de los anteriores, hacen uso del conocimiento a priori acerca de los parámetros del modelo, este conocimiento se presentan en forma de una función de densidad de probabilidad. Primerose formularán muy brevemente los métodos Bayesianos junto con sus alcances generales y después se presentarán algunos de estos métodos para determinar el orden de un modelo ARMA(p,q).

Considérese una muestra aleatoria de tamaño n de una variable aleatoria continua $Y = (y_1, y_2, \dots, y_n)'$ y supóngase

que se tienen dos posibles modelos que proporcionan una representación estacionaria para este conjunto de datos. Más aún, supóngase que estos modelos están caracterizados por las hipótesis H_1 y H_2 , que a su vez están representadas por la función de densidad de probabilidad $f(Y|\bar{\beta}_i)$, en donde $\bar{\beta}_i$, $i = 1, 2$ es un vector de parámetros independientes y desconocidos, con $\bar{\beta}_i \in R^K$. En el contexto Bayesiano la elección entre H_1 y H_2 puede hacerse usando la razón a posteriori.

$$K_{12} = \frac{P(H_1|Y)}{P(H_2|Y)} = \frac{P(H_1)P(Y|H_1)}{P(H_2)P(Y|H_2)} \quad (2.3.1)$$

en donde $P(H_i)$ denota la probabilidad de que H_i sea verdadera; ver por ejemplo, Zellner (1971, cap 10).

Nótese que K_{12} es igual al cociente a priori multiplicado por el cociente del promedio de las verosimilitudes, en donde los promedios se obtienen usando las probabilidades a priori de los parámetros. Evidentemente, éste procedimiento para contrastar hipótesis difiere del procedimiento conocido como cociente de verosimilitudes, en donde para contrastar entre las hipótesis H_1 y H_2 se usan estimadores de máxima verosimilitud, como si estos fueran los valores teóricos de los parámetros desconocidos. La razón a posteriori es una medida de la tendencia a creer en las hipótesis H_1 y H_2 dada una evidencia muestral. Para elegir un modelo con la probabilidad a posteriori más alta, se minimiza la pérdida esperada asociada

con el rechazo de una de las dos hipótesis. Esto es, si $K_{12} > 1$ se elige H_1 y si $K_{12} < 1$ se elige H_2 . Cuando $K_{12} = 1$, no se puede concluir nada en términos de la pérdida esperada. Obsérvese que estos resultados son óptimos sólo en el caso en que los modelos tengan funciones de pérdida simétricas.

De lo expuesto anteriormente es claro que dada una función de pérdida simétrica, el modelo que debe ser elegido es aquel que tenga la probabilidad a posteriori $P(H_1|Y)$ más alta, de tal forma que $P(H_1)P(H_1|Y)$ sea máxima. Desafortunadamente, en la práctica resulta difícil especificar completamente la probabilidad a priori de que dado un modelo éste sea el verdadero. En algunos casos $P(H_1)$ debe ser reemplazada por alguna probabilidad de que un modelo sea el más adecuado. En el último de los casos y sin entrar en problemas de interpretación, el concepto de probabilidad a priori, vaga ó difusa, puede ser empleado para evitar el problema de la determinación de la probabilidad a priori. La dificultad con estos conceptos puede ser evitada siguiendo la sugerencia de Jeffreys (1961, pag 117) de tomar probabilidades a priori $P(H_1)$ iguales para cada modelo. Por lo tanto, el modelo que se debe elegir es aquél que tenga la probabilidad $P(Y|H_1)$ más grande.

Mediante un teorema de Jeffreys (1961, pag 139), citado por Zellner (1971, pag 31-33), Chow (1981) demuestra que con una información mínima a priori de los datos, el $\log P(Y|H_1)$ evaluado en $\hat{\beta}$ está dado por

$$\log P(Y|H_1) = \ln(Y; \hat{\beta}_1) - 1/2 k_1 \log n - 1/2 \log \det I(\hat{\beta}_1) + \quad (2.3.1)$$

$$1/2 K_1 \log(2\pi) + \log f_{H_1}(\hat{\beta}) + O_p(n^{-1/2})$$

en donde Ln denota la función de verosimilitud, $f_{H_1}(\hat{\beta})$ es la densidad a priori cuando el modelo H_1 es verdadero, $\hat{\beta}$ es el estimador de máxima verosimilitud de β_1 y $I(\hat{\beta}_1)$ es la matriz de información evaluada en $\beta_1 = \hat{\beta}_1$.

2.3.1 EL CRITERIO S

Desafortunadamente la ecuación (2.3.1) presenta problemas muy difíciles, por ejemplo la elección de la densidad $f_{H_1}(\beta)$ para cada modelo H_1 . Schwarz (1978) ha superado este problema reteniendo solamente los dos primeros términos del lado derecho de (2.3.1). Para muestras grandes Schwarz (1978) propone como elección del modelo aquél para el cual:

$$S = \text{Ln}(Y, \hat{\beta}_1) - 1/2 k_1 \log n \quad (2.3.2)$$

es un máximo.

Un análisis crítico del criterio S es presentado por Akaike (1981), en dicho análisis Akaike argumenta que el uso de (2.3.2) sólo es posible cuando se cuenta con una función de densidad a priori $f_{H_1}(\beta)$ claramente definida. Ésta en realidad sería una situación en donde la aplicación de los resultados de la inferencia Bayesiana pueden ser usados para resolver el problema de seleccionar el orden de un modelo ARMA(p,q). Dependiendo de la naturaleza de las probabilidades a priori, Chow (1981) demuestra que el criterio de Schwarz no necesariamente es una buena aproximación del logaritmo de la

probabilidad a posteriori para muestras de tamaño finito. En un estudio en donde se comparan los criterios AIC y S, Stone (1979) también cuestiona la validez de este último procedimiento para valores finitos de n . Stone considera que el trabajo de Schwarz debe entenderse como un caso especial de un primer esfuerzo sobre la discriminación de modelos presentado por Jeffreys (1967).

Para seleccionar el modelo más apropiado en un conjunto de modelos ARMA(p, q) con $p, q = 0, 1, 2, \dots, L$, la ecuación (2.3.2) es asintóticamente equivalente a minimizar la función

$$S(p, q) = n \log \hat{\sigma}_a^2 + (p+q) \log n \quad (2.3.3)$$

sobre todos los valores de p y q .

Varias formas similares al criterio S han aparecido en la literatura, por ejemplo: Atkinson (1980, 1981) sugiere una partición Y_1 de n_1 observaciones de un total de n observaciones $Y = (Y_1, Y_2)$ para obtener la densidad a priori $f_{H_1}(\bar{\beta}|Y_1)$ para $i = 1, 2$. Para la muestra inicial Y_1 , denótese por $\hat{S}_1(Y_1)$ la suma de cuadrados residuales del i -ésimo modelo y por $\hat{S}_1(Y)$ la del complemento de la muestra. Ahora el modelo especificado por $f_{H_1}(\bar{\beta}|Y_1)$ y por $f(Y|\bar{\beta}_1)$ satisface la siguiente probabilidad a posteriori

$$\log P(H_1|Y) \propto 1/2 \sigma_a^{-2} \{\hat{S}_1(Y_1) - \hat{S}_1(Y)\} - 1/2 k_1 \log(n/n_1) + C \quad (2.3.4)$$

en donde se supone que todos los modelos son igualmente

probables después de la muestra inicial de las observaciones, C es una constante y α denota la proporcionalidad. Si σ_a^2 es desconocida se puede usar un estimador de máxima verosimilitud de σ_a^2 para el i -ésimo modelo.

Atkinson sugiere que el modelo a elegir es aquél para el cual (2.3.4) alcanza un mínimo. Para muestras suficientemente grandes los efectos de la muestra inicial de n_1 observaciones son insignificantes por lo que (2.3.4) se reduce a

$$-2 \log P(H_1|Y) \propto \hat{S}_1(Y)/\sigma_a^2 + k_1 \log n \quad (2.3.5)$$

Nótese que el criterio (2.3.4) puede ser visto como un caso especial del criterio de Schwarz (2.3.3).

2.3.2 EL CRITERIO BIC.

Minimizando la entropía negativa de la distribución especificada por la verosimilitud con respecto a la densidad a priori de $\bar{\beta}$, Akaike (1977, 1978a, 1979) determina un mínimo para el valor de la función de pérdida esperada

$$E_{H_1} [\log f(Y|\bar{\beta})]$$

Este valor mínimo alcanzable es conocido como el Criterio de Información Bayesiana (BIC) y toma la forma:

$$BIC = (n - k_1) \log \{ \hat{S}_1(Y) / (n - k_1) \} + 1/2 \log \{ (\hat{\sigma}_Y^2 - \hat{S}_1(Y)/k_1) \} \quad (2.3.6)$$

en donde $\hat{\sigma}_y^2$ es un estimador de la varianza muestral. Akaike (1977) pone de manifiesto que el uso del método de máxima entropía evita un mal ajuste de la probabilidad a priori de los datos bajo observación.

Se puede demostrar que para los modelos ARMA(p,q) la relación (2.3.6) se transforma en:

$$\begin{aligned} \text{BIC}(p,q) = & n \log \hat{\sigma}_a^2 - (n-p-q) \log(1-(p+q)/n) + (p+q) \log n \\ & + (p+q) \log\{(p+q)^{-1} (\hat{\sigma}_y^2 / \hat{\sigma}_a^2 - 1)\} \end{aligned} \quad (2.3.7)$$

en donde al igual que antes $\hat{\sigma}_a^2$ representa el estimador de máxima verosimilitud de σ_a^2 . El conjunto de valores $\{p,q\}$ con $p, q = 0, 1, 2, \dots, L$ que produce el mínimo valor para el BIC es elegido como el orden del modelo ARMA.

Cuando $n \gg (p+q)$, se puede usar la aproximación $-(n-p-q) \log \{1-(p+q)/n\} \cong (p+q)$ para encontrar una relación entre los criterios BIC y AIC de la siguiente forma:

$$\begin{aligned} \text{BIC} & \cong \text{AIC}(p,q) + (p+q)(\log n - 1) + (p+q) \log\{(p+q)^{-1} (\hat{\sigma}_y^2 / \hat{\sigma}_a^2 - 1)\} \\ & \cong \text{AIC}(p,q) + (p+q)(\log F_{p+q,n} - 1) \end{aligned} \quad (2.3.8)$$

en donde $F_{p+q,n} = n(\hat{\sigma}_y^2 - \hat{\sigma}_a^2) / \{(p+q)\hat{\sigma}_a^2\}$. De (2.3.8) se puede observar que la penalización del criterio BIC en la sobreparametrización, es mayor que la del criterio AIC sólo si $F_{p+q,n} > e$. Por otra parte, si el modelo teórico es un ruido

blanco o si tanto el tamaño de la muestra y los parámetros AR y MA son pequeños, entonces $F_{p+q,n}$ es pequeño, en este caso la penalización del criterio AIC en la sobreparametrización es mayor que la del criterio BIC.

Rissanen (1978) sugiere obtener el orden de un modelo ARMA(p,q) minimizando la función

$$\text{BIC}^*(p,q) = \log \hat{\sigma}_a^2 + (p+q)(\log n)/n ; p,q = 0,1, \dots, L \quad (2.3.9)$$

Nótese que (2.3.9) es esencialmente igual al criterio de Schwarz. La principal diferencia de la aproximación entre los criterios BIC^* y S es que la derivación de (2.3.9) no necesita la hipótesis Bayesiana de que los parámetros son obtenidos a partir de experimentos aleatorios con distribuciones similares.

Rissanen (1980) demuestra que el criterio BIC^* proporciona estimadores consistentes del orden de un modelo AR(p). Este resultado fue extendido por Hannan (1980) y Hannan y Rissanen (1982) a modelos ARMA.

CAPÍTULO 3

MÉTODOS BASADOS EN EL ERROR DE PREDICCIÓN UN PASO HACIA ADELANTE.

3.1 CRITERIO DEL ERROR FINAL DE PREDICCIÓN.

Akaike (1960, 1970a) propuso un método que permite elegir el orden p de una autorregresión de modo que el cuadrado del valor esperado del error final de predicción sea mínimo, considerando los errores debidos a las inexactitudes cuando un orden pequeño es seleccionado y los errores debidos al incremento de la varianza residual (causados por un decrecimiento del número de grados de libertad) cuando se selecciona un orden grande.

Considérese una serie de tiempo estacionaria generada por un proceso $AR(p)$, en donde p es finito y acotado por algún entero L . Si se supone que los parámetros de este proceso son conocidos, el pronóstico con error cuadrático medio mínimo de la observación Z_{n+1} , a partir del tiempo $t = n$ (un paso hacia adelante) es

$$\hat{Z}_n(1) = \phi_1 Z_n + \dots + \phi_p Z_{n-p+1} \quad (3.1.1)$$

así que el error de predicción un paso hacia adelante $Z_{n+1} - \hat{Z}_n(1)$, tiene error cuadrático medio igual a

$$E [(Z_{n+1} - \hat{Z}_n(1))^2] = \sigma_a^2 ;$$

ver Box-Jenkins (1976, sección 5.1.

Cuando los parámetros ϕ_i 's son remplazados por sus estimaciones de MCO y cuando n es grande, el error cuadrático medio de la estimación del error de predicción un paso hacia adelante es aproximadamente igual a

$$E [(Z_{n+1} - \hat{Z}_n(1))^2] = \sigma_a^2 (1 + p/n) \quad (3.1.2)$$

ver Yamamoto (1976). Un estimador insesgado de σ_a^2 está dado por

$$n \hat{\sigma}_a^2 / (n-p)$$

De tal forma que sustituyendo dicho estimador en (3.1.2) Akaike (1970a) le llama estadística del error final de predicción (EFP) y se define como

$$EFP(p) = \hat{\sigma}_a^2 (n+p)/(n-p) \quad (3.1.3)$$

El orden para un proceso AR se elige entonces calculando el EFP(p) para $p = 0, 1, \dots, L$ y tomando como el valor de p aquel en el cual el EFP alcanza su valor mínimo. Nótese que en este método no existen elementos subjetivos en la definición del EFP(p), excepto para la determinación del límite superior L y la elección del criterio de predicción un paso hacia adelante. En la práctica, cuando el proceso es puramente AR, usualmente se fija el valor de L suficientemente grande, de tal forma que se asegure que éste excede el orden real del proceso.

Generalmente, si el orden del ajuste se incrementa, σ_a^2 tiende a decrecer a un nivel cercano al valor teórico σ_a^2 , mientras que el factor extra del lado derecho de (3.1.3)

penaliza un incremento de p elevando el valor del EFP. Al buscar el mínimo del EFP, se selecciona un valor de p que balancee el riesgo entre ambos términos de manera óptima. Debe notarse sin embargo, que cuando $n \gg p$, el término $(n+p)/n-p$ es apenas afectado por un cambio en p , lo cual hace que en general resulte difícil elegir el orden correcto del modelo con un alto grado de consistencia (Gresch y Sherpe; 1972).

Las propiedades del EFP para seleccionar el orden de un modelo AR fueron estudiadas mediante simulaciones por Bhansali (1973) quien concluyó que el EFP tiende a sobreestimar el orden teórico de los procesos generados por un proceso AR de aproximadamente 50 retrasos. Jones (1975) revisa los resultados de Bhansali y encuentra que éste utilizó una definición incorrecta del EFP, Akaike (1974) también hace algunos comentarios acerca del trabajo de Bhansali.

Akaike (1971) extiende el método del EFP para determinar el orden de un proceso AR multivariado. Söderström (1977) demuestra que (3.1.3) puede ser usada independientemente de la estructura del modelo. Así, reemplazando p en (3.1.3) por el número de parámetros estimados, dicha relación puede ser aplicada a todos los modelos de series de tiempo, incluso a los modelos mixtos ARMA. Una generalización del método del EFP es obtenida por Bhansali y Downham (1977), quienes sugieren que la elección del orden de un modelo de series de tiempo puede basarse en el mínimo de la función

$$EFP_{\delta}(p) = \hat{\sigma}_a^2 (1 + \delta p/n) \quad \text{para } p = 0, 1, \dots, L \quad (3.1.4)$$

en donde δ es una constante positiva.

Se observa que cuando $\delta > 2$, se da un incremento en la penalización de la sobreparametrización. En las simulaciones de Bhansali y Downham (1977) se consideran valores de δ entre 1 y 4. Sus resultados muestran que, si n es muy grande ($n \geq 300$) el modelo teórico es ajustado más frecuentemente para $\delta > 2$ que para $\delta = 2$, pero, si n es pequeño ($n = 50$) este puede no necesariamente ser el caso. Sin embargo, resultados de simulaciones realizadas por Akaike (1979), indican que varios casos de sobreestimación del orden teórico pueden esperarse para modelos de orden grande, cuando $\delta > 2$. Este resultado es confirmado por Atkinson (1980), quien afirma que sobre las bases del error cuadrático medio de predicción, la elección de $\delta = 2$ debe de ser adoptada en (3.1.4).

Para mayor información acerca del método del EFP, se tienen los siguientes artículos Akaike (1970b), Bhansali (1978) y McClave (1975, 1978). Un criterio similar al método del EFP $_{\delta}$, ha sido desarrollado por Bhansali (1983b) para determinar el orden de un modelo MA. Por otro lado aplicaciones del método del EFP pueden consultarse en Caines, Keny y Sethi (1981).

3.2 CRITERIO DE VALIDACIÓN CRUZADA.

De la presentación del método anterior, aparentemente la habilidad de poder predecir un paso hacia adelante valores de $\{Z_t\}$, es medida con los mismos datos usados para estimar los parámetros del modelo. Para eliminar algunas de estas situaciones irreales, varios autores han propuesto y defendido

algunos métodos basados en el criterio de validación cruzada. Aunque estos métodos originalmente fueron obtenidos dentro del contexto de modelos de regresión lineal, pueden ser utilizados para seleccionar el orden de un modelo puramente autorregresivo AR.

El criterio de validación cruzada se construye de la siguiente manera: Dada la especificación particular de un modelo, dentro de la clase de modelos $AR(p)$ con $p = 0, 1, \dots, L$, el modelo es estimado n veces, cada vez se elimina una observación de la muestra y se pronostica (en el sentido del error cuadrático medio mínimo) esta observación. Nótese que el criterio de validación cruzada está basado en la suma de cuadrados de los errores de predicción que se muestra a continuación

$$\sum_{t=1}^n w_{t(1)} \{Z_t - \hat{Z}_{t(1)}\}^2, \quad (3.2.1)$$

en donde $w_{t(1)}$ es alguna función de peso fija y $\hat{Z}_{t(1)}$ es la predicción de la t -ésima observación de la autorregresión formada por todas las observaciones, excluyendo la t -ésima observación. El orden apropiado del modelo se encuentra al minimizar la relación (3.2.1) para $p = 0, 1, \dots, L$. Obsérvese que, en contraste con el criterio del EFP, este criterio elimina a la observación en la predicción de ella misma.

Al usar diferentes funciones de peso $w_{t(1)}$, se producen diferentes métodos para determinar el orden de un modelo. Allen (1971) propone un método llamado la Predicción de Suma de Cuadrados (PSC), en el cual se le da el mismo peso a cada observación que se elimina. El criterio de la PSC fue aplicado

por Bessler y Binkley (1980) para determinar el orden de algunas series de tiempo económicas de los Estados Unidos. Estos autores también comparan los modelos AR obtenidos mediante este criterio con los obtenidos con el criterio del EFP. Dichas comparaciones sugieren que el criterio del EFP tiende a sobreestimar el orden de los modelos autorregresivos encontrados mediante la aplicación del criterio de la PSC. Hjorth y Holmqvist (1981) aplican el criterio de la PSC, para construir modelos AR multivariados en datos meteorológicos. Otras funciones de peso se pueden consultar en Schmidt (1974), Stone (1975) y Hjorth (1982).

3.3 CRITERIO PARA UNA FUNCIÓN DE AUTOCORRELACIÓN TRANSFERIDA.

Otro método para determinar el orden de un modelo AR, fue propuesto por Parzen (1974, 1977), dicho método está basado en la teoría de aproximación espectral. Parzen (1974) supone que la serie de tiempo observada es generada por un proceso AR de orden infinito, descrito por

$$\phi_{\infty}(B)Z_t = a_t \quad \text{con } \text{var}(a_t) = \sigma_a^2,$$

y que este proceso puede ser aproximado por un proceso AR de orden finito, $AR(p)$ con varianza igual a $\sigma_a^2(p)$. Basándose en la función de densidad espectral de la serie de tiempo observada, Parzen (1974) propone el siguiente criterio de

distancia entre dos procesos AR

$$J_p = (2\pi)^{-1} \int_{-\pi}^{\pi} |(\hat{\phi}_p(e^{iw}) - \phi_{\omega}(e^{iw}))/\phi_{\omega}(e^{iw})|^2 dw$$

Además Parzen demuestra que para un valor fijo de p , el mínimo de J_p es igual a

$$1 - \sigma_{\omega}^2 / \sigma_a^2(p)$$

y que este valor se alcanza cuando los parámetros autorregresivos son estimados por las ecuaciones de Yule-Walker. Por otra parte Kramer (1969) demuestra que el valor esperado de J_p es igual a

$$E \left[(2\pi)^{-1} \int_{-\pi}^{\pi} |(\hat{\phi}_p(e^{iw}) - \phi_{\omega}(e^{iw}))/\phi_{\omega}(e^{iw})|^2 dw \right] = 1 - \frac{\sigma_{\omega}^2}{\sigma_a^2(p)} + \frac{p}{n} \quad (3.3.1)$$

en donde el término

$$1 - \frac{\sigma_a^2}{\sigma_a^2(p)}$$

representa el sesgo debido a la aproximación de $\phi_{\omega}(B)$ por $\hat{\phi}_p(B)$, mientras que p/n representa la varianza total de la estimación de los parámetros ϕ_1 's en

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

Originalmente Parzen propuso adoptar el valor de p para

el cual (3.3.1) se minimiza, como el orden más apropiado del proceso. Posteriormente reconoció que dicho criterio es muy difícil de usar en la práctica, pues se requiere de un estimador de σ_{∞}^2 y en Parzen (1975) modifica su criterio a una nueva versión que le llama "Criterio para una Función de Autocorrelación Transferida" (CFAT). Dicho criterio está dado por la siguiente función

$$CFAT(p) = \begin{cases} \left[n^{-1} \sum_{j=1}^p \hat{\sigma}_a^{-2}(j) \right] - \hat{\sigma}_a^{-2}(p) & \text{si } p = 1, 2, \dots, L \\ -(1+n)^{-1} & \text{si } p = 0 \end{cases} \quad (3.3.2)$$

en donde $\hat{\sigma}_a^{-2}(j)$ es un estimador de la varianza residual ajustado por los grados de libertad, cuando un modelo $AR(j)$ es ajustado a los datos.

Usualmente, la relación (3.3.2) se minimiza con respecto a $p = 1, 2, \dots, L$ para obtener el orden apropiado de un modelo AR. Sin embargo, Touy (1979) demuestra que la ausencia de $\hat{\sigma}_a^{-2}(0)$ en (3.3.2) puede conducir a una sobreestimación del orden teórico del proceso y sugiere la siguiente modificación del CFAT.

$$CFAT(p) = \left[n^{-1} \sum_{j=0}^p \hat{\sigma}_a^{-2}(j) \right] - \hat{\sigma}_a^{-2}(p) \quad \text{para } p=0, 1, \dots, L \quad (3.3.3)$$

Los criterios (3.3.2) y (3.3.3) son discutidos y aplicados por Parzen (1975, 1978, 1979, 1980) y Parzen y Pagano (1979), mientras que Beamish y Priestley (1981) comparan los

criterios del EFP y el CFAT mediante simulaciones, en las cuales determinan el orden de varios procesos AR. Una extensión del CFAT para procesos AR múltiples es presentada por Parzen (1977).

CAPÍTULO 4

EJEMPLOS ILUSTRATIVOS DE LA IDENTIFICACIÓN

En la práctica, el manejo de los métodos y conceptos expuestos en las secciones 1.2, 1.3, 1.4, 1.5 y 1.6, se ilustran con la identificación de modelos para las series i) $\{Z_t\}$, en donde $\{Z_t\}$ es una serie simulada con 200 observaciones, ii) la serie "C", cuyos datos se muestran en Box y Jenkins (1970) y iii) la serie $\{INPC_t\}$, en donde $\{INPC_t\}$ denota a la serie del Índice Nacional de Precios al Consumidor en México.

Es conveniente mencionar que los arreglos R, S, S^* y el arreglo para la FACPG, así como el estadístico D se calcularon mediante el programa que se muestra en el Apéndice 2. Mientras que los arreglos de la FAME, así como la estimación de los parámetros del polinomio autorregresivo y la tabla del método de la esquina, se calcularon mediante el paquete "THE SCA STATISTICAL SYSTEM".

Es relevante manifestar que para probar la veracidad del programa mostrado en el Apéndice 2, se reprodujeron los ejemplos 2 y 10 de Gray, Kelley y McIntire (1978), obteniéndose los mismos resultados.

Ejemplo 1 .- Sea $\{Z_t\}$ una serie simulada con 200 observaciones generadas según el proceso

$$Z_t - 1.32 Z_{t-1} + 0.68 Z_{t-2} = a_t - 0.8 a_{t-1} \quad ; \quad a_t \sim N(0,1)$$

cuya gráfica se muestra en la figura 4.1.1. Observando dicha gráfica se puede concluir que la serie $\{Z_t\}$ es estacionaria en

lo que a nivel se refiere. Esto mismo queda corroborado por los valores de las desviaciones estándar muestrales de la serie

$$S(0) = 1.2936 \quad S(1) = 1.3511$$

ya que $S(0)$ proporciona el mínimo de estos valores (ver Guerrero 1983, cap 4). Dichos valores indican que la serie $\{Z_t\}$ es estacionaria y que no se requiere de ninguna diferenciación de esta. Por lo tanto para identificar el orden del modelo habrá que estudiar entonces a las FAC y FACP muestrales de $\{Z_t\}$, las cuales se presentan en el cuadro 4.1.1

CUADRO 4.1.1
FAC Y FACP MUESTRALES PARA $\{Z_t\}$

FAC MUESTRAL											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.48	.02	-.28	-.43	-.36	-.11	.10	.17	.13	.02	-.07
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	-.13	-.10	.02	.10	.16	.10	.01	-.12	-.14	-.07	.02
FACP MUESTRAL											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.48	-.27	-.22	-.24	-.13	.03	-.03	-.06	-.07	-.07	-.03
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	-.09	-.04	.04	-.00	.06	-.03	.01	-.07	.02	.03	.01

Al inspeccionar este cuadro se observa que no hay un "corte" notorio en las autocorrelaciones para un cierto retraso q , así pues, no es del todo razonable proponer como un

lo que a nivel se refiere. Esto mismo queda corroborado por los valores de las desviaciones estándar muestrales de la serie

$$S(0) = 1.2936 \quad S(1) = 1.3511$$

ya que $S(0)$ proporciona el mínimo de estos valores (ver Guerrero 1983, cap 4). Dichos valores indican que la serie $\{Z_t\}$ es estacionaria y que no se requiere de ninguna diferenciación de esta. Por lo tanto para identificar el orden del modelo habrá que estudiar entonces a las FAC y FACP muestrales de $\{Z_t\}$, las cuales se presentan en el cuadro 4.1.1

CUADRO 4.1.1
FAC Y FACP MUESTRALES PARA $\{Z_t\}$

FAC MUESTRAL											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.48	.02	-.28	-.43	-.36	-.11	.10	.17	.13	.02	-.07
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	-.13	-.10	.02	.10	.16	.10	.01	-.12	-.14	-.07	.02
FACP MUESTRAL											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.48	-.27	-.22	-.24	-.13	.03	-.03	-.06	-.07	-.07	-.03
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	-.09	-.04	.04	-.00	.06	-.03	.01	-.07	.02	.03	.01

Al inspeccionar este cuadro se observa que no hay un "corte" notorio en las autocorrelaciones para un cierto retraso q , así pues, no es del todo razonable proponer como un

modelo factible a un $MA(q)$. De hecho, se aprecia un comportamiento irregular en las dos primeras autocorrelaciones y después una convergencia a cero. Por otra parte, al inspeccionar la FACP, se observa que las dos primeras autocorrelaciones parciales son sensiblemente grandes en comparación con las otras autocorrelaciones parciales, sería entonces razonable postular a un modelo $AR(2)$ para representar a la serie, pero como se puede apreciar en la figura 4.1.2 las autocorrelaciones parciales en los retrasos 3 y 4 no se pueden considerar estadísticamente iguales a cero, por lo que se rechaza la posibilidad del modelo $AR(2)$. Sin embargo, podría pensarse que la convergencia a cero de la FAC está dictada por la ecuación

$$(1 - \phi B - \phi B^2) \hat{\rho}_m = 0 \quad \text{para } m > 1$$

además obsérvese la convergencia a cero de la FACP. Por lo anterior, es quizá razonable pensar en un modelo mixto $ARMA(2,1)$ como posible modelo generador de la serie $\{Z_t\}$, esto es:

$$ARMA(2,1): (1 - \phi_1 B - \phi_2 B^2) Z_t = (1 - \theta_1 B) a_t$$

aunque no es del todo claro que ésta sea la única posibilidad a considerar y quizá un modelo con términos AR y MA hasta de orden 4 pudieran considerarse necesarios, pero en la etapa de verificación del modelo se podrían cancelar los parámetros innecesarios (véase Guerrero, 1983, cap. 4).

FIGURA 4.11
GRAFICA DE LA SERIE SIMULADA $\langle Z_t \rangle$

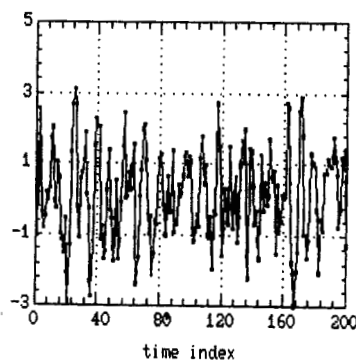
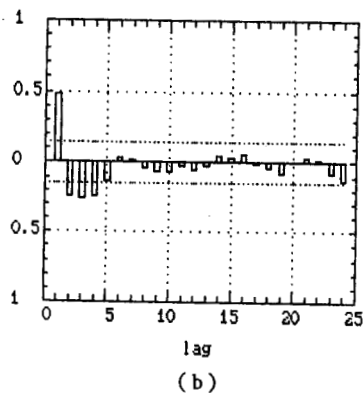
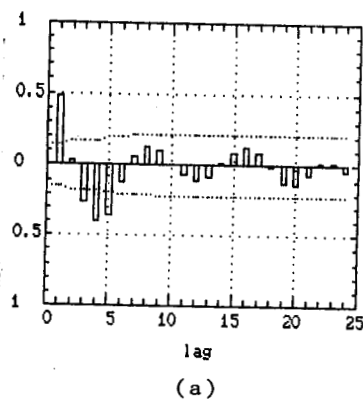


FIGURA 4.12
FAC MUESTRAL DE $\langle Z_t \rangle$ (a) Y FAC MUESTRAL DE $\langle Z_t \rangle$ (b)



ii) Arreglos R y S

Las tablas 4.1.1 y 4.1.2 muestran los arreglos S y R para los datos de la serie $\{Z_t\}$ usando $f_{\square} = (-1)^{\square} \hat{\rho}_{\square}$ y $f_{\square} = \hat{\rho}_{\square}$ respectivamente. La manera de identificar el orden de un modelo a partir de los patrones mostrados por los arreglos de las tablas (4.1.1) y (4.1.2), es la siguiente: Primero se inspeccionan los arreglos S con $f_{\square} = (-1)^{\square} \hat{\rho}_{\square}$ y $f_{\square} = \hat{\rho}_{\square}$ y se selecciona el caso en el cual se manifieste más claramente el comportamiento dictado por los teoremas (1.3.2) y (1.3.3) en una de sus columnas seguida de otra con alta variabilidad. De esta forma se identifica el valor de p. Este comportamiento se manifiesta claramente en la segunda columna de la tabla (4.1.1), identificándose con esto un valor para $p=2$. En esta misma tabla se puede observar que $q-p+1 = 0$, de donde se sigue que $q=1$. Nótese que los valores de p y q pueden ser identificados en varios lugares de este patrón, por ejemplo, el hecho de que $p=2$ y $q=1$ resulta claro a partir de observar los valores subrayados de la columna $S_2(m)$, los cuales son diferentes de los comportamientos constantes que se aprecian en esta columna, por lo que $2q=2$, entonces $q=1$. También obsérvese que esta elección de p y q queda corroborada por el hecho de que $|S_3(-4)| = 62.835$, que en magnitud es el valor más grande que aparece en esta tabla y $S_3(-1) = -3.3804$, lo cual concuerda con el comportamiento mostrado en la tabla (1.3.3) de la sección 1.3, esto es, $|S_3(-4)| = \infty$ mientras que $S_3(-1) = -C_1 = -S_2(0)$. Nótese además que la elección de $p=2$ se confirma al observar la tercera columna de los arreglos R mostrados en la tabla (4.1.1), dicha columna manifiesta un comportamiento

cercano a cero salvo los valores que van de $R_3(-3)$ a $R_3(-1)$. Por consiguiente al aplicar conjuntamente los resultados de los teoremas (1.3.2) a (1.3.5) y el estadístico D que se muestra en el cuadro 4.1.2, a los resultados de la tabla (4.1.1), ninguna otra elección de p y q parece ser posible, por lo que el modelo identificado para representar a la serie $\{Z_t\}$ es un modelo:

$$\text{ARMA}(2,1): (1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta B)a_t.$$

Finalmente, del teorema (1.3.5) se tiene que

$$\phi_p = (-1)^{p+1} \frac{C_3}{C_4}$$

Si C_3 es estimada mediante $S_p((-1)^{q-p+1} \hat{\rho}_{q-p+1})$ y C_4 por $S_p((-1)^{-q-p} \hat{\rho}_{-q-p})$, entonces en este ejemplo como $p=2$ y $q=1$ se tiene que una estimación de ϕ_2 está dada por

$$\hat{\phi} = - \frac{3.01}{4.7082} = -0.639$$

TABLA 4.1.1

$$(f_n = (-1)^n \hat{\rho}_n)$$

VALORES $R(f_n)$				VALORES $S(f_n)$			
1	2	3	m	1	2	3	
0.1711	0.1043	0.0144	-8	-1.5732	4.5342	10.5342	
-0.0981	4.8648	0.0479	-7	0.0993	4.5476	-2.7487	
-0.1078	-0.1842	0.0088	-6	-4.3817	3.3644	-20.4749	
0.3646	0.1016	-0.0404	-5	-2.1682	3.6547	-2.9804	
-0.4259	-0.1233	-0.0156	-4	-1.6509	4.8507	-62.8357	
0.2772	0.5295	<u>0.1921</u>	-3	-0.9164	4.7082	6.5459	
0.0232	0.4155	<u>0.3748</u>	-2	-21.8576	<u>6.8856</u>	-1.4669	
-0.4833	-0.5167	<u>0.0840</u>	-1	-3.0690	<u>1.8906</u>	-3.3804	
1.0000	0.1419	-0.0103	0	-1.4833	3.0100	<u>5.3947</u>	
-0.4833	-0.2656	-0.0282	1	-1.0479	3.1271	-4.4361	
0.0232	-0.3414	0.0087	2	10.9638	3.3852	11.0554	
0.2772	0.1142	0.0353	3	-2.5363	3.6443	-1.8813	
-0.4259	-0.1101	0.0085	4	-1.8560	2.4776	-2.7881	
0.3646	0.1003	-0.0003	5	-1.2957	2.6857	75.6529	
-0.1078	-2.8815	-0.0094	6	-0.0903	2.6860	-1.3873	
-0.0981	-0.0614	-0.0052	7	-2.7445	3.0982	5.7121	
0.1711	0.0468	-0.0160	8	-1.7589	2.7534	0.2616	

TABLA 4.1.2

$$(f_n = \hat{\rho}_n)$$

VALORES $R(f_n)$				VALORES $S(f_n)$			
0.1711	0.3844	0.0773	-8	-0.4268	0.8424	-3.9458	
0.0981	0.2301	-0.1937	-7	-2.0993	1.1252	-0.6343	
-0.1078	0.3389	0.0612	-6	2.3817	0.4820	1.9053	
-0.3646	1.3101	0.2924	-5	0.1682	0.5045	-0.5098	
-0.4259	-0.5833	-0.3000	-4	-0.3491	0.2516	2.5950	
-0.2772	-0.4478	-2.1529	-3	-1.0836	0.4200	-3.5973	
0.0232	-0.4573	1.0760	-2	19.8576	2.3098	-0.8061	
0.4833	-1.4833	-0.2411	-1	1.0690	0.6585	-0.1396	
1.0000	0.4073	-0.1158	0	-0.5167	0.2686	-0.9227	
0.4833	0.2924	0.5430	1	-0.9521	0.1622	-0.4128	
0.0232	0.2887	0.0633	2	-12.9638	0.4673	-2.5514	
-0.2772	0.5403	-0.2462	3	0.5363	0.5221	-0.6980	
-0.4259	-1.4191	0.0343	4	-0.1440	0.6127	-0.4736	
-0.3646	-0.1845	0.0017	5	-0.7043	0.4968	-15.0462	
-0.1078	-0.1363	-0.0514	6	-1.9097	0.4926	-0.9182	
0.0981	-0.2265	0.0267	7	0.7445	0.6044	-2.3031	
0.1711	0.3414	-0.0676	8	-0.2411	0.6517	-1.0131	

CUADRO 4.1.2

ESTADISTICO D PARA LA SERIE $\{Z_t\}$ CON $n = 3$ Y $m = 3$

(n, m)	
(1,0)	0.012
(1,1)	0.002
(1,2)	1.043
(1,3)	1.044
(2,0)	0.815
(2,1)	112.445
(2,2)	0.044
(3,0)	0.004
(3,1)	0.002
(3,2)	0.049
(3,3)	0.007

iii) Función de Autocorrelación Parcial y de Autocorrelación Parcial Generalizada.

La tabla 4.1.3 muestra el arreglo S^* para los datos de la serie $\{Z_t\}$. En una primera inspección de esta tabla se aprecia que los comportamientos de $\pm \infty$ y de $(-1)^i C_1$ del teorema (1.4.1) ocurren, para el primer caso en las columnas 3, 4, 5, 6, 7, y 8 y en el retraso $-q-1 = -2$, dicho comportamiento se manifiesta con los números -62.835, 20.183, 5.914, -0.5343, 45.391 y 147.864, mientras que en el retraso $q = 1$ y en las columnas 3, 4, 5, 6, 7 y 8 las cantidades -3.380, 2.552, -1.116, 0.0499, -4.186 y 3.2192 corresponden al comportamiento de $(-1)^i C_1$, lo cual concuerda con el comportamiento mostrado en la tabla (1.4.3), esto es, $S_{2+i}^*(-2) = \pm \infty$ para $i = 1, 2, \dots$ mientras que $S_{2+i}^*(1) = (-1)^i C_1$. Estas observaciones junto con el comportamiento dictado por los teoremas (1.3.4) y (1.3.5) que se aprecia en la columna $S_2^*(m)$, identifican como posible modelo generador de la serie $\{Z_t\}$ a un modelo estacionario:

$$\text{ARMA}(2,1): (1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta B)a_t$$

con

$$\hat{\phi}_2 = -\frac{3.01}{4.7082} = -0.639$$

(por el mismo argumento que se da al estimar ϕ_2 en (i)).

La tabla 4.1.4 muestra el arreglo de la FACPG para los datos de la serie $\{Z_t\}$ basados en los resultados de la tabla (4.1.3). Inspeccionando dicha tabla se puede observar que $|\phi_{2+1,2+1}^{(1)}| \leq 0.188$ para $i = 1, 2, \dots$ y $\hat{\phi}_{2,2}^{(1+1)} = -0.639$ para $i = 0, 1, 2, \dots$. Por lo tanto, de acuerdo al patrón mostrado por la tabla 4.1.4, el posible modelo generador de la serie $\{Z_t\}$ identificado mediante este método es un:

$$\text{ARMA}(2,1): (1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta B)a_t$$

$$\text{con } \hat{\phi}_2 = -0.639$$

TABLA 4.1.3

ARREGLO S^* PARA LA SERIE $\{Z_t\}$ CON $f_{\square} = (-1)^{\square} \hat{\rho}_{\square}$

m.n.	1	2	3	4	5	6	7	8
-8	-1.573	4.558	-4.378	6.528	-1.959	10.494	-9.934	0.0
-7	0.099	4.534	-123.777	-5.857	-14.220	13.893	-2.937	0.02
-6	-4.381	4.547	10.635	30.514	-14.819	167.882	-0.332	-19.136
-5	-2.168	3.364	-2.748	5.326	-2.300	-11.915	16.967	-16.786
-4	-1.650	3.654	-20.474	7.990	18.836	-12.790	39.257	-27.971
-3	-0.916	4.850	-2.980	29.991	2.290	-6.731	32.092	242.440
-2	-21.857	4.708	-62.835	20.183	5.914	-0.534	45.391	147.864
-1	-3.069	6.885	6.545	7.609	12.498	-45.931	44.437	26.656
0	-1.483	1.890	-1.466	1.817	-1.586	1.533	-1.482	1.569
1	-1.047	3.010	-3.380	2.552	-1.116	0.049	-4.186	3.219
2	10.963	3.127	5.394	4.925	-0.502	4.137	-4.207	3.543
3	-2.536	3.385	-4.436	7.075	-2.729	3.324	-4.012	-27.236
4	-1.856	3.644	11.055	3.887	1.571	2.705	-10.989	-8.209
5	-1.295	2.476	-1.881	2.067	-3.227	3.272	-0.088	0.783
6	-0.090	2.685	-2.788	-3.418	-2.983	-44.768	-0.797	-0.053
7	-2.744	2.686	75.652	4.212	-1.055	-2.973	-7.461	0.000
8	-1.758	3.098	-1.387	-2.462	2.418	-0.420	0.000	0.000

TABLA 4.1.4
ARREGLO DE LA FACPG PARA LA SERIE $\{Z_t\}$

ORDEN DEL POLINOMIO DE PROMEDIOS MOVILES	ORDEN DEL POLINOMIO AUTORREGRESIVO					
	1	2	3	4	5	6
0	0.483	-0.274	-0.224	-0.238	-0.126	0.033
1	0.047	-0.639	0.053	-0.126	-0.188	0.019
2	-11.963	-0.644	-1.809	-0.164	-0.219	0.614
3	1.536	-0.680	0.216	-0.885	-0.144	0.259
4	0.856	-0.980	-4.022	-0.720	-0.683	0.227
5	0.295	-0.610	-0.176	-0.067	0.217	-0.019

iv) El Método de la Esquina.

Como se sabe de la sección 1.5, el método de la esquina presenta algunas desventajas con respecto a los métodos presentados en las secciones 1.3 y 1.4, pues mientras dichos métodos presentan un criterio para identificar cuando un proceso es no estacionario, el método de la esquina no presenta ninguna caracterización de este tipo, por lo que para usarlo habrá que asegurarse de que el proceso en cuestión es estacionario. La manera de averiguar si el proceso es no estacionario y en su caso remover la no estacionariedad, puede ser mediante el criterio presentado en el método de Box-Jenkins.

Por otro lado, dado que existen restricciones sobre la prueba de hipótesis (1.5.3) para el caso de la configuración rectangular que en estos tres ejemplos no se satisfacen, sólo se hará una inspección visual de la caracterización de la esquina.

Del inciso i) se sabe que la serie $\{Z_t\}$ es estacionaria, por lo que el método de la esquina se puede aplicar

directamente. La tabla 4.1.5 presenta el arreglo- $\hat{A}$ para los datos de la serie $\{Z_t\}$. Inspeccionando dicha tabla se puede observar que $|\hat{A}(m,n)| \leq 0.013$ para $m \geq 2$ y $n \geq 3$, más aún se puede apreciar que sólo $\hat{A}(2,4)$ y $\hat{A}(3,3)$ son considerablemente distintos del resto de los valores, a los que se puede considerar prácticamente igual a cero. Por consiguiente, de acuerdo al patrón mostrado por la tabla (1.5.1) el modelo identificado mediante este método es un:

$$\text{ARMA}(2,1): (1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta B)a_t$$

TABLA 4.1.5

ARREGLO- $\hat{A}$ PARA LA SERIE $\{Z_t\}$; m ES EL ORDEN DE PROMEDIOS MOVILES, n ES EL ORDEN AUTORREGRESIVO.

$m \cdot n$	1	2	3	4	5	6
1	0.483	0.211	-0.124	0.090	-0.028	-0.002
2	0.023	0.134	-0.007	0.013	0.004	0.000
3	-0.277	0.087	0.013	0.002	0.000	0.000
4	-0.425	0.084	0.002	0.001	0.000	0.000
5	-0.364	0.082	0.008	0.000	0.000	0.000
6	-0.107	0.048	0.000	0.000	0.000	0.000
7	0.098	0.028	0.000	0.000	0.000	0.000

v) El método de la FAME.

La FAME y la tabla de símbolos correspondiente para la serie $\{Z_t\}$ se muestran en las tablas 4.1.6 y 4.1.7 respectivamente. Inspeccionando estas tablas se aprecia un punto de corte triangular (2,1), indicando hacia un modelo ARMA(2,1), salvo para unos valores en la cuarta columna. No

obstante, estos valores no son sustancialmente superiores a $2\sigma = 0.16$. Además en el cuadro 4.1.3 se muestran los estimadores $\hat{\phi}_{1(2)}^{(j)}$ y $\hat{\phi}_{2(2)}^{(j)}$ para las primeras iteraciones, observándose que para $j \geq 1$ los valores están todos próximos a 1.37 y -0.64 respectivamente. Por consiguiente, el modelo obtenido mediante este método es un:

$$\text{ARMA}(2, 1): (1 - \phi_1 B - \phi_2 B^2)Z_t = (1 - \theta B)a_t$$

con $\hat{\phi}_1 = 1.37$ y $\hat{\phi}_2 = -0.64$.

TABLA 4.1.6

LA FUNCION DE AUTOCORRELACION MUESTRAL EXTENDIDA DE $\{Z_t\}$

AR ⁺ NA	0	1	2	3	4	5	6	7	8	9
0	.48	.02	-.28	-.43	-.36	-.11	.10	.17	.13	.02
1	.45	.03	-.16	-.21	-.30	-.16	.12	.12	.13	.02
2	-.49	-.06	.11	.02	-.16	.05	-.01	-.04	.03	-.04
3	-.49	-.32	.08	.04	-.18	.04	-.01	-.10	-.02	-.03
4	-.41	.26	-.23	.10	-.18	-.04	-.08	-.04	-.02	-.06
5	.30	-.09	-.27	-.26	-.20	-.00	.07	-.07	.02	-.08
6	.49	-.11	-.10	-.08	-.18	-.00	.05	-.07	-.06	-.03

TABLA 4.1.6

TABLA DE SIMBOLOS (NIVEL DE SIGNIFICANCIA $\alpha = 0.05$)

[illegible]

CUADRO 4.1.3
ESTIMADORES DEL POLINOMIO $\Phi_2(B)$

J	$\hat{\phi}_{1(2)}$	$\hat{\phi}_{2(2)}$
0	.62	-.28
1	1.37	-.64
2	1.50	-.65
3	1.47	-.94
4	1.57	-1.10
5	.92	-.54
6	1.09	-.59
7	1.08	-.60

Ejemplo 2.- Considérese ahora la serie C de Box y Jenkins (1970) que está constituida por 226 mediciones de temperatura cada minuto en un determinado proceso químico. Una gráfica de los datos se presenta en la figura 4.2.1.

i) Método de Box-Jenkins.

La tabla 4.2.1 presenta las autocorrelaciones muestrales de la serie C y de sus dos primeras diferencias, mientras que en la tabla 4.2.2 se presentan las correspondientes autocorrelaciones parciales muestrales. En las gráficas de la figura 4.2.2 se muestran las autocorrelaciones parciales muestrales de la serie C y de sus dos primeras diferencias, aquí puede apreciarse que la serie original no es estacionaria y que al menos una diferencia es necesaria para estacionarizar la serie.

Esto mismo queda corroborado por los valores de las desviaciones estándar muestrales (ver Guerrero, 1983, cap. 4).

$$S(0) = 2.0540$$

$$S(1) = 0.2306$$

$$S(2) = 0.1408$$

$$S(3) = 0.2062$$

En una primera inspección de las FAC y FACP muestrales de la primera diferencia de la serie C se aprecia básicamente que la caída de la FAC muestral es bastante suave, así como el valor sensiblemente alto de la primera autocorrelación parcial en comparación con las otras autocorrelaciones parciales, esto desde luego indica la posibilidad de que el proceso generador de la serie sea un esquema autorregresivo de primer orden $ARI(1,1)$. Por otra parte, al observar las FAC y FACP muestrales de la segunda diferencia de la serie C se podría pensar en un modelo $IMA(2,0)$. Las propuestas anteriores quedan corroboradas al observar las gráficas de la figura 4.2.3, pues mientras la gráfica de la FACP de la serie C sugiere un modelo $AR(2)$, la gráfica de la FACP de la primera diferencia de la serie C sugiere un modelo $ARI(1,1)$ y la gráfica de la FACP de la segunda diferencia sugiere un ruido blanco no correlacionado $IMA(2,0)$. Por consiguiente surgen como posibles los siguientes modelos:

$$1) \text{ } ARI(1,1,0): \quad (1-\phi B)(1-B)Z_t = a_t$$

$$2) \text{ } IMA(0,2,0) \quad (1-B)^2 Z_t = a_t$$

en donde, en particular el modelo 1) consideraría como caso especial al $IMA(2,0)$ si el valor estimado de ϕ resulta ser 1.

CUADRO 4.2.1

FAC MUESTRAL DE LA SERIE C Y DE SUS DOS PRIMERAS DIFERENCIAS

<u>SERIE C</u>											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.98	.94	.90	.85	.80	.75	.69	.64	.58	.52	.47
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.41	.36	.31	.25	.20	.15	.10	.05	-.00	-.05	-.10
<u>∇ SERIE C</u>											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.81	.65	.53	.44	.38	.32	.26	.19	.14	.14	.10
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.09	.07	.07	.07	.09	.05	.04	.04	.04	.04	.05
<u>∇^2 SERIE C</u>											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	-.08	-.07	-.12	-.06	.01	-.02	.05	-.05	-.12	.12	-.12
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.07	-.08	.03	-.01	-.06	.17	-.10	-.01	-.02	-.01	.14

191
CUADRO 4.2.2

FACP MUESTRAL DE LA SERIE C Y DE SUS DOS PRIMERAS DIFERENCIAS

SERIE C											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	.98	-.26	-.16	-.09	-.06	-.05	-.01	-.04	-.02	-.01	-.04
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	-.04	-.04	-.02	-.04	-.03	-.03	-.07	-.05	-.02	-.03	-.06
V SERIE C											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	.81	.01	-.01	.05	.03	-.02	-.01	-.08	.02	.12	-.14
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	.09	-.03	.03	.01	.01	.04	-.12	.06	-.01	.04	-.02
V ² SERIE C											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	-.08	-.07	-.13	-.09	-.02	-.05	.02	-.06	-.14	.09	-.14
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	.02	-.09	-.01	-.04	-.07	.13	-.10	-.01	-.05	.00	.08

FIGURA 4.2.1

GRAFICA DE LA SERIE C

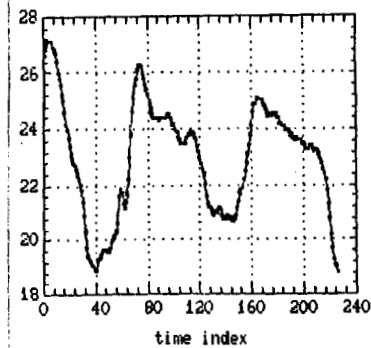


FIGURA 4.22
FAC MUESTRAL DE: (a) SERIE C, (b) $\sqrt{C}$ SERIE C, (c) $\sqrt{C^2}$ SERIE C

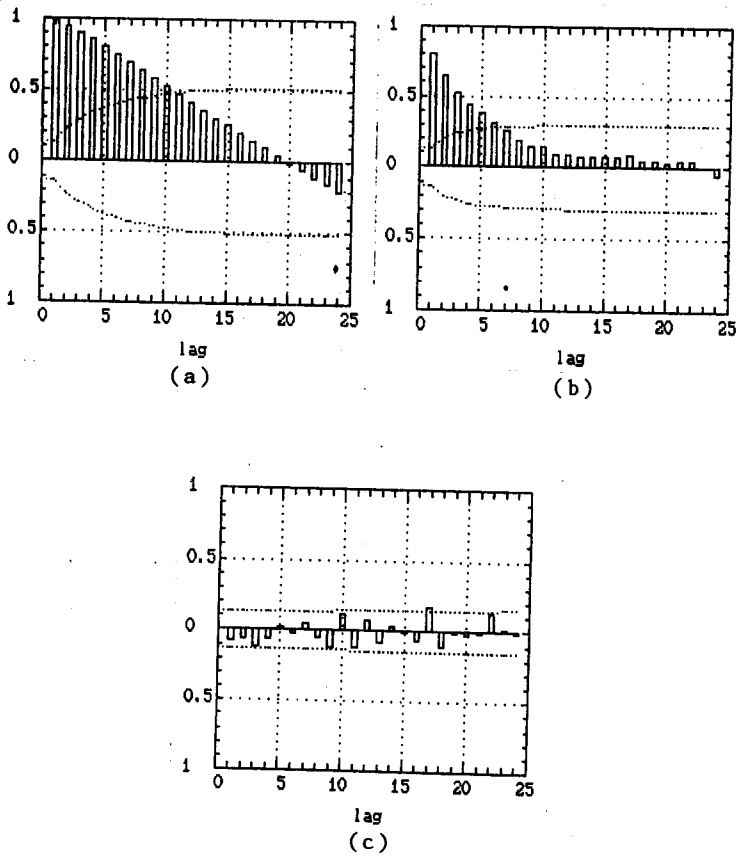
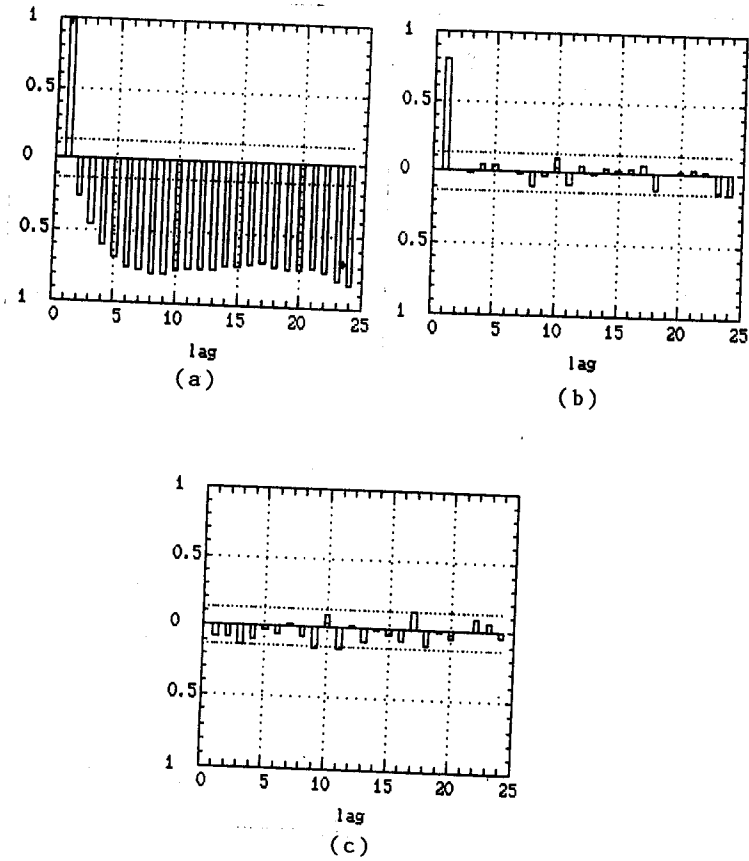


FIGURA 4.23

FACP MUESTRAL DE: (a) SERIE C, (b) $\sqrt{V}$ SERIE C, (c) V^2 SERIE C



11) Arreglos R y S.

Las tablas 4.2.1 y 4.2.2 muestran los arreglos S para los datos de la serie C y de su primera diferencia respectivamente. En una primera inspección de la tabla 4.2.1 se aprecia un comportamiento constante en la primera columna, esto es, $S_1(m) \approx -2$ para $m \in \mathbb{Z}$ que de acuerdo con el teorema (1.3.7), indica que el proceso es puramente no estacionario, debido a la existencia de una raíz unitaria. en el polinomio autorregresivo. No obstante, la segunda y tercera columna de dicha tabla parecen indicar que el proceso generador de la serie es un esquema autorregresivo y de promedios móviles ARMA(2,1). Sin embargo, la no estacionariedad tiene que ser removida antes de la identificación final del modelo. La tabla 4.2.2 presenta los arreglos S para la primera diferencia de la serie C. Inspeccionando dicha tabla se observa que las dos primeras columnas de esta identifican un modelo AR(1), pues se aprecia que el comportamiento dictado por los teoremas (1.3.4) y (1.3.5) se manifiesta en la primera columna de esta tabla, identificandose con esto un valor $p = 1$. En esta misma tabla se puede observar que $q-p+1 = 0$, de donde se sigue que $q = 0$. La elección de p y q está justificada por el hecho de que $|S_2(-2)| = 170.280$ y $S_2(-1) = 1.786$, que concuerda con el comportamiento mostrado en la tabla (1.3.3), esto es, $|S_2(-2)| = \infty$ mientras que $S_2(-1) = -C_1 = -S_1(0)$. Además la elección de $p = 1$ se confirma al observar la segunda columna de los arreglos R (tabla 4.2.3), que es visiblemente próxima a cero. Por consiguiente al aplicar los teoremas (1.3.2) - (1.3.5) a los resultados de la tabla 4.22, ninguna otra elección

de p y q parece ser posible, por lo que el modelo identificado para representar a la serie C es un modelo:

$$\text{ARI}(1,1): (1-B)(1-\phi B)Z_t = a_t$$

con

$$\hat{\phi} \approx \frac{1.8055}{2.2415} = 0.81$$

Esto último demuestra claramente que la aparente existencia de un término de promedios móviles indicado en la tabla (4.2.1), es inducido por la no estacionariedad de la serie, lo que se manifiesta en el comportamiento constante de la primera columna de dicha tabla. Este resultado concuerda con lo expuesto en el teorema (1.3.8), pues la autocorrelación ρ_m^* correspondiente al modelo $\text{ARI}(1,1)$ satisface una ecuación en diferencias de primer orden.

Nótese además que se han obtenido estimadores de los parámetros autorregresivos, pues del comportamiento de la primera columna de la tabla (4.2.1) y del teorema (1.3.7) se obtuvo el factor $(1-B)$, mientras que de la tabla (4.2.2) y del teorema (1.3.5) se obtuvo la estimación del parámetro ϕ .

TABLA 4.2.1

ARREGLO S PARA LA SERIE C CON $f = (-1)^m \hat{\rho}_m$

m.n.	1	2	3	4
-8	-2.089	3.943	-1.985	-2.542
-7	-2.080	4.755	-2.464	3.554
-6	-2.073	4.314	-5.288	-618.796
-5	-2.064	4.471	5.590	-1277.194
-4	-2.056	4.497	567.959	24.625
-3	-2.046	4.449	13.380	2.296
-2	-2.035	9.576	-2.100	3.582
-1	-2.022	2.492	-3.544	24.214
0	-1.977	3.604	-1.989	9.326
1	-1.965	3.572	-23.995	-1.477
2	-1.955	3.589	4.549	20.870
3	-1.946	3.704	2.909	2.610
4	-1.939	3.414	-1.719	0.655
5	-1.931	4.060	-4.178	7.832
6	-1.925	3.660	31.451	3.774
7	-1.918	3.604	-2.994	10.463
8	-1.911	4.210	-5.950	0.000

TABLA 4.2.2

ARREGLO S PARA {V SERIE C} CON $f = (-1)^m \hat{\rho}_m$

m.n.	1	2	3	4
-8	-2.403	2.607	2.269	6.803
-7	-2.217	5.250	-7.401	36.703
-6	-2.192	-0.212	-7.330	-59.369
-5	-2.163	7.000	-1.330	-48.318
-4	-2.180	1.911	162.057	-33.277
-3	-2.240	-0.989	246.094	1.683
-2	-2.234	-170.280	-1.773	2.631
-1	-2.241	1.786	10.769	3.580
0	-1.805	0.546	6.193	3.798
1	-1.810	-10.704	-3.660	-39.277
2	-1.806	2.706	-3.039	-5.826
3	-1.840	0.164	-3.450	-2.161
4	-1.859	3.127	8.671	5.838
5	-1.838	11.435	-4.335	-5.625
6	-1.821	1.314	-3.354	0.668
7	-1.712	12.831	-0.602	0.000
8	-1.745	-0.240	-0.618	0.000

TABLA 4.2.3

ARREGLO R PARA {Vserie c} CON $f_n = (-1)^n \hat{\rho}_n$

$m \cdot n$	1	2	3
-8	0.186	0.020	-0.003
-7	-0.261	-0.003	-0.005
-6	0.318	0.005	-0.185
-5	-0.379	0.005	0.008
-4	0.441	-0.012	0.002
-3	-0.528	-0.001	-0.441
-2	0.652	-0.002	0.196
-1	-0.805	-0.194	0.001
0	1.000	-0.002	0.029
1	-0.805	-0.001	0.012
2	0.652	-0.009	-0.004
3	-0.528	0.004	0.076
4	0.441	0.004	-0.007
5	-0.379	-0.003	-0.013
6	0.318	0.015	0.031
7	-0.261	0.003	0.023
8	0.186	-0.023	0.082

iii) Función de Autocorrelación Parcial y de Autocorrelación Parcial Generalizada.

La tabla 4.2.5 muestra el arreglo S^* para los datos de la serie C. En una primera inspección de dicha tabla se aprecia un comportamiento constante en la primera columna igual al que se observó en la primera columna de la tabla (4.2.1) de los arreglos S, que de acuerdo con el teorema (1.3.7) indica que el proceso es no estacionario, debido a la existencia de por lo menos una raíz unitaria en el polinomio autorregresivo. No obstante, al inspeccionar las columnas 3, 4, 5 y 6 de dicha tabla se aprecia que los comportamientos de $\pm \infty$ y de $(-1)^i C_1$ del teorema (1.4.1), se aprecian para el primer caso en el retraso $-q-1 = -2$ este comportamiento se manifiesta con los números 567.959, -1277.194, -250.663 y 94.968, mientras que en el retraso $q = 1$ y en las columnas 3, 4, 5 y 6 las cantidades -3.544, 3.582, -3.927 y 2.845 corresponden al comportamiento $(-1)^i C_1$. Estas observaciones, junto con el comportamiento dictado por los teoremas (1.3.4) y (1.3.5), que se aprecia en la columna $S_2^*(m)$, identifican al modelo generador de la serie $\{Z_t\}$ como un modelo ARMA(2,1).

La tabla 4.2.6 muestra el arreglo de la FACPG basándose en los resultados de la tabla (4.2.5) y el teorema (1.4.2). Inspeccionando dicha tabla se puede observar que $|\hat{\phi}_{2+1,2+1}^{(1)}| \leq 0.029$ para $i = 1, 2, \dots$ y que $\hat{\phi}_{2,2}^{(1+1)} \approx 0.80$ para $i = 0, 1, 2, \dots$ identificándose con esto un valor para $p = 2$ y $q = 1$. Por consiguiente, el valor obtenido mediante este arreglo es un ARMA(2,1). Sin embargo, la no estacionariedad que se manifiesta en la columna 1 de la tabla (4.2.5) tiene que ser removida antes de la identificación final del modelo. La.

tabla 4.2.7 muestra el arreglo S^* para los datos de la primera diferencia de la serie C, inspeccionando esta tabla se aprecia que los comportamientos de $\pm \infty$ y de $(-1)^1 C_1$ del teorema (1.4.1) se aprecian para el primer caso en la segunda columna y en el retraso $-q-1 = 1$, dicho comportamiento se manifiesta con los números -170.28, 246.094, -33.277, -63.846, 91.289 y 131.733, mientras que en el retraso $q = 0$ y en las columnas 2, 3, 4, 5, 6, y 7 las cantidades 1.786, -1.773, 1.683, -1.729, 1.762 y -1.739 corresponden al comportamiento de $(-1)^1 C_1$. Estas observaciones, junto con el comportamiento dictado por los teoremas (1.3.4) y (1.3.5) que se aprecia en la columna $S_1^*(m)$, identifican al modelo generador de la serie $\{Z_t\}$ como un modelo AR(1) con

$$\hat{\phi} \approx \frac{1.805}{2.241} = 0.81$$

Por consiguiente el posible modelo generador de la serie es un:

$$ARI(1,1): (1-B)(1-0.81B)Z_t = a_t$$

Por otra parte, la tabla 4.2.8 muestra el arreglo de la FACPG para los datos de la primera diferencia de la serie C basándose en los resultados de la tabla (4.2.8).

Inspeccionando dicha tabla se puede observar que $|\hat{\phi}_{1+1,1+1}^{(0)}| \approx 0.05$ para $i = 1, 2, \dots$ y $\hat{\phi}_{1,1}^{(1)} \approx 0.81$ para $i = 0, 1, 2, \dots$

Identificándose con esto un modelo AR(1). Nótese que la identificación del orden de este modelo puede obtenerse inspeccionando solamente el primer renglón de la tabla (4.2.8), pues como se explicó en la sección 1.4, en dicho renglón se encuentra la FACP muestral. Sin embargo, el comportamiento constante que se aprecia en la primera columna es un patrón más

visible. En cualquier caso, la combinación de ambas formas de identificación proporcionan más información que una sola. Por lo tanto el modelo obtenido mediante el arreglo de la FACPG es un:

$$\text{ARI}(2,1) : (1-B)(1-\phi B)Z_t = a_t$$

con

$$\hat{\phi} \approx 0.81$$

TABLA 4.2.5

ARREGLO S^* PARA LA SERIE C CON $f_{\square} = (-1)^{\square} \hat{\rho}_{\square}$

m. n.	1	2	3	4	5	6	7	8
-8	-2.089	4.370	-31.841	0.649	-6.289	-7.913	-2.347	0.0
-7	-2.080	3.943	3.494	5.817	-7.048	-0.355	-2.973	-393.977
-6	-2.073	4.755	-1.985	2.163	-2.836	2.671	-14.257	108.930
-5	-2.064	4.314	-2.464	-2.542	-3.151	1708.965	-97.102	101.852
-4	-2.056	4.471	-5.288	3.554	81.545	-192.711	-67.838	-92.973
-3	-2.046	4.497	5.590	-618.796	06.335	-405.387	-31.992	114.139
-2	-2.035	4.449	567.959	-1277.194	-250.663	94.968	-72.223	120.679
-1	-2.022	9.576	13.380	24.625	37.599	49.771	182.195	61.815
0	-1.977	2.492	-2.100	2.296	-2.164	2.263	-2.235	2.319
1	-1.965	3.604	-3.544	3.582	-3.972	2.845	-9.154	3.647
2	-1.955	3.572	-1.989	24.214	4.665	-2.883	-1.464	3.988
3	-1.946	3.589	-23.995	9.326	1.415	20.022	-2.870	2.135
4	-1.939	3.704	4.549	-1.477	-11.550	11.024	-3.693	11.060
5	-1.931	3.414	2.909	20.870	-11.003	-61.313	-0.515	4.750
6	-1.925	4.060	-1.719	2.610	-3.767	0.419	-4.128	3.984
7	-1.918	3.660	-4.178	0.655	-8.734	4.621	-4.014	0.000
8	-1.911	3.604	31.451	7.832	-8.294	17.546	0.000	0.000

TABLA 4.2.6

ARREGLO DE LA FACPG PARA LA SERIE C

ORDEN DEL POLINOMIO DE PROMEDIOS MOVILES	ORDEN DEL POLINOMIO AUTORREGRESIVO					
	1	2	3	4	5	6
0	0.977	-0.260	-0.157	-0.093	-0.057	-0.045
1	0.965	-0.810	-0.006	0.002	0.015	-0.029
2	0.955	-0.794	-0.355	0.039	0.022	-0.007
3	0.946	-0.802	4.537	-2.623	-0.017	0.103
4	0.939	-0.858	-1.845	-0.580	3.665	-0.006
5	0.935	-0.798	-1.465	-9.647	3.879	22.947

TABLA 4.2.7

ARREGLO S PARA {Vserie C} CON $f_n = (-1)^n \hat{\rho}_n$

m.n.	1	2	3	4	5	6	7	8
-8	-2.403	-7.333	-3.742	8.797	-4.849	5.786	-5.988	0.0
-7	-2.217	2.607	-2.264	0.753	-3.596	6.781	52.081	-97.843
-6	-2.192	5.250	2.269	4.648	-3.378	-4.653	103.209	-2022.032
-5	-2.163	-0.212	-7.401	6.803	-1.595	-74.550	192.141	147.786
-4	-2.190	7.000	-7.330	36.703	50.897	-37.540	6.444	43.321
-3	-2.240	1.911	-1.330	-59.369	-133.236	21.284	-51.619	40.853
-2	-2.234	-0.969	162.057	-48.318	-8.393	77.856	-119.681	17.247
-1	-2.241	-170.280	246.094	-33.277	-63.846	91.289	131.733	23.490
0	-1.805	1.786	-1.773	1.683	-1.729	1.762	-1.739	1.878
1	-1.810	0.546	10.769	2.631	-0.530	2.945	-12.415	1.437
2	-1.806	-10.704	6.193	3.580	-2.627	4.338	8.536	4.424
3	-1.840	2.706	-3.660	3.798	22.821	-5.016	1.165	6.086
4	-1.859	0.164	-3.039	-39.277	-1.784	5.625	-6.991	6.699
5	-1.838	3.127	-3.450	-5.826	-5.746	4.279	-5.409	17.028
6	-1.821	11.435	8.671	-2.161	-7.385	14.715	-16.158	-294.667
7	-1.712	1.314	-4.335	5.838	-6.912	18.263	137.816	0.000
8	-1.745	12.831	-3.354	-5.625	-2.955	-4.651	0.000	0.000

TABLA 4.2.8

ARREGLO DE LA FACPG PARA {V SERIE C}

ORDEN DEL POLINOMIO DE PROMEDIOS MOVILES	ORDEN DEL POLINOMIO AUTORREGRESIVO					
	1	2	3	4	5	6
0	0.805	0.010	-0.007	0.050	0.027	-0.019
1	0.810	-0.563	0.066	0.054	0.063	-0.037
2	0.806	5.599	-4.654	0.060	0.019	-0.203
3	0.839	-0.386	0.499	0.103	0.448	0.133
4	0.859	0.773	0.410	-5.773	1.118	0.075
5	0.838	-0.595	-1.520	-1.253	1.700	-0.919

iv) El Método de la Esquina.

Del inciso i) se sabe que la serie original no es estacionaria, pero que una primera diferencia sí parece volverla estacionaria. Por lo que el método de la esquina se aplica a los datos de la serie {V serie c}; la tabla 4.2.9 presenta el arreglo- $\hat{A}$ para los datos de dicha serie. Inspeccionando esta tabla se puede observar que $|\hat{A}(m,n)| \leq 0.007$ para $m \geq 1$ y $n \geq 2$, más aún se puede apreciar que sólo $\hat{A}(4,2)$, $\hat{A}(1,2)$ y $\hat{A}(2,2)$ son considerablemente distintos del resto de los valores, a los que se puede considerar prácticamente igual a cero. Por consiguiente, de acuerdo al patrón mostrado por la tabla (4.2.9), el modelo identificado mediante este método es un:

$$\text{ARI}(1,1): (1-B)(1-\phi B)Z_t = a_t$$

TABLA 4.2.9

ARREGLO-A PARA LA SERIE {Vserie C}; m ES EL ORDEN DE
PROMEDIOS MOVILES, n ES EL ORDEN AUTORREGRESIVO.

m \ n	1	2	3	4	5	6
1	0.805	0.006	0.003	0.000	0.000	0.000
2	0.652	0.006	0.000	0.000	0.000	0.000
3	0.526	0.005	0.000	0.000	0.000	0.000
4	0.441	0.007	0.000	0.000	0.000	0.000
5	0.379	0.003	0.000	0.000	0.000	0.000
6	0.318	0.003	0.000	0.000	0.000	0.000
7	0.261	0.006	0.000	0.000	0.000	0.000

v) El Método de la FAME.

La FAME y la tabla de símbolos correspondiente para la serie C se muestran en las tablas 4.2.10 y 4.2.11 respectivamente. Manifiestamente se aprecia un punto de corte triangular (2,0), indicando hacia un modelo AR(2).

Observando los estimadores $\hat{\phi}_{1(2)}^{(0)}$; $i = 1, 2$, del cuadro 4.2.3, se deduce para el polinomio $\hat{\phi}_2^{(0)}(B)$ una factorización del tipo

$$\hat{\phi}_2^{(0)}(B) \approx (1-B)(1-0.81B)$$

Esta factorización destaca la presencia de una diferenciación regular del modelo. Nótese que a diferencia de los métodos anteriores, la identificación del orden del modelo mediante la FAME no requiere de la diferenciación de la serie antes de especificar el orden del modelo. Por lo tanto, el modelo obtenido mediante la FAME es un:

$$\text{ARI}(2,1,0): (1-B)(1-0.81B)Z_t = a_t$$

TABLA 4.2.10

LA FUNCION DE AUTOCORRELACION MUESTRAL EXTENDIDA DE LA SERIE C

AR. NA	0	1	2	3	4	5	6	7	8	9
0	.98	.94	.90	.85	.80	.75	.69	.64	.58	.52
1	.81	.66	.55	.48	.43	.38	.34	.28	.25	.25
2	-.04	-.03	-.12	-.06	.02	-.01	.07	-.04	-.12	.13
3	-.50	.01	-.07	-.11	-.01	-.00	.03	-.03	-.10	.01
4	-.25	-.27	-.05	-.11	-.01	.03	.00	-.02	-.09	-.01
5	-.48	.28	-.29	-.07	.04	-.05	-.00	-.01	-.08	.07
6	-.08	-.32	.14	.04	-.03	-.04	-.01	-.03	-.08	.07
7	-.18	-.11	.13	-.00	-.32	.02	-.00	.03	-.07	.02

TABLA 4.2.11

TABLA DE SIMBOLOS (NIVEL DE SIGNIFICANCIA $\alpha = 0.05$)

AR. NA	0	1	2	3	4	5	6	7	8	9
0	X	X	X	X	X	X	X	X	X	X
1	X	X	X	X	X	X	X	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	X	0	0	0	0	0	0	0	0	0
4	X	X	0	0	0	0	0	0	0	0
5	X	X	X	0	0	0	0	0	0	0
6	0	X	0	0	0	0	0	0	0	0
7	X	0	0	0	X	0	0	0	0	0

CUADRO 4.2.3

ESTIMADORES DEL POLINOMIO $\phi_2(B)$

J	$\hat{\phi}_{1(2)}$	$\hat{\phi}_{2(2)}$
0	1.81	-.82
1	1.83	-.84
2	1.84	-.85
3	1.89	-.90
4	1.92	-.93
5	1.92	-.92
6	1.92	-.93
7	1.88	-.89

Ejemplo 3. - Considérese la serie del Índice Nacional de Precios al Consumidor (INPC) que aparece en el cuadro 4.3.1. Dicha serie cubre el período que va de Enero de 1981 a Diciembre de 1987, por lo que el tamaño de la muestra manejada ha sido de 84 observaciones. Una gráfica de los datos se presenta en la figura 4.3.1.

1) Método de Box-Jenkins.

Para eliminar la heteroscedasticidad se eligió la transformación Z_t^λ con $\lambda = 0$, así pues $\{Z_t\}$ será la serie logarítmica correspondiente.

El cuadro 4.3.2 presenta las autocorrelaciones muestrales de la serie $\{\log(\text{INPC}_t)\}$ y de sus dos primeras diferencias, mientras que en el cuadro 4.3.2 se presentan las correspondientes autocorrelaciones parciales muestrales. La figura 4.3.2 muestra las gráficas de las FAC y FACP muestrales de la serie $\{\log(\text{INPC}_t)\}$ y de sus dos primeras diferencias, aquí puede apreciarse que la serie original no es estacionaria, pero que a lo más una segunda diferencia sí parece volverla estacionaria. Esto mismo queda corroborado por los valores de las desviaciones estándar muestrales

$$S(0) = 1.1733$$

$$S(1) = 0.0230$$

$$S(2) = 0.0179$$

$$S(3) = 0.0276$$

ya que $S(2)$ proporciona el mínimo de estos valores. Así pues, para identificar un modelo para esta serie se requiere exclusivamente conocer el comportamiento de las FAC y FACP muestrales para la segunda diferencia de la serie $\{\log(\text{INPC}_t)\}$.

Inspeccionando estas funciones se aprecia básicamente que éstas presentan el patrón.

de un proceso MA(1). Por consiguiente el posible modelo generador de la serie es un:

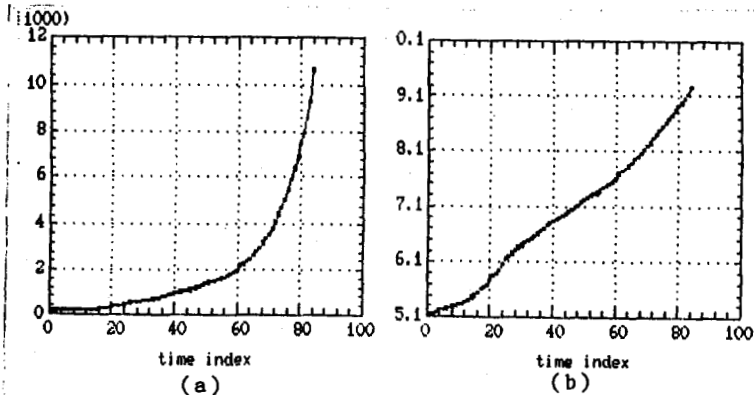
$$\text{IMA}(2,1): (1-B)^2 Z_t = (1-\theta B) a_t$$

CUADRO 4.2.1

INDICE NACIONAL DE PRECIOS AL CONSUMIDOR

MES . AÑO	1981	1982	1983	1984	1985	1986	1987
ENE.	171.0	223.7	469.9	814.8	1309.8	2173.3	4440.9
FEB.	175.2	232.6	495.1	857.8	1364.2	2269.9	4761.3
MAR.	178.9	241.0	519.1	894.5	1417.1	2375.4	5076.3
ABR.	182.9	254.1	552.0	933.2	1460.7	2499.4	5520.1
MAY.	185.7	268.4	575.9	964.1	1495.3	2638.3	5936.2
JUN.	188.3	281.3	597.7	999.0	1532.8	2807.6	6365.7
JUL.	191.6	295.8	627.3	1031.8	1586.2	2947.7	6881.3
AGO.	195.6	329.0	651.6	1061.1	1655.5	3182.7	7443.7
SEP.	199.2	346.6	671.7	1092.7	1721.6	3373.7	7934.1
OCT.	203.6	364.6	694.0	1130.9	1787.0	3566.5	8595.2
NOV.	207.5	382.9	734.0	1169.7	1869.5	3807.6	9277.0
DIC.	213.1	423.8	766.1	1219.4	1996.7	4108.2	10647.2

FIGURA 4.2.1

(a) SERIE $\{\text{INPC}_t\}$, (b) SERIE $\{\log(\text{INPC}_t)\}$ 

CUADRO 4.3.1

FAC MUESTRAL DE LA SERIE $\{\log(\text{INPC}_t)\}$ Y DE SUS DOS PRIMERAS
DIFERENCIAS

SERIE $\{\log(\text{INPC})\}$											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.96	.92	.89	.85	.81	.77	.74	.70	.66	.62	.59
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.55	.51	.48	.44	.41	.38	.34	.31	.28	.25	.22
SERIE $\{\nabla \log(\text{INPC})\}$											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	.60	.47	.44	.48	.38	.26	.27	.22	.09	.09	.15
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.13	-.03	-.10	-.10	-.12	-.17	-.15	-.15	-.17	-.15	-.14
SERIE $\{\nabla^2 \log(\text{INPC})\}$											
m	1	2	3	4	5	6	7	8	9	10	11
$\hat{\rho}_m$	-.24	-.19	-.07	-.18	.03	-.18	.07	.09	-.15	.11	.13
m	12	13	14	15	16	17	18	19	20	21	22
$\hat{\rho}_m$	.22	-.11	-.08	.04	-.03	-.01	.04	.00	-.06	.03	-.11

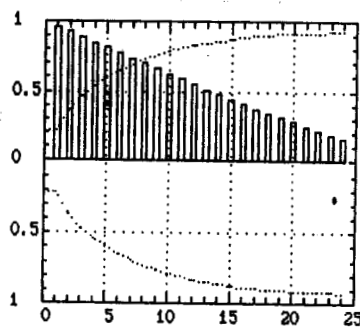
CUADRO 4.3.2

FACP MUESTRAL DE LA SERIE $\{\log(\text{INPC}_t)\}$ Y DE SUS DOS PRIMERAS DIFERENCIAS

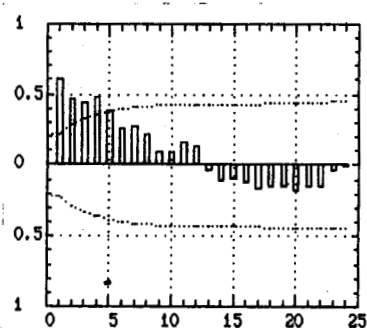
SERIE $\{\log(\text{INPC})\}$											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	.96	-.00	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	-.02	-.01	-.01	-.02	-.02	-.02	-.01	-.02	-.02	-.01	-.02
SERIE $\{\nabla \log(\text{INPC})\}$											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	.60	.17	.017	.21	-.04	-.10	.06	-.06	-.16	.07	.11
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	-.00	-.17	-.14	-.09	-.05	.01	.08	.00	.01	.07	-.08
SERIE $\{\nabla^2 \log(\text{INPC})\}$											
i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{11}$	-.24	-.26	-.21	.05	.06	-.12	.04	.06	-.16	-.16	.01
i	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{11}$	.18	.08	.07	.02	-.15	-.07	.05	-.04	-.04	.14	-.12

FIGURA 4.32

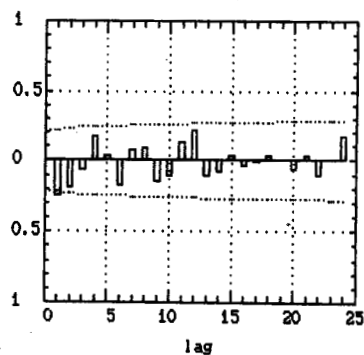
(a) FAC MUESTRAL DE: (i) $\langle \log(\text{INPC}_t) \rangle$, (ii) $\nabla \log(\text{INPC}_t) \rangle$ y $\nabla^2 \log(\text{INPC}_t) \rangle$. (b) FACF DE (i) $\langle \log(\text{INPC}_t) \rangle$, (ii) $\nabla \log(\text{INPC}_t) \rangle$ y $\nabla^2 \log(\text{INPC}_t) \rangle$



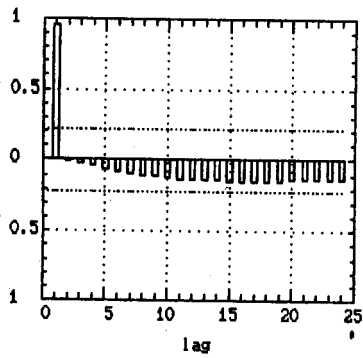
(i)



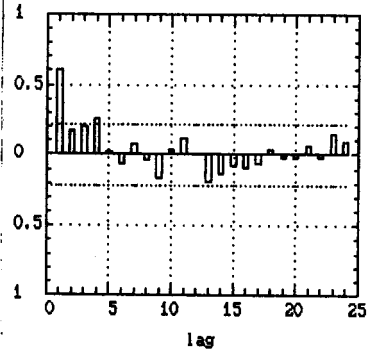
(ii)



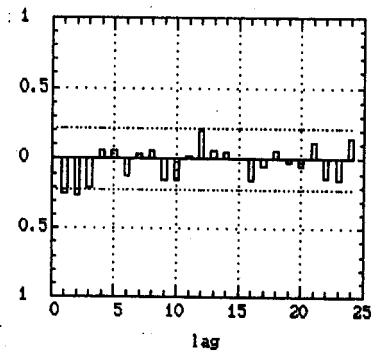
(iii)



(i)



(ii)



(iii)

ii) Arreglos R y S.

La tabla 4.3.1 muestra el arreglo S para los datos de la serie $\{\log(\text{INPC}_t)\}$. En una primera inspección de esta tabla se aprecia un comportamiento constante en su primera columna, esto es, $S_m(1) \approx -2$ para $m = -8, -7, \dots, 0, 1, 2, \dots, 8$ que de acuerdo con el teorema (1.3.7) indica que el proceso no es estacionario, debido a la existencia de al menos una raíz unitaria en el polinomio autorregresivo. Una vez removida la no estacionariedad, que en este caso se logra después de una segunda diferenciación de la serie, se procede a inspeccionar los arreglos R y S para los datos de la serie $\{\nabla^2 \log(\text{INPC}_t)\}$, que se muestran en las tablas 4.3.2 y 4.3.3 respectivamente. Inspeccionando la primera columna de los arreglos R, que como se sabe proporciona la FAC de la serie, que en este caso está regida por $(-1)^m \hat{\rho}_m$, se puede apreciar que dicha autocorrelación presenta las propiedades de un ruido blanco de orden uno, esto es, un MA(1). Lo anterior queda corroborado al mostrar que $\hat{\rho}_m = 0$ para $m \geq 2$, pues como se sabe (ver Guerrero, 1983, cap. 4).

$$\text{var}(\hat{\rho}_m) = \frac{1}{N} \sum_{j=-\infty}^{\infty} (\rho_j^2 + \rho_{j+m} \rho_{j-m} - 4 \rho_m \rho_j \rho_{j-m} + \rho_j^2 \rho_m^2)$$

por lo que

$$\text{var}(\hat{\rho}_2) = \frac{1}{82} [1 + 2(.24)^2] = 0.136$$

y

$$2\sqrt{\text{var}(\hat{\rho}_2)} = 0.233$$

de donde se sigue que $|\hat{\rho}_m| \leq 0.233$ para $m \geq 2$. De acuerdo a lo anterior los valores de $(-1)^m \hat{\rho}_m$ para $m \geq 2$ son

estadísticamente iguales a cero. Por lo tanto, el modelo identificado para representar a la serie $\{\nabla^2 \log(\text{INPC}_t)\}$ es un MA(1). Por consiguiente, el posible modelo generador de la serie $\{\log(\text{INPC}_t)\}$ es un:

$$\text{IMA}(2,1): (1-B)^2 Z_t = (1-\theta B)a_t$$

Sin embargo, si se observa el arreglo S para los datos de la serie $\{\nabla^2 \log(\text{INPC}_t)\}$ que se muestra en la tabla 4.3.4, se aprecia que la segunda y tercera columna de dicha tabla parecen identificar que el proceso generador de la serie es un esquema ARMA(2,3) con.

$$\hat{\phi} = (-1)^{p+1} \frac{C_3}{C_4} = - \frac{1.723}{1.940} = -0.888 \quad (4.1.1)$$

en donde C_3 es estimada mediante $S_p((-1)^{q-p+1} \hat{\rho}_{q-p+1})$ y C_4 por $S_p((-1)^{q-p} \hat{\rho}_{q-p+1})$. Por otra parte, se puede demostrar que la estimación de ϕ_2 mediante (4.3.1) es igual a la estimación de ϕ_2 mediante las ecuaciones de Yule-Walker para q dada (la demostración es análoga a la del teorema (1.4.1)). Ahora, la estimación de ϕ_2 mediante las ecuaciones de Yule-Walker es igual a

$$\hat{\phi}_2 = \frac{\hat{\rho}_2 - (\hat{\rho}_1)^2}{1 - (\hat{\rho}_1)^2} = -0.134$$

observándose con esto una diferencia significativa entre ambas estimaciones. Lo anterior conduce a concluir que los arreglos S son muy sensibles para valores pequeños de la FAC muestral y podrían presentarse patrones engañosos como este, por lo que se recomienda al usuario de este método un análisis previo de las autocorrelaciones de la serie que se presentan en la columna uno del arreglo R.

TABLA 4.3.1

ARREGLO S PARA LA SERIE $\{\log(\text{INPC}_t)\}$ con $f = (-1)^m \hat{\rho}_m$

m · n	1	2	3	4
-8	-2.053	3.901	-8.639	7.599
-7	-2.051	4.124	10.864	-3.456
-6	-2.048	4.379	-1.119	-94.382
-5	-2.046	3.553	6.249	-609.908
-4	-2.044	4.175	238.610	107.054
-3	-2.042	2.428	99.546	1.965
-2	-2.040	533.570	-1.930	4.533
-1	-2.040	1.968	-3.793	-1.806
0	-1.961	12.277	-2.760	2.087
1	-1.961	3.832	1.686	3.240
2	-1.959	4.604	-2.763	4.015
3	-1.957	3.666	-7.872	10.375
4	-1.955	3.877	-0.952	8.666
5	-1.953	4.109	-11.733	5.394
6	-1.951	3.938	5.812	4.635
7	-1.949	3.620	-0.619	5.280
8	-1.946	4.506	-5.771	0.000

TABLA 4.3.2

ARREGLO R PARA $\{v^2 \log(\text{INPC}_t)\}$ con $f = (-1)^m \hat{\rho}_m$

m · n	1	2	3
-8	0.089	0.130	-0.029
-7	-0.072	0.278	-0.051
-6	-0.177	-0.219	-0.010
-5	-0.030	-0.159	0.181
-4	0.175	0.342	0.427
-3	0.067	0.071	0.469
-2	-0.185	-0.570	0.920
-1	0.242	-1.242	0.195
0	1.000	0.322	0.125
1	0.242	0.042	0.105
2	-0.185	-0.146	-0.009
3	0.067	-0.304	-0.037
4	0.175	0.155	-0.023
5	-0.030	0.199	0.039
6	-0.177	-0.200	-0.177
7	-0.072	-0.115	0.061
8	0.089	-0.545	0.040

TABLA 4.3.3

ARREGLO S PARA LA SERIE $\{V^2 \log(INPC_t)\}$ CON $f = (-1)^m \hat{\rho}_m$

m.n.	1	2	3	4
-8	-1.810	1.648	1.121	4.493
-7	1.453	1.475	-1.450	6.443
-6	-0.825	1.821	-35.782	0.153
-5	-6.678	1.940	4.018	-5.367
-4	-0.616	2.969	2.046	-15.099
-3	-3.746	20.853	3.530	0.715
-2	-2.309	3.680	-0.751	1.601
-1	3.125	0.953	-0.548	0.069
0	-0.757	1.534	-0.991	1.157
1	-1.763	6.133	-1.879	0.756
2	-1.364	1.723	5.269	2.032
3	1.604	1.777	-0.512	3.849
4	-1.176	1.344	-3.207	-8.731
5	4.741	1.187	-5.191	5.607
6	-0.592	0.951	-3.260	1.267
7	-2.233	2.419	-0.747	8.918
8	0.647	2.145	-1.362	0.000

iii) Función de Autocorrelación Parcial y de Autocorrelación Parcial Generalizada.

La tabla 4.3.4 muestra el arreglo S^* para los datos de la serie $\{\log(INPC_t)\}$. En una primera inspección de esta tabla se aprecia un comportamiento constante en su primera columna, esto es, $S_m(1) \approx -2$ para $m \in \mathbb{Z}$ que de acuerdo con el teorema (1.3.7), indica que el proceso es no estacionario debido a la existencia de al menos una raíz unitaria en el polinomio autorregresivo. Una vez removida la no

stacionariedad que en este caso se logra después de una segunda diferenciación de la serie, se procede a inspeccionar el arreglo R para los datos de la serie $\{V^2 \log(\text{INPC}_t)\}$ que se presenta en las tablas 4.3.4. Utilizando el mismo argumento del método anterior se puede concluir que la primera columna del arreglo R identifica un modelo MA(1) para representar a la serie $\{V^2 \log(\text{INPC}_t)\}$. Por consiguiente el posible modelo generador de la serie $\{\log(\text{INPC}_t)\}$ identificado mediante este método es un

$$\text{IMA}(2,1): (1-B)^2 Z_t = (1-\theta B)a_t$$

TABLA 4.3.4

ARREGLO S^* PARA $\{\log(\text{INPC}_t)\}$ CON $f_n = (-1)^n \hat{\rho}_n$

$m \cdot n$	1	2	3	4	5	6	7	8
-8	-2.053	4.060	-5.872	3.514	6.354	-11.527	-3.312	0.000
-7	-2.051	3.901	1.243	9.501	-0.321	4.859	18.099	-115.941
-6	-2.048	4.124	-8.629	11.311	-5.255	7.169	151.962	-641.720
-5	-2.046	4.379	10.864	7.599	49.662	-200.569	-11.232	-990.077
-4	-2.044	3.553	-1.119	-3.456	258.936	-2220.159	1007.063	-999.558
-3	-2.042	4.175	6.249	-94.382	370.751	-607.432	-238.056	1596.128
-2	-2.040	2.428	232.610	-609.908	405.655	-1596.500	-1444.389	2028.663
-1	-2.040	533.570	99.546	107.054	78.482	94.285	97.443	91.839
0	-1.961	1.968	-1.930	1.965	-1.917	1.957	-1.919	1.960
1	-1.961	12.277	-3.793	4.533	-3.587	3.775	-4.039	3.854
2	-1.959	3.832	-2.760	-1.806	-2.800	8.429	-1.110	11.677
3	-1.957	4.604	1.686	2.087	-3.284	3.942	-9.935	8.024
4	-1.955	3.666	-2.763	3.240	-3.482	-11.968	0.060	3.511
5	-1.953	3.877	-7.872	4.015	9.004	-1.136	-3.551	3.490
6	-1.951	4.109	-0.952	10.375	0.303	3.719	-2.663	-8.574
7	-1.949	3.938	-11.733	8.666	-3.473	3.379	7.235	0.000
8	-1.946	3.620	5.812	5.394	-1.516	-32.415	0.000	0.000

216
TABLA 4.3.5

ARREGLO R PARA $\{V^2 \log(INPC_t)\}$ CON $f = (-1)^m \hat{\rho}_n$

m \ n	1
-8	0.089
-7	-0.072
-6	-0.177
-5	-0.030
-4	0.175
-3	0.067
-2	-0.185
-1	0.242
0	1.000
1	0.242
2	-0.185
3	0.067
4	0.175
5	-0.030
6	-0.177
7	-0.072
8	0.089

iv) El Método de la Esquina.

Del inciso i) se sabe que la serie $\{\log(INPC_t)\}$ no es estacionaria, pero que una segunda diferencia sí parece volverla estacionaria. Por lo que el método de la esquina se aplica a los datos de la serie $\{V^2 \log(INPC_t)\}$. La tabla 4.3.6 presenta el arreglo-A para los datos de dicha serie. Inspeccionando esta tabla se puede observar que $|\hat{A}(m,n)| \leq 0.185$ para $m \geq 2$ y $n \geq 1$, más aún se puede apreciar que sólo $\hat{A}(2,1)$, $\hat{A}(4,1)$ y $\hat{A}(6,1)$ son considerablemente distintos del resto de los valores, a los que se puede considerar prácticamente iguales a cero. Por consiguiente, de acuerdo al patrón mostrado en la tabla (4.3.6), el modelo identificado mediante este método es un.

$$IMA(2,1): (1-B)^2 Z_t = (1-\theta B) a_t$$

217
TABLA 4.3.1

ARREGLO-A PARA LA SERIE $\{V^2 \log(\text{INPC}_t)\}$; n ES EL ORDEN DE PROMEDIOS MOVILES, m ES EL ORDEN AUTORREGRESIVO.

$m \cdot n$	1	2	3	4	5	6
1	-0.242	-0.050	-0.180	-0.033	0.031	0.059
2	-0.185	0.019	0.049	0.010	0.005	0.006
3	-0.067	0.039	-0.012	0.005	-0.002	0.000
4	0.175	0.034	0.000	0.000	0.000	0.000
5	0.030	0.033	0.002	0.000	0.000	0.000
6	-0.177	0.030	-0.001	0.000	0.000	0.000
7	0.072	0.021	-0.002	0.000	0.000	0.000

v) El método de la FAME.

La FAME y la tabla de símbolos correspondiente para la serie $\{\log(\text{INPC}_t)\}$, se muestran en las tablas 4.3.7 y 4.3.8 respectivamente. Inspeccionando estas tablas se aprecia un punto de corte triangular (2,1), el cual identifica un modelo ARMA(2,1).

Observando los estimadores $\hat{\phi}_{1(2)}^{(0)}$ para $i=1,2$ del cuadro 4.3.3 se puede apreciar que para $j \geq 1$ los valores están todos muy próximos a 2 y -1 respectivamente, además de que se deduce para el polinomio $\phi_2^{(0)}(B)$ una factorización del tipo.

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

Esta factorización destaca la presencia de un operador diferencia de segundo orden. Nótese que en contraste con los métodos anteriores, la identificación del orden del modelo mediante la FAME no requirió de la diferenciación de la serie $\{\log(\text{INPC}_t)\}$ antes de especificar el orden del mismo. Por lo tanto, el posible modelo generador de la serie obtenido mediante la FAME es un:

$$\text{IMA}(2,1): (1-B)^2 Z_t = a_t$$

TABLA 4.2.7

LA FUNCION DE AUTOCORRELACION MUESTRAL EXTENDIDA DE $\{\log(\text{IMPC})\}$

AR MA	0	1	2	3	4	5	6	7	8	9
0	.96	.92	.89	.85	.81	.77	.74	.70	.66	.62
1	.51	.35	.34	.41	.31	.18	.22	.19	.07	.10
2	-.24	-.19	-.07	.18	.03	-.18	.07	.09	-.15	.11
3	-.47	-.07	-.08	.16	.02	-.18	.11	.05	-.09	-.15
4	-.47	.05	-.24	-.03	.08	-.03	-.05	-.12	-.03	.01
5	-.04	.11	-.24	-.01	.08	-.05	.04	-.09	-.14	.03
6	.32	-.02	-.23	.02	-.06	-.07	.07	-.09	-.12	.03
7	.28	-.01	-.20	-.04	-.04	-.17	.08	-.10	.00	.06

TABLA 4.2.8

TABLA DE SIMBOLOS (NIVEL DE SIGNIFICANCIA $\alpha = 0.05$)

AR MA	0	1	2	3	4	5	6	7	8	9
0	X	X	X	X	X	X	X	X	X	X
1	X	X	X	X	0	0	0	0	0	0
2	X	0	0	0	0	0	0	0	0	0
3	X	0	0	0	0	0	0	0	0	0
4	X	0	X	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	X	0	0	0	0	0	0	0	0	0
7	X	0	0	0	0	X	0	0	0	0

CUADRO 4.3.3

ESTIMADORES DEL POLINOMIO $\phi_2(B)$

j	$\hat{\phi}_{1(2)}$	$\hat{\phi}_{2(2)}$
0	1.98	-.98
1	2.00	-1.00
2	2.02	-1.02
3	2.03	-1.03
4	2.01	-1.01
5	2.01	-1.00
6	2.02	-1.02
7	2.02	-1.02

APENDICE 1

DEMOSTRACION DEL TEOREMA 1.2.1

Antes de demostrar el teorema 1.2.1 se enuncia y demuestra el siguiente lema que resulta útil en la demostración de dicho teorema.

Lema A1.1. El proceso $\{Z_t\}$ admite una representación ARMA(p,q) (no necesariamente minimal) si y sólo si la sucesión de autocorrelaciones $\{\rho_n, n \in \mathbb{Z}\}$ satisface una ecuación en diferencias de orden p, para el rango q+1:

$$\rho_m - \phi_1 \rho_{m-1} - \dots - \phi_p \rho_{m-p} \begin{cases} = 0; \text{ si } m \geq q+1 \text{ con } \phi_p \neq 0 \\ \neq 0; \text{ si } m = q \end{cases}$$

(p no necesariamente es el orden mínimo).

Demostración. => Ver Doob (1953, pag 505)

<=> Doob (1953, pag 504-505) demuestra que si la sucesión de autocorrelaciones $\{\rho_m\}$ satisface las hipótesis del lema A1.1, el proceso $\{Y_t, t \in \mathbb{Z}\}$ definido por:

$$Y_t = Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p}$$

es tal que:

$$E [Y_t Y_{t+m}] \begin{cases} = 0; & \text{si } m \geq q+1 \\ \neq 0; & \text{si } m = q \end{cases}$$

Por otra parte, se sabe que esta última condición es una caracterización de un proceso MA(q) (Ver Ansley, Spireg y Woobleski, 1977). Por lo tanto el proceso $\{Z_t\}$ admite una representación ARMA(p,q) ■

Demostración del teorema 1.2.1 $\Rightarrow$) De la condición necesaria del lema A1.1 se sabe que si el proceso $\{Z_t\}$ admite una representación ARMA(p,q), la sucesión $\{\rho_m; m \in \mathbb{Z}\}$ satisface una ecuación en diferencias de orden p no es el orden mínimo de dicha ecuación, entonces existe un orden mínimo $p' < p$. Por lo tanto, de la condición suficiente del lema A1.1 se sigue que $\{Z_t\}$ admite una representación ARMA(p',q') con $p < p'$, lo cual contradice el hecho de que la representación ARMA(p,q) es minimal, entonces el orden mínimo de la ecuación en diferencias es p.

$\Leftarrow$) De la condición suficiente del lema A1.1 se sabe que si la sucesión $\{\rho_m; m \in \mathbb{Z}\}$ satisface una ecuación en diferencias de orden mínimo p, para el rango mínimo q+1, el proceso $\{Z_t\}$ admite una representación ARMA(p,q).

Supóngase que dicha representación no es minimal, entonces existe una representación ARMA(p',q') con $p' < p$ que sí lo es. Por lo tanto, de la condición necesaria del lema A1.1, se sigue que la sucesión $\{\rho_m; m \in \mathbb{Z}\}$ satisface una ecuación en diferencias de orden $p' < p$ contradiciendo el hecho de que p es el orden mínimo de dicha ecuación, por lo tanto la representación ARMA(p,q) es minimal. ■

```
(
                                APENDICE 2
PROGRAMA PARA CALCULAR LOS ARREGLOS R Y S, S_MOD Y EL ESTADISTICO D
```

```
PROGRAM TESIS;
USES
    CRT, PRINTER;
```

```
CONST
    COL      = 10;
    FONDO    = 11;
TYPE
    TIIMENU  = STRING;
    CONTMENU = ARRAY [0..8] OF STRING[30];
    QUINCE   = STRING[15];
    MAT      = ARRAY [1..50,1..30] OF REAL;
    VECT     = ARRAY [1..300] OF REAL;
```

```
VAR
    R,S,R_EST,S_EST      : MAT;
    G,STAT               : VECT;
    NOMBRE               : QUINCE;
    CAR                  : CHAR;
    MXROW,MROW,NEG,KOL,IRHO,IRS,IP,IQ,IPNT,JK : INTEGER;
    MONCOL               : BOOLEAN;
```

```
(*****
*****
**                                     **
**          ESTE PROCEDIMIENTO INICIA LAS MATRICES R Y S          **
**                                     **
*****
*****)
```

```
PROCEDURE INICIA(MROW,KOL : INTEGER;
                 VAR R,S : MAT);
```

```
VAR
    IR,IC : INTEGER;
BEGIN
    FOR IC := 1 TO KOL DO
        FOR IR := 1 TO MROW DO
            BEGIN
                R[IR,IC] := 0.0;
                S[IR,IC] := 0.0;
            END; (* DEL FOR *)
        END; (* DE INICIA *)
```

```

(*****
**
** ESTE PROCEDIMIENTO CALCULA LOS ARREGLOS R Y S SIGUIENDO LAS REL-
** RELACIONES PROPUESTAS POR PYE Y ATCHISON (1973).
**
*****)

```

```

PROCEDURE ARRAY_R_AND_S(MXROW,MROW,NEG,KOL,IRHO:INTEGER;

```

```

    G      : VECT;
    VAR R,S : MAT;
    IRS     : INTEGER);

```

```

VAR

```

```

    IWK      : VECT;
    KOLUM1,MN,I,IC,NEGI,J1,ICM1,IR,IRP1,IRN      : INTEGER;

```

```

LABEL      40,50,60;

```

```

BEGIN

```

```

    KOLUM1 := KOL-1;
    INICIA(MROW,KOL,R,S);
    MN := MROW - NEG;
    FOR I := 1 TO MN DO
        BEGIN
            NEGI := NEG + I;
            S[NEGI,1] := 1.0;
            R[NEGI,1] := G[I];
        END; (* DEL FOR *)
    FOR I := 1 TO NEG DO
        BEGIN
            NEGI := NEG + 2-I;
            S[I,1] := 1.0;
            R[I,1] := G[NEGI];
        END; (* FIN DEL FOR *)
    IF ( IRHO = 0 )
    THEN
        GOTO 40
    ELSE
        FOR I:= 1 TO MROW DO
            BEGIN
                NEGI := I-NEG-1;
                IF ( NEGI MOD 2 ) <> 0
                THEN
                    R[I,1] := (-1) * R[I,1]
                ELSE
                    R[I,1] := R[I,1];
            END; (* DEL FOR *)
        END;
    J1 := MROW;
    FOR IC := 2 TO KOL DO
        BEGIN

```

```

40:

```

```

      ICM1 := IC - 1;
      J1 := J1-1;
      FOR IR := 1 TO J1 DO
        BEGIN
          IRP1 := IR + 1;
          IF (ABS(R[IR,ICM1]) = 0.0 )
            THEN
              GOTO 50
            ELSE
              BEGIN
                S[IR,IC] := S[IRP1,ICM1]*((R[IRP1,ICM1]/R[IR,ICM1])-1.0);
                GOTO 60;
              END;
          S[IR,IC] := -S[IRP1,ICM1];
60:      END; (* DEL FOR *)
      J1 := J1-1;
      FOR IR := 1 TO J1 DO
        BEGIN
          IRP1 := IR + 1;
          R[IR,IC] := R[IRP1,ICM1]*((S[IRP1,IC]/S[IR,IC])-1.0)
          END;(* DEL FOR *)
        END; (* DE ARRAYS_R_AND_S *)

```

```

(*****
*****
** ESTE PROCEDIMIENTO CALCULA LOS ARREGLOS R Y S MODIFICADOS SIGUIEN- **
** DO LA MODIFICACION PROPUESTA EN LA SECCION 1.4 **
** **
*****
*****)

```

```

PROCEDURE ARRAY_S_EST(MXROW,MROW,NEG,KOL,IRHO:INTEGER;
      G : VECT;
      VAR R_EST,S_EST : MAT;
      IRS : INTEGER);

```

```

VAR
  IWK : VECT;
  KOLUM1,MN,I,IC,NEGI,J1,ICM1,IR,IRP1,IRN : INTEGER;
  AUX,AUX1,IRPO : INTEGER;

```

```

LABEL 40,50,60;

```

```

BEGIN
  KOLUM1 := KOL-1;
  INICIA(MROW,KOL,R,S);
  MN := MROW - NEG;
  FOR I := 1 TO MN DO
    BEGIN
      NEGI := NEG + I;
      S_EST[NEGI,1] := 1.0;
      R_EST[NEGI,1] := G[I];
    END;
  END;

```

```

      END; (* DEL FOR *)
FOR I := 1 TO NEG DO
  BEGIN
    NEG1      := NEG +2-I;
    S_EST[I,1] := 1.0;
    R_EST[I,1] := G[NEG1];
  END;(* FIN DEL FOR *)
IF ( IRHO = 0 )
  THEN
    GOTO 40
  ELSE
    FOR I:= 1 TO MROW DO
      BEGIN
        NEG1 := I-NEG-1;
        IF ( NEG1 MOD 2 ) <> 0
          THEN
            R_EST[I,1] := (-1) * R_EST[I,1]
          ELSE
            R_EST[I,1] := R_EST[I,1];
      END; (* DEL FOR *)
40:  J1      := MROW;
    FOR IC := 2 TO KOL DO
      BEGIN
        ICM1 := IC-1;
        J1    := J1-1;
        AUX   := IC-1;
        FOR IR := AUX TO J1 DO
          BEGIN
            IRP1 := IR +1;
            IRPO := IR -1;
            IF (ABS(R_EST[IR,ICM1]) = 0.0 )
              THEN
                GOTO 50
              ELSE
                BEGIN
                  S_EST[IR,IC] := S_EST[IR,ICM1]*((R_EST[IRP1,ICM1]/R_EST[IR,
                    GOTO 60;
                END;
50:  IF ((IC+1) MOD 2 ) <> 0
        THEN
          S_EST[IR,IC] := -1*S_EST[IRP1,ICM1]
        ELSE
          S_EST[IR,IC] := S_EST[IRP1,ICM1];
60:  END; (* DEL FOR *)
      AUX1 := IC;
      FOR IR := AUX1 TO J1 DO
        BEGIN
          IRPO := IR -1;
          R_EST[IR,IC] := R_EST[IR,ICM1]*((S_EST[IR,IC]/S_EST[IRPO,IC))-1.0
        END;(* DEL FOR *)
      END; (* DEL FOR *)
END; (* DE ARRAY_S_EST *)

{*****

```

```

*****
**
** ESTE PROCEDIMIENTO CREA UNA VENTANA DE UN COLOR ESPECIFICO
**
*****

```

```

PROCEDURE VENTANA(COLOR:INTEGER);
BEGIN
  CLRSCR;
  WINDOW(2,2,72,23);
  TEXTBACKGROUND(COLOR);
END;(* DE VENTANA *)

```

```

{*****
*****
**
** ESTE PROCEDIMIENTO CREA UNA VENTANA EN LA PARTE INFERIOR DE LA PAN-
** TALLA.
**
*****
}

```

```

PROCEDURE VENT;
BEGIN
  WINDOW(2,21,72,22);
  TEXTBACKGROUND(1);
  WRITE('PARA CONTINUAR OPRIME <RETURN> ');
END;(* DE VENT *)

```

```

{*****
*****
**
** ESTE PROCEDIMIENTO PROPORCIONA LA SALIDA EN PANTALLA DEL ESTADIS-
** TICO D.
**
*****
}

```

```

PROCEDURE IMPRI_ESTA(STAT:VECT;JK,IP,IQ:INTEGER);
VAR
  I,J,K,AUX,AUX1 : INTEGER;
BEGIN
  CLRSCR;
  WRITELN; WRITELN;
  WRITELN(' ',#186); WRITELN(' ','(',IP,',',',',IQ,') ',#186,' D(n,m)')
  WRITELN(' ',#186);
  FOR I := 1 TO 29 DO
    WRITE(#205);
  WRITELN;
  AUX := 1;
  AUX1 := IQ+1;
  K := 1;

```

```

FOR I := 1 TO IP {JK} DO
  BEGIN
    FOR J:= AUX TO AUX1 DO
      WRITELN(' ',I,',',J-K,') ',#186,STAT[J]);
      AUX := AUX1;
      AUX1 := AUX1 +IQ;
      K := IQ+1;
    END;(* DEL FOR *)
    VENT;
    READLN;
    VENTANA(4);
    CLRSCR;
  END;(* DE IMPRI_ESTA *)

{*****
**
** ESTE PROCEDIMIENTO PROPORCIONA LA SALIDA A IMPRESORA DEL ESTADIS- **
** TICO D. **
*****}

PROCEDURE IMPRI_LST(STAT:VECT;JK,IP,IQ:INTEGER);
VAR
  I,J,K,AUX,AUX1 : INTEGER;
BEGIN
  CLRSCR;
  WRITELN(LST); WRITELN(LST);
  WRITE(LST,' ',#186); WRITELN(LST,' ',(' ',IP,',',IQ,') ',#186);
  WRITELN(LST,' D(n,m)');
  WRITELN(LST,' ',#186);
  FOR I := 1 TO 29 DO
    WRITE(LST,#205);
  WRITELN;
  AUX := 1;
  AUX1 := IQ+1;
  K := 1;
  FOR I := 1 TO IP {JK} DO
    BEGIN
      FOR J:= AUX TO AUX1 DO
        WRITELN(LST,' ',I,',',J-K,') ',#186,STAT[J]);
        AUX := AUX1;
        AUX1 := AUX1 +IQ;
        K := IQ+1;
      END;(* DEL FOR *)
    END;(* DE IMPRI_LST *)

{*****
**
** ESTE PROCEDIMIENTO CALCULA EL ESTADISTICO D MEDIANTE LA RELACION **
** PROPORCIONADA EN LA DEFINICION 1.3.3 **
*****}

```

```

*****
PROCEDURE DSTAT(MXROW,MROW,NEG: INTEGER;
                R,S      : MAT;
                IP       : INTEGER;
                IQ,IPNT  : INTEGER);
VAR
  BOT,ARG,XXM      : REAL;
  JPP,JP,JQQ,IOP,I,J,ISUB1,ISUB2,ISUB3,NEGO      : INTEGER;
  JQ,IP1,IM1,JJ,NEGM,NEGP,NEGM1,NEGP1,NEG1,IMX,II : INTEGER;
  LABEL            20,35,50;
BEGIN
  JPP := IP + 1;
  JQQ := IQ + 1;
  IOP := JPP*JQQ;
  FOR I:= 1 TO JPP DO
    BEGIN
      JP := I-1;
      FOR J := 1 TO JQQ DO
        BEGIN
          JQ := J-1;
          JK := JQQ*(I-1)+J;
          ISUB2 := NEG-JP+JQ+2;
          ISUB1 := ISUB2-1;
          IP1 := I+1;
          IM1 := I-1;
          BOT := SQR((S[ISUB2,I] + S[ISUB1,IP1]));
          IF ( I > 1 )
          THEN
            GOTO 20
          ELSE
            BEGIN
              FOR JJ := 1 TO 3 DO
                BEGIN
                  NEGM := NEG - JP - JQ - JJ +1;
                  NEGP := NEG-JP+JQ+JJ+1;
                  BOT := BOT + SQR(R[NEGM,I]) + SQR(R[NEGP,I]);
                END; (* DEL FOR *)
              GOTO 35;
            END;
          ISUB3 := ISUB2+1;
          BOT := BOT/SQR(S[ISUB3,IM1] + S[ISUB2,I]);
          FOR JJ := 1 TO 3 DO
            BEGIN
              NEGM := NEG-JP-JQ-JJ+1;
              NEGP := NEG-JP+JQ+JJ+1;
              NEGM1 := NEGM + 1;
              NEGP1 := NEGP + 1;
              BOT := BOT + SQR(R[NEGM,I]/R[NEGM1,IM1]) + SQR(R[NEGP,I]/R[NEGP1,IM1]);
            END; (* DEL FOR *)
          35:  NEGO := NEG-JP-JQ;
              NEG1 := NEGO + 1;
              ARG := S[NEGO,IP1] / (BOT*S[NEG1,I]);
        END;
      END;
    END;
  END;

```

```

        STAT[JK] := ABS(ARG);
    END; (* DEL FOR *)
END; (* DEL FOR *)
IMPRI_ESTA(STAT,JK,IP,IQ);
IMX := 1;
XMX := STAT[1];
FOR II := 1 TO IOP DO
    BEGIN
        IF (STAT[II] < XMX )
            THEN
                GOTO 50
            ELSE
                BEGIN
                    IMX := II;
                    XMX := STAT[II];
                END;
    END;
50:  END; (* DEL FOR *)
    IP := ROUND(INT((IMX-1)/JQQ));
    IQ := IMX-(IP*JQQ)-1;
END; (* DE DSTAT *)

(.....)
**
** ESTE PROCEDIMIENTO REALIZA LA LECTURA DE UN ARCHIVO CON EXTENSION
** DAT, EL CUAL CONTIENE LA INFORMACION DE UN PROCESO.
**
(.....)

PROCEDURE LEE_ARCH(NOMBRE:QUINCE;VAR DATOS:VECT);
VAR
    I          : INTEGER;
    NUMERO     : REAL;
    ARCH       : TEXT;
BEGIN
    ASSIGN(ARCH,NOMBRE+'.DAT');
    RESET(ARCH); { HABRO EL ARCHIVO PARA LECTURA }
    I := 1;
    WHILE NOT EOF(ARCH) DO
        BEGIN
            READLN(ARCH,NUMERO);
            DATOS[I] := NUMERO;
            I := I+1;
        END; (* DEL WHILE *)
    CLOSE(ARCH);
END; (* DE LEE_ARCH *)

(.....)
**
** ESTE PROCEDIMIENTO PIDE INFORMACION AL USUARIO SOBRE LOS ARREGLOS
** R Y S A SER CALCULADOS.
**
(.....)

```

```

*****
PROCEDURE LECTURA(VAR MROW,NEG,KOL,IP,IQ: INTEGER;CAR:CHAR);
BEGIN
  VENTANA(4);
  GOTOXY(24,4);
  WRITE('LECTURAS ');
  GOTOXY(4,7);
  WRITE('DAME LA DIMENSION DE LOS ARREGLOS "R" Y "S" :');
  READLN(MXROW);
  GOTOXY(4,9);
  WRITE('NUMERO DE RENGLONES DE LOS ARREGLOS "R" Y "S"');
  GOTOXY(4,10);
  WRITE('A SER PROCESADOS. ');
  READLN(MROW);
  GOTOXY(4,12);
  WRITE('NUMERO DE RENGLONES CON RETRASO NEGATIVO. ');
  READLN(NEG);
  IF CAR <> 'D'
  THEN
    BEGIN
      GOTOXY(4,14);
      WRITE('NUMERO DE COLUMNAS DE LOS ARREGLOS "R" Y "S" ');
      READLN(KOL);
    END
  ELSE
    BEGIN
      GOTOXY(4,14);
      WRITE('ORDEN MAXIMO DE LA AUTORREGRESION ');
      READLN(IP);
      GOTOXY(4,16);
      WRITE('ORDEN MAXIMO DE PROMEDIOS MOVILES ');
      READLN(IQ);
    END;
  CLRSCR;
  TEXTBACKGROUND(1);
END; (* DE LECTURAS *)

{*****
*****
**
** ESTE PROCEDIMIENTO PROPORCIONA LA SALIDA A PANTALLA DE LOS ARREGLOS **
** R Y S SEGUN DESEE EL USUARIO. **
**
*****}

PROCEDURE IMPRIME(CAR:CHAR; R: MAT;NEG,MROW:INTEGER);
VAR
  I,J : INTEGER;
BEGIN
  CLRSCR;
  GOTOXY(19,4);
  IF CAR <> '*'
  THEN

```

```

        WRITELN('ARREGLO ',CAR)
    ELSE
        WRITELN('ARREGLO S_EST');
    WRITELN;
    FOR I := 1 TO MROW {100} DO
        BEGIN
            IF ((-1*NEG+I-1) < 0 ) AND (ABS(-1*NEG+I-1) < 10 )
            THEN
                WRITE(-1*NEG+I-1,' ',#186,' ')
            ELSE
                IF ((-1*NEG+I-1) < 0 )
                THEN
                    WRITE(-1*NEG+I-1,' ',#186,' ')
                ELSE
                    WRITE(' ',-1*NEG+I-1,' ',#186,' ');
            IF ( I MOD 20 ) = 0
            THEN
                CLRSCR;
            FOR J := 1 TO KOL {50} DO
                WRITE(R[I,J]:2,' ');
            WRITELN;
        END;
    VENT;
    READLN;
    VENTANA(4);
END; (* DE IMPRIME *)

```

```

{*****
*****
**
** ESTE PROCEDIMIENTO PROPORCIONA LA SALIDA A IMPRESORA DE LOS ARRE-
** GLOS R Y S SEGUN DESEE EL USUARIO.
**
*****
*****}

```

```

PROCEDURE IMP_R_S(CAR:CHAR; R: MAT;NEG,MROW:INTEGER);
VAR
    I,J      : INTEGER;
BEGIN
    CLRSCR;
    GOTOXY(19,4);
    IF CAR <> '*'
    THEN
        WRITELN(LST,'                                ARREGLO ',CAR)
    ELSE
        WRITELN(LST,'                                ARREGLO S_EST');
    WRITELN(LST);
    FOR I := 1 TO MROW {100} DO
        BEGIN
            IF ((-1*NEG+I-1) < 0 ) AND (ABS(-1*NEG+I-1) < 10 )
            THEN
                WRITE(LST,-1*NEG+I-1,' ',#186,' ')
            ELSE
                IF ((-1*NEG+I-1) < 0 )

```

```

        THEN
            WRITE(LST,-1*NEG+I-1,' ',#186,' ')
        ELSE
            WRITE(LST,' ',-1*NEG+I-1,' ',#186,' ');
        FOR J := 1 TO KOL {50} DO
            WRITE(LST,R[I,J]:2,' ');
        WRITELN(LST);
    END;
END; (* DE IMP_R_S *)

(*****
**
** ESTE PROCEDIMIENTO PIDE AL USUARIO LA AUTOCORRELACION UTILIZADA
** PARA LA SIMULACION.
**
*****),

PROCEDURE SELEC_FACM;
BEGIN
    CLRSCR;
    {WINDOW(35,9,72,15);}
    TEXTBACKGROUND(6);
    GOTOXY(12,7);
    WRITE('0.- CALCULO DE "R" Y "S" A PARTIR DE r(m) ');
    GOTOXY(12,10);
    WRITE('1.- CALCULO DE "R" Y "S" A PARTIR DE (-1)**m r(m) ');
    GOTOXY(12,15);
    WRITE('QUE AUTOCORRELACION QUIERES: ');
    READLN(IRHO);
    CLRSCR;
    TEXTBACKGROUND(1);
END; (* DE SELEC_FACM *)

(*****
**
** ESTE PROCEDIMIENTO PROPORCIONA LA VENTANA PARA PEDIR EL NOMBRE DEL
** ARCHIVO CON EL QUE SE REALIZARA LA SIMULACION.
**
*****),

PROCEDURE VENT_DATOS(VAR DATOS:VECT);
VAR
    NOMBRE : QUINCE;
BEGIN
    CLRSCR;
    TEXTBACKGROUND(6);
    GOTOXY(2,2);
    WRITE('NOMBRE DEL ARCHIVO: ');
    READLN(NOMBRE);
    LEE_ARCH(NOMBRE,DATOS);
    CLRSCR;

```

```
TEXTBACKGROUND(1);
END; (* DE VENT_DATOS *)
```

```
{*****
*****
**
** ESTE PROCEDIMIENTO CALCULA LA AUTOCORRELACION A PARTIR DE LA INFOR- **
** CION ENCONTRADA EN EL ARCHIVO DE TRABAJO. **
**
*****
*****}
```

```
PROCEDURE AUTOCORRELACION(DATOS:VECT;VAR G:VECT;N,RETRASO:INTEGER);
```

```
VAR
  K,T : INTEGER;
  NUM,RHOK,ZBAR,DEN : REAL;
BEGIN
  ZBAR := 0;
  FOR T:=1 TO N DO { CALCULA LA MEDIA }
    ZBAR := ZBAR + DATOS[T];
  ZBAR := ZBAR/N;
  FOR K:=0 TO RETRASO DO
    BEGIN
      NUM := 0;
      FOR T := 1 TO N-K DO
        NUM := NUM + (DATOS[T]-ZBAR)*(DATOS[T+K]-ZBAR);
      DEN:=0.0;
      FOR T:=1 TO N DO
        DEN := DEN+(DATOS[T]-ZBAR)*(DATOS[T]-ZBAR);
      G[K+1] := NUM/DEN;
    END;
  END;
END;(* DE AUTOCORRELACION *)
```

```
{*****
*****
**
** ESTE PROCEDIMIENTO CREA UNA VENTANA Y PIDE INFORMACION AL USUARIO **
** ACERCA DE LA FACH. **
**
*****
*****}
```

```
PROCEDURE VENT_FACH(DATOS:VECT);
```

```
VAR
  RETRASO,N : INTEGER;
BEGIN
  CLRSCR;
  TEXTBACKGROUND(5);
  GOTOXY(4,2);
  WRITE('DAME EL TAMANO DE LA MUESTRA : ');
  READLN(N);
  GOTOXY(4,4);
  WRITE('DAME EL NUMERO DE RETRASOS: ');
  READLN(RETRASO);
```

```

AUTOCORRELACION(DATOS,G,N,RETRASO);
CLRSCR;
TEXTBACKGROUND(1);
END; (* DE VENT_FACH *)

```

```

(*****
*****
**
** ESTE PROCEDIMIENTO PROPORCIONA AL USUARIO EL MENU DE LOS ARREGLOS
** R, S Y SMODIFICADO.
**
*****
*****
)

```

```

PROCEDURE MENU1;
VAR
  DATOS : VECT;
  OP : INTEGER;
BEGIN
  VENTANA(1);
  TEXTCOLOR(9);
  REPEAT
    GOTOXY(19,1);
    WRITE(' MENU ARREGLOS "R" , "S" Y "SM" ');
    GOTOXY(22,3);
    WRITE('1.- LECTURA DE DATOS ');
    GOTOXY(22,5);
    WRITE('2.- CALCULO DE LA FACH ');
    GOTOXY(22,7);
    WRITE('3.- SELECCION DE LA FACH ');
    GOTOXY(22,9);
    WRITE('4.- LECTURA ');
    GOTOXY(22,11);
    WRITE('5.- CALCULO DE R Y S ');
    GOTOXY(22,13);
    WRITE('6.- CALCULO DE SM ');
    GOTOXY(22,15);
    WRITE('7.- SALIDA A IMPRESORA DE R Y S ');
    GOTOXY(22,17);
    WRITE('8.- SALIDA A IMPRESORA DE SM');
    GOTOXY(22,19);
    WRITE('9.- SALIDA ');
    GOTOXY(15,21);
    WRITE('QUE OPCION QUIERES: ');
    READLN(OP);
    CASE OP OF
      1: VENT_DATOS(DATOS);
      2: VENT_FACH(DATOS);
      3: SELEC_FACH;
      4: LECTURA(MROW,NEG,KOL,IP,IG,'R'); { LEE MROW,NEG,KOL,IRHO,G ES VECTOR
      5: BEGIN
          ARRAY_R_AND_S(MXROW,MROW,NEG,KOL,IRHO,G,R,S,IRS);
          CAR := 'R';
          IMPRIME(CAR,R,NEG,MROW);

```

```

CAR := 'S';
IMPRIME(CAR,S,NEG,MROW);
CLRSCR;
TEXTBACKGROUND(1);
END;
6: BEGIN
  ARRAY_S_EST(MXROW,MROW,NEG,KOL,IRHO,G,R_EST,S_EST,IRS);
  CAR := 'S';
  IMPRIME(CAR,S_EST,NEG,MROW);
  CLRSCR;
  TEXTBACKGROUND(1);
END;
7: BEGIN
  CAR := 'R';
  IMP_R_S(CAR,R,NEG,MROW);
  GOTOXY(14,14);
  WRITE('FAVOR DE COLOCAR OTRA HOJA. ');
  GOTOXY(4,21);
  WRITE('PARA CONTINUAR OPRIMA <RETURN> ');
  READLN;
  CAR := 'S';
  IMP_R_S(CAR,S,NEG,MROW);
END;
8: BEGIN
  CAR := 'S';
  IMP_R_S(CAR,S_EST,NEG,MROW);
END;
9: CLRSCR;
END; (* DEL CASE *)
UNTIL (OP = 9)
END; (* DE MENU1 *)

{*****
*****
**
** ESTE PROCEDIMIENTO PROPORCIONA AL USUARIO EL MENU DEL ESTADISTICO D **
**
*****}

PROCEDURE MENU2;
VAR
  OP : INTEGER;
BEGIN
  VENTANA(1);
  TEXTCOLOR(9);
  REPEAT
    GOTOXY(19,3);
    WRITE(' MENU ESTADISTICO "D" ');
    GOTOXY(22,8);
    WRITE('1.- LECTURA ');
    GOTOXY(22,11);
    WRITE('2.- CALCULO DEL ESTADISTICO "D" ');
    GOTOXY(22,14);

```

```

WRITE('3.- SALIDA A IMPRESORA D(n,m) ');
GOTOXY(22,17);
WRITE('4.- SALIDA ');
GOTOXY(15,21);
WRITE('QUE OPCION QUIERES: ');
READLN(OP);
CASE OP OF
  1: LECTURA(MROW,NEG,KOL,IP,IQ,'D');
  2: DSTAT(MXROW,MROW,NEG,R,S,IP,IQ,IPNT);
  3: IMPRI_LST(STAT,JK,IP,IQ);
  4: CLRSCR;
END; (* DEL CASE *)
UNTIL ( OP = 4 )
END; (* DE MENU2 *)

```

```

{*****
*****
** ESTE PROCEDIMIENTO PROPORCIONA AL USUARIO EL MENU PRINCIPAL. **
**
*****
*****}

```

```

PROCEDURE MENU;
VAR
  OP : INTEGER;
BEGIN
  VENTANA(1);
  TEXTCOLOR(9);
  REPEAT
    GOTOXY(25,6);
    WRITE(' MENU PRINCIPAL');
    GOTOXY(20,9);
    WRITE('1.- ARREGLOS "R" , "S" Y "SM" ');
    GOTOXY(20,11);
    WRITE('2.- ESTADISTICO "D" ');
    GOTOXY(20,13);
    WRITE('3.- SALIDA ');
    GOTOXY(15,17);
    WRITE('QUE OPCION QUIERES: ');
    READLN(OP);
    CASE OP OF
      1: MENU1;
      2: MENU2;
    END;(* DEL CASE *)
  UNTIL (OP = 3 )
END; (* DE MENU *)

BEGIN (* DEL PROGRAMA PRINCIPAL *)
  MENU;
END.

```

APÉNDICE 3

MATRICES INFINITAS DE HENKEL DE RANGO FINITO

Sea $S_0, S_1, S_2, \dots$ una sucesión de números complejos. que determina una matriz simétrica infinita

$$S = \begin{bmatrix} S_0 & S_1 & S_2 & \dots \\ S_1 & S_2 & S_3 & \dots \\ S_2 & S_3 & S_4 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

a la cual usualmente se le llama MATRIZ DE HENKEL. Junto con la matriz infinita de Henkel, se considera también a la matriz finita de Henkel $S_n = || S_{i+k} ||_0^{n-1}$ y su forma de Henkel asociada (ver Gantmacher 1960, pag 205).

$$S_n(x, x) = \sum_{i,k=0}^{n-1} S_{i+k} x_i x_k$$

Los menores principales de S serán denotados por $D_1, D_2, \dots$

en donde $D_p = |S_{i+k}|_0^{p-1}$ para $p = 1, 2, \dots$

Las matrices infinitas pueden ser de rango finito o infinito. En el último caso, las matrices tienen menores diferentes de cero de orden arbitrariamente grande. El siguiente teorema da una condición necesaria y suficiente para que una sucesión de números $S_0, S_1, \dots$ genere una matriz infinita de Henkel $S = ||S_{i+k}||_0^\infty$ de rango finito.

Teorema A3.1 La matriz infinita $S = ||S_{i+k}||_0^\infty$ es de rango finito r si y sólo si existen r números $\alpha_1, \alpha_2, \dots, \alpha_r$ tales que

$$S_l = \sum_{i=1}^r \alpha_i S_{l-i} \quad \text{para } l = r, r+1, \dots \quad (\text{A.3.1})$$

además r es el entero más pequeño que tiene esta propiedad.

Demostración $\Rightarrow$) Si la matriz $S = ||S_{i+k}||_0^\infty$ tiene rango finito r , entonces, sus primeros $r+1$ renglones $R_1, R_2, \dots, R_{r+1}$ son linealmente dependientes. Por lo tanto, existe un número $h \leq r$ tal que $R_1, R_2, \dots, R_h$ son linealmente independientes y R_{h+1} es una combinación lineal de estos

$$R_{h+1} = \sum_{i=1}^h \alpha_i R_{h-i+1}$$

Ahora considérense los renglones $R_{j+1}, R_{j+2}, \dots, R_{j+h+1}$, en donde j es cualquier entero positivo. Obsérvese que a partir de la estructura de S , es claro que los renglones $R_{j+1}, R_{j+2}, \dots, R_{j+h+1}$ se obtienen a partir de los renglones $R_1, R_2, \dots, R_{h+1}$ mediante un proceso en el cual los elementos de las primeras j columnas son omitidos. Por lo tanto

$$R_{j+h+1} = \sum_{i=1}^h \alpha_i R_{j+h-i+1} \quad \text{para } j = 0, 1, 2, \dots$$

Así, todos los renglones de S que comienzan con $h+1$ pueden ser expresados linealmente en términos de los h renglones anteriores y por lo tanto, en términos de los primeros h renglones linealmente independientes. De aquí que el rango de S sea $r = h$. Entonces, de la dependencia lineal

$$R_{j+r+1} = \sum_{i=1}^r \alpha_i R_{j+r-i+1}$$

se sigue que existen r números $\alpha_1, \alpha_2, \dots, \alpha_r$ tales que

$$S_l = \sum_{i=1}^r \alpha_i S_{l-i} \quad \text{para } l = r, r+1, \dots$$

$\Leftarrow$) Ahora supóngase que (A.3.1) se cumple, entonces, todo renglón (columna) de S es una combinación lineal de los primeros r renglones (columnas). Por lo tanto, todos los menores de S cuyo orden sea mayor que r son iguales a cero y S es de rango a lo más r , pero el rango no puede ser menor que r ya que, como se ha demostrado, existiría una relación de la forma (A.3.1) con un valor más pequeño que r , lo cual contradice la segunda condición del teorema. Por lo tanto, el rango de S es igual a r ■

Corolario A.3.1 De la relación (A.3.1) se sigue que todos los renglones (columnas) de S son una combinación lineal de los primeros r -renglones (columnas). Por lo tanto, todo menor de S de orden r , puede ser expresado en la forma αD_r , en donde α es una constante. De donde se sigue que $D_r \neq 0$ ■

APÉNDICE 4

CONCEPTOS BASICOS

Definición A.4.1 Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias definidas sobre un espacio de probabilidad $(E, \mathcal{F}, p)$. Se dirá que X_n converge en probabilidad a la variable aleatoria X , si, para toda $\varepsilon > 0$

$$P \{ w \in E : |X_n(w) - X(w)| \geq \varepsilon \} \longrightarrow 0$$

cuando $n \longrightarrow \infty$. En este caso se escribe

$$X_n \xrightarrow{P} X.$$

Las siguientes definiciones de orden en probabilidad para sucesiones de variables aleatorias, fueron introducidas por Mann y Wald (1943). Sea $\{X_n\}$ una sucesión de variables aleatorias y $\{g_n\}$ una sucesión de números reales positivos.

Definición A.4.2 Se dirá que $\{X_n\}$ es de orden menor que $\{g_n\}$ en probabilidad y se denota como

$$X_n = o_p(g_n)$$

si

$$X_n/g_n \xrightarrow{P} 0$$

Definición A.4.3 Se dirá que $\{X_n\}$ es a lo más de orden $\{g_n\}$ en probabilidad y se denota como

$$X_n = O_p(g_n)$$

si, para toda $\epsilon > 0$, existe un número real positivo M_ϵ tal que

$$P \{ |X_n| \geq M_\epsilon g_n \} \leq \epsilon \quad \text{para toda } n$$

Si $X_n = O_p(g_n)$, algunas veces se dice que X_n es acotada en probabilidad por g_n .

Una vez mencionadas las definiciones anteriores, es fácil demostrar el siguiente lema

Lema A.4.1 Si $\{X_n\}$ es una sucesión de variables aleatorias y C es una constante tal que

$$X_n = C + O_p(n^{-\delta}) \quad \text{para } \delta > 0$$

entonces

$$X_n \xrightarrow{p} C$$

Demostración Se tiene que

$$X_n - C = O_p(n^{-\delta})$$

esto es

$$P [|X_n - C| \geq n^{-\delta} M_\epsilon] \leq \epsilon \quad \text{para toda } n$$

Ahora sea ζ dado (fijo) tal que

$$1/n^\delta < \zeta \quad \text{para toda } n \geq N$$

y sea $\varepsilon > 0$ dado tal que

$$M_\varepsilon / n^\delta < \zeta \quad \text{para toda } n \geq N_1$$

por lo tanto

$$[|X_n - C| \geq \zeta] \subseteq [|X_n - C| \geq M_\varepsilon / n^\delta] \quad \text{para toda } n \geq N_1$$

entonces

$$p[|X_n - C| \geq \zeta] \leq p[|X_n - C| \geq M_\varepsilon / n^\delta] \quad \text{para toda } n \geq N_1$$

lo cual implica que

$$p[|X_n - C| \geq \zeta] \xrightarrow{p} 0 \quad ; \quad n \longrightarrow \infty \quad \text{para toda } \zeta > 0$$

esto es

$$X_n \xrightarrow{p} C \quad \blacksquare$$

CONCLUSIONES

El principal objetivo de este trabajo ha sido presentar un estudio detallado de algunos métodos que actualmente se usan para determinar el orden de un modelo de series de tiempo ARMA. En este trabajo se puso bajo un mismo contexto de notación los métodos desarrollados en el capítulo 1; secciones 1.2, 1.3, 1.4, 1.5 y 1.6, facilitándose con esto una contrastación entre dichos métodos, misma que es manifestada en las aplicaciones del capítulo 4.

Las conclusiones a las que se llegó con este desarrollo son:

i) **Método de Box-Jenkins:** Se aplica a series de tiempo estacionarias y no estacionarias así como a series de tiempo estacionales, mostrando en teoría unicidad en la determinación del orden para modelos AR y MA, no así para modelos mixtos ARMA. Requiere de familiaridad con el reconocimiento de patrones de la FAC y de la FACP.

ii) **Arreglos R y S:** A diferencia del Método de Box-Jenkins estos arreglos no se aplican a series de tiempo estacionales; presentan un criterio para determinar si una serie es estacionaria o no (teorema 1.3.8), pues la no estacionariedad tiene que ser removida antes de la identificación final del modelo. Por otro lado, estos arreglos muestran en teoría, unicidad en la determinación del orden de los modelos AR, MA y ARMA. Más aún, se presenta un estadístico que condensa la información de los arreglos R y S, el cual proporciona una determinación automática del orden de un modelo ARMA en el caso en que la serie en cuestión sea estacionaria. Sin embargo, las propiedades distribucionales de dicho estadístico no están completamente determinadas.

iii) **Función de Autocorrelación Parcial y de Autocorrelación Parcial Generalizada:** Igual que los arreglos R y S , la FACPG y el arreglo S^* sólo se aplica a series de tiempo no estacionales. Dado que el arreglo S^* no es más que una ligera modificación del arreglo S , que facilita el uso práctico de éste, el arreglo S^* proporciona el mismo criterio que los arreglos R y S para determinar si una serie de tiempo es estacionaria o no, pues la no estacionariedad tiene que ser removida antes de la identificación final del modelo. Por otro lado, los arreglos de la FACPG y S^* muestran en teoría unicidad en la determinación del orden de los modelos AR, MA y ARMA, además dada la relación que existe entre el arreglo S^* y la FACPG (teorema 1.4.2), se puede observar que esta última, igual que el estadístico D de los arreglos R y S es una condensación de la información del arreglo S^* , que al ser usada junto con el arreglo S^* facilita la determinación del orden del modelo.

iv) **El Método de la Esquina:** Este método presenta algunas desventajas con respecto a los anteriores, pues mientras dichos métodos presentan un criterio para identificar cuando una serie es no estacionaria, el método de la esquina no presenta ninguna caracterización de este tipo, por lo que para usarlo habrá que asegurarse de que la serie en cuestión sea estacionaria. Más aún la prueba de hipótesis que presenta este método enfrenta una serie de restricciones que hacen imposible su uso práctico (subsección 1.5.4). Por otro lado, se demostró en la subsección 1.5.5 que el Método de la Esquina puede ser considerado como un caso particular del arreglo R .

v) El Método de la FAME: A diferencia de los métodos anteriores, el Método de la FAME se aplica directamente a series de tiempo estacionarias y no estacionarias, evitándose con esto una transformación de los datos de la serie para volverla estacionaria antes de especificar el orden del modelo. Una desventaja que presenta este método con respecto a los últimos tres mencionados, es el problema de la falta de unicidad, pues la teoría de la FAME sólo determina el comportamiento asintótico para $j \geq q$ y $m \geq p$, pero no impide la existencia de puntos de corte con $n < p$. Sin embargo, la discriminación entre varios puntos de corte posibles, puede lograrse mediante la etapa de estimación de los parámetros autorregresivos y las propiedades de convergencia que presentan (teorema 1.6.2). Este método no se aplica a series de tiempo estacionales.

De las conclusiones antes expuestas, es claro percatarse que las líneas de investigación a seguir en estos métodos son:

- a) Extender la aplicación de los últimos 4 métodos a series de tiempo estacionales.
- b) Determinar propiedades distribucionales del estadístico D
- c) Modificar o en su defecto proporcionar otra prueba de hipótesis para el Método de la Esquina, de tal forma que se le de un uso práctico.
- d) Analizar la robustez de estos métodos ante la presencia de outliers.

BIBLIOGRAPHY

- Abraham, B. & Ledolter, J. (1984) "A note on inverse autocorrelations". *Biometrika* 71, 609-14.
- Akaike, H. (1970a). "Statistical predictor identification". *Ann. Inst. Statist. Math.* 22, 203-217.
- Akaike, H. (1970b). "A fundamental relation between predictor identification and power spectrum estimation". *Ann. Inst. Statist. Math.* 22, 219-223.
- Akaike, H. (1971). "Autoregressive model fitting for control". *Ann. Inst. Statist. Math.* 23, 163-180.
- Akaike, H. (1974). "A new look at the statistical model identification". *IEEE Trans. Autom. Control.* 19, 716-723.
- Akaike, H. (1977). "On entropy maximization principle". In *Applications of Statistical*, Ed. P. R. Krishnaiah, pp. 27-41. Amsterdam North Holland.
- Akaike, H. (1978a). "A bayesian analysis of the minimum AIC procedure". *Ann. Inst. Statist. Math.* 30, 9-14.
- Akaike, H. (1978b). "On the likelihood of a time series model". *Statistician* 27, 217-235.
- Akaike, H. (1979). "A Bayesian extension of the minimum AIC procedure". *Biometrika* 66, 237-242.
- Akaike, H. (1981). "Likelihood of a model and information criteria". *J. Econometrics* 16, 3-14.
- Allen, D.M. (1971). "Mean square error of prediction as a criterion for selecting variables". *Technometrics* 13, 469-475.
- Atkinson, A. C. (1980). "A note on the generalized information criterion of choice of a model". *Biometrika* 67, 413-418.

- Atkinson, A. C. (1981). "Likelihood ratios, posterior odds and information criteria". *J. Econometrics* 16, 15-20.
- Bartlett, M. S. (1946). "On the theoretical specification and sampling properties of autocorrelated time-series". *J. R. Statist. Soc. B* 8, 27-41.
- Beamish, N. & Priestley, M. B. (1981). "A study of autoregressive and window spectral estimation". *Appl. Statist.* 30, 41-58.
- Beguin, J. M. Gouriéroux, C. & Monfort, A. (1980). "Identification of a mixed autoregressive-moving average process: the cornel method". In *Time Series*, Ed. O. D. Anderson, pp. 423-436. Amsterdam: North-Holland.
- Bessler, D. A., & Binkley, J. K. (1980). "Autoregressive filtering of some economic data using PRESS and FPE. In Proc. Bus". *Econ. Statist. Sect., Am. Statist. Assoc.*, pp 261-265.
- Bhansali, R. J. (1973). "A Monte Carlo comparison of the regression method and the spectral methods of prediction". *J. Am. Statist. Assoc.* 68, 621-625.
- Bhansali, R. J. (1980). "The autoregressive and the window estimate of the inverse correlation function". *Biometrika* 67, 511-566.
- Bhansali, R. J. (1983a). "The inverse partial correlation function of a time series and its application". *J. Mult. Anal.* 13, 310-327.
- Bhansali, R. J. (1983b). "Estimation of the order of a moving average model from autoregressive and window estimates of the inverse correlation function". *J. Time Series Anal.* 4, 137-162.
- Bhansali, R. J. & Downham, D. Y. (1977). "Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE criterion". *Biometrika* 64, 547-551.

- Box, G. E. P. & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and control*. San Francisco :Holden Day.
- Caines, P.E., Keng, C. W. & Sethi, S.P. (1981). "Causality analysis and multivariate autoregressive modelling with an application to supermarket sales analysis". *J. Econ. Dynamics Control* 3, 267-298.
- Chatfield, C. (1979). "Inverse autocorrelations". *J. R. Statist Soc. A* 142, 363-377.
- Chow, G. C. (1981). "A comparison of the information and posterior probability criteria for model selection". *J. Econometrics* 16, 21-33.
- Cliveland, W. S. (1972). "The inverse autocorrelations of a time series and their applications". *Technometrics* 14, 277-298.
- Fuller, W. A. (1976). *Introduction to Statistical Time Series*. John Wiley & Sons.
- Gantmacher, F.R. (1976). *Theorie des matrices*. Dunod, vol. 2.
- Gersch, W. & Sharpe, D. R. (1973). "Estimation of power spectra with finite-order autoregressive models". *IEEE Trans. Autom. Control* 18, 367-369.
- Gooljer, J. G. de & Heuts, R. M. J. (1981). "'The corner method: and investigation of an order determination procedure for general ARMA processes". *J. Oper. Res. Soc.* 32, 1039-1042.
- Gooljer, J. G. de , Abraham, B., Gould, A. & Robinson, L. (1985). " Methods for Determining the order of an Autoregressive-Moving Average Process: A Survey". *International Statistical Review.* 53, 3, pp. 301-329.
- Gray, H. L., Kelley, G. D. & McIntire, D. D. (1978). "A new approach to ARMA modelling". *Comm. Statist. B* 7, 1-77.

Gray, H. L., Houston, T. A., & Morgan, F. W. (1978). "On G-Spectral estimation. Proceedings of the 1976 Tulsa Symposium on Applied Time Series". New York: Academic Press.

Guerrero, G. V. M. (1983). *Analisis Estadístico de series de tiempo*. Libro por publicar.

Hannan, E. J. (1980). "The estimation of the order of an ARMA process". *Ann. Statist.* 8, 1071-1081.

Hannan, E. J. (1982b). "Estimating the dimension of a linear System. *J. Mult. Anal.* 11. 459-473.

Hannan, E. J. & Rissanen, J. (1982). "Recursive estimation of mixed autoregressive-moving average order". *Biometrika* 69, 81-94. Correction (1983), 70,303.

Hipel, K. W., McLeod, A. I. & Lennox, W. C. (1977). "Advances in Box-Jenkins modelling 1. Model construction". *Water Resources Res.* 13, 567-575.

Hjorth, U. & Holmqvist, L. (1981). "A model selection based on validation with applications to pressure and temperature prognosis". *Appl. Statist.* 30, 264-274.

Hjorth, U. (1982). "Model selection and forward validation". *Scand. J. Statist.* 9, 95-105.

Jeffreys, H. (1961). *Theory of probability*, 3rd ed. Oxford: Clarendon press.

Jeffreys, H. (1967). *Theory of probability*, 4th ed. Oxford University Press.

Jones, R. H. (1974). "Identification and autoregressive spectrum estimation". *IEEE. Trans. Autom. Control* 19, 894-897.

Jones, R. H. ((1975). "Fitting autoregression". *J. Am. Statist. Assoc.* 70, 590-592.

Kitagawa, G. (1977). "On a search procedure for the optimal AR-MA order". *Ann. Inst. Statist. Math.* 29, 319-332.

Kitagawa, G. (1981). "A nonstationary time series model and fitting by a recursive filter". *J. Time series Anal.* 2, 103-116.

Kitagawa, G. & Akaike, H. (1978). "A procedure for the modelling of nonstationary time series". *Ann. Inst. Statist. Math.* 30, 351-363.

Kozin, F. & Nakajima, F. (1980). "The order determination problem for linear time varying AR models". *IEEE Trans. Autom. Control.* 25, 250-257.

Kromer, R. E. (1969). "Asymptotic properties of the autoregressive spectral estimator". *Ph. D. Thesis, Stanford University.*

Kullback, S. (1959). "Information Theory and Statistics". *New York: Wiley.*

McClave, J. T. (1975). "Subset autoregression". *Technometrics* 17, 213-220.

McClave, J. T. (1978). "Estimating the order of autoregressive models: The $\max x^2$ method". *J. Am. Statist. Assoc.* 73, 122-128.

Neftci, S. N. (1982). "Specification of economic time series models using Akaike's criterion". *J. Am. Statist. Assoc.* 77, 537-540.

Ozaki, T. (1977). "On the order determination of ARIMA models". *Appl. Statist.* 26, 290-301.

Pandit, S. M. (1973). "Data dependent systems: Modelling, Analysis and Optimal Control via Time Series". *Ph. D. Thesis.*

Pandit, S. M. & Wu, S. M. (1983). "Time Series and System Analysis with Applications". Wiley.

Parzen, E. (1974). "Some recent advances in time series modelling". *IEEE Trans. Autom. Control* 19, 723-730.

Parzen, E. (1975). "Some solutions to the time series modelling and prediction problem". In *The Search for Oil*, Ed. D. B. Owen, pp. 1-16. New York: Marcel Dekker.

Parzen, E. (1977). "Multiple time series: Determining the order of approximating autoregressive schemes. In *Multivariate Analysis-IV*. Ed. P. R. Krishnaiah, pp. 283-295. Amsterdam: North-Holland.

Parzen, E. (1978). "Comments on 'Analysis and modelling of seasonal time series' by George E. P. Box, Steven C. Hillmer, and George C. Tiao". In *Seasonal Analysis of Econometric Time Series*, Ed. A. Zellner, pp. 338-341.

Parzen, E. (1979). "Forecasting and whitening filter estimation". In *Forecasting Studies in Management Sciences*, Ed. S. Makridakis and S. C. Wheelwright, pp. 149-165. Amsterdam: North-Holland.

Parzen, E. (1980). "Time series modelling, spectral analysis, and forecasting". In *Directions in Time Series*, Ed. D. R. Brillinger and G. C. Tiao, pp. 80-111. Hayward: California, Institute of mathematical Statistics.

Parzen, E. & Pagano, M. (1979). "An approach to modelling seasonally stationary time series". *J. Econometrics* 9, 137-153.

Rissanen, J. (1978). "Modelling by shortest data description". *Automatica* 14, 465-471.

Schmidt, P. (1974). "Choosing among alternative linear regression models". *Atl. Econ. J.* 1, 7-13.

Schwarz, G. (1978). "Estimating the dimension of a model". *Ann. Statist.* 6, 461-464.

Shibata, R. (1976). "Selection of the order of an autoregressive model by Akaike's information criterion". *Biometrika* 63, 117-126.

Soderstrom, T. (1977). "On model structure testing in systems identification". In *J. Control* 26, 1-18.

Stone, M. (1978). "Cross-validation: A review". *Math. Oper. Statist.* 9, 127-139.

Stone, M. (1979). "Comments on model selection criteria of Akaike and Schwarz". *J. R. Statist. Soc. B* 41, 267-278.

Tiao, G. C. & Tsay, R. S. (1983a). "Consistency properties of least squares estimates of autorregressive parameters in ARMA models". *Ann. Statist.* 11, 856-871.

Tong, H. (1977). "Some comments on the Canadian lynx data". *J. R. Statist. Soc. A* 140, 432-436.

Tsay, R. S. & Tiao, G. C. (1984). "Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models". *J. Am. Statist. Assoc.* 79, 84-96.

Woodward, W. A. & Gray, H. L. (1978). "New ARMA models for Wolfer's sunspot data". *Comm. Statist. B* 7, 97-115.

Woodward, W. A. & Gray, H. L. (1981). "On the relationship between the S array and the Box-Jenkins method of ARMA model identification". *J. Am. Statist. Assoc.* 76, 579-587.

Zellner, A. (1971). "An Introduction to Bayesian Inference in Econometrics". New York: Wiley.