

⑦ Sujent.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO  
FACULTAD DE CIENCIAS



INTRODUCCION AL ANALISIS DE SERIES DE TIEMPO  
MEDIANTE EL METODO DE BOX & JENKINS

T E S I S

QUE PARA OBTENER EL TITULO DE:  
A C T U A R I O  
P R E S E N T A:

HUGO ARRAZOLA LOPEZ



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## INDICE

	Pág.
PROLOGO -----	i
INTRUDUCCION -----	ii
CAPITULO	
I.- ESPACIO DE HILBERT -----	1
Espacios con Producto Interno -----	4
Espacios Lineales Normados -----	6
Espacio de Hilbert -----	8
II.- CONCEPTOS FUNDAMENTALES EN ANALISIS DE SERIES DE TIEMPO -----	11
Objetivos del Análisis de Series de Tiempo -	13
El Concepto de Estacionario -----	14
Autocorrelación -----	18
Concepto de No Estacionario -----	20
Modelos para Series de Tiempo Estacionarias	21
Procesos de Medias Móviles -----	25
Proceso de Medias Móviles de Primer Orden --	26
Proceso de Medias Móviles de Alto Orden ----	28
Procesos Autorregresivos -----	29
Proceso Autorregresivo de Primer Orden -----	31
Proceso Autorregresivo de Segundo Orden ----	36
Proceso Autorregresivo de Alto Orden -----	39
Invertibilidad de los Procesos de Medias Móviles -----	42

	Procesos Mixtos: Autorregresivo-Medias	
	Móviles -----	43
	Proceso ARMA (1,1) -----	44
	Procesos Mixtos de Alto Orden -----	52
III.	MODELOS PARA SERIES DE TIEMPO NO ESTACIONARIAS -----	57
	Forma de la Ecuación Diferencia de los Procesos ARIMA -----	66
	Forma: Cambio Aleatorio de los Procesos ARIMA -----	67
	Forma Invertida del Proceso ARIMA -----	69
IV.	IDENTIFICACION, ESTIMACION Y PREDICCION ----	70
	Estimación de la Función de Autocorrelación	70
	Determinación del Grado Apropiado de Diferenciación -----	75
	Identificación del Orden de un Proceso Autorregresivo y la Función de Autocorrelación Parcial -----	76
	Identificación de Procesos Mixtos -----	81
	Estimación -----	84
	Predicción -----	98
	Bibliografía -----	104

## PROLOGO

De las pláticas sostenidas con el grupo de trabajo de pro babilidad de la Facultad de Ciencias, se llegó a la conclusión de que en la mayoría de los cursos impartidos de probabilidad, estadística y econometría siempre quedaba un tema de suma importancia dentro de esta área sin ser tratado y que es el de análisis de series de tiempo, esto motivó al Dr. Joaquín Cu- riel a impartir un seminario en el cual se tratará dicho pro- blema a manera de introducción para con esto motivar a los es- tudiantes a introducirse al tema; sin embargo debido a la am- plitud del tema se consideró suficiente tratar el método desa- rrollado por Box & Jenkins y que finalmente motivó el desarro- llo de este trabajo.

## INTRODUCCION

Este trabajo fué desarrollado con el objetivo principal de introducir al lector de una manera intuitiva en los conceptos fundamentales del método desarrollado por Box & Jenkins para el análisis de series de tiempo, para esto, fueron seleccionados los temas mas importantes de dicho método para ser tratados en tres capítulos básicos (II, III, IV) de tal forma que con conocimientos básicos en matemáticas y estadística, el lector podrá obtener una visión general del método. Asi mismo, el trabajo presentado de esta forma, da la opción de profundizar en un tema específico que quizá sirva para desarrollar futuras tesis.

La forma en que fueron desarrollados los temas, parte del primer capítulo en el cual se trata el concepto de espacios de Hilbert de una manera muy superficial, ya que el objetivo de este capítulo es simplemente mostrar al lector que existe una parte puramente matemática en el análisis de series de tiempo que se desarrolla en dichos espacios; esto es, se dice que una serie de tiempo por tener varianza finita pertenece a un espacio de Hilbert, es claro que el lector podrá hacer caso omiso de este capítulo sin desviarse del objetivo principal.

En el segundo capítulo se estudian los conceptos fundamentales del análisis de series de tiempo de una manera sim-

ple, de tal forma que el lector tenga las herramientas suficientes para entender el problema de series de tiempo, así mismo, dentro de este capítulo se estudian los modelos desarrollados para series de tiempo estacionarias. En el capítulo tercero se trata el problema de no estacionalidad y los modelos adecuados para su estudio para finalmente en el cuarto, tratar el problema de identificación, estimación y predicción de dichos modelos y de esta forma cumplir con el objetivo principal de este trabajo.

CAPITULO I



## ESPACIO DE HILBERT

Partiendo de las nociones elementales de geometría analítica, podremos describir un espacio de Hilbert como un posible espacio euclidiano de dimensión infinita; por lo tanto en esta primera parte recordaremos algunas características del espacio euclidiano el cual usaremos para introducir en forma axiomática la notación de espacio de Hilbert como una generalización del espacio euclidiano.

1. Denotaremos por  $\mathbb{R}$  el campo de los números reales, por  $\mathbb{R}^2$  el plano euclidiano de dos dimensiones y  $\mathbb{R}^n$  el espacio euclidiano en  $n$  dimensiones, empezaremos recordando algunas características del espacio euclidiano  $\mathbb{R}^2$ , partiendo de su definición

$$\mathbb{R}^2 = \{a = (\alpha_1, \alpha_2) / \alpha_1 \in \mathbb{R}, \alpha_2 \in \mathbb{R}\} \quad 1.1$$

si los vectores  $a = (\alpha_1, \alpha_2)$ ,  $b = (\beta_1, \beta_2)$  y el real  $\lambda$  están dados entonces  $a+b$  y  $\lambda a$  están definidos por:

$$\begin{aligned} a+b &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2) \\ \lambda a &= (\lambda \alpha_1, \lambda \alpha_2) \end{aligned} \quad 1.2$$

$\mathbb{R}^2$  es un espacio lineal,  $\alpha, \beta$  números reales y  $a, b, c$  vectores.

$$(a+b)+c = a+(b+c)$$

$$a+0 = 0+a = a$$

$$a+(-a) = (-a)+a = 0$$

$$a+b = b+a$$

$$\alpha(\beta a) = (\alpha\beta)a$$

$$1 \cdot a = a$$

$$\alpha(a+b) = \alpha a + \alpha b \quad 1.3$$

$$(\alpha+\beta)a = \alpha a + \beta a$$

si  $a = (\alpha_1, \alpha_2)$  entonces  $-a$  está definido por  $-a = (-\alpha_1, -\alpha_2)$  si  $a = (\alpha_1, \alpha_2)$  y  $b = (\beta_1, \beta_2)$  entonces la longitud de  $a$  está dada por:

$$\|a\| = \sqrt{\alpha_1^2 + \alpha_2^2} \quad 1.4$$

mientras que el producto interno está definido por

$$\langle a, b \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 = \|a\| \|b\| \frac{\alpha_1\beta_1 + \alpha_2\beta_2}{\sqrt{\alpha_1^2 + \alpha_2^2} \sqrt{\beta_1^2 + \beta_2^2}} = \|a\| \|b\| \cos \hat{a}\hat{b} \quad 1.5$$

De 1.4 y 1.5 deducimos

$$\langle a, a \rangle = \|a\|^2$$

$$|\langle a, b \rangle| \leq \|a\| \|b\| \quad (\text{Desigualdad de Cauchy}) \quad 1.6$$

Las siguientes propiedades también se cumplen para todo número real  $\alpha$  y para (todo vector  $a, b$ )  $a$  y  $b$  cualesquiera vectores.

$$\|a\| \geq 0$$

$$\|a\| = 0 \iff a = 0$$

$$\|\alpha a\| = |\alpha| \|a\| \quad 1.7$$

$$\|a+b\| \leq \|a\| + \|b\| \quad (\text{Desigualdad del Triángulo})$$

$$\langle a+b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$$

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle$$

$$\langle a, b \rangle = \langle b, a \rangle$$

$$\langle a, a \rangle \geq 0$$

$$\langle a, a \rangle = 0$$

- La distancia de dos vectores  $a, b$  en  $\mathbb{R}^2$  está dada por:

$$\|b-a\| = \|a-b\|$$

- Una sucesión  $\{a_n\}_{n=1}^{\infty}$  en  $\mathbb{R}^2$  converge a un vector  $a \in \mathbb{R}^2$  si la distancia  $\|a - a_n\|$  converge a 0 cuando  $n \rightarrow \infty$

- Una sucesión  $\{a_n\}_{n=1}^{\infty}$  es fundamental o de Cauchy si las distancias  $\|a_m - a_n\|$  se vuelven pequeñas para  $m$  y  $n$  grande, es decir;

$$\lim_{m, n \rightarrow \infty} \|a_m - a_n\| = 0$$

Es importante el hecho que cada sucesión fundamental en  $\mathbb{R}^2$  converge a algún vector en  $\mathbb{R}^2$ , este hecho es también expresado diciendo que  $\mathbb{R}^2$  es completo.

Finalmente, consideremos el conjunto  $U$  de todos los vectores  $a = (a_1, a_2)$ ,  $a_1, a_2 \in \mathbb{Q}$ , por la teoría de conjuntos sabemos que  $U$  es un conjunto numerable, es decir, dado cualquier vector  $b \in \mathbb{R}^2$  y cualquier número real  $\epsilon > 0$  podemos encontrar un vector  $a \in U$  tal que  $\|b - a\| < \epsilon$  ó en otras palabras el conjunto  $U$  es denso en  $\mathbb{R}^2$ , el

hecho que  $\mathbb{R}^2$  contenga un conjunto denso en todas partes es también expresado diciendo que  $\mathbb{R}^2$  es separable.

## 1.2 ESPACIOS CON PRODUCTO INTERNO

Hasta ahora hemos considerado  $\mathbb{R}^2$  como un espacio lineal sobre el campo de los números reales, sin embargo, a partir de ahora también admitiremos números complejos como escalares, de esta forma denotaremos por  $\mathbb{C}$  el campo de los números complejos y por  $F$  a cualquier otro campo si no queremos restringir a un caso particular como  $\mathbb{R}$  o  $\mathbb{C}$ .

Definición 1.2.1 Un conjunto  $L$  forma un grupo conmutativo bajo la adición, si para cada par ordenado  $(a, b)$  de elementos de  $L$  hay asociado un único elemento  $a+b \in L$ , llamado la suma de  $a$  y  $b$  de tal forma que las siguientes condiciones sean cumplidas:

Sean  $a, b, c$  elementos arbitrarios de  $L$

- i)  $(a+b)+c = a+(b+c)$  ley asociativa
- ii)  $a+0 = 0+a = a$  existencia del cero
- iii) para cada elemento  $a \in L$  existe  $a' \in L \rightarrow$   
 $a+a' = a'+a = 0$  existencia del inverso
- iv)  $a+b = b+a$  ley conmutativa

Definición 1.2.2 Un conjunto  $L$  es un espacio lineal

sobre  $F$  si  $L$  es un grupo conmutativo bajo la suma y si una multiplicación de los elementos de  $L$  con los elementos de  $F$ , llamada multiplicación escalar, es definida en tal forma que las siguientes condiciones se cumplan:

Sean  $a, b$  elementos arbitrarios de  $L$  y  $\alpha, \beta$  escalares arbitrarios en  $F$

- i)  $\alpha(\beta a) = (\alpha\beta)a$
- ii)  $1 \cdot a = a$
- iii)  $\alpha(a+b) = \alpha a + \alpha b$
- iv)  $(\alpha+\beta)a = \alpha a + \beta a$

Un espacio lineal sobre  $\mathbb{C}$  o  $\mathbb{R}$  es llamado espacio lineal complejo o real respectivamente.

Definición 1.2.3 Los elementos  $f_k \in L$  ( $1 \leq k \leq n$ ) son linealmente dependientes si existen  $\alpha_k$  escalares en  $F$  ( $1 \leq k \leq n$ ) no todos iguales a cero tales que  $\sum_{k=1}^n \alpha_k f_k = 0$  y los elementos  $f_k$  son linealmente independientes si  $\sum_{k=1}^n \alpha_k f_k = 0$  para  $\alpha_k = 0$  con  $1 \leq k \leq n$

Definición 1.2.4 Un espacio lineal  $L$  sobre  $F$  es un espacio con producto interior sobre  $F$  si para cada par ordenado  $(f, g)$  de elementos de  $L$  hay asociado un único escalar  $\langle f, g \rangle$  llamado el producto interior de  $f$  y  $g$  de tal forma que las siguientes condiciones se cumplan.

$$i) \quad \langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle$$

- ii)  $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$
- iii)  $\langle g, f \rangle = \langle f, g \rangle$
- iv)  $\langle f, f \rangle \geq 0$
- v)  $\langle f, f \rangle = 0 \iff f = 0$

Definición 1.2.5 Para cada vector  $f \in L$  el número real no negativo  $\|f\| = \sqrt{\langle f, f \rangle}$  es llamada la norma de  $f$ . Un vector de norma 1 es llamado vector unitario.

Definición 12.6 Se dice que  $f$  es ortogonal a  $g$ , si  $\langle f, g \rangle = 0$ . Una familia  $E = \{f_{\sigma}\} \subset L$  es ortogonal si  $f_{\sigma} \perp f_{\sigma_i}$  para  $\sigma_i \neq \sigma_j$ . La familia  $E$  es llamada ortonormal si  $E$  es ortogonal y  $\|f_{\sigma}\| = 1$  para todo  $\sigma$ .

### 1.3 ESPACIOS LINEALES NORMADOS

Definición 1.3.1 Un espacio lineal  $L$  sobre  $\bar{F}$  es llamado un espacio lineal normado, sobre  $\bar{F}$  si para cada vector  $f \in L$  hay asociado un único número  $\|f\| \in \mathbb{R}$ , llamado la norma de  $f$ , de tal forma que las siguientes condiciones se cumplan:

- i)  $\|f\| \geq 0$
- ii)  $\|f\| = 0 \iff f = 0$
- iii)  $\|f+g\| \leq \|f\| + \|g\|$

$$iv) \|\lambda f\| = |\lambda| \|f\|$$

Definición 1.3.2 Para cualquier  $f_0 \in L$  y para cualquier  $\varepsilon > 0$  el conjunto  $\{f: \|f - f_0\| < \varepsilon\}$  es llamado la bola abierta de radio  $\varepsilon$ .

Un subconjunto  $U \subset L$  es llamado abierto si para toda  $f \in U$  existe  $\varepsilon > 0$  tal que la bola abierta de radio  $\varepsilon$  con centro en  $f_0$  está contenida en  $U$ .

Definición 1.3.3 Un vector  $f_1$  es llamado un punto de acumulación de un subconjunto  $U \subset L$  si toda bola abierta con centro en  $f_1$  contiene al menos un vector  $f \in U$  diferente de  $f_1$ . Un subconjunto  $U \in L$  es cerrado si contiene todos sus puntos de acumulación.

Definición 1.3.4 Una sucesión  $\{f_n\}_{n=1}^{\infty} \subset L$  converge a un vector  $f$  llamado su límite si para toda  $\varepsilon > 0$  existe un índice  $n(\varepsilon)$  tal que:

$$\|f - f_n\| < \varepsilon \quad \forall n \geq n(\varepsilon)$$

Definición 1.3.5 Una sucesión  $\{f_n\}_{n=1}^{\infty} \subset L$  es llamada fundamental si para toda  $\varepsilon > 0$  existe un índice  $n(\varepsilon)$  tal que

$$\|f_m - f_n\| < \varepsilon \quad \forall m, n \geq n(\varepsilon)$$

Definición 1.3.6 Un espacio lineal normado  $L$  es llamado completo si toda sucesión fundamental en  $L$  con

verge a algún elemento de  $L$ . Un espacio lineal norma do completo sobre  $F$  es llamado espacio de Banach sobre  $F$ .

Después de haber enunciado las definiciones anteriores, ahora estamos en condiciones de estudiar el concepto de espacio de Hilbert.

#### 1.4 ESPACIO DE HILBERT

Definición 1.4.1 Si un espacio  $X$  con producto interno es completo con respecto a la norma inducida por el producto interno es llamado ESPACIO DE HILBERT.

##### 1.4.1 DESIGUALDAD DE BESSEL

Teorema: Sea  $X$  un espacio con producto interno,  $A$  un conjunto ortonormal de vectores en  $X$  y  $Y$  un vector arbitrario en  $X$ . Entonces

Para todo  $x_1, x_2, \dots, x_n \in A$

$$\sum_{i=1}^n |(Y, x_i)|^2 \leq \|Y\|^2$$

Demostración: Sea  $\alpha_i = (Y, x_i)$  es claro que

$$0 \leq \left( Y - \sum_{i=1}^n \alpha_i x_i, Y - \sum_{i=1}^n \alpha_i x_i \right)$$



$$\begin{aligned}
&= \|y\|^2 - \sum_{i=1}^n \bar{\alpha}_i (y, x_i) - \sum_{i=1}^n \alpha_i (x_i, y) + \sum_{i=1}^n \sum_{j=1}^n \alpha_i \bar{\alpha}_j (x_i, x_j) \\
&= \|y\|^2 - \sum_{i=1}^n |\alpha_i|^2 - \sum_{i=1}^n |\alpha_i|^2 + \sum_{i=1}^n |\alpha_i|^2 \\
&= \|y\|^2 - \sum_{i=1}^n |\alpha_i|^2 \quad \circ \quad \|y\|^2 \geq \sum_{i=1}^n |(y, x_i)|^2
\end{aligned}$$

#### 1.4.2 CONJUNTOS ORTONORMALES COMPLETOS

Definición: Sea  $A$  un conjunto ortonormal en el espacio  $X$  con producto interno. Se dice que  $A$  es completo si no existe otro conjunto ortonormal conteniendo a  $A$ ; Esto es,  $A$  debe ser un conjunto ortonormal maximal completo.

Criterio: Un conjunto  $A$  ortonormal es completo si y solo si,  $X$  tal que  $x \perp A \Rightarrow x=0$

Demostración: Suponga que  $A$  es completo y  $x$  es distinto de cero tal que  $x \perp A$ , claramente esto es contradictorio porque el conjunto ortonormal

$$A \cup \left\{ \frac{x}{\|x\|} \right\}$$

contiene  $A$  propiamente y contradice la maximalidad de  $A$ . recíprocamente, suponga que la condición se cumple, esto es,

$x \perp A$  implica  $x = 0$ . Si  $A$  no es completo, debe existir algún conjunto ortonormal  $B$  tal que  $B \supset A$  propiamente. En este caso, sea  $x \in B - A$ , por tanto  $\|x\| = 1$  y  $x \perp A$ , la suposición de que tal conjunto ortonormal  $B$  existe, debe ser falsa y  $A$  debe ser completo.

CAPITULO II

## 2. CONCEPTOS FUNDAMENTALES EN ANALISIS DE SERIES DE TIEMPO

El punto principal del análisis de series de tiempo, es el concepto, esto es: ¿Qué es una serie de tiempo y Para qué sirve?, será entonces la respuesta a esta pregunta nuestro punto de partida.

Una serie de tiempo es una colección de observaciones hechas secuencialmente en el tiempo.

Gran parte de la teoría estadística esta enfocada en el estudio de muestras aleatorias de observaciones independientes; la especial característica del análisis de series de tiempo es el hecho que las observaciones sucesivas usualmente son NO independientes y que el análisis debe hacerse dentro del orden en que aparecen en el tiempo las observaciones.

Si una serie de tiempo puede ser predecida exactamente, se dice que es determinística, pero las que nos interesan y que son la mayoría, son las estocásticas, en las que solamente parte del futuro esta determinada por observaciones pasadas.

Para series estocásticas, predicciones exactas son imposibles y deben ser reemplazadas por la idea de que los valores futuros tienen una distribución de probabilidad, la cual

esta condicionada por el conocimiento de observaciones pasadas, esto es, para nuestros fines, analizaremos un conjunto de observaciones formando una serie de tiempo, como una realización de variables aleatorias distribuidas conjuntamente; es decir, la secuencia de observaciones  $z_1, \dots, z_n$  tomadas en espacios de tiempo discretos e iguales  $1, \dots, n$  son pensadas que provienen de una distribución de probabilidad

$$f_{1, \dots, n}(z_1, \dots, z_n)$$

donde  $f(z_1, \dots, z_n)$  es una función de densidad de probabilidad, los subíndices  $1, \dots, n$  sobre  $f$  indican que la distribución esta asociada con esos períodos de tiempo y las variables aleatorias en cuestión son  $z_1, \dots, z_n$

Nuestro último objetivo será usar esta distribución conjunta para hacer pronosticos acerca de futuras observaciones.

Por ejemplo, supongamos que conociamos la función de distribución conjunta para  $N = T+1$  y que estabamos en el tiempo  $T$  habiendo observado  $z_1, \dots, z_T$ . Entonces del conocimiento de  $f_{1, \dots, T+1}(z_1, \dots, z_{T+1})$  y de nuestro conocimiento de valores observados de  $z_1, \dots, z_T$ , podríamos construir la función de distribución condicional para la futura observación  $z_{T+1}$ , esto es:  $f_{T+1/1, \dots, T}(z_{T+1}/z_1, \dots, z_T)$

En otras palabras, la información que tenemos acerca de

la relación entre  $Z_1$  ----  $Z_T$  y  $Z_{T+1}$  de su función de distribución conjunta, nos permite usar  $Z_1$  ----  $Z_T$  para hacer predicciones acerca del resultado más probable de  $Z_{T+1}$ .

## 2.1 OBJETIVOS DEL ANALISIS DE SERIES DE TIEMPO

Es importante notar que el pasado histórico de las series de tiempo nos sirve para realizar dos funciones:

1°) Debe informarnos acerca del particular mecanismo, el cual describe su evolución a través del tiempo.

2°) Permitir poner este mecanismo para usarlo en predicciones futuras.

Dentro de todos los posibles objetivos en el análisis de series de tiempo podremos considerar como los más importantes:

a) Descripción b) Explicación c) Predicción c) Control

a) Cuando nos encontramos con una serie de tiempo, el primer paso en el análisis es usualmente graficar los datos y obtener medidas descriptivas simples de las principales propiedades de las series, como en la gráfica.



VENTAS DE UNA CIERTA COMPAÑIA DE INGENIERIA EN MESES SUCESIVOS.

Podemos observar que existe un efecto estacional regular con las ventas altas en invierno y bajas en el verano, también podemos observar que las ventas anuales tienden a incrementarse.

Para algunas series la variación esta determinada por tan obvia característica, y un modelo simple, el cual solamente intenta describir tendencia y variación estacional, puede ser perfectamente adecuado para describir la variación en las series de tiempo.

b) Cuando las observaciones son tomadas sobre dos o más variables, puede ser posible usar la variación en una serie de tiempo para explicar la variación en otra serie.

c) Dada una serie de tiempo observada, quisieramos predecir los valores futuros de las series. Esto es una importante tarea en predicción de ventas y en el análisis de series de tiempo económicas e industriales.

d) Cuando una serie de tiempo es generada, la cual mide la calidad del proceso de manufactura, la dirección del análisis puede ser controlar el proceso.

## 2.2 EL CONCEPTO DE ESTACIONARIO

Antes de empezar con la definición matemática de esta-

cionario, es conveniente introducir la idea desde un punto de vista intuitivo, así podemos pensar en una serie de tiempo estacionaria si:

- i) No hay cambios sistemáticos en la media (sin tendencia)
- ii) No hay cambios sistemáticos en la varianza y
- iii) Si las variaciones periódicas han sido removidas estrictamente.

Ahora consideremos el problema de describir los primeros dos momentos de la función de distribución conjunta, las medidas, varianzas y covarianzas de las variables aleatorias  $Z_1, \dots, Z_n$ .

Sabemos que las medias pueden ser representadas como un vector de  $N$  individuales valores esperados, es decir;

$$(E Z_1, \dots, E Z_n)$$

así mismo las varianzas y covarianzas pueden ser representadas en una matriz simétrica de  $N^2$  elementos, esto es:

$$\begin{pmatrix} \text{Var}(Z_1) & \text{Cov}(Z_1, Z_2) & \dots & \text{Cov}(Z_1, Z_n) \\ \text{Cov}(Z_2, Z_1) & \text{Var}(Z_2) & \dots & \text{Cov}(Z_2, Z_n) \\ \vdots & & & \\ \text{Cov}(Z_n, Z_1) & & & \text{Var}(Z_n) \end{pmatrix}$$

Es claro que para describir tan solo los dos primeros momentos de la distribución conjunta, necesitamos calcular



$N$  esperanzas y  $\frac{1}{2}(N^2 + N)$  varianzas y covarianzas.

Sabemos que en la práctica donde las observaciones  $Z_1, \dots, Z_N$  están observadas y deseamos inferir su distribución conjunta, estimar los  $\frac{3}{2}N$  y  $\frac{1}{2}N^2$  es demasiado complicado.

Una considerable simplificación es lograda si pedimos que la distribución conjunta sea invariante con respecto a el desplazamiento en el tiempo, es decir:

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

Donde  $t$  es cualquier punto en el tiempo y  $k$  y  $m$  son cualquier par de enteros.

La propiedad definida anteriormente es conocida como estacionaria.

Ahora veremos todas las consecuencias al aplicar esta propiedad, para esto, tomemos  $k = 0$ , esto es:

$$f(Z_t) = f(Z_{t+m}) \quad m = \pm 1, \pm 2, \dots$$

esto es, la función de distribución marginal para cualesquiera dos observaciones son las mismas, de aquí podemos seguir que sus esperanzas son las mismas

$$E(Z_t) = E(Z_{t+m})$$

y sus varianzas son las mismas

$$\text{Var}(Z_t) = \text{Var}(Z_{t+m})$$

Similarmente para  $k=1$ , tenemos

$$f(z_t, z_{t+1}) = f(z_{t+m}, z_{t+m+1}) \quad m = \pm 1, \pm 2, \dots$$

Esto significa que la covarianza entre  $z_t$  y  $z_{t+1}$  y entre  $z_{t+m}$  y  $z_{t+m+1}$  debe ser la misma. Estas covarianzas pueden ser denotadas simplemente por  $\gamma_1$  puesto que su valor depende solamente sobre el hecho que las observaciones en cuestión son separadas por un período. Esto es:

$$\text{Cov}(z_t, z_{t+1}) = \text{Cov}(z_{t+m}, z_{t+m+1}) \quad m = \pm 1, \pm 2, \dots$$

Similarmente, para cualquier par de observaciones separadas por  $j$  períodos, tenemos:

$$f(z_t, z_{t+j}) = f(z_{t+m}, z_{t+m+j})$$

y por lo tanto la covarianza entre cualquier par, solamente depende de  $j$ , esto es:

$$\text{Cov}(z_t, z_{t+j}) = \text{Cov}(z_{t+m}, z_{t+m+j}) = \gamma_j$$

De esta forma nos referimos a  $\gamma_j$  como la autocovarianza de atraso  $j$ .

Si denotamos  $E(z_t)$  en general por  $\mu$  es decir

$$\mu = E(z_t) = \int_{-\infty}^{\infty} z f(z) dz$$

y la  $\text{Var}(z_t)$  por  $\gamma_0$ , es decir:

$$\gamma_0 = E[(z_t - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 f(z) dz$$

entonces el primer y segundo momento se convierte en

$$(\mu, \dots, \mu) \quad y$$

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & & \\ \vdots & & & \\ \gamma_{N-1} & & & \gamma_0 \end{pmatrix}$$

El cual involucra ahora solamente  $N+1$  distintos términos, los cuales aun son muchos, y cuyo problema será tratado más adelante.

### 2.3 AUTOCORRELACION

Una consecuencia de estacionalidad la cual tiene una importante interpretación en términos del comportamiento de una serie de tiempo deriva del hecho que la autocovarianza entre dos cualesquiera observaciones depende solamente del número de períodos de tiempo que los separa.

De la definición de autocovarianza tenemos

$$\gamma_j = \text{COV}(Z_t, Z_{t+j}) = \begin{aligned} & E[(Z_t - E(Z_t))(Z_{t+j} - E(Z_{t+j}))] \\ & E[(Z_t - \mu)(Z_{t+j} - \mu)] \end{aligned}$$

lo cual nos dice que la covarianza entre observaciones  $Z_t$  y  $Z_{t+j}$  es el producto esperado de sus desviaciones de la media del proceso. Esto es, si una observación más alta que

el promedio tiende a ser seguida por otra observación  $\dot{\phantom{z}}$  más alta que el promedio períodos más tarde, así mismo para observaciones más bajas que el promedio; la autocovarianza entre  $Z_t$  y  $Z_{t+j}$  es positiva, pero si una observación más alta que el promedio tiende a ser seguida por una observación  $\dot{\phantom{z}}$  abajo del promedio períodos más tarde y viceversa, entonces la autocovarianza es negativa.

Es claro que la estructura de la autocovarianza de una serie de tiempo juega un papel importante en el análisis para efectos de predicción.

Para series tales como series de ventas, tasa de desempleo, como se muestran en las siguiente gráfica.



Claramente predecimos que  $Z_{t+1}$  estaría sobre la media  $\mu$ , si hubieramos observado que  $Z_t$  estaba por debajo de la media.

El hecho que la autocovarianza  $\gamma_j$  parezca determinar la apariencia de una serie de tiempo sugiere que un proceso estacionario presentará el mismo patrón de comportamiento sin importar cuando lo observamos.

Entonces parecería apropiado caracterizar un simple proceso presentando el conjunto de covarianzas  $\gamma_0, \gamma_1, \dots$ , para propósitos de comparar diferentes series, sin embargo, esto no es enteramente satisfactorio ya que una diferencia en la dispersión del proceso, quizá causada por diferentes escalas de medidas, conducirá a muy diferentes autocovarianzas. Puesto que la varianza es una medida de dispersión, la comparabilidad es realizada si nosotros estandarizamos las autocovarianzas dividiéndolas entre  $\gamma_0$ , esto es, transformándolas a correlaciones; tales correlaciones son conocidas como autocorrelaciones.

Si denotamos la correlación entre  $Z_t$  y  $Z_{t+j}$  por  $\rho_j$  entonces el conjunto de autocorrelaciones conocido conjuntamente como la función de autocorrelación, es dada por

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1, \quad \rho_1 = \frac{\gamma_1}{\gamma_0}, \quad \rho_2 = \frac{\gamma_2}{\gamma_0}, \quad \dots$$

Así mismo, una gráfica de la función de autocorrelación es llamada correlograma.

## 2.4 CONCEPTO DE NO ESTACIONARIO

Es de hacer notar que muchas series de tiempo de interés, para propósitos de predicción son obviamente no estacionarias, por ejemplo, los precios del mercado no presentan

afinidad alrededor de la media, en lugar de esto, estas parecen perderse libremente en ambas direcciones ( $\downarrow \uparrow$ ). Afortunadamente las diferencias, esto es, cambios sucesivos en muchas series no estacionarias son estacionarias. Por ejemplo el modelo del camino aleatorio explica muy bien el comportamiento de los precios del mercado, y las diferencias  $Z_t - Z_{t-1}$  son estacionarias como se verá más adelante.

## 2.5 MODELOS PARA SERIES DE TIEMPO ESTACIONARIAS

### Procesos estocasticos lineales discretos

Hablaremos de un proceso lineal estocastico si cada observación  $Z_t$  puede ser expresada de la siguiente forma:

$$Z_t = \mu + \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots \quad (2.5.1)$$

$$\circ \quad Z_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

donde las constantes  $\psi_j \quad j=0,1,2,\dots$  son usualmente llamadas "ponderadores" y  $\mu$  es una constante que determina el nivel del proceso. Usualmente  $\psi_0 = 1$ , una forma alternativa de escribir 2.5.1 es en terminos del operador de cambio hacia atrás  $B$ , definido de la siguiente forma:

$$B \epsilon_t = \epsilon_{t-1}$$

En general

$$B^j \epsilon_t = \epsilon_{t-j}$$

Entonces la ecuación 2.5.1 podemos describirla como

$$Z_t = \mu + (\psi_0 B^0 + \psi_1 B^1 + \psi_2 B^2 + \dots) \varepsilon_t$$

o

$$Z_t = \mu + \psi(B) \varepsilon_t$$

donde

$$\psi(B) = \psi_0 B^0 + \psi_1 B^1 + \psi_2 B^2 + \dots \quad \text{y } \psi_0 = 1$$

Regresando al modelo 2.5.1 la serie de tiempo .....

$\varepsilon_{t-1}, \varepsilon_t, \dots$  es una secuencia de errores aleatorios independientes e idénticamente distribuidos, con media cero y varianza  $\sigma^2$ , usualmente suponemos que la distribución del  $\{\varepsilon_i\}$  es normal y entonces la secuencia de variables aleatorias  $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$  es llamado Proceso de Ruido Blanco.

Es importante hacer notar que las observaciones sucesivas en la serie de tiempo  $\{Z_t\}$  son dependientes, porque ellas son determinadas de las mismas realizaciones previas de  $\{\varepsilon_i\}$  y también si los  $\varepsilon_i$  se distribuyen normal, entonces las  $Z_t$  se distribuyen normal también.

La forma del modelo 2.5.1 es usualmente llamado filtro lineal, es entonces que podemos definir un modelo de series de tiempo como una función que transforma un proceso de ruido blanco a una serie de tiempo.

Cuando tenemos un proceso lineal, el primer paso a dar

es checar si el proceso es estacionario, esto es, debemos re-  
cordar que el proceso será estacionario si la media y la matriz  
de varianzas y covarianzas existe y si es invariante respec-  
to al tiempo.

Entonces, dado un particular proceso, *es decir* valores par-  
ticulares de los parámetros, calcularemos sus dos primeros  
momentos.

Sea el proceso  $Z_t = \mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$   
entonces 
$$\begin{aligned} E(Z_t) &= E(\mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots) \\ E(Z_t) &= E\left(\mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right) \\ &= \mu + E\left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right) \end{aligned}$$

aunque sabemos que la esperanza es un operador lineal, tene-  
mos una suma infinita y por lo tanto necesitamos que:

$$\sum_{j=0}^{\infty} \psi_j \text{ converge con } \psi_0 = 1 \quad (2.5.2)$$

entonces si (2.5.2) se cumple, tomaremos la esperanza de  
 $\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$  termino a termino, obteniendo  $\sum_{j=0}^{\infty} \psi_j E(\varepsilon_{t-j}) = 0$   
Entonces la media del proceso será  $E(Z_t) = \mu$

Notamos que la media del proceso no depende del tiempo.

Para calcular la varianza del proceso simplemente parti-  
mos de la definición,  
utilizando la notación convencional.



$$\begin{aligned} \text{Var}(z_t) &= \gamma_0 = \mathbb{E}(z_t - \mathbb{E}(z_t))^2 \\ &= \mathbb{E}(z_t - \mu)^2 = \mathbb{E}(\mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots - \mu)^2 \\ &= \mathbb{E}(\varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots)^2 \end{aligned}$$

entonces

$$\text{Var}(z_t) = \mathbb{E}(\varepsilon_t^2 + \psi_1^2 \varepsilon_{t-1}^2 + \psi_2^2 \varepsilon_{t-2}^2 + \dots) + \mathbb{E}(2\varepsilon_t \psi_1 \varepsilon_{t-1} + 2\varepsilon_t \psi_2 \varepsilon_{t-2} + \dots)$$

ya que la media de cualquier  $\varepsilon_{t-i}$  es cero y  $\mathbb{E}(\varepsilon_{t-i}^2)$  es la varianza de  $\varepsilon_{t-i}$ , es decir,  $\sigma_\varepsilon^2$  y la  $\mathbb{E}(\varepsilon_{t-i} \varepsilon_{t-j})$  cuando  $i \neq j$  es la covarianza entre  $\varepsilon_{t-i}$  y  $\varepsilon_{t-j}$  que es cero puesto que sabemos que los errores son independientes, obtenemos que:

$$\gamma_0 = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i^2$$

obviamente  $\gamma_0$  existirá si la media del proceso existe y la  $\sum_{i=0}^{\infty} \psi_i^2$  converja.

Ahora la covarianza, estará determinada por

$$\text{Cov}(z_t, z_{t-j}) = \gamma_j = \mathbb{E}[z_t - \mathbb{E}(z_t)][z_{t-j} - \mathbb{E}(z_{t-j})]$$

Así la autocovarianza es solamente la covarianza de dos variables aleatorias; el prefijo auto solamente implica que nos estamos refiriendo a la covarianza de dos observaciones cualesquiera en la misma serie de tiempo con  $K$  períodos de tiempo de separación, entonces,

$$\begin{aligned} \gamma_j &= \mathbb{E}[(\varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots)(\varepsilon_{t-j} + \psi_1 \varepsilon_{t-j-1} + \psi_2 \varepsilon_{t-j-2} + \dots)] \\ &= \mathbb{E}[(\psi_j \varepsilon_{t-j}^2) + (\psi_1 \psi_{j+1} \varepsilon_{t-j-1}^2) + \dots] + \mathbb{E}[\text{Terminos de los productos cruzados}] \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j}, \text{ esto solo existirá si la } Z \text{ converge} \end{aligned}$$

Como un simple ejemplo, consideremos:

$$Z_t = \mu + \varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \dots$$

tal que  $|\phi| < 1$

procedemos a verificar si el proceso es estacionario, i.e.

$$\sum_{i=0}^{\infty} \phi^i \text{ converge?}$$

sabemos que por ser una progresión geométrica converge a:

$$\sum_{i=0}^{\infty} \phi^i = \frac{1}{1-\phi}$$

esto implica que la media del proceso es  $\mu$ .

así mismo con los resultados obtenidos anteriormente, sabemos que la varianza  $\gamma_0$ , esta dada por:  $\sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} \phi_i^2$

habría que checar que la  $\sum$  converja, esto es:

$$\sum (\phi^i)^2 = \sum (\phi^2)^i = \frac{1}{1-\phi^2}$$

entonces  $\gamma_0 = \frac{\sigma_{\varepsilon}^2}{1-\phi^2}$

de igual manera las autocovarianzas son

$$\begin{aligned} \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+j} &= \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} \phi^i \phi^{i+j} = \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} \phi^{2i} \phi^j = \phi^j \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} \phi^{2i} \\ \gamma_j &= \frac{\phi^j \sigma_{\varepsilon}^2}{1-\phi^2} \end{aligned}$$

## 2.6 PROCESOS DE MEDIAS MOVILES

Consideremos el caso del proceso lineal en forma general, es decir, como el modelo (2.5.1), pero ahora solamente con los primeros  $q$  ponderadores distintos de cero, esto es

$$Z_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots + \psi_q \varepsilon_{t-q}$$

es decir  $\psi_i = 0 \quad \forall i > q$ .

Así, este modelo es conocido como proceso de medias móviles de orden  $q$  ya que las observaciones son un promedio móvil en los errores por  $q$  períodos de atraso.

Para distinguir el caso de medias móviles del proceso lineal en general, se conviene un pequeño cambio en los ponderadores de la siguiente manera:

$$Z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

tal que el signo menos es introducido por convención y una manera de referirse a este de una manera fácil será:  $MA(q)$

En términos del operador de cambio hacia atrás, el proceso  $MA(q)$  puede ser expresado así:

$$\begin{aligned} Z_t &= \mu + (1 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t \\ &= \mu + \Theta_q(B) \varepsilon_t \end{aligned}$$

donde

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

### 2.6.1 PROCESO DE MEDIAS MOVILES DE PRIMER ORDEN

Consideremos  $MA(1)$  es decir:

$$Z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Veamos ahora los momentos

$$\begin{aligned} E(Z_t) &= \mu + E(\varepsilon_t - \theta_1 \varepsilon_{t-1}) \\ &= \mu + E(\varepsilon_t) - \theta_1 E(\varepsilon_{t-1}) \\ &= \mu \end{aligned}$$

$$\begin{aligned}
\text{Var}(z_t) &= \gamma_0 = \mathbb{E} (z_t - \mathbb{E}(z_t))^2 \\
&= \mathbb{E} (\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \mu)^2 \\
&= \mathbb{E} (\varepsilon_t - \theta_1 \varepsilon_{t-1})^2 \\
&= \mathbb{E} (\varepsilon_t^2 - 2\varepsilon_t \theta_1 \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2) \\
&= \mathbb{E} (\varepsilon_t^2) - 2\theta_1 \mathbb{E} (\varepsilon_t \varepsilon_{t-1}) + \theta_1^2 \mathbb{E} (\varepsilon_{t-1}^2) \\
&= \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 \\
&= \sigma_\varepsilon^2 (1 + \theta_1^2)
\end{aligned}$$

la autocovarianza de período de atraso 1

$$\begin{aligned}
\text{Cov}(z_t, z_{t-1}) &= \gamma_1 = \mathbb{E} [(z_t - \mathbb{E}(z_t))(z_{t-1} - \mathbb{E}(z_{t-1}))] \\
&= \mathbb{E} [(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2})] \\
&= \mathbb{E} [\varepsilon_t \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} \varepsilon_t - \theta_1 \varepsilon_{t-1}^2 + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-2}] \\
&= \mathbb{E} (\varepsilon_t \varepsilon_{t-1}) - \theta_1 \mathbb{E} (\varepsilon_{t-2} \varepsilon_t) - \theta_1 \mathbb{E} (\varepsilon_{t-1}^2) + \theta_1^2 \mathbb{E} (\varepsilon_{t-1} \varepsilon_{t-2}) \\
&= -\theta_1 \sigma_\varepsilon^2
\end{aligned}$$

Veremos que pasa períodos de atraso más grandes que 1

$$\begin{aligned}
\text{Cov}(z_t, z_{t-j}) &= \gamma_j = \mathbb{E} [(z_t - \mathbb{E}(z_t))(z_{t-j} - \mathbb{E}(z_{t-j}))] \\
&= \mathbb{E} [(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-j} - \theta_1 \varepsilon_{t-j-1})] \\
&= \mathbb{E} [(\varepsilon_t \varepsilon_{t-j} - \theta_1 \varepsilon_t \varepsilon_{t-j-1} - \theta_1 \varepsilon_{t-1} \varepsilon_{t-j} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-j-1})] \\
&= \mathbb{E} (\varepsilon_t \varepsilon_{t-j}) - \theta_1 \mathbb{E} (\varepsilon_t \varepsilon_{t-j-1}) - \theta_1 \mathbb{E} (\varepsilon_{t-1} \varepsilon_{t-j}) + \theta_1^2 \mathbb{E} (\varepsilon_{t-1} \varepsilon_{t-j-1})
\end{aligned}$$

vemos que para  $j > 1$   $\gamma_j = 0$

Ahora la función de autocorrelación es:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 (1 + \theta_1^2)} = \frac{-\theta_1}{(1 + \theta_1^2)}$$

Obviamente para períodos de atraso mayores que 1

$$\rho_j = 0 \quad \forall j > 1$$

## 2.6.2 PROCESO DE MEDIAS MOVILES DE ALTO ORDEN

Sabemos que el modelo es de la forma:

$$Z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Sus momentos:

$$\begin{aligned} \mathbb{E}(Z_t) &= \mathbb{E}(\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}) \\ &= \mu \end{aligned}$$

la varianza:

$$\begin{aligned} \text{Var}(Z_t) &= \gamma_0 = \mathbb{E}(Z_t - \mathbb{E}(Z_t))^2 \\ &= \mathbb{E}(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})^2 \\ &= \mathbb{E}(\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + \dots + \theta_q^2 \varepsilon_{t-q}^2) + \mathbb{E}[\text{Terminos de los productos cruzados}] \\ &= \mathbb{E}(\varepsilon_t^2) + \theta_1^2 \mathbb{E}(\varepsilon_{t-1}^2) + \dots + \theta_q^2 \mathbb{E}(\varepsilon_{t-q}^2) \\ &= \sigma_\varepsilon^2 \sum_{i=0}^q \theta_i^2 \end{aligned}$$

con la convención  $\theta_0 = 1$

La autocovarianza estará dada por:

$$\begin{aligned} \gamma_j &= \mathbb{E}[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-j} - \theta_1 \varepsilon_{t-j-1} - \theta_2 \varepsilon_{t-j-2} - \dots - \theta_q \varepsilon_{t-j-q})] \\ \gamma_j &= \mathbb{E}[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_{j-1} \varepsilon_{t-j+1} - \theta_j \varepsilon_{t-j} - \theta_{j+1} \varepsilon_{t-j-1} - \dots - \theta_q \varepsilon_{t-j-q}) \times \\ &\quad (\varepsilon_{t-j} - \theta_1 \varepsilon_{t-j-1} - \theta_2 \varepsilon_{t-j-2} - \dots - \theta_{q-j} \varepsilon_{t-q} - \dots - \theta_q \varepsilon_{t-j-q})] \\ \gamma_j &= \mathbb{E}[-\theta_j \varepsilon_{t-j}^2 + \theta_1 \theta_{j+1} \varepsilon_{t-j-1}^2 + \dots + \theta_{q-j} \theta_q \varepsilon_{t-q}^2] \\ \gamma_j &= -\theta_j \mathbb{E}(\varepsilon_{t-j}^2) + \theta_1 \theta_{j+1} \mathbb{E}(\varepsilon_{t-j-1}^2) + \dots + \theta_{q-j} \theta_q \mathbb{E}(\varepsilon_{t-q}^2) \\ \gamma_j &= \sigma_\varepsilon^2 (-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q) \quad j=1, 2, \dots, q \\ \gamma_j &= 0 \quad \forall j > q \end{aligned}$$

Entonces la función de autocorrelación será:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\sigma_\varepsilon^2 (-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_{q-j} \theta_q)}{\sigma_\varepsilon^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)}$$

es decir

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_{q-j} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} \\ 0 \end{cases}$$

174

## 2.7 PROCESOS AUTORREGRESIVOS

Para empezar a hablar de procesos autoregresivos, recordemos el modelo lineal general, es decir

$$Z_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots \quad 2.7.1$$

Una forma alternativa de escribir este modelo será expresar  $Z_t$  en términos de los errores y todas las observaciones pasadas, para demostrar esto, recordaremos el modelo general haciendo que  $\varepsilon_t$  aparezca del lado izquierdo de la ecuación y todo lo demás del lado derecho, esto es:

$$\varepsilon_t = Z_t - \mu - \psi_1 \varepsilon_{t-1} - \psi_2 \varepsilon_{t-2} \dots \quad 2.7.2$$

Ahora como esto se cumple para cualquier índice, obtenemos:

$$\varepsilon_{t-1} = Z_{t-1} - \mu - \psi_1 \varepsilon_{t-2} - \psi_2 \varepsilon_{t-3} \dots \quad 2.7.3$$

Sustituyendo 2.7.2 en 2.7.3 podemos eliminar  $\varepsilon_{t-1}$  obteniendo:

$$\varepsilon_t = Z_t - \mu - \psi_1 (Z_{t-1} - \mu - \psi_1 \varepsilon_{t-2} - \dots) - \psi_2 \varepsilon_{t-2} \dots$$

$$E_t = Z_t - \mu - \psi_1 Z_{t-1} + \psi_1 \mu + \psi_1^2 E_{t-2} + \psi_1 \psi_2 E_{t-3} + \dots - \psi_2 E_{t-2} \dots$$

reordenando

$$Z_t = E_t + \mu + \psi_1 Z_{t-1} - \psi_1 \mu - \psi_1^2 E_{t-2} - \psi_1 \psi_2 E_{t-3} - \dots + \psi_2 E_{t-2} + \dots$$

$$Z_t = \mu(1 - \psi_1) + \psi_1 Z_{t-1} + E_t + (\psi_2 - \psi_1^2) E_{t-2} + (\psi_3 - \psi_1 \psi_2) E_{t-3} + \dots$$

Similarmente, podríamos substituir para  $E_{t-2}, E_{t-3}, \dots$ , hasta obtener una expresión para  $Z_t$  de la forma:

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + \zeta + E_t$$

donde  $\pi_i$  como coeficientes de las observaciones pasadas son funciones de los ponderadores  $\psi_i$  y  $\zeta$  es una constante la cual es también función de  $\psi_i$  y  $\mu$

Como nuestro interés será desarrollar modelos "parsimoniosos", esto es, modelos los cuales describan adecuadamente la serie de tiempo con poco parámetros, si este es el caso,  $\pi_i = 0 \quad \forall i > p$  entonces el proceso se transformará en:

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + \pi_p Z_{t-p} + \zeta + E_t$$

el cual es conocido como:

Proceso autorregresivo de orden  $p$  ó  $AR(p)$ , para ser consistente en la literatura que se refiere al proceso autorregresivo los coeficientes son denotados  $\phi_i$  en lugar de  $\pi_i$ , es decir;

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

El proceso  $AR(p)$  puede ser escrito en términos del operador de cambio hacia atrás como

$$Z_t = \zeta + (\phi_1 B^1 + \phi_2 B^2 + \dots + \phi_p B^p) Z_t + \varepsilon_t$$

o

$$(1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = \zeta + \varepsilon_t$$

o simplemente

$$\Phi_p(B) Z_t = \zeta + \varepsilon_t$$

con

$$\Phi_p(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$$

Box y Jenkins demuestran que si las raíces del polinomio  $\Phi_p(B) = 0$  están fuera del círculo unitario, el proceso es estacionario. Esta condición se deriva del hecho que  $\sum_{i=0}^{\infty} \psi_i$  debe converger para que el proceso  $\{Z_t\}$  sea estacionario.

Más adelante veremos que condiciones implica en términos de valores admisibles para  $\phi_j$  en el proceso  $AR(p)$

### 2.7.1 PROCESO AUTORREGRESIVO DE PRIMER ORDEN

Si tomamos  $p=1$  en el proceso  $AR(p)$  obtenemos:

$$Z_t = \zeta + \phi_1 Z_{t-1} + \varepsilon_t$$

el cual solo depende de la observación al tiempo  $t-1$  y es conocido como proceso de Markov. Para que el proceso sea estacionario se requiere que las raíces de  $\Phi(B) = 1 - \phi_1 B^1 = 0$  se hallen fuera del círculo unitario, esto es equiva-



lente a pedir que:

$$|\phi_1| < 1$$

La media del proceso será:

$$\begin{aligned}\mathbb{E}(Z_t) = \mu &= \mathbb{E}(\gamma + \phi_1 Z_{t-1} + \varepsilon_t) \\ &= \gamma + \phi_1 \mathbb{E}(Z_{t-1}) \\ &= \gamma + \phi_1 \mathbb{E}[\gamma + \phi_1 Z_{t-2} + \varepsilon_{t-1}] \\ &= \gamma + \phi_1 \gamma + \phi_1^2 \mathbb{E}(Z_{t-2}) \\ &= \gamma + \phi_1 \gamma + \phi_1^2 \mathbb{E}[\gamma + \phi_1 Z_{t-3} + \varepsilon_{t-2}] \\ &= \gamma + \phi_1 \gamma + \phi_1^2 \gamma + \phi_1^3 \mathbb{E}(Z_{t-3})\end{aligned}$$

y así sucesivamente, es decir:

$$\mu = \gamma \sum_{i=0}^{\infty} \phi_1^i \quad \text{y ya que } |\phi_1| < 1 \Rightarrow \mu = \frac{\gamma}{1 - \phi_1}$$

o de otra manera, sabiendo que el proceso es estacionario.

$$\begin{aligned}\mu &= \mathbb{E}(Z_t) = \mathbb{E}(\gamma + \phi_1 Z_{t-1} + \varepsilon_t) & \mu(1 + \phi_1) &= \gamma \\ \mu &= \gamma + \phi_1 \mathbb{E}(Z_{t-1}) & \mu &= \frac{\gamma}{1 + \phi_1} \\ \mu &= \gamma + \phi_1 \mu\end{aligned}$$

Para calcular la varianza y autocovarianza de una manera fácil, haremos uso de los resultados obtenidos anteriormente, simplemente mostrando que si escribimos el proceso AR(1) en términos de los errores pasados substituyendo  $Z_{t-1}, Z_{t-2}, \dots$  etc., sucesivamente, esto es:

$$Z_t = \phi_1 Z_{t-1} + \gamma + \varepsilon_t \quad Z_{t-1} = \phi_1 Z_{t-2} + \gamma + \varepsilon_{t-1}$$

entonces

$$Z_t = \phi_1 [\phi_1 Z_{t-2} + \gamma + \varepsilon_{t-1}] + \gamma + \varepsilon_t$$

$$= \phi_1^2 z_{t-2} + \phi_1 \eta + \phi_1 \varepsilon_{t-1} + \eta + \varepsilon_t$$

$$z_{t-2} = \phi_1 z_{t-3} + \eta + \varepsilon_{t-2}$$

$$= \phi_1^2 [\phi_1 z_{t-3} + \eta + \varepsilon_{t-2}] + \phi_1 \eta + \phi_1 \varepsilon_{t-1} + \eta + \varepsilon_t$$

$$= \phi_1^3 z_{t-3} + \eta \phi_1^2 + \phi_1^2 \varepsilon_{t-2} + \phi_1 \eta + \phi_1 \varepsilon_{t-1} + \eta + \varepsilon_t$$

$$= \phi_1^3 z_{t-3} + \phi_1^2 \eta + \phi_1 \eta + \eta + \phi_1^2 \varepsilon_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

⋮

en el paso  $n = \phi_1^n z_{t-n} + \phi_1^n \eta + \dots + \phi_1 \eta + \eta + \phi_1^n \varepsilon_{t-n} + \dots + \phi_1 \varepsilon_{t-1} + \varepsilon_t$

$$= \phi_1^n z_{t-n} + \eta [\phi_1^n + \dots + \phi_1 + 1] + \phi_1^n \varepsilon_{t-n} + \dots + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

cuando  $n \rightarrow \infty$

$$z_t = \frac{\eta}{1 - \phi_1} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots$$

notemos que AR(1) escrito en estos términos, es visto como un proceso de medias móviles de orden infinito, cuyo caso ya lo habíamos analizado, simplemente tomando

$$\phi_1 = \phi \quad \text{y} \quad \mu = \frac{\eta}{1 - \phi_1} \quad \text{y es entonces que podemos}$$

hacer uso de los resultados obtenidos para nuestro proceso, tomando:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi^2}, \quad \gamma_j = \frac{\phi^j \sigma_\varepsilon^2}{1 - \phi_1^2}$$

La función de autocorrelación para nuestro proceso, será entonces:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\frac{\phi^j \sigma_\varepsilon^2}{1 - \phi_1^2}}{\frac{\sigma_\varepsilon^2}{1 - \phi_1^2}} = \phi_1^j$$

Vemos que la función de autocorrelación para el proceso decae exponencialmente cuando  $\phi_1$  es positivo y decae expo-

nencialmente, pero oscila en signo cuando  $\phi_1$  es negativo.

Otra forma de calcular los momentos del proceso AR(1) la cual utilizaremos posteriormente es de la siguiente manera:

$$\begin{aligned} \text{Tenemos: } Z_t &= \phi_1 Z_{t-1} + \zeta + \varepsilon_t \\ \mathbb{E}(Z_t) &= \mathbb{E}(\phi_1 Z_{t-1} + \zeta + \varepsilon_t) \\ &= \phi_1 \mathbb{E}(Z_{t-1}) + \zeta \end{aligned}$$

como sabemos que el proceso es estacionario

$$\mathbb{E}(Z_t) = \frac{\zeta}{1 - \phi_1}$$

como lo habíamos mostrado anteriormente.

Para obtener la varianza haremos un pequeño cambio para obtenerla de una manera fácil, es decir;

$$\tilde{Z}_t = Z_t - \frac{\zeta}{1 - \phi_1}$$

esto es, al proceso le restamos su media y lo cual nos conduce a:

$$\begin{aligned} Z_t &= \phi_1 Z_{t-1} + \zeta + \varepsilon_t \\ Z_t &= \phi_1 Z_{t-1} + \frac{\zeta}{1 - \phi_1} (1 - \phi_1) + \varepsilon_t \\ Z_t &= \phi_1 Z_{t-1} - \phi_1 \left( \frac{\zeta}{1 - \phi_1} \right) + \frac{\zeta}{1 - \phi_1} + \varepsilon_t \\ Z_t &= \phi_1 \left[ Z_{t-1} - \frac{\zeta}{1 - \phi_1} \right] + \frac{\zeta}{1 - \phi_1} + \varepsilon_t \\ Z_t - \frac{\zeta}{1 - \phi_1} &= \phi_1 \left[ Z_{t-1} - \frac{\zeta}{1 - \phi_1} \right] + \varepsilon_t \\ \tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \varepsilon_t \end{aligned}$$

que no es más que un reordenamiento del proceso, sin la constante  $\xi$  y cuya esperanza es cero.

La varianza de  $\tilde{z}_t$  es entonces.

$$\begin{aligned} \text{Var}(z_t) = \gamma_0 &= \mathbb{E}(\tilde{z}_t^2) = \mathbb{E}[\tilde{z}_t(\phi_1 \tilde{z}_{t-1} + \varepsilon_t)] \\ &= \phi_1 \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) + \mathbb{E}(\tilde{z}_t \varepsilon_t) \\ &= \phi_1 \gamma_1 + \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t) \varepsilon_t] \\ &= \phi_1 \gamma_1 + \phi_1 \mathbb{E}(\tilde{z}_{t-1} \varepsilon_t) + \mathbb{E}(\varepsilon_t^2) \\ &= \phi_1 \gamma_1 + \sigma_\varepsilon^2 \dots \dots \dots A \end{aligned}$$

Dado que la varianza involucra autocovarianza de período de atraso 1 veremos la forma en que está dada:

$$\begin{aligned} \text{Cov}(\tilde{z}_t \tilde{z}_{t-1}) &= \gamma_1 = \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) \\ &= \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t)(\tilde{z}_{t-1})] \\ &= \phi_1 \mathbb{E}(\tilde{z}_{t-1}^2) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-1}) \\ &= \phi_1 \gamma_0 \dots \dots \dots B \end{aligned}$$

Ahora tenemos un sistema de ecuaciones que puede ser resuelto simplemente substituyendo B en A, es decir;

$$\begin{aligned} \gamma_0 &= \phi_1 [\phi_1 \gamma_0] + \sigma_\varepsilon^2 \\ &= \phi_1^2 \gamma_0 + \sigma_\varepsilon^2 \\ &= \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} \quad \therefore \gamma_1 = \phi_1 \left( \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} \right) \end{aligned}$$

Ahora, autocovarianza de períodos de atraso más grandes, son calculados de igual manera, esto es;

$$\gamma_j = \mathbb{E}(\tilde{z}_t \tilde{z}_{t-j})$$

$$\begin{aligned}
&= E[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t)(\tilde{z}_{t-j})] \\
&= \phi_1 E(\tilde{z}_{t-1} \tilde{z}_{t-j}) + E(\varepsilon_t \tilde{z}_{t-j}) \\
&= \phi_1 \gamma_{j-1} \quad \text{para } j=1,2,\dots
\end{aligned}$$

esto es, para cualquier  $j=1,2,\dots$

Tenemos  $\gamma_j = \phi_1^j \gamma_0$

entonces la función de autocorrelación, es:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\phi_1^j \gamma_0}{\gamma_0} = \phi_1^j$$

### 2.7.2 PROCESO AUTORREGRESIVO DE SEGUNDO ORDEN

Si  $p=2$  en el proceso AR(p) obtenemos

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \zeta + \varepsilon_t$$

Para que este proceso AR(2) sea estacionario, requerimos que las raíces de la ecuación característica  $(1 - \phi_1 B - \phi_2 B^2) = 0$  se hallen fuera del círculo unitario en el plano complejo, esto es equivalente a pedir que los parámetros  $\phi_1$  y  $\phi_2$  sean tal que:

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

Si estas condiciones se cumplen, es fácil encontrar la media y varianza.

$$\begin{aligned}
Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \gamma + \varepsilon_t \\
E(Z_t) &= \phi_1 E(Z_{t-1}) + \phi_2 E(Z_{t-2}) + \gamma + E(\varepsilon_t) \\
\mu &= \phi_1 \mu + \phi_2 \mu + \gamma \\
\gamma &= \mu - \phi_1 \mu - \phi_2 \mu \\
\gamma &= \mu (1 - \phi_1 - \phi_2) \\
\mu &= \frac{\gamma}{1 - \phi_1 - \phi_2}
\end{aligned}$$

Si hacemos un cambio en el proceso como en AR(1) es decir;

$$\tilde{Z}_t = Z_t - \frac{\gamma}{1 - \phi_1 - \phi_2}$$

lo cual nos conduce a:

$$\begin{aligned}
Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \gamma + \varepsilon_t \\
Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \gamma \frac{(1 - \phi_1 - \phi_2)}{(1 - \phi_1 - \phi_2)} + \varepsilon_t \\
Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \gamma \left[ \frac{1}{1 - \phi_1 - \phi_2} - \phi_1 \left( \frac{\gamma}{1 - \phi_1 - \phi_2} \right) - \phi_2 \left( \frac{\gamma}{1 - \phi_1 - \phi_2} \right) \right] + \varepsilon_t \\
Z_t &= \phi_1 \left[ Z_{t-1} - \frac{\gamma}{1 - \phi_1 - \phi_2} \right] + \phi_2 \left[ Z_{t-2} - \frac{\gamma}{1 - \phi_1 - \phi_2} \right] + \frac{\gamma}{1 - \phi_1 - \phi_2} + \varepsilon_t \\
Z_t - \frac{\gamma}{1 - \phi_1 - \phi_2} &= \phi_1 \left[ Z_{t-1} - \frac{\gamma}{1 - \phi_1 - \phi_2} \right] + \phi_2 \left[ Z_{t-2} - \frac{\gamma}{1 - \phi_1 - \phi_2} \right] + \varepsilon_t \\
\tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \varepsilon_t
\end{aligned}$$

lo cual, como sabemos anteriormente es solamente un reordenamiento del proceso, que nos facilitará los cálculos de varianza y autocovarianza dado que sabemos que  $E(\tilde{Z}_t) = 0$

$$\begin{aligned}
\text{Var}(Z_t) &= \gamma_0 = E(\tilde{Z}_t^2) \\
&= E\left[\tilde{Z}_t (\phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \varepsilon_t)\right]
\end{aligned}$$

$$\begin{aligned}
&= \phi_1 \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) + \phi_2 \mathbb{E}(\tilde{z}_t \tilde{z}_{t-2}) + \mathbb{E}(\tilde{z}_t \varepsilon_t) \\
&= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \mathbb{E}(\tilde{z}_t \varepsilon_t)
\end{aligned}$$

por otro lado  $\mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \varepsilon_t)(\varepsilon_t)]$

$$= \phi_1 \mathbb{E}(\tilde{z}_{t-1} \varepsilon_t) + \phi_2 \mathbb{E}(\tilde{z}_{t-2} \varepsilon_t) + \mathbb{E}(\varepsilon_t^2) = \sigma_\varepsilon^2$$

entonces:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\varepsilon^2$$

ahora recordemos que  $\gamma_1$  y  $\gamma_2$  están calculados de la forma ya conocida

$$\begin{aligned}
\gamma_1 &= \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) = \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \varepsilon_t) \tilde{z}_{t-1}] \\
&= \phi_1 \mathbb{E}(\tilde{z}_{t-1}^2) + \phi_2 \mathbb{E}(\tilde{z}_{t-2} \tilde{z}_{t-1}) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-1}) \\
&= \phi_1 \gamma_0 + \phi_2 \gamma_1
\end{aligned}$$

$$\begin{aligned}
\gamma_2 &= \mathbb{E}(\tilde{z}_t \tilde{z}_{t-2}) = \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \varepsilon_t) \tilde{z}_{t-2}] \\
&= \phi_1 \mathbb{E}(\tilde{z}_{t-1} \tilde{z}_{t-2}) + \phi_2 \mathbb{E}(\tilde{z}_{t-2}^2) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-2}) \\
&= \phi_1 \gamma_1 + \phi_2 \gamma_0
\end{aligned}$$

De esta forma obtenemos el sistema de ecuaciones

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\varepsilon^2$$

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

$$\gamma_2 = \phi_1 \gamma_1 + \phi_2 \gamma_0$$

que pueden ser resueltas para  $\gamma_0$ ,  $\gamma_1$  y  $\gamma_2$  conociendo  $\phi_1$ ,  $\phi_2$  y  $\sigma_\varepsilon^2$ . Para períodos de atraso mayores que 2,  $\gamma_j$  esta dada por

$$\gamma_j = \mathbb{E}(\tilde{z}_t \tilde{z}_{t-j})$$

$$\begin{aligned}
&= \mathbb{E} \left[ (\phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \varepsilon_t) (\tilde{z}_{t-j}) \right] \\
&= \phi_1 \mathbb{E}(\tilde{z}_{t-1} \tilde{z}_{t-j}) + \phi_2 \mathbb{E}(\tilde{z}_{t-2} \tilde{z}_{t-j}) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-j}) \\
&= \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} \quad \forall j \geq 2
\end{aligned}$$

Para calcular la función de autocorrelación, simplemente dividimos por  $\gamma_0$  y obtenemos

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\phi_1 \gamma_0 + \phi_2 \gamma_1}{\gamma_0} = \phi_1 + \phi_2 \rho_1$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{\phi_1 \gamma_1 + \phi_2 \gamma_0}{\gamma_0} = \phi_1 \rho_1 + \phi_2$$

las cuales son conocidas como ecuaciones de Yule-Walker, si damos valores de  $\phi_1$  y  $\phi_2$  podemos resolver para  $\rho_1$  y  $\rho_2$

El cálculo para autocorrelaciones de alto orden son calculadas de forma recursiva de la fórmula.

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} \quad j \geq 2$$

Es importante hacer notar que para ambos procesos  $AR(1)$  y  $AR(2)$  la función de autocorrelación sigue la misma relación.

### 2.7.3. PROCESO AUTORREGRESIVO DE ALTO ORDEN

Como una generalización de los procesos  $AR(1)$  y  $AR(2)$  ahora consideremos el caso de un proceso autoregresivo de or



den arbitrario  $p$ , esto es:

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

La estacionalidad de este proceso puede ser determinada como una generalización de las condiciones para AR(1) y AR(2), es decir, las raíces de la ecuación característica  $(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) = 0$  deben estar fuera del círculo unitario.

Para calcular sus momentos procedemos como sigue:

$$\mathbb{E}(Z_t) = \phi_1 \mathbb{E}(Z_{t-1}) + \dots + \phi_p \mathbb{E}(Z_{t-p}) + \eta + \mathbb{E}(\varepsilon_t)$$

$$\mu = \phi_1 \mu + \dots + \phi_p \mu + \eta$$

$$\eta = \mu (1 - \phi_1 - \phi_2 - \dots - \phi_p)$$

$$\mu = \frac{\eta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

Haciendo el cambio acostumbrado, esto es:

$$Z_t - \frac{\eta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

obtenemos que:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \eta \frac{(1 - \phi_1 - \dots - \phi_p)}{(1 - \phi_1 - \dots - \phi_p)} + \varepsilon_t$$

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \frac{\eta}{1 - \phi_1 - \dots - \phi_p} - \frac{\eta \phi_1}{(1 - \phi_1 - \dots - \phi_p)} - \dots - \frac{\eta \phi_p}{1 - \phi_1 - \dots - \phi_p} + \varepsilon_t$$

$$Z_t = \phi_1 \left[ Z_{t-1} - \frac{\eta}{(1 - \phi_1 - \dots - \phi_p)} \right] + \dots + \phi_p \left[ Z_{t-p} - \frac{\eta}{(1 - \phi_1 - \dots - \phi_p)} \right] + \frac{\eta}{(1 - \phi_1 - \phi_2 - \dots - \phi_p)} + \varepsilon_t$$

$$Z_t - \frac{\eta}{(1 - \phi_1 - \dots - \phi_p)} = \phi_1 \left[ Z_{t-1} - \frac{\eta}{(1 - \phi_1 - \dots - \phi_p)} \right] + \dots + \phi_p \left[ Z_{t-p} - \frac{\eta}{(1 - \phi_1 - \dots - \phi_p)} \right] + \varepsilon_t$$

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

por otro lado:

$$\begin{aligned} \gamma_1 &= E(\tilde{z}_t \tilde{z}_{t-1}) = E[(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \epsilon_t)(\tilde{z}_{t-1})] \\ &= \phi_1 E(\tilde{z}_{t-1}^2) + \dots + \phi_p E(\tilde{z}_{t-p} \tilde{z}_{t-1}) + E(\tilde{z}_{t-1} \epsilon_t) \\ &= \phi_1 \gamma_0 + \phi_2 \gamma_1 + \dots + \phi_p \gamma_{p-1} \end{aligned}$$

⋮

⋮

⋮

⋮

⋮

$$\begin{aligned} \gamma_p &= E(\tilde{z}_t \tilde{z}_{t-p}) = E[(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \epsilon_t)(\tilde{z}_{t-p})] \\ &= \phi_1 E(\tilde{z}_{t-1} \tilde{z}_{t-p}) + \dots + \phi_p E(\tilde{z}_{t-p}^2) \\ &= \phi_1 \gamma_{p-1} + \dots + \phi_p \gamma_0 \end{aligned}$$

y obtenemos un sistema de ecuaciones lineales, que puede ser resuelto conociendo los parámetros  $\phi_i$  y  $\sigma_\epsilon^2$ .

Para periodos de atraso más grandes que  $p$ , la **COVARIANZA** es calculada recursivamente, es decir,

$$\gamma_i = \phi_1 \gamma_{i-1} + \dots + \phi_p \gamma_{i-p} \quad i > p$$

Ahora como una extensión de los métodos utilizados en AR (1) y AR (2) para el cálculos de la función de autocorrelación, ésta puede ser calculada directamente dividiendo las

p ecuaciones lineales  $\gamma_1, \dots, \gamma_p$  entre  $\gamma_0$ , y así obtenemos las ecuaciones del Yule-Walker para AR (p).

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{1-p}$$

⋮

$$\rho_p = \phi_1 \rho_{p-1} + \dots + \phi_p$$

de igual manera, el cálculo para la función de autocorrelación para períodos de atraso más grandes que p obtenemos:

$$\rho_i = \phi_1 \rho_{i-1} + \dots + \phi_p \rho_{i-p} \quad i > p$$

#### 2.7.4. INVERTIBILIDAD DE LOS PROCESOS DE MEDIAS MOVILES

Esta es una interesante propiedad entre los procesos de Medias móviles y autorregresivo; por ejemplo, consideremos el proceso MA (1).

$$\tilde{z}_t = \epsilon_t - \theta_1 \epsilon_{t-1}$$

o utilizando el operador de cambio hacia atrás

$$\tilde{z} = (1 - \theta_1 B) \epsilon_t$$

cuya solución para  $\epsilon_t$  está dada por:

$$\epsilon_t = (1 - \theta_1 B)^{-1} \tilde{z}_t$$

Ahora si  $|\theta_1| < 1$  podemos escribir esta ecuación como

$$\epsilon_t = \left( \sum_{i=0}^{\infty} \theta_1^i B^i \right) \tilde{z}_t$$

o

$$\epsilon_t = (1 + \theta_1 B^1 + \theta_1^2 B^2 + \dots) \tilde{z}_t$$

el cual reconocemos como un proceso autorregresivo de orden infinito con ponderadores  $\phi_i = \theta_1^i$ , es entonces, que hemos invertido el proceso NA (1) para obtener un proceso AR ( $\infty$ ), por tanto la condición  $|\theta_1| < 1$  es conocida como condición de invertibilidad para el proceso MA (1).

Ahora, en general, para cualquier proceso MA (q) sea invertible a un proceso AR ( $\infty$ ) necesitamos que las raíces del polinomio  $\Theta_q(B) = 0$  estén fuera del círculo unitario, lo cual se traduce en el caso del proceso MA (2) en:

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$|\theta_2| < 1$$

\* Box y Jenkins  
Time Series Analysis

Es de hacer notar que las condiciones para los parámetros en el proceso MA (q) son exactamente las mismas condiciones de estacionalidad para un proceso AR(q).

Resumiendo:

1. El proceso MA (q) es estacionario sin hacer caso de los valores de los ponderadores  $\{\theta_i\}$  pero es invertible solamente si las raíces de  $\Theta_q(B) = 0$  están fuera del círculo unitario.
2. El proceso AR (p) es estacionario solamente si las raíces de  $\Phi_p(B) = 0$  se hallan fuera del círculo unitario, pero es invertible para todos los valores de  $\{\phi_i\}$ .

2.7.5. PROCESOS MIXTOS: AUTORREGRESIVO -MEDIAS MOVILES.

Para muchas series encontradas en la práctica, la inclusión de ambos términos: autorregresivos y medias móviles, resulta en un modelo que conducirá a un modelo más parsimonioso o que tiene menos parámetros, de los que hubiera sido necesario para un modelo satisfactorio de forma pura ya sea AR o MA.

Así, el modelo mixto autorregresivo-medias móviles de orden  $p, q$  es de la forma:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \eta + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

o en términos del operador de cambio hacia atrás.

$$\Phi_p(B) Z_t = \eta + \Theta_q(B) \epsilon_t$$

Las condiciones de estacionalidad e invertibilidad para este proceso son consecuencia de cada proceso, esto es, ARMA  $(p, q)$  es estacionario si las raíces de  $\Phi_p(B) = 0$  están fuera del círculo unitario, y es invertible si las raíces de  $\Theta_q(B) = 0$  están fuera del círculo unitario.

#### 2.7.6. PROCESO ARMA (1,1)

Un caso especial muy usual es el proceso ARMA (1,1) es decir,

$$Z_t = \phi_1 Z_{t-1} + \eta + \epsilon_t - \theta_1 \epsilon_{t-1}$$

el cual puede ser escrito en forma pura de medias móviles,

sustituyendo secuencialmente,  $Z_{t-1}, Z_{t-2} \dots$  esto es:

$$Z_{t-1} = \phi_1 Z_{t-2} + \rho + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2}$$

sustituyendo  $Z_{t-1}$  en  $Z_t$  obtenemos

$$\begin{aligned} Z_t &= \phi_1 [\phi_1 Z_{t-2} + \rho + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2}] + \rho + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ &= \phi_1^2 Z_{t-2} + \phi_1 \rho + \phi_1 \varepsilon_{t-1} - \phi_1 \theta_1 \varepsilon_{t-2} + \rho + \varepsilon_t - \theta_1 \varepsilon_{t-1} \end{aligned}$$

agrupando términos obtenemos:

$$= \phi_1^2 Z_{t-2} + (\rho + \phi_1 \rho) + \varepsilon_t + \varepsilon_{t-1} (\phi_1 - \theta_1) - \theta_1 \phi_1 \varepsilon_{t-2}$$

recursivamente, sustituyendo  $Z_{t-2}$  obtenemos:

$$Z_t = \phi_1^2 [\phi_1 Z_{t-3} + \rho + \varepsilon_{t-2} - \theta_1 \varepsilon_{t-3}] + (\rho + \phi_1 \rho) + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} - \theta_1 \phi_1 \varepsilon_{t-2}$$

$$Z_t = \phi_1^3 Z_{t-3} + \phi_1^2 \rho + \phi_1^2 \varepsilon_{t-2} - \phi_1^2 \theta_1 \varepsilon_{t-3} + (\rho + \phi_1 \rho) + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} - \theta_1 \phi_1 \varepsilon_{t-2}$$

agrupando términos:

$$Z_t = \phi_1^3 Z_{t-3} + (\rho + \phi_1 \rho + \phi_1^2 \rho) + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} + \phi_1 (\phi_1 - \theta_1) \varepsilon_{t-2} - \phi_1^2 \theta_1 \varepsilon_{t-3}$$

si sustituimos sucesivamente obtenemos finalmente

$$Z_t = \frac{\rho}{1 - \phi_1} + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} + \phi_1 (\phi_1 - \theta_1) \varepsilon_{t-2} + \phi_1^2 (\phi_1 - \theta_1) \varepsilon_{t-3}$$

Es claro que si el proceso es estacionario la suma de los coeficientes  $\sum_{i=0}^{\infty} \phi_1^i (\phi_1 - \theta_1)$  debe converger, esto es,  $|\phi_1| < 1$  como en el caso AR(1)

Es importante hacer notar que el proceso ARMA (1,1) es un proceso MA de orden infinito, pero podríamos tratar de aproximarlo con un proceso MA de orden finito eliminando los términos después del punto donde los coeficientes  $\phi_i^i (\phi_1 - \theta_1)$  se vuelvan más pequeños que una cantidad arbitraria. Es claro que un proceso de medias móviles de alto orden sería necesario para aproximar este proceso ARMA (1,1).

Esto es una importante ilustración del parsimonioso modelo realizado por el modelo mixto en el sentido de que el modelo mixto tiene solamente dos coeficientes, pero la aproximación mediante MA requiere muchos coeficientes.

De igual manera, el proceso ARMA (1,1) puede ser escrito en forma autorregresiva simplemente sustituyendo recursivamente  $\epsilon_{t-1}, \epsilon_{t-2}, \dots$  etc, esto es:

$$Z_t = \phi_1 Z_{t-1} + \eta + \epsilon_t - \theta_1 \epsilon_{t-1} \dots \text{I}$$

$$\Rightarrow \epsilon_t = Z_t + \theta_1 \epsilon_{t-1} - \eta - \phi_1 Z_{t-1}$$

$$\gamma \epsilon_{t-1} = Z_{t-1} + \theta_1 \epsilon_{t-2} - \eta - \phi_1 Z_{t-2}$$

Ahora sustituyendo  $\epsilon_{t-1}$  en I.

$$Z_t = \phi_1 Z_{t-1} + \eta + \epsilon_t - \theta_1 [Z_{t-1} + \theta_1 \epsilon_{t-2} - \eta - \phi_1 Z_{t-2}]$$

$$Z_t = \phi_1 Z_{t-1} + \eta + \epsilon_t - \theta_1 Z_{t-1} - \theta_1^2 \epsilon_{t-2} + \theta_1 \eta + \theta_1 \phi_1 Z_{t-2}$$

agrupando obtenemos:

$$Z_t = \epsilon_t + (\eta + \theta_1 \eta) + (\phi_1 - \theta_1) Z_{t-1} + \theta_1 \phi_1 Z_{t-2} - \theta_1^2 \epsilon_{t-2} \dots \text{II}$$

Ahora sustituyendo  $\varepsilon_{t-2}$ , obtenemos

$$Z_t = \varepsilon_t + (\gamma + \theta_1 \gamma) + (\phi_1 - \theta_1) Z_{t-1} + \theta_1 \phi_1 Z_{t-2} - \theta_1^2 [Z_{t-2} + \theta_1 \varepsilon_{t-3} - \gamma - \phi_1 Z_{t-3}]$$

$$Z_t = \varepsilon_t + (\gamma + \theta_1 \gamma) + (\phi_1 - \theta_1) Z_{t-1} + \theta_1 \phi_1 Z_{t-2} - \theta_1^2 Z_{t-2} + \theta_1^3 \varepsilon_{t-3} + \theta_1^3 \gamma + \theta_1^2 \phi_1 Z_{t-3}$$

$$Z_t = \varepsilon_t + (\gamma + \theta_1 \gamma + \theta_1^2 \gamma) + (\phi_1 - \theta_1) Z_{t-1} + \theta_1 (\phi_1 - \theta_1) Z_{t-2} + \theta_1^2 \phi_1 Z_{t-3} + \theta_1^3 \varepsilon_{t-3}$$

sustituyendo sucesivamente  $\varepsilon_{t-i}$  obtenemos

$$Z_t = (\phi_1 - \theta_1) Z_{t-1} + \theta_1 (\phi_1 - \theta_1) Z_{t-2} + \theta_1^2 (\phi_1 - \theta_1) Z_{t-3} + \dots + \frac{\gamma}{1 - \theta_1} + \varepsilon_t$$

Como en el caso anterior, invertibilidad requiere que la suma:

$$\sum_{i=1}^{\infty} \theta_1^{i-1} (\phi_1 - \theta_1) \text{ converja, esto es, que } |\theta_1| < 1$$

Esta forma autorregresiva de aproximar el proceso ARMA (1,1) sería obtenida por truncación de los coeficientes don de se volvieran lo suficientemente pequeños después de una cierta cantidad arbitraria.

Otra vez esto es otra importante ilustración del modelo parsimonioso ARMA(1,1) ya que la aproximación autorregresiva requerirá muchos coeficientes, comparados con los dos de ARMA (1,1).

Después de haber analizado el proceso ARMA (1,1), ahora procederemos a calcular sus momentos:

$$E(Z_t) = \phi_1 E(Z_{t-1}) + \gamma + E(\varepsilon_t) - \theta_1 E(\varepsilon_{t-1})$$

$$\mu = \phi_1 \mu + \gamma$$

$$\mu = \frac{\gamma}{1 - \phi_1}$$



$$\mathbb{E}(z_t) = \phi_1 \mathbb{E}(z_{t-1}) + \gamma + \mathbb{E}(\varepsilon_t) - \theta_1 \mathbb{E}(\varepsilon_{t-1})$$

$$\mu = \phi_1 \mu + \gamma$$

$$\mathbb{E}(z_t) = \mu = \frac{\gamma}{1 - \phi_1}$$

para calcular la varianza, recordemos la transformación

$$\tilde{z}_t = z_t - \frac{\gamma}{1 - \phi_1} \quad \text{entonces,} \quad \tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Así la varianza del proceso esta dada por:

$$\begin{aligned} \text{Var}(z_t) = \gamma_0 &= \mathbb{E}(\tilde{z}_t^2) = \mathbb{E}[\tilde{z}_t (\phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})] \\ &= \phi_1 \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) + \mathbb{E}(\tilde{z}_t \varepsilon_t) - \theta_1 \mathbb{E}(\tilde{z}_t \varepsilon_{t-1}) \end{aligned}$$

por otra parte

$$\begin{aligned} \mathbb{E}(\tilde{z}_t \varepsilon_{t-1}) &= \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) (\varepsilon_{t-1})] \\ &= \phi_1 \mathbb{E}(\tilde{z}_{t-1} \varepsilon_{t-1}) + \mathbb{E}(\varepsilon_t \varepsilon_{t-1}) - \theta_1 \mathbb{E}(\varepsilon_{t-1}^2) \\ &= \phi_1 \sigma_\varepsilon^2 - \theta_1 \sigma_\varepsilon^2 \end{aligned}$$

finalmente obtenemos que:

$$\text{Var}(z_t) = \gamma_0 = \mathbb{E}(\tilde{z}_t^2) = \phi_1 \gamma_0 + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2$$

Es de hacer notar que la Esperanza para el proceso ARMA (1,1) es exactamente de la misma que la del proceso AR (1), sin embargo en la varianza difiere ya que el factor  $\varepsilon_{t-1}$  es un componente de  $\tilde{z}_t$ . Ahora la autocovarianza de período de atraso 1 es:

$$\begin{aligned} \gamma_1 &= \mathbb{E}(\tilde{z}_t \tilde{z}_{t-1}) = \mathbb{E}[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) (\tilde{z}_{t-1})] \\ &= \phi_1 \mathbb{E}(\tilde{z}_{t-1}^2) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-1}) - \theta_1 \mathbb{E}(\varepsilon_{t-1} \tilde{z}_{t-1}) \\ &= \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2 \end{aligned}$$

que también difiere de AR (1) porque  $\varepsilon_{t-1}$  es componente de  $\tilde{z}_t$ , de igual manera la autocovarianza de período de atraso 2 es:

$$\begin{aligned}\gamma_2 &= E(\tilde{z}_t \tilde{z}_{t-2}) = E[(\phi_1 \tilde{z}_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) \tilde{z}_{t-2}] \\ &= \phi_1 E(\tilde{z}_{t-1} \tilde{z}_{t-2}) + E(\tilde{z}_{t-2} \varepsilon_t) - \theta_1 E(\varepsilon_{t-1} \tilde{z}_{t-2}) \\ &= \phi_1 \gamma_1\end{aligned}$$

Entonces podemos deducir que también las autocovarianzas, de períodos ----- de atraso mas grandes que 1 son idénticas a las del modelo AR (1) porque la componente de medias móviles del ARMA (1,1) llega hasta solamente un período de atraso. Entonces para períodos de atraso mayores que 1 obtenemos:

$$\gamma_i = \phi_1 \gamma_{i-1} \quad i=2,3,\dots$$

Para calcular las autocovarianzas dados los parámetros del proceso simplemente sustituimos  $\gamma_i$  en  $\gamma_0$  y viceversa, esto es:

$$\begin{aligned}\gamma_0 &= \phi_1 [\phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2] + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2 \\ &= \phi_1^2 \gamma_0 - \phi_1 \theta_1 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 - \theta_1 \phi_1 \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 \\ &= \phi_1^2 \gamma_0 - 2\phi_1 \theta_1 \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= \phi_1^2 \gamma_0 + \sigma_\varepsilon^2 (-2\phi_1 \theta_1 + \theta_1^2 + 1) \\ \gamma_0 - \phi_1^2 \gamma_0 &= \sigma_\varepsilon^2 (-2\phi_1 \theta_1 + \theta_1^2 + 1) \\ \gamma_0 (1 - \phi_1^2) &= \sigma_\varepsilon^2 (-2\phi_1 \theta_1 + \theta_1^2 + 1) \\ \gamma_0 &= \frac{(-2\phi_1 \theta_1 + \theta_1^2 + 1) \sigma_\varepsilon^2}{1 - \phi_1^2}\end{aligned}$$

$$\begin{aligned}
\gamma_1 &= \phi_1 [\phi_1 \gamma_1 + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2] - \theta_1 \sigma_\varepsilon^2 \\
&= \phi_1^2 \gamma_1 + \phi_1 \sigma_\varepsilon^2 - \theta_1 \phi_1^2 \sigma_\varepsilon^2 + \phi_1 \theta_1^2 \sigma_\varepsilon^2 - \theta_1 \sigma_\varepsilon^2 \\
&= \phi_1^2 \gamma_1 + \sigma_\varepsilon^2 (\phi_1 - \theta_1 \phi_1^2 + \phi_1 \theta_1^2 - \theta_1) \\
&= \phi_1^2 \gamma_1 + \sigma_\varepsilon^2 [(1 - \phi_1 \theta_1) (\phi_1 - \theta_1)] \\
\gamma_1 - \phi_1^2 \gamma_1 &= \sigma_\varepsilon^2 [(1 - \phi_1 \theta_1) (\phi_1 - \theta_1)] \\
\gamma_1 (1 - \phi_1^2) &= \sigma_\varepsilon^2 [(1 - \phi_1 \theta_1) (\phi_1 - \theta_1)] \\
\gamma_1 &= \frac{(1 - \phi_1 \theta_1) (\phi_1 - \theta_1) \sigma_\varepsilon^2}{1 - \phi_1^2}
\end{aligned}$$

y sabemos que

$$\gamma_i = \phi_1 \gamma_{i-1} \quad i \geq 2$$

Entonces la función de autocorrelación de período de atraso 1 esta dada por:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi_1 \theta_1) (\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1 \theta_1}$$

y para  $i \geq 2$

$$\rho_i = \phi_1 \rho_{i-1}$$

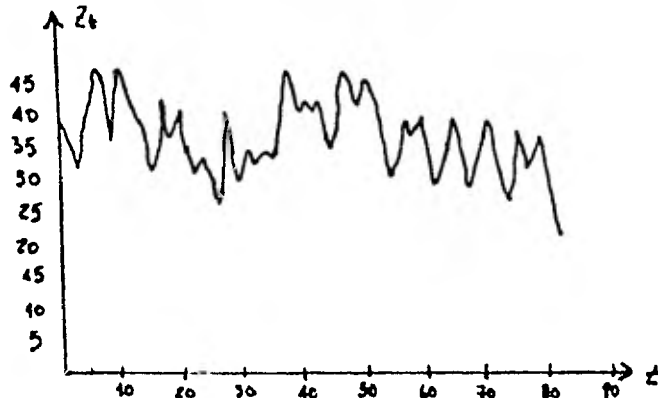
Es entonces importante notar que las componentes de medias móviles entran solamente en la determinación de  $\rho_1$ , la parte restante de la función de autocorrelación esta determinada solamente por la parte autoregresiva del modelo, esto es, la función de autocorrelación decae exponencialmente de  $\rho_1$ .

Como un ejemplo, consideremos el modelo

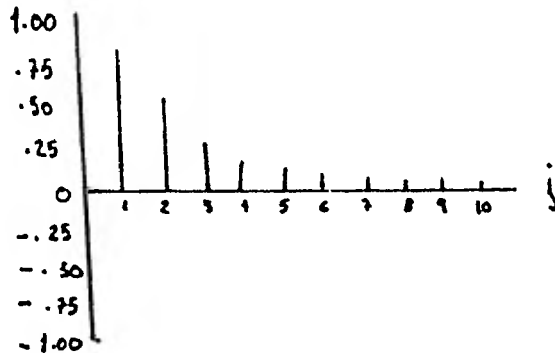
$$Z_t = 10 + 0.6Z_{t-1} + \varepsilon_t + 0.9\varepsilon_{t-1} \quad \text{con } \sigma_{\varepsilon}^2 = 4$$

cuya gráfica es:

PROCESO  
ARMA(1,1)



y que tiene como función de autocorrelación



En esta gráfica podemos ver que el efecto que produce el término de medias móviles altera el valor de  $\rho_1$ , pero a partir de ahí la función de autocorrelación decae exponencialmente como si se tratara de un proceso AR(1) con parámetro  $\phi_1 = 0.6$

### 2.7.7 PROCESOS MIXTOS DE ALTO ORDEN

Como una generalización de proceso ARMA (1,1) podemos pensar en el modelo ARMA (p,q). Esto es, cualquier proceso ARMA puede ser escrito como un proceso MA de orden infinito cuya condición de estacionalidad es satisfecha si las raíces de la ecuación característica  $(1 - \phi_1 B \dots - \phi_p B^p) = 0$  se hallan fuera del círculo unitario. De igual manera, el proceso ARMA (p,q) puede ser escrito en forma autoregresiva de orden infinito y cuya condición de invertibilidad es satisfecha si las raíces de la ecuación característica  $(1 - \theta_1 B - \dots - \theta_q B^q) = 0$  se hallan fuera del círculo unitario.

El modelo es entonces de la forma:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \eta + \epsilon_t - \dots - \theta_q \epsilon_{t-q}$$

y la esperanza del proceso es:

$$E(Z_t) = \phi_1 E(Z_{t-1}) + \dots + \phi_p E(Z_{t-p}) + \eta + E(\epsilon_t) - \dots - \theta_q E(\epsilon_{t-q})$$

$$E(Z_t) = \phi_1 \mu + \dots + \phi_p \mu + \eta$$

$$\mu (1 - \phi_1 - \dots - \phi_p) = \eta$$

$$\mu = \frac{\eta}{1 - \phi_1 - \dots - \phi_p}$$

para calcular la varianza, haremos nuevamente uso de la transformación

$$\tilde{Z}_t = Z_t - \frac{\eta}{1 - \phi_1 - \dots - \phi_p}$$

entonces

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

Así

$$\begin{aligned} \text{Var}(z_t) = \gamma_0 &= E(\tilde{z}_t^2) = E\left[\tilde{z}_t(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p}) + \varepsilon_t - \dots - \theta_q \varepsilon_{t-q}\right] \\ &= \phi_1 E(\tilde{z}_t \tilde{z}_{t-1}) + \phi_2 E(\tilde{z}_t \tilde{z}_{t-2}) + \dots + E(\tilde{z}_t \varepsilon_t) - \theta_1 E(\tilde{z}_t \varepsilon_{t-1}) - \dots - \theta_q E(\tilde{z}_t \varepsilon_{t-q}) \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_\varepsilon^2 - \theta_1 E(\tilde{z}_t \varepsilon_{t-1}) - \dots - \theta_q E(\tilde{z}_t \varepsilon_{t-q}) \end{aligned}$$

por otra parte

$$\begin{aligned} E(\tilde{z}_t \varepsilon_{t-1}) &= E\left[(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-1})\right] \\ &= \phi_1 E(\tilde{z}_{t-1} \varepsilon_{t-1}) + \dots + \phi_p E(\tilde{z}_{t-p} \varepsilon_{t-1}) + E(\varepsilon_t \varepsilon_{t-1}) - \dots - \theta_q E(\varepsilon_{t-q} \varepsilon_{t-1}) \\ &= \phi_1 \sigma_\varepsilon^2 - \theta_1 \sigma_\varepsilon^2 = (\phi_1 - \theta_1) \sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} E(\tilde{z}_t \varepsilon_{t-2}) &= E\left[(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}) \varepsilon_{t-2}\right] \\ &= \phi_1 E(\tilde{z}_{t-1} \varepsilon_{t-2}) + \phi_2 E(\tilde{z}_{t-2} \varepsilon_{t-2}) + \dots + E(\varepsilon_t \varepsilon_{t-2}) - \dots - \theta_2 E(\varepsilon_{t-2}^2) - \theta_q E(\varepsilon_{t-q} \varepsilon_{t-2}) \\ &= \phi_1 [(\phi_1 - \theta_1) \sigma_\varepsilon^2] + (\phi_2 - \theta_2) \sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} E(\tilde{z}_t \varepsilon_{t-3}) &= E\left[(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-3})\right] \\ &= \phi_1 E(\tilde{z}_{t-1} \varepsilon_{t-3}) + \phi_2 E(\tilde{z}_{t-2} \varepsilon_{t-3}) + \phi_3 E(\tilde{z}_{t-3} \varepsilon_{t-3}) + \dots \\ &\quad \dots + E(\varepsilon_t \varepsilon_{t-3}) - \dots - \theta_3 E(\varepsilon_{t-3}^2) - \dots - \theta_q E(\varepsilon_{t-q} \varepsilon_{t-3}) \\ &= \phi_1 [\phi_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2 + (\phi_2 - \theta_2) \sigma_\varepsilon^2] + \phi_2 [(\phi_1 - \theta_1) \sigma_\varepsilon^2] + \\ &\quad + (\phi_3 - \theta_3) \sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\tilde{z}_t \varepsilon_{t-4}) &= \mathbb{E}\left[\left(\phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}\right) \varepsilon_{t-4}\right] \\
&= \phi_1 \mathbb{E}(\tilde{z}_{t-1} \varepsilon_{t-4}) + \phi_2 \mathbb{E}(\tilde{z}_{t-2} \varepsilon_{t-4}) + \phi_3 \mathbb{E}(\tilde{z}_{t-3} \varepsilon_{t-4}) + \\
&\quad + \phi_4 \mathbb{E}(\tilde{z}_{t-4} \varepsilon_{t-4}) + \dots + \mathbb{E}(\varepsilon_t \varepsilon_{t-4}) - \dots - \\
&\quad - \theta_4 \mathbb{E}(\varepsilon_{t-4}^2) - \dots - \theta_q \mathbb{E}(\varepsilon_{t-q} \varepsilon_{t-4})
\end{aligned}$$

$$\begin{aligned}
&= \phi_1 \left\{ \phi_1 \left\{ \phi_1 \left[ (\phi_1 - \theta_1) \sigma_\varepsilon^2 \right] + (\phi_2 - \theta_2) \sigma_\varepsilon^2 \right\} + \right. \\
&\quad \left. + \phi_2 \left\{ (\phi_1 - \theta_1) \sigma_\varepsilon^2 \right\} + (\phi_3 - \theta_3) \sigma_\varepsilon^2 \right\} + \\
&\quad + \phi_2 \left\{ \phi_1 \left[ \phi_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2 \right] + (\phi_2 - \theta_2) \sigma_\varepsilon^2 \right\} + \\
&\quad + \phi_3 \left\{ (\phi_1 - \theta_1) \sigma_\varepsilon^2 \right\} + (\phi_4 - \theta_4) \sigma_\varepsilon^2
\end{aligned}$$

En general: 
$$\mathbb{E}(\tilde{z}_t \varepsilon_{t-q}) = \phi_1 \mathbb{E}(\tilde{z}_t \varepsilon_{t-(q-1)}) + \phi_2 \mathbb{E}(\tilde{z}_t \varepsilon_{t-(q-2)}) + \phi_3 \mathbb{E}(\tilde{z}_t \varepsilon_{t-(q-3)}) + \dots + \phi_p \mathbb{E}(\tilde{z}_t \varepsilon_{t-(q-p)}) + (\phi_p - \theta_p) \sigma_\varepsilon^2$$

finalmente la varianza estará dada por

$$\text{Var}(z_t) = \gamma_0 = \mathbb{E}(\tilde{z}_t^2) = \phi_1^2 \gamma_1 + \phi_2^2 \gamma_2 + \dots + \phi_p^2 \gamma_p - \theta_1 \mathbb{E}(\tilde{z}_t \varepsilon_{t-1}) - \dots - \theta_q \mathbb{E}(\tilde{z}_t \varepsilon_{t-q})$$

Este cálculo de la varianza, podemos representarlo dentro de una fórmula que a la vez nos ayudará para representar las autocovarianzas de orden  $j$ , es decir,

$$\gamma_j = \mathbb{E}(\tilde{z}_t \tilde{z}_{t-j}) = \phi_1 \mathbb{E}(\tilde{z}_{t-1} \tilde{z}_{t-j}) + \dots + \phi_p \mathbb{E}(\tilde{z}_{t-p} \tilde{z}_{t-j}) + \mathbb{E}(\varepsilon_t \tilde{z}_{t-j}) - \theta_1 \mathbb{E}(\varepsilon_{t-1} \tilde{z}_{t-j}) - \dots - \theta_q \mathbb{E}(\varepsilon_{t-q} \tilde{z}_{t-j})$$

si quisieramos la representación de la varianza, simplemente consideremos  $j=0$ , y obtendremos la representación antes mencionada y cuyos cálculos ya hemos desarrollado.

Es de hacer notar que si  $j < q$ , entonces los términos involucran  $\tilde{z}_{t-j}$  y los errores serán distintos de cero, es decir,  $\tilde{z}_{t-j}$  está correlacionado con todos los errores que ocurren en el período  $t-j$ , un ejemplo claro de esto, es cuando calculamos las  $\mathbb{E}(\tilde{z}_{t-j} \varepsilon_{t-q})$ , note que el caso  $j=0$  y  $q=4$  ( $\mathbb{E}(\tilde{z}_t \varepsilon_{t-4})$ ,  $j < q$ ) todos los errores que ocurren en el período menor a  $t-4$  están correlacionados, es decir  $\varepsilon_{t-4}, \varepsilon_{t-3}, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t$

De igual manera si  $j > q$ , sabemos que el proce-



so se comporta como un proceso autoregresivo cuyas covarianzas sabemos que son de la forma:

$$\gamma_i = \phi_1 \gamma_{i-1} + \dots + \phi_p \gamma_{i-p} \quad i > p$$

y la función de autocorrelación

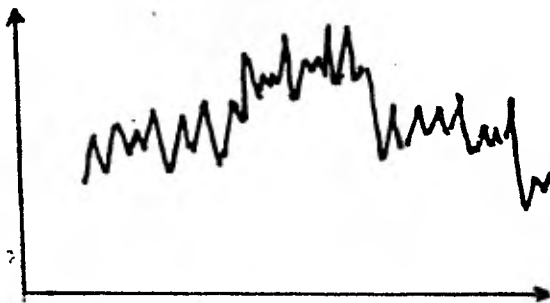
$$\rho_i = \phi_1 \rho_{i-1} + \dots + \phi_p \rho_{i-p} \quad i > p$$

CAPITULO III

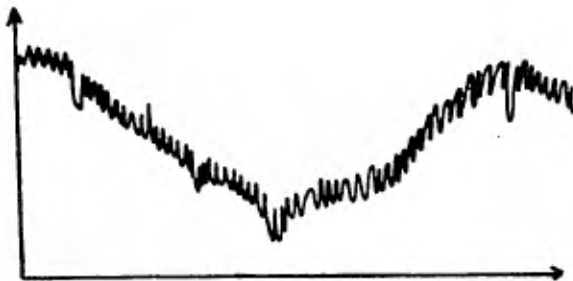
### 3. MODELOS PARA SERIES DE TIEMPO NO ESTACIONARIAS

Muchas series de tiempo encontradas en la práctica tales como las que se presentan en economía, como son: los precios de la bolsa de valores, el producto nacional bruto, ventas de una compañía o gasto público, no presentan afinidad en el valor de la media, es decir, en cualquier segmento local del tiempo aparecen observaciones junto al valor medio y en cualquier otro segmento del tiempo aparecen observaciones separadas del valor medio, tales series son conocidas como no estacionarias en la media. De igual manera, es frecuente encontrar series no estacionarias tanto en la media como en la pendiente, esto es, series sin afinidad en el valor de la media y sin tendencia definida.

Para ilustrar los dos casos mencionados, veamos las siguientes gráficas:



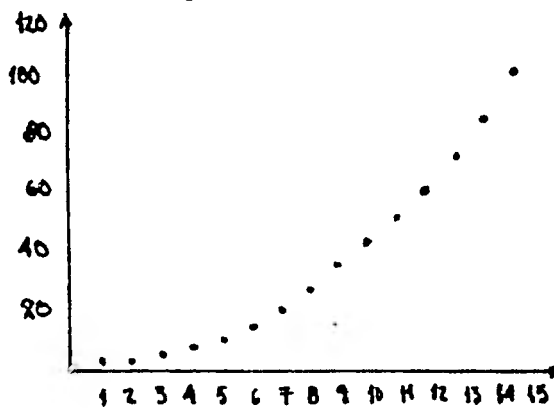
SERIE DE TIEMPO NO ESTACIONARIA EN LA MEDIA.



SERIE DE TIEMPO NO ESTACIONARIA EN LA MEDIA Y PENDIENTE

El problema es ahora, ¿Qué modelo o modelos pueden explicar "satisfactoriamente" tales series? Antes de responder a esta pregunta, analizaremos como es el patrón de no estacionalidad.

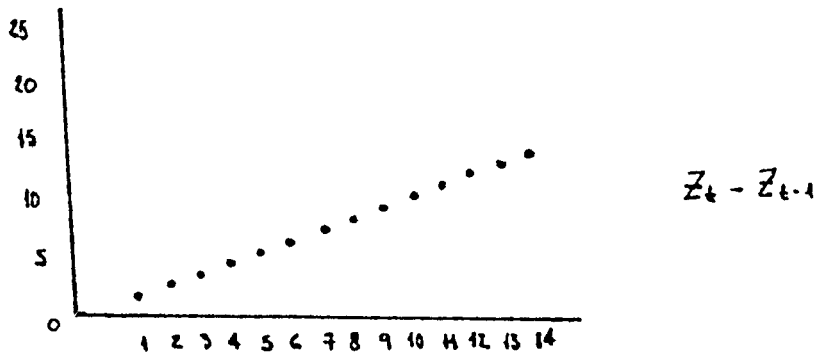
La forma en que se presenta la no estacionalidad en tales series puede ser pensada como homogénea en el sentido que aunque las series se mueven libremente sin afinidad para un particular punto, su comportamiento a diferentes períodos en el tiempo, es esencialmente el mismo y en consecuencia podemos pensar que, la no estacionalidad homogénea proviene de series cuyas diferencias sucesivas son estacionarias; para aclarar esta idea, consideramos el caso discreto y determinístico en el siguiente diagrama



Esta serie de tiempo presenta un comportamiento no estacionario tanto en la media como en la pendiente, sin embargo po

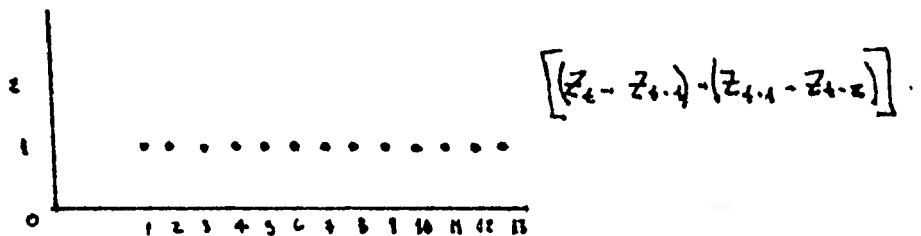
demos considerar dicho comportamiento como homogéneo.

Ahora consideremos las diferencias  $Z_t - Z_{t-1}$  en la siguiente gráfica:



Esta serie de tiempo es no estacionaria solamente en la media:

De igual manera, consideremos ahora las segundas diferencias  $[(Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2})] = Z_t - 2Z_{t-1} + Z_{t-2}$  mostradas en la siguiente gráfica



Finalmente esta serie es estacionaria.

Entonces ahora podemos dar respuesta a nuestra pregunta planteada inicialmente; la forma en que trabajaremos las series no estacionarias, será tratar de incorporar estas series

mediante las diferencias sucesivas , dentro del marco de series estacionarias y hacer uso de los modelos desarrollados para éstas simplemente trabajando con sus diferencias.

Así mismo, podemos concluir que si una serie de tiempo no estacionaria puede ser reducida a una serie estacionaria mediante las diferencias sucesivas, diremos que la serie original es homogéneamente no estacionaria.

Para empezar con el problema de incorporar las series no estacionarias a estacionarias, consideremos el caso del proceso autorregresivo de primer orden

$$Z_t = \phi_1 Z_{t-1} + \varepsilon_t$$

Sabemos que si  $|\phi_1| < 1$  el proceso es estacionario, como ya lo hemos estudiado ; sin embargo, ahora si consideramos el caso  $\phi_1 = 1$  obtendremos el modelo  $Z_t = Z_{t-1} + \varepsilon_t$  que es conocido como caminata aleatoria y que es no estacionaria ya que cada cambio sucesivo en la variable  $Z_t$  es independiente de una distribución de probabilidad con media cero, es decir, podemos representar el proceso de la siguiente forma

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

donde  $\varepsilon_t$  es una variable aleatoria con media cero e

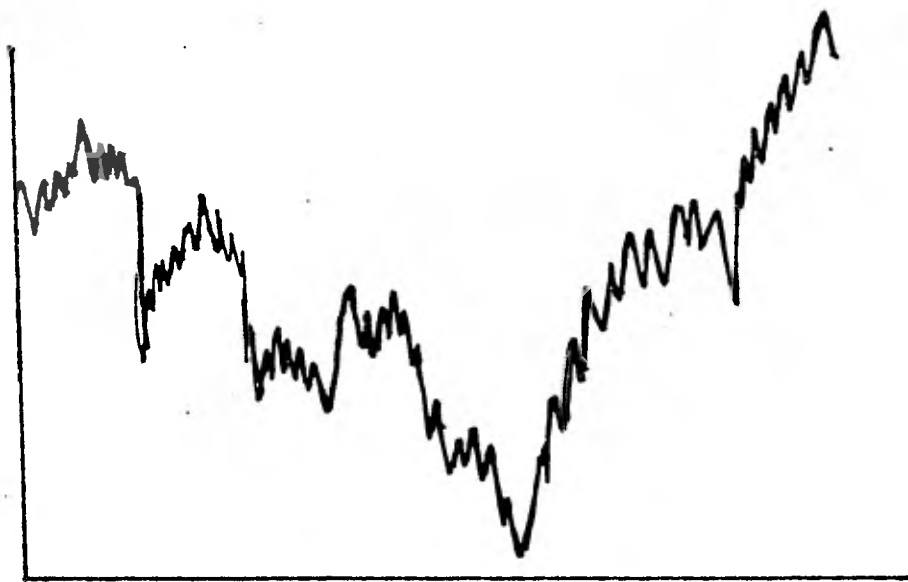
independiente en cada periodo  $t$  esto produce que cada cambio en  $Z_t$  sea aleatorio. Así, podemos pensar que  $\varepsilon_t$  es generada por el lanzamiento de una moneda con

$$\varepsilon_t = \begin{cases} +1 & \text{con probabilidad } 1/2 \\ -1 & \text{con probabilidad } 1/2 \end{cases}$$

esto es, si empezamos el proceso en alguno  $Z_0$ , obtendremos

$$\begin{aligned} Z_1 &= Z_0 + \varepsilon_1 \\ Z_2 &= Z_0 + \varepsilon_1 + \varepsilon_2 \\ &\vdots \\ Z_t &= Z_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t \end{aligned}$$

Una gráfica representativa de dicho proceso, bajo la suposición de que los  $\varepsilon_t$  se distribuyen como una normal, es la siguiente:



Aquí podemos observar claramente que dicho proceso no es estacionario; sin embargo, podemos observar que es homogéneo en su comportamiento, ya que la distribución de cambios o diferencias en el proceso es invariante, esto es, la serie de

tiempo o diferencias es estacionaria, porque las diferencias son solamente  $Z_t - Z_{t-1} = \epsilon_t$  y la distribución de  $\epsilon_t$  es fija.

Ahora, sabiendo ya que estas diferencias son estacionarias, podemos incorporar dichas diferencias dentro de una clase de procesos ya conocidos como son los procesos ARMA, definiendo.

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

entonces el modelo sería de la forma:

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

o escribiendo la serie directamente

$$Z_t = Z_{t-1} + \phi_1 (Z_{t-1} - Z_{t-2}) + \dots + \phi_p (Z_{t-p} - Z_{t-p-1}) + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

Así este proceso es conocido como proceso integrado autorregresivo medias móviles ARIMA. Pero como sabemos que en algunos casos las primeras diferencias pueden ser no estacionarias (la caminata aleatoria no es el caso) entonces las segundas diferencias pueden ya serlo; y como ya sabemos, las segundas diferencias son las diferencias de las primeras diferencias.

Esto es, podemos definir

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

$$y_t = [(Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2})] = Z_t - 2Z_{t-1} + Z_{t-2}$$

Si denotamos el grado de diferencia por  $d$ , entonces el



proceso ARIMA puede ser descrito por las dimensiones p,d y q esto es, ARIMA (p,d,q) así mismo un proceso ARIMA que no contenga la parte de medias móviles podemos representarlo como ARI (p,d) y de igual manera si el proceso ARIMA no contiene la parte autorregresiva podemos representarlo como IMA (d,q)

Si definimos ahora el operador diferencia  $\nabla$  como:

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

es posible expresar dicho operador en términos del operador de cambio hacia atrás B como:

$$\nabla = 1 - B$$

De esto, cualquier grado de las diferencias puede ser expresado bajo esta igualdad, por ejemplo, consideremos el grado de diferencias  $d=1$  y  $d=2$

si  $d=1$  obtenemos

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

si  $d=2$  obtenemos

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

que es exactamente lo mismo que habíamos encontrado anteriormente cuando habíamos definido  $y_t = W_t - W_{t-1}$

En general.

$$\nabla^d = (1 - B)^d$$

lo cual nos permite representar los procesos ARIMA (p,d,q) como

$$\Phi_p(B) \nabla^d z_t = \Theta_q(B) \varepsilon_t$$

o

$$\Phi_p(B) w_t = \Theta_q(B) \varepsilon_t$$

veamos el desarrollo para los modelos ARIMA (1,1,1) y ARIMA (2,1,0), ARIMA (1,1,1),

esto es:

$$(1 - \phi_1 B) \nabla z_t = (1 - \theta_1 B) \varepsilon_t$$

$$\nabla z_t - \phi_1 B \nabla z_t = \varepsilon_t - \theta_1 B \varepsilon_t$$

$$z_t - z_{t-1} - \phi_1 B (z_t - z_{t-1}) = \varepsilon_t - \theta_1 B \varepsilon_t$$

$$z_t - z_{t-1} - \phi_1 z_{t-1} + \phi_1 z_{t-2} = \varepsilon_t - \theta_1 B \varepsilon_t$$

$$z_t = z_{t-1} + \phi_1 z_{t-1} - \phi_1 z_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$z_t = (1 + \phi_1) z_{t-1} - \phi_1 z_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

note que este modelo lo podemos ver como un caso particular del proceso ARMA (2,1) con  $\phi_1^* = (1 + \phi_1)$  y  $\phi_2^* = -\phi_1$  el cual no satisface la condición de estacionalidad.

ARIMA (2,1,0)

$$(1 - \phi_1 B - \phi_2 B^2) \nabla z_t = \varepsilon_t$$

$$\nabla z_t - \phi_1 B \nabla z_t - \phi_2 B^2 \nabla z_t = \varepsilon_t$$

$$\begin{aligned}
Z_t - Z_{t-1} - \phi_1 B(Z_t - Z_{t-1}) - \phi_2 B^2(Z_t - Z_{t-1}) &= \epsilon_t \\
Z_t - Z_{t-1} - \phi_1 B Z_t + \phi_1 B Z_{t-1} - \phi_2 B^2 Z_t + \phi_2 B^2 Z_{t-1} &= \epsilon_t \\
Z_t - Z_{t-1} - \phi_1 Z_{t-1} + \phi_1 Z_{t-2} - \phi_2 Z_{t-2} + \phi_2 Z_{t-3} &= \epsilon_t \\
Z_t = Z_{t-1} + \phi_1 Z_{t-1} - \phi_1 Z_{t-2} + \phi_2 Z_{t-2} - \phi_2 Z_{t-3} + \epsilon_t \\
Z_t = Z_{t-1} + \phi_1 (Z_{t-1} - Z_{t-2}) + \phi_2 (Z_{t-2} - Z_{t-3}) + \epsilon_t
\end{aligned}$$

Es importante hacer notar que el término  $\gamma$  en todos los procesos ARIMA no aparece; veamos porque,

En ausencia del término constante, la media del proceso estacionario generado por las diferencias  $W_t$  es cero, esto es: consideremos el proceso ARIMA (p,1,q)

$$\begin{aligned}
W_t &= \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \\
E(W_t) &= \phi_1 E(W_{t-1}) + \dots + \phi_p E(W_{t-p}) + E(\epsilon_t) - \theta_1 E(\epsilon_{t-1}) - \dots - \theta_q E(\epsilon_{t-q})
\end{aligned}$$

como las diferencias son estacionarias

$$\begin{aligned}
\mu &= \phi_1 \mu + \dots + \phi_p \mu \\
0 &= \phi_1 \mu + \dots + \phi_p \mu - \mu \\
0 &= (1 - \phi_1 - \dots - \phi_p) \mu \\
\Rightarrow \mu &= 0
\end{aligned}$$

Si analizamos el proceso sin tomar las diferencias, el hecho de que no aparezca el término constante  $\gamma$  significa que aunque el proceso no presente afinidad para un valor medio, tampoco presentará tendencias ya sea en dirección positi

va o negativa.

Veamos ahora el caso en que aparezca el término constante  $\rho$ . Así la media del proceso de las diferencias  $W_t$  está dada por:

sea el modelo:

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \rho$$

entonces

$$\mathbb{E}(W_t) = \phi_1 \mathbb{E}(W_{t-1}) + \dots + \phi_p \mathbb{E}(W_{t-p}) + \mathbb{E}(\varepsilon_t) - \theta_1 \mathbb{E}(\varepsilon_{t-1}) - \dots - \theta_q \mathbb{E}(\varepsilon_{t-q}) + \rho$$

$$\mu = \phi_1 \mu + \dots + \phi_p \mu + \rho$$

$$\mu - \phi_1 \mu - \dots - \phi_p \mu = \rho$$

$$\mu(1 - \phi_1 - \dots - \phi_p) = \rho$$

$$\mu = \frac{\rho}{1 - \phi_1 - \dots - \phi_p}$$

lo cual significa que la diferencia promedio sobre un largo período de tiempo será distinta de cero. Por ejemplo, consideremos el caso  $d=1$ , si es positiva, entonces la diferencia promedio será positiva y la serie  $Z_t$  tendrá tendencia positiva, aunque hay que tener cuidado en el sentido de que la presencia del término constante no implica que la serie siga un determinado camino a través del tiempo.

### 3.1. FORMA DE LA ECUACION DIFERENCIA DE LOS PROCESOS ARIMA.

La ecuación diferencia, es simplemente una forma de reescribir los procesos ARIMA en términos de las observaciones pasadas y errores pasados de tal manera que facilite los cálculos para predecir mediante métodos computacionales. Esto

es:

Nosotros, originalmente habíamos escrito la serie de la siguiente forma

$$Z_t = Z_{t-1} + \phi_1 (Z_{t-1} - Z_{t-2}) + \dots + \phi_p (Z_{t-p} - Z_{t-p-1}) + \epsilon_t - \dots - \theta_q \epsilon_{t-q}$$

Si ahora factorizamos las observaciones pasadas, obtenemos:

$$Z_t = (1 + \phi_1)Z_{t-1} + (\phi_2 - \phi_1)Z_{t-2} + \dots + (\phi_p - \phi_{p-1})Z_{t-p} - \phi_p Z_{t-p-1} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

Esta forma de presentar el proceso es conocido como la ecuación diferencia del proceso ARIMA (p,1,q).

Veamos la ecuación diferencia de proceso ARIMA (0,1,1) ó IMA (1,1),

$$Z_t = Z_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

la forma de ver que este proceso es no estacionaria es simplemente, notando que  $\phi_1 = 1$

### 3.2. FORMA: CAMBIO ALEATORIO DE LOS PROCESOS ARIMA

Esta es otra forma de escribir los procesos ARIMA, en términos de los errores pasados, solamente

$$Z_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

Dicha forma es conocida como cambio aleatorio del proceso; la forma de obtener esta expresión, es simplemente sustituyendo las observaciones pasadas usando la ecuación diferencia del proceso.

Para ilustrar el proceso, consideremos el proceso IMA

$$(1,1) \quad Z_t = Z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \dots A$$

por otro lado  $Z_{t-1} = Z_{t-2} + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} \dots B$

ahora si sustituimos  $Z_{t-1}$  en (A) obtenemos

$$Z_t = (Z_{t-2} + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2}) + \varepsilon_t - \theta_1 \varepsilon_{t-1} \dots C$$

si sustituimos  $Z_{t-2}$  en (C)

$$\begin{aligned} Z_t &= [(Z_{t-3} + \varepsilon_{t-2} - \theta_1 \varepsilon_{t-3}) + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2}] + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ &= \varepsilon_t + (1-\theta_1)\varepsilon_{t-1} + (1-\theta_1)\varepsilon_{t-2} - \theta_1 \varepsilon_{t-3} + Z_{t-3} \end{aligned}$$

Si sustituimos  $Z_{t-i}$  sucesivamente finalmente obtenemos

$$Z_t = \varepsilon_t + (1-\theta_1)\varepsilon_{t-1} + (1-\theta_1)\varepsilon_{t-2} + (1-\theta_1)\varepsilon_{t-3} + \dots$$

entonces los  $\psi_i$  en este caso serán de la forma

$$\psi_i = (1-\theta_1) \quad i=1,2,\dots$$

que claramente es no estacionario.

De esta forma de expresar los procesos ARIMA, surge el concepto de nivel que es definido como:

$$\bar{Z}_t = (1-\theta_1)\varepsilon_{t-1} + (1-\theta_1)\varepsilon_{t-2} + \dots$$

entonces podemos expresar  $Z_t$  en función del nivel más el error, esto es:

$$Z_t = \bar{Z}_t + \varepsilon_t$$

Note que el nivel es por sí mismo un proceso no estacionario ya que

$$\bar{Z}_t = \bar{Z}_{t-1} + (1-\theta_1)\varepsilon_{t-1}$$

que podemos interpretarlo como la caminata aleatoria, simplemente tomando

$$(1-\theta_1)\varepsilon_{t-1} = \varepsilon_t'$$

#### 4.3. FORMA INVERTIDA DEL PROCESO ARIMA

Para expresar el proceso ARIMA en forma invertida, partimos de la ecuación diferencia del proceso y una vez así, sustituimos sucesivamente los  $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$  de tal forma que  $Z_t$  pueda ser expresada solamente en términos del error  $\varepsilon_t$  y las observaciones pasadas, esto es:

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + \varepsilon_t$$

Así, esta forma es conocida como forma invertida del proceso. Como ejemplo, consideremos el proceso IMA (1,1).

$$Z_t = Z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

por otro lado

$$\varepsilon_{t-1} = Z_{t-1} - Z_{t-2} + \theta_1 \varepsilon_{t-2}$$

si sustituimos  $\varepsilon_{t-1}$  en la ecuación original

$$Z_t = Z_{t-1} + \varepsilon_t - \theta_1 [Z_{t-1} - Z_{t-2} + \theta_1 \varepsilon_{t-2}]$$

$$Z_t = (1 - \theta_1) Z_{t-1} + \theta_1 Z_{t-2} - \theta_1^2 \varepsilon_{t-2} + \varepsilon_t$$

de igual manera para  $\varepsilon_{t-2}$

$$Z_t = (1 - \theta_1) Z_{t-1} + \theta_1 Z_{t-2} - \theta_1^2 [(Z_{t-2} - Z_{t-3}) + \theta_1 \varepsilon_{t-3}] + \varepsilon_t$$

$$Z_t = (1 - \theta_1) Z_{t-1} + \theta_1 Z_{t-2} - \theta_1 (1 - \theta_1) Z_{t-2} - \theta_1^2 Z_{t-3} - \theta_1^3 \varepsilon_{t-3} + \varepsilon_t$$

si sustituimos sucesivamente  $\varepsilon_{t-i}$ , finalmente obtenemos:

$$Z_t = (1 - \theta_1) Z_{t-1} + \theta_1 (1 - \theta_1) Z_{t-2} + \theta_1^2 (1 - \theta_1) Z_{t-3} + \dots + \varepsilon_t$$

entonces las  $\pi_i$  estarán dadas por

$$\pi_i = \theta_1^{i-1} (1 - \theta_1) \quad i = 1, 2, \dots$$

CAPITULO IV



#### 4. IDENTIFICACION, ESTIMACION Y PREDICCIÓN.

La pregunta que surge inmediatamente después de haber estudiado los modelos para los procesos ARIMA es: Si tenemos un conjunto de observaciones hechas secuencialmente en el tiempo, ¿Cómo podemos saber de qué proceso se trata, es decir, qué valores de  $p, d, q$  son los más apropiados? De igual manera que en el capítulo precedente, la respuesta a esta pregunta será el objetivo de estudio en este capítulo.

Una identificación tentativa de un modelo de series de tiempo ARIMA es hecha a través del análisis histórico de los datos [Se recomienda un mínimo de 50 observaciones para identificar el modelo apropiado].

La primera herramienta con que contamos para la identificación del proceso, es la función de autocorrelación, ya que conocemos la particular forma en que se comporta dicha función para cada uno de los modelos estudiados y para las estimaciones necesarias solamente necesitaremos de los datos.

##### 4.1 ESTIMACION DE LA FUNCION DE AUTOCORRELACION

Sabemos que los coeficientes de autocorrelación  $\rho_i$  están dados por el cociente de la autocovarianza de atraso

$j$  y la varianza del proceso  $\gamma_0$ .

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

por otro lado

$$\gamma_j = \mathbb{E} [(Z_t - \mu)(Z_{t+j} - \mu)]$$

Un estimador natural de  $\gamma_j$ , es el producto promedio de las desviaciones de  $Z_t$  y  $Z_{t+j}$  de la media muestral, denotado dicho estimador por  $C_j$

Así

$$C_j = \frac{1}{T} \sum_{t=1}^{T-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})] \quad j=1,2,\dots$$

donde  $T$  es la longitud de la serie de tiempo bajo estudio y

$$\bar{Z} = \frac{1}{T} \sum_{t=1}^T Z_t$$

entonces el estimador de  $\rho_j$  podemos expresarlo de la siguiente forma:

$$\hat{\rho}_j = \frac{C_j}{C_0}$$

Un recurso de suma importancia en la identificación es el correlograma muestral, es decir, la gráfica de las autocorrelaciones muestrales, con esto trataríamos de reconocer el comportamiento típico de un proceso ARIMA.

Dado que estamos trabajando con estimadores, un punto de interés sería el saber qué tan buenos son dichos estimadores, pero ya que no es el interés principal de nuestro es

tudio, solamente enunciaremos como ejemplo, la varianza y covarianza de los estimadores propuestos por Bartlett para un proceso MA(4)

$$\text{Var}(\hat{\gamma}_j) \approx \frac{1}{T} \left\{ 1 + 2 \sum_{i=1}^q \rho_i^2 \right\} \quad i > q$$

y

$$\text{Cov}(\hat{\gamma}_j, \hat{\gamma}_{j+s}) \approx \frac{1}{T} \sum_{i=-\infty}^{\infty} \rho_i \rho_{i+s}$$

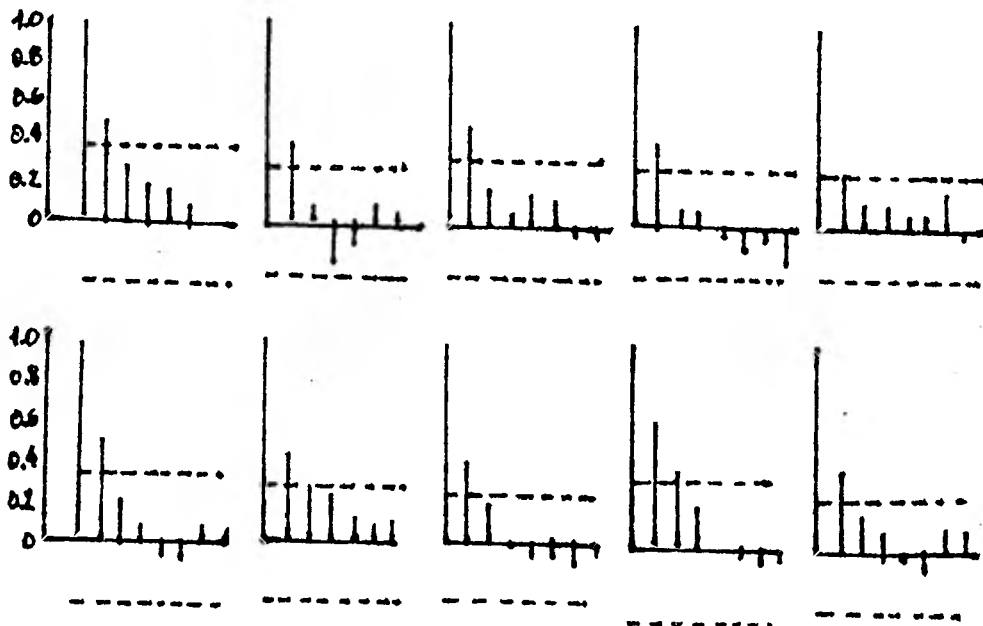
A manera de un intento de justificación para el hecho de recomendar un mínimo de 50 observaciones, basta notar que la varianza decrecerá tanto como el tamaño de la muestra crezca.

Es importante que el observador analice las características generales del correlograma muestral y no analice detalladamente cada  $\hat{\gamma}_j$ , por ejemplo, autocorrelaciones muestrales producidas por el proceso autorregresivo de primer orden decrecerán paulatinamente, pero no precisamente en forma exponencial; para aclarar esta idea, Nelson presenta un ejemplo de un proceso autorregresivo de primer orden AR(1)

$$Z_t = .5Z_{t-1} + \varepsilon_t$$

con autocorrelaciones muestrales de 10 independientes realizaciones del proceso, cada uno conteniendo 100 observaciones generadas por una computadora con el uso de un generador de números aleatorios, y cuyos correlogramas aparecen a conti-

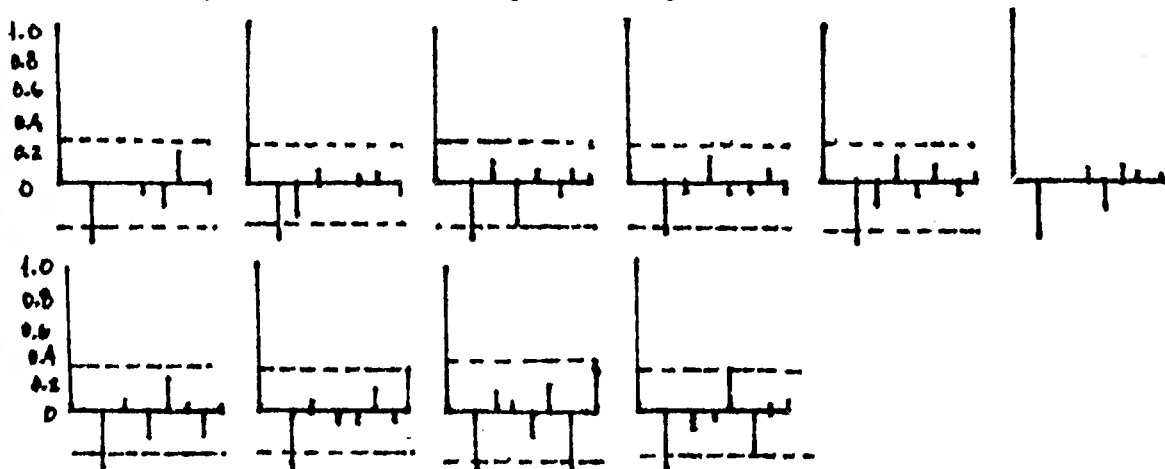
nuación.



Otro ejemplo presentado por Nelson es el de un proceso de medias móviles,

$$Z_t = \varepsilon_t - .5 \varepsilon_{t-1}$$

con autocorrelaciones muestrales de 10 independientes realizaciones de 100 observaciones cada una y cuyos correlogramas aparecen en la siguiente gráfica:



Es importante notar que dado que trabajamos con estimadores de  $\rho_i$  puede resultar que aunque  $\rho_i = 0$  para  $i = 2, 3, \dots$  los valores de  $\gamma_i$  sean grandes en períodos de atraso mayores que  $i$ , como es el caso de las gráficas, para distinguir que  $\gamma_i$ 's, son significativos del correlograma, se propone una prueba de significancia estadística para las autocorrelaciones muestrales, basada en la aproximación de Bartlett para la varianza de  $\gamma_i$ . Para esto, es importante notar que la varianza de  $\gamma_i$   $i > q$  está dado por la fórmula de Bartlett bajo la suposición de que el orden verdadero del proceso es  $q$ .

Así la prueba estadística consiste en:

Dado que la distribución de  $\gamma_i$  para  $i > q$  es aproximadamente normal, se puede considerar una autocorrelación muestral más grande en valor absoluto que 1.96 veces la desviación estandar como significativamente diferente de cero al 95% de confianza, esto es

$$|\gamma_i| > (1.96) \frac{1}{\sqrt{T}} \left(1 + 2 \sum_{i=1}^q \rho_i^2\right)^{1/2} \quad i > q$$

sustituyendo el estimador de  $\rho_i$  podemos obtener un criterio aproximado de la siguiente forma:

$$\text{si } |\gamma_i| > 2 \frac{1}{\sqrt{T}} \left(1 + 2 \sum_{i=1}^q \gamma_i^2\right) \quad i > q$$

decimos que  $\gamma_i$  es considerablemente distinto de cero, rechazamos la hipótesis  $H_0 : \gamma_i = 0$ .

Regresando al ejemplo presentado por Nelson ( $Z_t = \xi_t - 0.5 \xi_{t-1}$ ) si observamos detenidamente los correlogramas, podemos notar que en cada realización del proceso aparece para  $i > 1$  muy pocos  $\gamma_i$  significativamente distintos de cero (las líneas punteadas marcan el criterio aproximado antes mencionado), esto nos conduce a tener cuidado al aplicar la prueba de significancia ya que si calculamos demasiados coeficientes de autocorrelación es factible encontrar algunos significativamente distintos de cero, debido quizá a causas como el redondeo o truncamiento.

Resumiendo, si la identificación tentativa del proceso es que  $q$  es, digamos  $q'$ , entonces sería interesante probar la significancia de  $\gamma_{q'+1}$ , ya que si  $\gamma_{q'+1}$  resulta ser significativo en un  $t$  particular, sería mucho mejor tratar de identificar un modelo  $MA(q'+1)$

#### 4.2 DETERMINACION DEL GRADO APROPIADO DE DIFERENCIACION

Como mencionamos en el capítulo anterior, en el caso de economía, los datos que presentan dichas series, exhiben no-estacionalidad y en dicho caso, ¿Cómo podemos determinar el grado apropiado de diferenciación?

Si la serie es no-estacionaria, hay que hacer notar que

la función de autocorrelación del proceso está indefinida, pero de cualquier forma siempre podemos calcular las autocorrelaciones de un conjunto de datos. La práctica ha mostrado que series no estacionarias producen autocorrelaciones muestrales que permanecen grandes incluso en períodos de atraso grandes debido a que cualquier realización de la serie tenderá a estar sobre uno o el otro lado de la media muestral por muchos períodos. Si los datos presentan no-estacionariedad, entoncés las autocorrelaciones muestrales de la primera diferencia son examinadas para tratar de encontrar un modelo estacionario apropiado, rara vez en datos económicos las primeras diferencias serán no estacionarias, en dicho caso, las segundas diferencias serán examinadas.

Es importante tener en cuenta que solamente necesitamos identificar el nivel más bajo de diferencias para la cual un modelo estacionario sea razonable, ya que diferencias más altas de cualquier serie estacionaria son también estacionarias.

#### 4.3 IDENTIFICACION DEL ORDEN DE UN PROCESO AUTORREGRESIVO Y LA FUNCION DE AUTOCORRELACION PARCIAL

El problema en este caso se plantea de la siguiente for

ma:

Sabemos que la función de autocorrelación para un proceso autorregresivo se comporta de tal forma que podemos decir que decrece exponencialmente conforme el período de atraso crece, pero ¿Cómo podemos saber de que orden se trata? es decir, ¿Cómo poder distinguir si se trata de un proceso  $AR(p)$  o  $AR(p^*)$ .

Para hacer una adecuada identificación del modelo, necesitamos hacer uso de una función conocida como función de autocorrelación parcial cuya interpretación es la siguiente:

Considere tres variables aleatorias  $X, Y, Z$  y sea su función de densidad conjunta  $f(x, y, z)$ , entonces la distribución condicional de  $X$  y  $Y$  dado  $Z$  es

$$h(x, y | z) = \frac{f(x, y, z)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dx dy}$$

Así, el coeficiente de correlación entre  $X$  y  $Y$  en la distribución condicional  $h(X, Y | Z)$  es conocido como coeficiente de correlación parcial.

En términos de una serie de tiempo, es conveniente interpretar a la función de autocorrelación parcial de atraso  $j$  como la correlación entre  $Z_t$  y  $Z_{t+j}$  con los efectos de intervención de las observaciones  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+j-1}$

Si denotamos por  $\phi_{pi}$ , el  $j$ -ésimo coeficiente



en un proceso autorregresivo de orden  $p$ , es decir  $\phi_{pp}$  es el último coeficiente; las ecuaciones de Yule-Walker

$$\begin{aligned} e_1 &= \phi_1 + \phi_2 e_1 + \dots + \phi_p e_{p-1} \\ &\vdots \\ e_p &= \phi_1 e_{p-1} + \phi_2 e_{p-2} + \dots + \phi_p \end{aligned}$$

pueden ser escritas de la siguiente forma:

$$\begin{bmatrix} 1 & e_1 & e_2 & \dots & e_{p-1} \\ e_1 & 1 & e_1 & \dots & e_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e_{p-1} & e_{p-2} & e_{p-3} & & 1 \end{bmatrix} \begin{bmatrix} \phi_{p1} \\ \phi_{p2} \\ \vdots \\ \phi_{pp} \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

o simplemente

$$P_k \phi_k = e_k$$

resolviendo estas ecuaciones para  $p = 1, 2, 3, \dots$

obtenemos:

$$p=1 \quad \phi_{11} = e_1$$

$$p=2 \quad \phi_{22} = \frac{\begin{vmatrix} 1 & e_1 \\ e_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & e_1 \\ e_1 & 1 \end{vmatrix}}$$

$$p=3 \quad \phi_{33} = \frac{\begin{vmatrix} 1 & e_1 & e_1 \\ e_1 & 1 & e_2 \\ e_2 & e_1 & e_3 \end{vmatrix}}{\begin{vmatrix} 1 & e_1 & e_2 \\ e_1 & 1 & e_1 \\ e_2 & e_1 & 1 \end{vmatrix}}$$

Así, la cantidad  $\phi_{pp}$   $p=1, \dots$  vista como una función de atraso  $p$ , es conocida como función de autocorrelación parcial. Entonces, para nuestro propósito identificación sabemos que en un proceso autorregresivo de orden  $p$  la función de autocorrelación parcial  $\phi_{pp}$  será distinta de cero para  $j$  menor o igual que  $p$  y cero para períodos más grandes que  $p$ , es decir, la función de autocorrelación parcial de un proceso autorregresivo de orden  $p$  se trunca después del período de atraso  $p$ .

Una forma de proceder para la identificación es la siguiente:

Supongamos por el momento la hipótesis de que  $p=1$ , entonces un estimador de  $\phi_{11}$  puede ser obtenido resolviendo las ecuaciones de Yule-Walker para  $p=1$ , en la cual  $\rho_1$  es reemplazado por el estimador  $\gamma_1$ , esto es,

$$\gamma_1 = \hat{\phi}_{11}$$

ahora, si  $\hat{\phi}_{11}$  es significativamente diferente de cero, podremos concluir que estamos tratando al menos con un proceso de orden 1; siguiendo el mismo procedimiento, supongamos ahora la hipótesis de que  $p=2$ , entonces un estimador de  $\phi_{22}$  puede ser obtenido resolviendo las ecuaciones de Yule-Walker, esto es:

$$\begin{aligned} Y_1 &= \hat{\phi}_{21} + \hat{\phi}_{22} Y_1 \\ Y_2 &= \hat{\phi}_{21} Y_1 + \hat{\phi}_{22} \end{aligned}$$

si el estimador resultante de  $\hat{\phi}_{22}$  resulta ser significativamente distinto de cero, podremos concluir que estamos tratando al menos con un proceso de orden 2.

Siguiendo el mismo procedimiento, supongamos que el verdadero valor de  $p$  es  $p'$ , entonces cuando resolvamos el sistema de ecuaciones para  $p = p + 1$  el valor de  $\hat{\phi}_{p,p+1}$  será aproximadamente cero ya que es un estimador de  $\phi_{p,p+1}$  el cual es cero, es decir:

$$\hat{\phi}_{ii} \approx 0 \quad i > p'$$

Como en el caso de identificación de un proceso MA ( $q$ ), es conveniente tener un estimador de la varianza del estimador  $\hat{\phi}_{ii}$  que nos ayude a decidir que tan "significativamente" es distinto de cero un  $\hat{\phi}_{ii}$ .

Este estimador fue propuesto por Quenoville y está basado en la hipótesis de que si el proceso es autorregresivo de orden  $p$ , la autocorrelación parcial estimada de orden  $p+1$  y más altas son aproximadamente independientemente distribuidas, de tal forma que si el tamaño de la serie es  $T$  entonces:

$$\text{Var} [\hat{\phi}_{ii}] \approx \frac{1}{T} \quad i > p+1$$

obteniendo de esta forma el error estandar

$$EE [\hat{\phi}_{ii}] \approx \frac{1}{\sqrt{T}} \quad i > p+1$$

Como una nota interesante de estudio, recordemos que cual

quier proceso de medias móviles finito puede ser escrito en forma autorregresiva de orden infinito, así la función de autocorrelación parcial que presentaría un proceso MA (q) deberá decrecer conforme los valores de j crezcan pero no truncarse a un determinado período de atraso, como en el caso autorregresivo de orden finito.

#### 4.4 IDENTIFICACION DE PROCESOS MIXTOS

Ahora supongamos que los datos obtenidos ha sido generados por un proceso mixto ¿Cómo esperaríamos que se computara la función de autocorrelación y la de autocorrelación parcial?

Como estudiamos en el capítulo anterior, sabemos que las primeras q autocorrelaciones están determinadas tanto por los parámetros de la parte autorregresiva como de la parte de medias móviles a través de una relación la cual es muy compleja aun si p y q son pequeños, pero tambien sabemos que para períodos de atraso más grandes que q la autocorrelación está únicamente determinada por la parte autorregresiva, es decir,

$$r_i = \phi_1 r_{i-1} + \dots + \phi_p r_{i-p} \quad i > p$$

Esto nos conduce a pensar que la función de autocorrelación deberá tener un comportamiento regular después del período de atraso q.

Para aclarar esta idea considere el siguiente ejemplo:

sea el proceso ARMA (1,1) con  $\phi_1 = .5$ ,  $\theta_1 = -.9$  y  $\sigma^2 = 1$

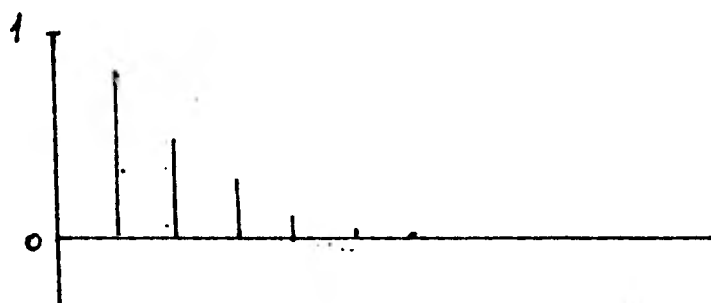
esto es  $Z_t = 0.5 Z_{t-1} + \varepsilon_t + 0.9 \varepsilon_{t-1}$

Una vez hechos los cálculos obtenemos:

$$\rho_1 = .75$$

$$\rho_2 = .375$$

$$\rho_i = (.5)^{i-1} (.75) \quad i=2, 3, \dots$$



Observamos que únicamente  $\rho_1$  es alterada por el término de medias móviles, pero para  $i > 2$  la función de autocorrelación mantiene un comportamiento regular, como lo hubiera sido, para un proceso AR (1) con  $\phi_1 = .5$

Recordemos del capítulo anterior que un proceso ARMA puede ser escrito como un proceso autorregresivo de orden infinito, de tal forma que las autocorrelaciones parciales como estimadores aproximativos de esos coeficientes autorregresivos decrecerán paulatinamente en lugar de truncarse. Esto implica que un modelo mixto tenderá a decrecer gradual

mente tanto, en la función de autocorrelación así como en la de autocorrelación parcial.

A manera de resumen, Montgomery y J. (véase [9], pag. 208)\* presentan una tabla de comportamiento de la función de autocorrelación y autocorrelación parcial para las tres clases de modelos

CLASE DE PROCESO	AUTOCORRELACION	AUTOCORRELACION PARCIAL
Medias Móviles, MA (q)	Barras durante los períodos de atraso 1 hasta q desapareciendo despues.	Decae en forma exponencial o geométrica aproximadamente, con un número relativamente grande de valores distintos de cero.
Autorregresivo AR(p)	Decae en forma exponencial o geométrica aproximadamente, con un número relativamente grande de valores distintos de cero.	Barras durante los períodos de atraso 1 hasta q desapareciendo despues.
Autorregresivo-Medias Móviles ARMA (p,q)	Decae en forma exponencial o geométrica aproximadamente, con un número relativamente grande de valores distintos de cero.	Decae en forma exponencial o geométrica aproximadamente, con un número relativamente grande de valores distintos de cero.

## ESTIMACION

Dentro del análisis de series de tiempo el siguiente paso a dar; una vez hecha una identificación tentativa, será estimar los parámetros que involucrados componen el modelo; como sabemos, existe una gran variedad de formas para estimar parámetros, sin embargo, estamos interesados en aquellos que de alguna manera sean los mejores y más eficientes; por otro lado sabemos que los estimadores que cumplen ser suficientes y consistentes son los de máxima verosimilitud.

Una razón de suma importancia dentro del análisis de series de tiempo para usar dichos estimadores es la parte en la que definimos que es una serie de tiempo y como iba a ser interpretada, esto es:

Teniendo un conjunto de observaciones  $z_1, \dots, z_t$  y un modelo identificado que suponemos han generado esas observaciones, sabemos que esas observaciones antes de ser generadas, eran variables aleatorias que fueron extraídas de una distribución conjunta definida por un particular modelo ARIMA, esto es, las observaciones fueron extraídas de una distribución conjunta  $f(w/\phi, \theta, \rho, \sigma_\epsilon^2)$  donde  $w$  denota el vector de observaciones sobre las diferencias estacionarias de las  $z_t$ 's,  $\phi, \theta$  los vectores de los parámetros y  $\rho, \sigma_\epsilon^2$  las constantes del modelo, ahora si hacemos uso de la interpre

tación clásica de máxima verosimilitud, esto es:

¿Qué valores de los parámetros son los más probables para haber generado dichas observaciones? en este caso;

$$L(\phi, \theta, \rho, \sigma_\varepsilon^2 / W)$$

Estaremos haciendo uso de la interpretación de series de tiempo utilizando los estimadores máximo verosímiles.

Es claro que si utilizamos los EMV (estimadores máximo verosímiles), debemos hacer una suposición acerca de la forma de la función de distribución conjunta. Una suposición práctica para empezar nuestro análisis será trabajar bajo la hipótesis de que los  $\varepsilon_t$  tienen una distribución normal, es decir;

$$(1) \quad f(\varepsilon_t / \sigma_\varepsilon^2) = \frac{1}{\sqrt{2\pi} \sigma_\varepsilon} e^{-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}}$$

Por ser independientes

$$(2) \quad f(\varepsilon_1, \dots, \varepsilon_T / \sigma_\varepsilon^2) = \frac{1}{(2\pi)^{T/2} (\sigma_\varepsilon^2)^{T/2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^T \varepsilon_t^2}$$

por otro lado, sabemos que la forma de un modelo ARIMA es:

$$(3) \quad W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \varepsilon_t + \rho - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

ahora, podemos expresar  $\varepsilon_t$  en términos de los parámetros, esto es

$$(4) \quad \varepsilon_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \rho + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

si sustituimos (4) en (2) obtenemos

$$f(W / \phi, \theta, \rho, \sigma_\varepsilon^2) = \frac{1}{(2\pi)^{T/2} (\sigma_\varepsilon^2)^{T/2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^T (W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \rho + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q})^2}$$

Entonces la función de verosimilitud será:



$$L(\phi, \theta, \rho, \sigma_\varepsilon^2 / W) = \frac{1}{(2\pi)^{T/2} (\sigma_\varepsilon^2)^{T/2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^T \hat{\varepsilon}(\phi, \theta, \rho)_t^2}$$

donde  $\hat{\varepsilon}(\phi, \theta, \rho)_t$  es el valor estimado de  $\varepsilon_t$  como función de los parámetros y las observaciones.

Siguiendo con la técnica para obtener los EMV consideremos ahora el logaritmo natural de la función, es decir:

$$\ln L(\phi, \theta, \rho, \sigma_\varepsilon^2 / W) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^T \hat{\varepsilon}(\phi, \theta, \rho)_t^2$$

Dado que los parámetros  $\phi, \theta, \rho$  entran solamente dentro de la suma de cuadrados, para maximizar la función de verosimilitud bastará minimizar la suma de los cuadrados.

Calculemos primeramente el EMV de  $\sigma_\varepsilon^2$ .

Tenemos que

$$\frac{\partial \ln L(\phi, \theta, \rho, \sigma_\varepsilon^2 / W)}{\partial \sigma_\varepsilon^2} = -\frac{T}{2\sigma_\varepsilon^2} + \frac{1}{2(\sigma_\varepsilon^2)^2} \sum_{t=1}^T \hat{\varepsilon}(\phi, \theta, \rho)_t^2$$

Multiplicando por  $2\sigma_\varepsilon^2$  obtenemos

$$-T + \frac{1}{\sigma_\varepsilon^2} \sum_{t=1}^T \hat{\varepsilon}(\hat{\phi}, \hat{\theta}, \hat{\rho})_t^2 = 0$$

$$\Rightarrow \hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^T \hat{\varepsilon}(\hat{\phi}, \hat{\theta}, \hat{\rho})_t^2}{T} \quad \text{donde } \hat{\phi}, \hat{\theta}, \hat{\rho} \text{ son los E.M.V. de } \phi, \theta, \rho.$$

Consideremos el caso ahora de un proceso AR (p).

La suma de cuadrados estará dada por:

$$\sum_{t=1}^T \hat{\varepsilon}(\phi, \rho)_t^2 = \sum_{t=1}^T (z_t - \rho - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \dots - \phi_p z_{t-p})^2$$

De donde

$$\begin{aligned} \frac{\partial \sum_{t=1}^T \hat{\varepsilon}(\phi_i, \rho)_t^2}{\partial \phi_i} &= -2 \sum_{t=1}^T (z_t - \rho - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p}) z_{t-1} \\ &= -2 \sum_{t=1}^T (z_t - \rho - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p}) z_{t-2} \\ &\vdots \\ &= -2 \sum_{t=1}^T (z_t - \rho - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p}) z_{t-p} \end{aligned}$$

$$\frac{\partial \sum_{t=1}^T \hat{\varepsilon}(\hat{\phi}_i, \hat{\rho})^2}{\partial \rho} = -2 \sum_{t=1}^T (z_t - \hat{\rho} - \hat{\phi}_1 z_{t-1} - \dots - \hat{\phi}_p z_{t-p})$$

Iguando a cero las derivadas obtenemos:

$$-2 \sum_{t=1}^T (z_t - \hat{\rho} - \hat{\phi}_1 z_{t-1} - \dots - \hat{\phi}_p z_{t-p}) z_{t-1} = 0$$

$$-2 \sum_{t=1}^T (z_t - \hat{\rho} - \hat{\phi}_1 z_{t-1} - \dots - \hat{\phi}_p z_{t-p}) z_{t-2} = 0$$

⋮

$$- 2 \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \dots - \hat{\phi}_p z_{t-p}) z_{t-p} = 0$$

$$- 2 \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \dots - \hat{\phi}_p z_{t-p}) = 0$$

obtenemos un sistema de  $p+1$  ecuaciones lineales con  $p+1$  parámetros desconocidos, y cuyo sistema puede ser resuelto por algunos de los métodos conocidos; como un simple ejemplo, consideremos el caso de un proceso AR (2), esto es:

$$z_t = \eta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \varepsilon_t$$

o

$$\varepsilon_t = z_t - \eta - \phi_1 z_{t-1} - \phi_2 z_{t-2}$$

y cuya suma de cuadrados es:

$$\Sigma = \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \hat{\phi}_2 z_{t-2})^2$$

entonces:

$$\frac{\partial \Sigma}{\partial \hat{\phi}_1} = - 2 \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \hat{\phi}_2 z_{t-2}) z_{t-1} = 0$$

$$\frac{\partial \Sigma}{\partial \hat{\phi}_2} = - 2 \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \hat{\phi}_2 z_{t-2}) z_{t-2} = 0$$

$$\frac{\partial \Sigma}{\partial \hat{\eta}} = - 2 \sum_{t=1}^T (z_t - \hat{\eta} - \hat{\phi}_1 z_{t-1} - \hat{\phi}_2 z_{t-2}) = 0$$

$$-\sum_{t=1}^T z_t z_{t-1} + \hat{\rho} \sum_{t=1}^T z_{t-1} + \phi_1 \sum z_{t-1}^2 + \phi_2 \sum z_{t-2} z_{t-1} = 0$$

$$-\sum z_t z_{t-2} + \hat{\rho} \sum z_{t-2} + \phi_1 \sum z_{t-1} z_{t-2} + \phi_2 \sum z_{t-2}^2 = 0$$

$$-\sum z_t + T \hat{\rho} + \phi_1 \sum z_{t-1} + \phi_2 \sum z_{t-2} = 0$$

despejando

$$\sum_{t=1}^T z_t z_{t-1} = \hat{\rho} \sum z_{t-1} + \phi_1 \sum z_{t-1}^2 + \phi_2 \sum z_{t-2} z_{t-1}$$

$$\sum z_t z_{t-2} = \hat{\rho} \sum z_{t-2} + \phi_1 \sum z_{t-1} z_{t-2} + \phi_2 \sum z_{t-2}^2$$

$$\sum z_t = \hat{\rho} T + \phi_1 \sum z_{t-1} + \phi_2 \sum z_{t-2}$$

$$\Delta_{\text{SISTEMA}} = \begin{vmatrix} \sum z_{t-1} & \sum z_{t-1}^2 & \sum z_{t-2} z_{t-1} \\ \sum z_{t-2} & \sum z_{t-1} z_{t-2} & \sum z_{t-2}^2 \\ T & \sum z_{t-1} & \sum z_{t-2} \end{vmatrix}$$

$$\Delta_{\hat{\rho}} = \begin{vmatrix} \sum z_t z_{t-1} & \sum z_{t-1}^2 & \sum z_{t-2} z_{t-1} \\ \sum z_t z_{t-2} & \sum z_{t-1} z_{t-2} & \sum z_{t-2}^2 \\ \sum z_t & \sum z_{t-1} & \sum z_{t-2} \end{vmatrix}$$

$$\Delta \hat{\phi}_1 = \begin{vmatrix} \sum z_{t-1} & \sum z_t z_{t-1} & \sum z_{t-2} z_{t-1} \\ \sum z_{t-2} & \sum z_t z_{t-2} & \sum z_{t-2}^2 \\ T & \sum z_t & \sum z_{t-2} \end{vmatrix}$$

$$\Delta \hat{\phi}_2 = \begin{vmatrix} \sum z_{t-1} & \sum z_{t-1}^2 & \sum z_t z_{t-1} \\ \sum z_{t-2} & \sum z_{t-1} z_{t-2} & \sum z_t z_{t-2} \\ T & \sum z_{t-1} & \sum z_t \end{vmatrix}$$

$$\hat{\rho} = \frac{\Delta \hat{\phi}}{\Delta s}$$

$$\hat{\phi}_1 = \frac{\Delta \phi_1}{\Delta s}$$

$$\hat{\phi}_2 = \frac{\Delta \phi_2}{\Delta s}$$

otra forma más práctica de representar el proceso AR (2), es escribiendo en forma matricial, es decir;

Suponiendo que tenemos N observaciones;

$$\begin{pmatrix} z_3 \\ z_4 \\ \vdots \\ z_{n-1} \\ z_n \end{pmatrix} = \begin{pmatrix} 1 & z_2 & z_1 \\ 1 & z_3 & z_2 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & z_{n-2} & z_{n-3} \\ 1 & z_{n-1} & z_{n-2} \end{pmatrix} \begin{pmatrix} \rho \\ \phi_1 \\ \phi_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_3 \\ \varepsilon_4 \\ \vdots \\ \varepsilon_{n-1} \\ \varepsilon_n \end{pmatrix}$$

esto es;

$$\underline{z} = \underline{Z} \underline{\phi} + \underline{\varepsilon}$$

Los estimadores por mínimos cuadrados de los parámetros, son:

$$\hat{\underline{\phi}} = (\underline{z}' \underline{z})^{-1} \underline{z}' \underline{z}$$

Nota: Como sabemos que maximizar la función de verosimilitud se reduce a minimizar la suma de cuadrados, podemos hablar directamente de los estimadores por mínimos cuadrados ya que coinciden con los de máxima verosimilitud.

Note que el caso AR (2) usamos solamente N-2 observaciones ya que  $z_0$  y  $z_{-1}$ ,<sup>ie</sup> decir, al tiempo  $t=1$  y  $t=2$  los valores para el proceso no existen, la solución a este problema es

suponer que  $Z_0 = Z_{-1} = 0$  ya que para series de tiempo "grandes", la elección de los valores iniciales tendrá poco efecto en los estimadores de  $\phi$ , [Box y Jenkins, capítulo 7.]

En general para un proceso AR(p) la matriz  $Z$  será  $(N-P) \times (p+1)$ .

Consideremos ahora el caso de un proceso MA (1), esto es:

$$Z_t = \epsilon_t - \theta_1 \epsilon_{t-1}$$

o

$$\epsilon_t = Z_t + \theta_1 \epsilon_{t-1}$$

cuya suma de cuadrados está determinada por:

$$\sum_{t=1}^T (Z_t + \theta_1 \hat{\epsilon}(\theta_1)_{t-1})^2$$

o

$$\sum_{t=1}^T \hat{\epsilon}_t(\theta_1)^2$$

note que en este caso, si aplicamos directamente la técnica de máxima verosimilitud no podremos obtener los estimadores ya que desconocemos los valores de  $\epsilon_t, t=1, 2, \dots$ , entonces procederemos de la siguiente forma:

Si queremos minimizar  $\sum_{t=1}^T \hat{\epsilon}_t(\theta_1)^2$  sobre todos los valores de  $\theta_1$  primero debemos evaluar  $\hat{\epsilon}(\theta_1)_t$ , es decir, si empezamos con  $\hat{\epsilon}(\theta_1)_1$  obtenemos:

$$\hat{\epsilon}(\theta_1)_1 = Z_1 + \theta_1 \hat{\epsilon}(\theta_1)_0$$

lo cual representa un problema ya que desconocemos el valor

de  $\hat{\epsilon}(\theta_1)_0$ , la solución a este problema es suponer que  $\hat{\epsilon}(\theta_1)_0 = 0$  ya que sabemos que  $E(\epsilon_t) = 0$ , y cuya suposición se hará de igual forma si se tratara de un modelo con  $q$  parámetros simplemente tomando

$$\hat{\epsilon}(\theta_1, \dots, \theta_q)_0 = \hat{\epsilon}_{-1} = \dots = \hat{\epsilon}_{-q+1} = 0$$

Entonces como  $\hat{\epsilon}(\theta_1)_0 = 0$  tenemos:

$$\hat{\epsilon}(\theta_1)_1 = z_1 + \theta_1(0)$$

$$\hat{\epsilon}(\theta_1)_2 = z_2 + \theta_1 \hat{\epsilon}(\theta_1)_1 = z_2 + \theta_1 z_1$$

$$\hat{\epsilon}(\theta_1)_3 = z_3 + \theta_1 \hat{\epsilon}(\theta_1)_2 = z_3 + \theta_1 z_2 + \theta_1^2 z_1$$

$$\hat{\epsilon}(\theta_1)_4 = z_4 + \theta_1 \hat{\epsilon}(\theta_1)_3 = z_4 + \theta_1 z_3 + \theta_1^2 z_2 + \theta_1^3 z_1$$

⋮

$$\hat{\epsilon}(\theta_1)_t = z_t + \theta_1 \hat{\epsilon}(\theta_1)_{t-1} = z_t + \dots + \theta_1^{t-1} z_1$$

Ahora solamente depende de  $\theta_1$  y de las observaciones pasadas lo cual nos permite calcular la suma de cuadrados,  $\sum_{t=1}^T \hat{\epsilon}_t(\theta_1)^2$

siguiendo con la técnica de máxima verosimilitud para encontrar el mínimo de  $\theta_1$ , el siguiente paso sería evaluar la

$$\frac{\partial \sum_{t=1}^T (z_t + \dots + \theta_1^{t-1} z_1)^2}{\partial \theta_1} = 0$$



la cual resulta una función que no es lineal en  $\theta_1$ ; Una discusión de mínimos cuadrados no lineales, es dada por Box y Jenkins, cap. 7 y Draper y Smith en su libro Applied Regression Analysis; nosotros trataremos este problema de una manera muy simple; recordemos que  $\theta_1$  requiere para ser invertible el proceso encontrarse dentro del intervalo  $-1 < \theta_1 < 1$ , bajo esta condición buscaremos el mínimo sobre una red de valores de  $\theta_1$  entre 1 y -1 auxiliándonos de una computadora, de esta manera podríamos empezar a evaluar la suma de cuadrados, digamos para  $\theta_1 = -.9, .5$  y  $.9$  lo cual nos daría una idea del intervalo en el cual se encuentra el mínimo, y de esta manera, haciendo interacciones cada vez más finas podremos dar una aproximación bastante buena del mínimo, es claro que una vez hecha esta aproximación, el estimador de la varianza del proceso por máxima verosimilitud estará determinada por:

$$V_{\epsilon}^2 = \frac{\sum_{t=1}^T \hat{\epsilon}(\theta_1)_t^2}{T}$$

Veamos ahora el caso de la estimación de un proceso ARIMA, por facilidad consideremos el caso ARIMA (1,d,1)

$$W_t = \phi_1 W_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

o

$$\epsilon_t = W_t - \phi_1 W_{t-1} + \theta_1 \epsilon_{t-1}$$

cuya suma de cuadrados está determinada por:

$$\sum_{t=1}^T (W_t - \phi_1 W_{t-1} + \theta_1 \epsilon_{t-1})^2$$

o

$$\sum_{t=1}^T \hat{\epsilon}(\phi_1, \theta_1)_t^2$$

Note que en este caso, la suma de cuadrados involucra más de un parámetro desconocido, y como en el caso del proceso MA (1) no es fácil estimar los parámetros ya que al obtener los derivados parciales con respecto a  $\phi_1$  y  $\theta_1$ , no conoceremos los valores de  $\hat{\epsilon}(\phi_1, \theta_1)_t$  y puede resultar una función no lineal en los parámetros; como en el caso del proceso MA (1) seguiremos el mismo procedimiento, es decir:

$$\hat{\epsilon}(\phi_1, \theta_1)_1 = w_1 - \phi_1 w_0 + \theta_1 \hat{\epsilon}(\phi_1, \theta_1)_0$$

en este caso volvemos a tener el mismo problema del valor inicial  $\hat{\epsilon}(\phi_1, \theta_1)_0$  y además que el valor de  $w_0$  no la podemos obtener, la solución a este problema será empezar nuestras iteraciones al tiempo  $t=1$  y suponiendo que  $\hat{\epsilon}_1 = 0$ , es decir;

$$\hat{\epsilon}(\phi_1, \theta_1)_2 = w_2 - \phi_1 w_1 + \theta_1 \hat{\epsilon}_1$$

si seguimos este método, es evidente que se tienen que cumplir las condiciones de invertibilidad del proceso, es decir, para darnos una idea de donde se encuentra el mínimo de los parámetros, tendremos que elegir algunos dentro de la región permitida y de esta manera empezar a reducir el intervalo para dar una buena aproximación; una forma adecuada de escoger los parámetros iniciales será mediante el método de Box y Jenkins conocido como ESTIMADORES PRELIMINARES DE LOS PARÁMETROS, estos parámetros son obtenidos resolviendo las ecuaciones de Yule-Walker, por ejemplo, para un proceso autorregresivo, en dichas ecuaciones reemplazaremos los  $\rho_j$  con  $\gamma_j$ ,

esto es:

$$\begin{aligned}
 Y_1 &= \hat{\phi}_1 + \dots + \hat{\phi}_p Y_{p-1} \\
 \vdots \\
 Y_p &= \hat{\phi}_1 Y_{p-1} + \dots + \hat{\phi}_p
 \end{aligned}$$

y el cálculo de los  $\hat{\phi}_i$ 's es solamente operacional, la forma más simple sería el caso del proceso AR (1), es decir:

$$\hat{\phi}_1 = Y_1$$

Veamos el caso de un proceso MA (1), sabemos que:

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

entonces un estimador de  $\theta_1$  puede ser obtenido de

$$Y_1 = \frac{-\hat{\theta}_1}{1 + \hat{\theta}_1^2}$$

esta ecuación es cuadrática en  $\theta_1$  y tiene dos soluciones, es decir;

$$\hat{\theta}_1 = -\frac{1}{2Y_1} \pm \sqrt{\left(\frac{1}{(2Y_1)^2} - 1\right)}$$

claramente será escogida aquella que satisfaga la condición de invertibilidad:  $|\theta_1| < 1$

para un proceso MA de orden más alto, las q ecuaciones

$$\begin{aligned}
 Y_1 &= \frac{-\hat{\theta}_1 + \hat{\theta}_1 \hat{\theta}_2 + \dots + \hat{\theta}_{q-1} \hat{\theta}_q}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2} \\
 \vdots \\
 Y_q &= \frac{-\hat{\theta}_q}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2}
 \end{aligned}$$

pueden ser resueltas, y la forma de escoger los estimadores serán aquellos que satisfagan la condición de invertibilidad.

Regresando a nuestro problema del proceso ARIMA (1,d,1),

la forma de elegir los parámetros iniciales es como sigue, sabemos que de las ecuaciones

$$\ell_{q+1} = \phi_1 \ell_q + \dots + \phi_p \ell_{q-p+1}$$

⋮

$$\ell_{q+p} = \phi_1 \ell_{q+p-1} + \dots + \phi_p \ell_q$$

los estimadores de  $\phi_1 \dots \phi_p$  pueden ser obtenidos si reemplazamos  $\gamma_i$  en lugar de  $\ell_i$ , ahora, de la relación entre  $\ell_1 \dots \ell_p$  los  $\phi$ 's y las  $\theta$ 's, podremos resolver para  $\hat{\theta}_1 \dots \hat{\theta}_q$  reemplazando los  $\gamma_i$ 's en lugar de los  $\ell_i$  y los estimadores ya calculados de las  $\phi$ 's, por ejemplo, si tenemos un proceso ARIMA (1,d,1), tenemos

$$\hat{\phi}_1 = \frac{\gamma_2}{\gamma_1}$$

y de

$$\gamma_1 = \frac{(1 - \hat{\theta}_1 \hat{\phi}_1)(\hat{\theta}_1 - \hat{\theta}_1)}{1 - \hat{\theta}_1^2 - 2\hat{\phi}_1 \hat{\theta}_1}$$

$\hat{\theta}_1$  puede ser obtenido, el cual resultará con valores alternativos, pero escogeremos aquel que cumpla la condición de invertibilidad.

Una vez obtenidos los estimadores iniciales, siguiendo nuestro método podremos dar una buena aproximación a los verdaderos valores de los parámetros.

## PREDICCIÓN

Una vez hecha la identificación tentativa del modelo y la estimación de sus parámetros, el último paso a dar, y que al final de cuentas es el objetivo principal del análisis de series de tiempo, es utilizar el modelo propuesto para predecir en algún particular tiempo  $t$ .

Sin embargo, es importante hacer notar que las predicciones que se hagan, estarán sujetas a errores de identificación así como de estimación de los parámetros ya que no debemos olvidar que el modelo con el cual vamos a predecir es solamente una aproximación del verdadero proceso.

Para ser coherentes con nuestro objetivo principal de utilizar la historia de los datos observados para predecir, empezaremos nuestra discusión a partir de la esperanza condicional de cada observación futura ya que además la esperanza condicional \* tiene la propiedad de ser la predicción de mínimo error cuadrático medio; en otras palabras, no hay otra predicción extrapolar que produzca errores cuyos cuadrados tengan valor esperado más pequeño. Así denotaremos por  $\hat{z}_t(l)$  a la esperanza de  $Z_{t+l}$  dada la serie hasta el tiempo  $t$  y la historia del proceso hasta  $t$ , por  $H_t$ , ahora, sea  $F$  cualquier otra predicción de  $Z_{t+l}$ , entonces el error cuadrático medio para  $F$ , será:

$$\mathbb{E} \left[ (Z_{t+k} - F)^2 / H_t \right] = \mathbb{E} \left[ (Z_{t+k} - (\hat{Z}_t(k) + g))^2 / H_t \right]$$

Donde  $g$  es el error entre  $F$  y la esperanza condicional de la predicción, por tanto:

$$\begin{aligned} \mathbb{E} \left[ (Z_{t+k} - F)^2 / H_t \right] &= \mathbb{E} \left[ (Z_{t+k} - \hat{Z}_t(k))^2 / H_t \right] + 2g \mathbb{E} \left[ Z_{t+k} - \hat{Z}_t(k) / H_t \right] + g^2 \\ &= \mathbb{E} \left[ (Z_{t+k} - \hat{Z}_t(k))^2 / H_t \right] + g^2 \end{aligned}$$

que es solamente el error cuadrático medio para la predicción condicional esperada más una constante  $g$  que debe ser positiva si  $F$  difiere de  $\hat{Z}_t(k)$ , ahora, si  $g=0$  concluimos que  $\hat{Z}_t(k)$  es la predicción con error cuadrático mínimo.

Ahora veamos la forma de calcular la esperanza condicional de una predicción un período de tiempo adelante; para esto consideremos el proceso en la forma de ecuación de diferencia, una vez hecho esto, supongamos que estamos en el tiempo  $t$ , conociendo  $H_t$  y deseamos predecir un período de tiempo adelante, es decir,  $\hat{Z}_t(k)$ , el cual es la esperanza condicional  $\mathbb{E}(Z_{t+k} / H_t)$ . Cuando es observado el siguiente período  $Z_{t+1}$  estará dado por:

$$Z_{t+1} = \Phi Z_t + \dots + \Phi_{p+1} Z_{t-(p+1)+1} + \theta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q+1}$$

Note que los coeficientes  $\Phi_i$  están determinados por  $\phi_i$  y el valor de  $\theta$  en función de la forma de la ecuación en diferencias; dado que todos los subíndices en las variables hasta el tiempo  $t$  ya han sido realizadas y  $\mathbb{E}(\varepsilon_{t+1} / H_t) = 0$  implica que

$\hat{z}_t(1)$  está dada solamente por:

$$\hat{z}_t(1) = E(z_{t+1} | H_t) = \Phi_1 z_t + \dots + \Phi_{p+d} z_{t-(p+d)+1} + \theta_1 \varepsilon_t - \dots - \theta_q \varepsilon_{t-q+1}$$

Es evidente que para la evaluación numérica tenemos un problema ya que los errores incluidos en la ecuación no son observables, sin embargo, notemos que  $\varepsilon_{t-i}$  es el error en la predicción  $z_{t-i}$  hecha en el período  $t-i-1$ . Para ver esto mas claramente, considere el ejemplo de calcular predicciones para un modelo  $MA(2)$ ; en base a este modelo sabemos que la predicción para  $\hat{z}_t(1)$  está dada por:

$$\hat{z}_t(1) = \rho - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-1}$$

Obviamente para evaluar esta expresión necesitamos los valores de  $\varepsilon_t$  y  $\varepsilon_{t-1}$ , sin embargo nuestra predicción previa fué  $\hat{z}_{t-1}(1)$  y el error producido por esta fue:

$$\begin{aligned} z_t - \hat{z}_{t-1}(1) &= (\rho + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}) - (\rho - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}) \\ &= \varepsilon_t \end{aligned}$$

De igual manera  $\varepsilon_{t-1}$  esta dado por el error producido en el período  $z_{t-1} - \hat{z}_{t-2}$ , es decir  $z_{t-1} - \hat{z}_{t-2}(1)$ , entonces de esta forma podemos evaluar  $\hat{z}_t(1)$ , simplemente

$$\hat{z}_t(1) = \rho - \theta_1 [z_t - \hat{z}_{t-1}(1)] - \theta_2 [z_{t-1} - \hat{z}_{t-2}(1)]$$

siguiendo este procedimiento, es claro, que cuando llegemos a la primera predicción  $\hat{z}_0$  nos encontramos con el problema de no poder evaluar los errores pasados, entonces para calcular la primera predicción, debemos hacer alguna suposi-

ción sobre los valores iniciales de los errores, y la forma mas adecuada de hacer dicha suposición es utilizar su esperanza, es decir, suponiéndolos igual a cero. Por supuesto estos errores iniciales no fueron cero y por lo tanto la primera y subsecuentes predicciones se distorsionarán debido a esa suposición. Sin embargo es importante notar que debido a la condición de invertibilidad del proceso sobre los parámetros, esta cantidad decrecerá conforme nos alejemos del período en el cual se hizo la primera predicción, debido a esto es importante que el analista, en la práctica empiece calculando las predicciones al inicio de la serie de datos para tener disponible la secuencia mas larga posible de errores pasados para calcular la predicción de verdadero interés. Para aclarar lo dicho anteriormente, considere el proceso  $MA(1)$  teniendo los datos observados  $z_1, \dots, z_t$ , así la primera predicción sería

$$\hat{z}_0(1) = -\theta_1 \epsilon_0$$

Entonces bajo la suposición antes expuesta ( $\epsilon_0 = 0$ ), pero para ver como decrece conforme nos alejamos del inicio, asignemosle  $\hat{\epsilon}_0$ , entonces

$$\hat{z}_0(1) = -\theta_1 \hat{\epsilon}_0$$

y las subsecuentes serían

$$\hat{z}_1(1) = -\theta_1 \hat{\epsilon}_1 = -\theta_1 [z_1 - \hat{z}_0(1)] = -\theta_1 z_1 + \theta_1^2 \hat{\epsilon}_0$$



$$\hat{z}_2(1) = -\theta_1 \hat{\epsilon}_2 = -\theta_1 [z_2 - \hat{z}_1(1)] = -\theta_1 z_2 - \theta_1^2 z_1 + \theta_1^3 \hat{\epsilon}_0$$

⋮

$$\begin{aligned} \hat{z}_t(1) &= -\theta_1 \hat{\epsilon}_t = -\theta_1 [z_t - \hat{z}_{t-1}(1)] \\ &= -\theta_1 z_t - \theta_1^2 z_{t-1} - \dots - \theta_1^t z_1 + \theta_1^{t+1} \hat{\epsilon}_0 \end{aligned}$$

Note que  $\hat{\epsilon}_0$  podemos pensar que consiste de dos partes, el verdadero valor de  $\epsilon_0$  pero desconocido y la diferencia entre la suposición y el verdadero valor, esto es,  $\hat{\epsilon}_0 - \epsilon_0$ .

Así 
$$\theta_1^{t+1} \hat{\epsilon}_0 = \theta_1^{t+1} \epsilon_0 + \theta_1^{t+1} (\hat{\epsilon}_0 - \epsilon_0)$$

dado que  $\theta_1$  satisface la condición de invertibilidad, es decir,  $|\theta_1| < 1$  entonces el término  $\theta_1^{t+1} (\hat{\epsilon}_0 - \epsilon_0)$  decrece conforme  $t$  crece, es por esto que el analista debería reconstruir la secuencia de predicciones desde el inicio de los datos aunque el solo estuviera interesado en algún particular valor de  $z_t$ .

Como último punto de interés, y parte final de este trabajo, veremos el problema de predicción, pero ahora, para predicciones  $l$  tiempos adelante, es decir

$$\hat{z}_t(l)$$

note que podemos tener directamente

$$\begin{aligned} \hat{z}_t(l) &= E(z_{t+l} | H_t) = \theta_1 E(z_{t+l-1} | H_t) + \dots \\ &\dots + \theta_1^{l-1} E(z_{t+1} | H_t) + \theta_1^l \epsilon_t - \dots - \theta_1^l \epsilon_t - (4.4) \end{aligned}$$

ya que todos los errores futuros tienen esperanza cero y los errores pasados desaparecen completamente si  $l > q$ , entonces la secuencia de cálculos es:

$$\hat{Z}_t(1) = \Phi_1 Z_t + \dots + \Phi_{p+d} Z_{t-(p+d)} + \rho - \theta_1 \epsilon_t - \dots - \theta_q \epsilon_{t-q+1}$$

$$\hat{Z}_t(2) = \Phi_1 \hat{Z}_t(1) + \Phi_2 Z_t + \dots + \Phi_{p+d} Z_{t-(p+d)+2} + \rho - \theta_2 \epsilon_t - \dots - \theta_q \epsilon_{t-q+2}$$

en general para  $l > q$

$$\hat{Z}_t(l) = \Phi_1 \hat{Z}_t(l-1) + \dots + \Phi_{p+d} \hat{Z}_t(l - \{p+d\}) + \rho \quad l > q$$

donde  $\hat{Z}_t(l-i)$  significa la observación  $Z_{t+l-i}$  si  $i \geq l$ .

Es de hacer notar que en la última ecuación desaparecen los términos de la parte de medias móviles ya que esta parte del proceso interviene solamente dentro del período 1 hasta  $q$ , estableciendo únicamente los valores iniciales para determinar las predicciones sucesivas.

## BIBLIOGRAFIA

- Anderson, O. D. Time Series Analysis and Forecasting, the Box-Jenkins approach, ed. Butterworth 1979.
- Bowerman, B. L. and O'Connell, R. T. Forecasting and Time Series ed. Duxbury 1979.
- Box, G. E. P. and Jenkins, G. M. Time Series Analysis: forecasting and control, ed. Holden-day 1976.
- Chatfield, C. The Analysis of Time Series: Theory and Practice ed. Chapman and Hall 1978.
- Makridakis, S. E. and Wheelwright, S. C. Forecasting Methods and Applications ed. Wiley 1978.
- Montgomery, D. C. and Johnson, L. A. Forecasting and Time Series Analysis, ed. McGraw-Hill 1976.
- Nelson, C. R. Applied Time Series Analysis for managerial forecasting, ed. Holden-day 1973
- Pindyck, R. S. and Rubinfeld D. L. Econometric Models and Economic Forecasts 1981.
- Thomopoulos, N. T. Applied Forecasting Methods, ed. Prentice-Hall 1980.