

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN" 30

DIVISION DE MATEMATICAS E INGENIERIA  
PROGRAMA DE ACTUARIA Y M.A.C.



SR. RUBEN MOLINA CRUZ  
Alumno de la carrera de M.A.C.  
P r e s e n t e .

Por acuerdo a su solicitud presentada con fecha 23 de -  
noviembre de 1995, me complace notificarle que esta Je--  
fatura tuvo a bien asignarle el siguiente tema de tesis:  
"ANALISIS DE FOURIER APLICADO A SERIES DE TIEMPO", el -  
cual se desarrollará como sigue:

INTRODUCCION  
CAP. I Análisis de Fourier.  
CAP. II Series de Tiempo.  
CAP. III Teoría Espectral.  
CAP. IV Estimación del Espectro.  
CAP. V Presentación del Programa AST.  
CONCLUSIONES.  
ANEXOS.  
BIBLIOGRAFIA.

Asimismo, fue asignada como Asesor de Tesis: ACT. MARIA  
DEL CARMEN GONZALEZ VIDEGARAY, profesor de esta Escuela.

Ruego a usted tomar nota que en cumplimiento de lo espe-  
cificado en la Ley de Profesiones, deberá presentar ser-  
vicio social durante un tiempo mínimo de seis meses como  
requisito básico para sustentar Examen Profesional, así-  
como de la disposición de la Coordinación de la Adminis-  
tración Escolar en el sentido de que se imprima en lugar  
visible de los ejemplares el título del trabajo de tesis  
realizado. Esta comunicación deberá imprimirse en el in-  
terior de la misma.

A T E N T A M E N T E  
"POR MI RAZA HABLARA EL ESPIRITU"  
Acatlán, Edo. Méx. abril 9 de 1996.

ACT. LAURA MARÍA RIVERA BECERRA  
Jefe del Programa de Actuarial  
y M.A.C.



cg'

**TESIS CON  
FALLA DE ORIGEN**

**TESIS CON  
FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

# **ANALISIS DE FOURIER APLICADO A SERIES DE TIEMPO**

**Tesis presentada por:  
Rubén Molina Cruz.**

**Para obtener el título de:  
Licenciado en Matemáticas Aplicadas y Computación.**

**México 1995.**

**Quisiera dedicar este trabajo a mis padres como reconocimiento a los esfuerzo y sacrificios que me hicieron estar donde estoy y mirar con esperanza el futuro.**

**A José Rubén y María Elena**

**Con todo mi amor**

**Rubén**

## **Agradecimientos:**

**A MariCarmen González:** Por haber dirigido este trabajo y darme consejo y apoyo en toda ocasión.

**A mi esposa Ana Lilia:** Por ayudarme a revisar el trabajo siendo como otro director de tesis; y por su "paciencia" ante mi insensatez.

**A mi hija Mariana:** Por alegrar día con día mi existencia con sus juegos y risas.

**A Manuel Valadez:** Por ser maestro, consejero y amigo.

**Objetivo:**

Presentar una metodología alterna en el análisis de series de tiempo basada en el espectro muestral de la serie; así como desarrollar un programa de software que ayude a aplicarla a problemas reales.

**Hipótesis:**

Existe una relación entre la función de autocovarianzas de una serie y su espectro muestral.

## TABLA DE CONTENIDOS

TABLA DE CONTENIDOS .....	3
INTRODUCCION .....	5
<b>Capítulo 1 Análisis de Fourier</b> .....	<b>7</b>
Introducción .....	8
1.1 Series de Fourier para funciones continuas .....	9
1.1.1 Funciones periódicas .....	9
1.1.2 La Función Armónica .....	10
1.1.3 Series y polinomios trigonométricos .....	14
1.1.4 Funciones y sistemas ortogonales .....	16
1.1.5 Series de Fourier basadas en sistemas ortogonales .....	18
1.1.6 Error medio cuadrático (la desigualdad de Bessel y el Teorema de Parseval) .....	23
1.1.7 Transformada de Fourier. ....	26
1.2 Series de Fourier para funciones discretas .....	29
1.2.1 Funciones ortogonales .....	29
1.2.2 Serie de Fourier para una sucesión finita .....	31
1.2.3 Series de Fourier para sucesiones periódicas .....	33
1.2.4 Transformada de Fourier .....	35
<b>Capítulo 2 Series de Tiempo</b> .....	<b>38</b>
Introducción .....	39
2.1 Series de tiempo y procesos estocásticos .....	40
2.1.1 Media y varianza de un proceso estacionario .....	41
2.1.2 Funciones de autocovarianza y autocorrelación .....	42
2.1.3 Ejemplo .....	46
2.1.4 Representación de procesos estocásticos .....	48
2.2 Series no estacionarias .....	52
2.3 Series estacionales .....	56
<b>Capítulo 3 Teoría Espectral</b> .....	<b>59</b>
Introducción .....	60
3.1 El Espectro .....	61
3.1.1 Propiedades del espectro .....	61
3.1.2 Representación espectral de la función de autocovarianza .....	64
3.1.3 Descomposición de Wold .....	67

3.2 El espectro de algunos procesos .....	68
3.2.1 El espectro y la función generadora de autocovarianzas .....	68
3.2.2 Espectro de modelos ARMA .....	68
<b>Capítulo 4. Estimación del espectro .....</b>	<b>77</b>
Introducción .....	78
4.1 El periodograma .....	79
4.1.1. Definición .....	79
4.1.2 Propiedades muestrales del espectro .....	81
4.1.3 Pruebas para componentes periódicos .....	83
4.2 El espectro muestral .....	87
4.3 El espectro suavizado .....	90
4.3.1 La ventana espectral (suavizando el dominio de las frecuencias) .....	90
4.3.2 La ventana de retardo (suavizando el dominio del tiempo) .....	92
4.3.3 Ventanas más utilizadas .....	94
4.3.4 Intervalos de confianza de las ordenadas espectrales .....	101
4.4 Estimación espectral ARMA .....	104
<b>Capítulo 5 Presentación del programa AST .....</b>	<b>106</b>
Introducción .....	107
5.1 Ejemplo .....	108
<b>Conclusiones .....</b>	<b>123</b>
<b>ANEXO I (Pruebas Matemáticas) .....</b>	<b>124</b>
Sistemas Completos. Convergencia en la media .....	125
Propiedades importantes de los sistemas completos .....	127
<b>ANEXO II (Guía del usuario de AST) .....</b>	<b>130</b>
A. Comandos de edición .....	131
B. Operaciones generales .....	139
C. Análisis .....	145
D. Estimación .....	148
D. Misceláneos .....	151
<b>Bibliografía .....</b>	<b>153</b>

## INTRODUCCION

El estudio de las series de tiempo ha venido cobrando mayor importancia en los últimos años debido a que cada día los individuos tienen que tomar decisiones que determinan el éxito o fracaso de un proyecto o empresa.

Estas decisiones se basan en la esperanza de que un fenómeno de la naturaleza siga cierto patrón y acorde a éste se toma una acción que maximice la ganancia.

De esta forma, a mayor certeza en el comportamiento futuro del fenómeno, mayor será la ganancia obtenida.

Dado lo anterior, se han inventado varios mecanismos que describen a los fenómenos o bien predicen de alguna forma su comportamiento futuro. Entre estos se encuentran los métodos que se basan en series de tiempo desarrolladas por Box y Jenkins.

Los métodos basados en las series de tiempo requieren que el fenómeno a estudiar tenga las siguientes cualidades:

- 1) Se presentan en el tiempo. Es decir, van sucediendo conforme transcurren los instantes. Más aún la cantidad de tiempo entre cada ocurrencia es igual.
- 2) Se pueden cuantificar de alguna manera. Es decir, se pueden medir.

La presente tesis tiene como objetivo apoyar a los alumnos que cursan la materia de "Modelos de Decisión con Procesos Estocásticos II" en el estudio de las series de tiempo, presentando una metodología alterna a la de Box-Jenkins que es la que se ha venido revisando hasta el momento.

Para lograr este objetivo se ha dividido esta tesis en 5 capítulos:

El capítulo 1 da las bases del estudio de Fourier, en éste capítulo se busca encontrar la representación de una serie de datos por medio de una función denominada polinomio de Fourier el cual está basado en la suma un conjunto de funciones periódicas independientes.

El objetivo del capítulo 2 es presentar a las series de tiempo. En él se define a las series de tiempo como un proceso estocástico, se dan una serie de estadísticos que describen a la serie; se establece una representación para los procesos estocásticos y se mencionan los mecanismos que permiten que un proceso estacionario se estabilice.

En el capítulo 3 se define el espectro el cual relaciona a las series de tiempo (mediante la función de autocovarianzas) con el polinomio de Fourier (cap 1 y 2). De igual forma se presentan los espectros de algunos procesos estocásticos teóricos.

El capítulo 4 describe al espectro muestral el cual se basa en una serie de datos. Además introduce el concepto de ventana, la cual se utiliza para suavizar el espectro muestral de tal forma que se pueda comparar el espectro muestral con el espectro teórico para determinar cual es el comportamiento de un conjunto de datos.

Finalmente, el capítulo 5 presenta un programa de cómputo que condensa los temas desarrollados en los capítulos anteriores de tal forma que el estudiante pueda probar bajo distintas condiciones los conceptos y comprenda mejor el material aquí presentado.

# Capítulo 1 Análisis de Fourier

## Introducción

Los temas que se desarrollarán en este capítulo se basan en los estudios realizados por el matemático francés J. B. J. Fourier, quien en 1807 encontró la manera de representar cualquier función periódica como una serie de funciones senoidales armónicamente relacionadas y, posteriormente, obtuvo también una representación para funciones aperiódicas.

La idea de Fourier puede expresarse como sigue:

Es posible representar cualquier función mediante una combinación lineal (o suma) de un conjunto (base) de funciones elementales.

Para iniciar el estudio del análisis de Fourier se darán, primeramente, algunos principios básicos, y después se procederá con la parte correspondiente a las series y transformadas de Fourier, tanto para el caso de funciones continuas como para funciones discretas (para las que este análisis fue desarrollado con posterioridad).

Es importante hacer notar que se desarrollará de una manera más completa la serie de Fourier en el caso de funciones continuas debido a que estamos más relacionados con el manejo de las integrales que con el de sumas y a que, ya que la integral no es sino una suma infinita, se puede decir que ambas presentan un comportamiento análogo. Al mismo tiempo, este desarrollo pretende ser una base para una mejor comprensión de la segunda parte de este capítulo, que es lo que realmente interesa en el estudio del análisis espectral aplicado a las series de tiempo.

## 1.1 Series de Fourier para funciones continuas

### 1.1.1 Funciones periódicas

La importancia de las funciones periódicas radica en la amplia gama de problemas matemáticos en los que tienen aplicación (sobre todo en las ramas de la econometría, ingeniería y física). Las funciones periódicas más usuales son las funciones trigonométricas (sen  $x$ , cos  $x$ , tan  $x$ , etc.).

Se dice que una función (continua o discreta) es **periódica** si existe una constante  $T > 0$  para la cual

$$f(x+T) = f(x), \quad 1.1.1.1$$

para cualquier  $x$  en el dominio de la función  $f$ . A la constante  $T$  se le conoce como **período de la función  $f$** . Gráficamente una función periódica es una curva que, una vez recorrido un intervalo de longitud  $T$ , se repite nuevamente. Resulta claro que la suma, diferencia o multiplicación de funciones con un período  $T$  es, nuevamente una función del mismo período.

Una de las características más importantes de las funciones periódicas es la posibilidad de definir las únicamente para un intervalo de longitud  $T$  y, a partir de esta definición, extenderlas por todo su dominio. Además los números  $2T, 3T, \dots$  son también períodos de ésta, es decir, cumplen con:

$$f(x) = f(x+T) = f(x+2T) = f(x+3T) = \dots,$$

lo cual se obtiene de aplicar la definición 1.1.1.1 repetidamente. Así, si el período de una función es  $T$ , entonces cualquier múltiplo entero de  $T$  ( $kT$ ) también será período de la función.

Otra de las propiedades de cualquier función continua de período  $T$  es que si ésta es integrable en algún intervalo de longitud  $T$ , entonces es integrable en todo intervalo de longitud análoga, y el valor de la integral será igual, i.e.,

$$\int_a^{a+T} f(x) dx = \int_b^{b+T} f(x) dx, \quad 1.1.1.2$$

Esta propiedad es una consecuencia de la interpretación de la integral como un área (véase fig. 1).

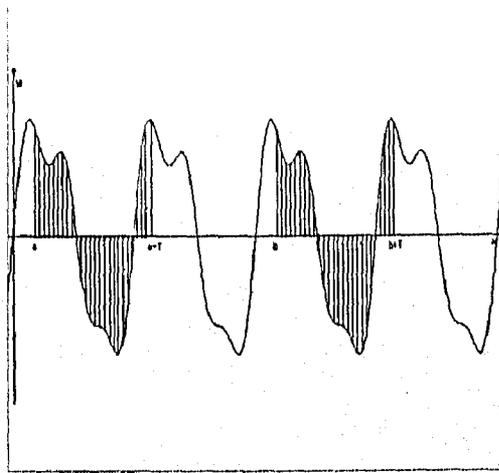


Fig. 1

### 1.1.2 La Función Armónica

La función periódica más simple, y una de las más importantes por su aplicación, es:

$$y = A \sin(\omega x + \varphi),$$

también llamada **función armónica**.

La función armónica surge del siguiente problema de física que involucra la clase más sencilla del movimiento oscilatorio: **el movimiento armónico simple**:

Supóngase que un punto **M** de masa  $m$  se mueve a lo largo de una línea recta bajo la acción de una fuerza de restitución **F** que es proporcional a la distancia de **M** con respecto a un origen fijo **O** y que siempre impulsa a **M** hacia **O** (véase fig. 2).

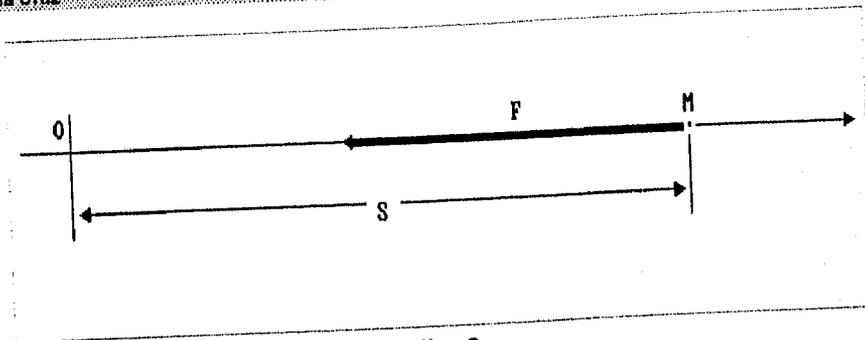


Fig. 2

Considerando la distancia al origen  $s$  positivo si  $M$  está a la derecha de  $O$  y negativo en caso contrario, entonces  $F = -ks$ , donde  $k > 0$  representa una constante de proporcionalidad. Por lo que:

$$m \frac{d^2 s}{dt^2} = -ks$$

o

$$\frac{d^2 s}{dt^2} + \omega^2 s = 0,$$

donde  $\omega^2 = k/m$ , o sea que  $\omega = \sqrt{k/m}$ .

El resultado de esta ecuación diferencial es la función  $s = A \text{sen}(\omega x + \phi)$  donde  $A$  y  $\phi$  son constantes que pueden ser calculadas a partir de la posición y velocidad del punto  $M$  en el momento inicial  $t=0$ . Además  $s$  es una función armónica que es periódica con respecto al tiempo con  $T = 2\pi/\omega$ . Así, bajo la fuerza de restitución  $F$ , el punto  $M$  se rige por un movimiento oscilatorio. El valor absoluto de  $A$  (o amplitud) indica la máxima desviación del punto  $M$  a partir de  $O$ , y  $1/T$  representa el número de oscilaciones en un intervalo que contiene  $2\pi$  unidades de tiempo, lo que explica la frecuencia. Por último  $\phi$ , llamada fase inicial, es la posición inicial del punto, ya que para  $t=0$  se tiene que  $s_0 = \text{sen}\phi$ .

El comportamiento de la curva dada por  $y = A \text{sen}(\omega x + \phi)$  se puede entender mejor si se analiza parte a parte.

Primeramente se puede hacer  $\omega = A = 1$  y  $\phi = 0$ ; esto da como resultado la curva  $y = \text{sen}(x)$  (véase fig. 3). Ahora considérese la armónica  $y = \text{sen}(\omega x)$  y sea  $z = \omega x$ , con lo que obtenemos  $y = \text{sen}(z)$ , que es una curva seno ordinaria. Entonces la gráfica  $y = \text{sen}(\omega x)$

esta dada mediante una compresión uniforme de la curva  $y=\text{sen}(x)$  a lo largo del eje  $x$  por un factor  $\omega$  si  $\omega > 1$  o a una expansión uniforme sobre el mismo eje si  $\omega < 1$  (en la fig. 4 se muestra este fenómeno para  $\omega = 3$ ).

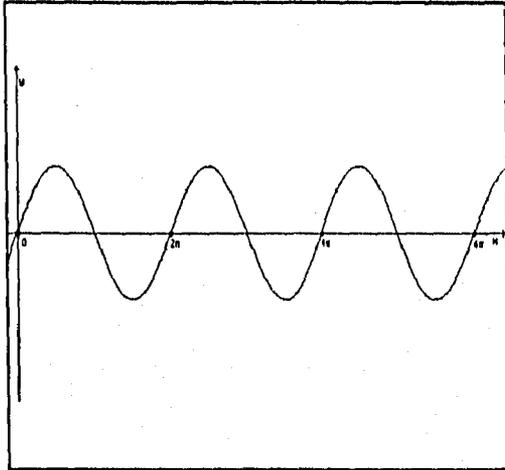


Fig. 3

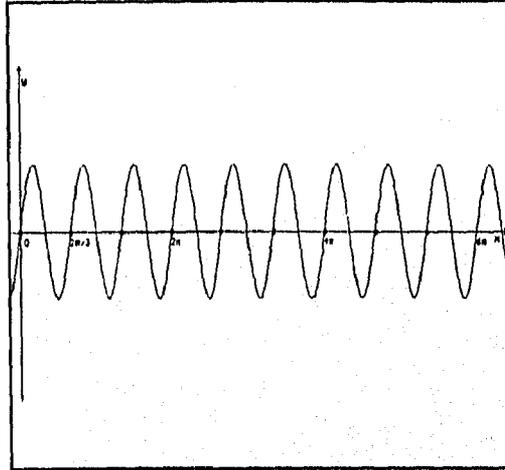


Fig. 4

Considérese ahora  $y=\text{sen}(\omega x + \varphi)$  y  $\omega x + \varphi = \omega z$ , entonces  $x = z - \varphi/\omega$ . Ya se conoce el comportamiento de la gráfica de  $y=\text{sen}(z)$  por lo que la curva  $y=\text{sen}(\omega x + \varphi)$  resulta al recorrer la gráfica  $y=\text{sen}(\omega x)$  moviéndola  $-\varphi/\omega$  unidades sobre el eje  $x$ . En la fig. 5 se contempla este caso para  $\omega = 3$  y  $\varphi = \pi/3$ , i.e.

$$y = \text{sen}\left(3x + \frac{\pi}{3}\right)$$

Finalmente, la gráfica de la armónica  $y = A \text{sen}(\omega x + \varphi)$  se obtiene al multiplicar todas las ordenadas de la curva  $y = \text{sen}(\omega x + \varphi)$  por un número  $A$ . La fig. 6 ejemplifica este caso cuando  $A = 2$ ,  $\omega = 3$  y  $\varphi = \pi/3$ , i.e.

$$y = 2 \text{sen}\left(3x + \frac{\pi}{3}\right).$$

lo cual modifica la amplitud de la función en  $A$  unidades

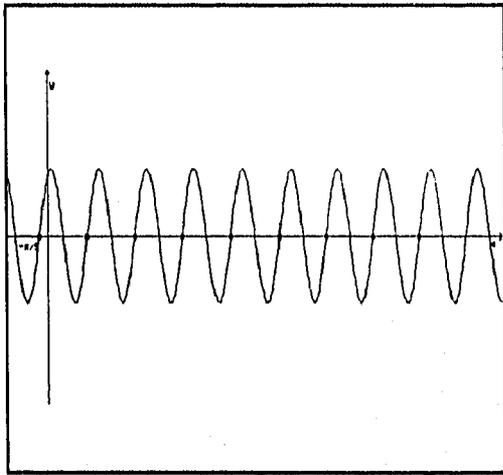


Fig. 5

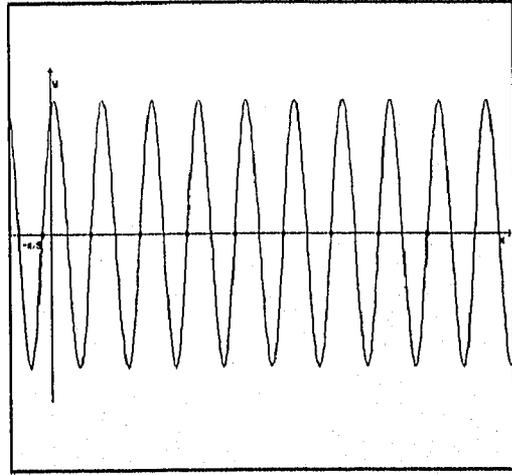


Fig.6

Por otro lado, usando identidades trigonométricas, es posible escribir

$$A \operatorname{sen}(\omega x + \varphi) = A(\cos \omega x \operatorname{sen} \varphi + \operatorname{sen} \omega x \cos \varphi).$$

Haciendo

$$a = A \operatorname{sen} \varphi, \quad b = A \cos \varphi, \quad 1.1.2.1$$

se llega a que cualquier armónica puede representarse como

$$a \cos \omega x + b \operatorname{sen} \omega x, \quad 1.1.2.2$$

de aquí, resolviendo en 1.1.2.2, se obtiene que

$$A = \sqrt{a^2 + b^2}, \quad \operatorname{sen} \varphi = \frac{a}{A} = \frac{a}{\sqrt{a^2 + b^2}}, \quad \cos \varphi = \frac{b}{A} = \frac{b}{\sqrt{a^2 + b^2}},$$

para la que  $\varphi$  se calcula fácilmente.

Un método para calcular el valor de  $\varphi$  utilizando una calculadora consiste en:

Comparar  $\operatorname{sen} \varphi = a$  con 0. Si  $a$  es mayor o igual a 0, entonces  $\varphi = \cos^{-1} a$ , en caso contrario ( $a$  negativo), entonces  $\varphi = \operatorname{sen}^{-1} a$ ,

Es conveniente introducir explícitamente el período  $T$  en 1.1.2.2. Si  $T = 2l$ , entonces:

$$\omega = \frac{2\pi}{T} = \frac{\pi}{l},$$

por lo tanto, toda armónica con período  $T = 2l$  es de la forma

$$a \cos \frac{\pi x}{l} + b \operatorname{sen} \frac{\pi x}{l}.$$

### 1.1.3 Series y polinomios trigonométricos

Sea  $T=2l$  un período dado, y sean las armónicas

$$a_k \cos \frac{\pi k x}{l} + b_k \operatorname{sen} \frac{\pi k x}{l} \quad (k=1,2,\dots) \quad 1.1.3.1$$

con frecuencias  $\omega_k = \pi k/l$  y períodos  $T_k = 2\pi/\omega_k = 2l/k$ . Debido a que

$$T = 2l = kT_k,$$

Lo que implica que  $T=2l$  es período simultáneo de todas las armónicas 1.1.3.1, ya que todo múltiplo entero de un período de una función es, nuevamente, un período. Entonces, se define la función

$$s_n(x) = A + \sum_{k=1}^n (a_k \cos \frac{\pi k x}{l} + b_k \operatorname{sen} \frac{\pi k x}{l}),$$

donde  $A$  es una constante, como el **polinomio trigonométrico de orden  $n$** .

Es común encontrar otra representación del polinomio trigonométrico obligando a que el índice  $k$  de la suma comience en 0, utilizando el hecho de que  $\operatorname{sen}(0)=0$  y  $\operatorname{cos}(0)=1$ , de tal forma que

$$s_n(x) = \sum_{k=0}^n (a_k \cos \frac{\pi k x}{l} + b_k \operatorname{sen} \frac{\pi k x}{l}),$$

El polinomio trigonométrico es una función con período  $2l$ , ya que es resultado de una suma de funciones periódicas, que, aunque está formado por funciones de comportamiento sencillo, puede llegar a presentar un comportamiento verdaderamente complejo (dependiendo de la manera en la que se escojan las constantes  $A, a_1, b_1, a_2, b_2, \dots$ ).

A continuación se muestra la gráfica del polinomio trigonométrico

$$y = 2 + \frac{3}{4} \cos 2x + 3 \operatorname{sen} x - \frac{7}{10} \operatorname{sen} 4x$$

para ilustrar lo anterior.

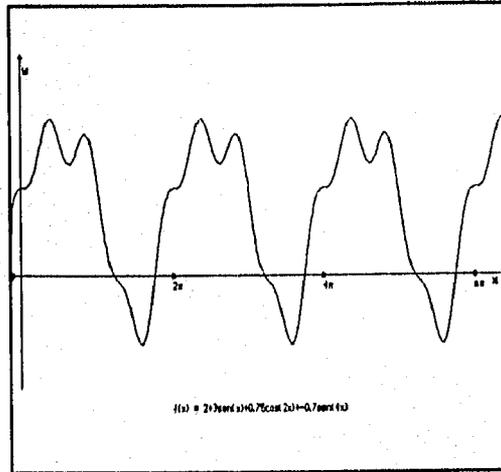


Fig. 7

Se define como **serie trigonométrica infinita** al polinomio trigonométrico infinito,

i.e.,

$$s_{\infty}(x) = A + \sum_{k=1}^{\infty} \left( a_k \cos \frac{\pi k x}{l} + b_k \operatorname{sen} \frac{\pi k x}{l} \right).$$

Si esta serie converge tendrá como período a  $T=2l$  nuevamente.

A partir de lo anterior se tratará de encontrar una representación de cualquier función periódica  $y=f(x)$  a partir de una superposición de funciones armónicas (las condiciones y particularidades de las funciones cuya representación en series trigonométricas es posible se proporcionan en el anexo I).

En este trabajo sólo se presentará el desarrollo concerniente a funciones con período  $2\pi$  ya que si

$$f(x) = A + \sum_{k=1}^{\infty} \left( a_k \cos \frac{\pi k x}{l} + b_k \operatorname{sen} \frac{\pi k x}{l} \right),$$

para cualquier función con período  $2l$ , podemos hacer  $\pi x/l = t$  o  $x = tl/\pi$  y construir una nueva función

$$\varphi(t) = f\left(\frac{tl}{\pi}\right) = A + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \operatorname{sen} kt),$$

cuyas armónicas se repiten cada  $2\pi$  unidades del dominio.

### 1.1.4 Funciones y sistemas ortogonales

Se dice que dos funciones,  $\varphi(x)$  y  $\psi(x)$  son **ortogonales** en un intervalo  $[a, b]$  si cumplen:

$$\int_a^b \varphi(x)\psi(x)dx = 0.$$

De igual forma, un **sistema de funciones reales**

$$\varphi_0(x), \varphi_1(x), \varphi_2(x), \dots, \varphi_n(x), \dots$$

1.1.4.1

es **ortogonal** en un intervalo  $[a, b]$  si

$$\int_a^b \varphi_n(x)\varphi_m(x) dx = 0 \quad (n \neq m, n, m = 0, 1, 2, \dots).$$

1.1.4.2

Y además se cumple

$$\int_a^b \varphi_n^2(x) dx \neq 0 \quad (n = 0, 1, 2, \dots).$$

En particular, el sistema (1.1.4.1) se dice que es **ortonormal** si

$$\int_a^b \varphi_n^2(x) dx = 1 \quad (n = 0, 1, 2, \dots).$$

Entonces cualquier sistema ortogonal puede normalizarse multiplicando cada una de las funciones del sistema por constantes  $\mu_n$  que deberán satisfacer

$$\int_a^b \mu_n^2 \varphi_n^2(x) dx = \mu_n^2 \int_a^b \varphi_n^2(x) dx = 1 \quad (n = 0, 1, 2, \dots),$$

lo que implica que

$$\mu_n = \frac{1}{\sqrt{\int_a^b \varphi_n^2(x) dx}}$$

Ahora se introduce la siguiente notación:

$$\|\varphi_n\| = \sqrt{\int_a^b \varphi_n^2(x) dx} \quad (n = 0, 1, 2, \dots),$$

para indicar la **norma** de la función  $\varphi_n(x)$ . Si el sistema 1.1.4.1 es ortonormal, entonces, obviamente

$$\|\varphi_n\| = 1 \quad (n = 0, 1, 2, \dots).$$

Por otro lado, se define el **sistema trigonométrico básico** como el conjunto

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx, \dots\}$$

de funciones que comparten el período  $2\pi$ . (aunque  $\sin 2n\pi$  y  $\cos 2n\pi$  tengan un período menor  $2\pi/n$ ).

Una de las propiedades del sistema trigonométrico básico es la de ser un sistema de funciones ortogonal.

Prueba:

Para cualquier  $n$  distinto de 0, se cumple

$$\int_{-\pi}^{\pi} \cos nx \, dx = \left[ \frac{\sin nx}{n} \right]_{x=-\pi}^{x=\pi} = 0, \quad 1.1.4.3$$

$$\int_{-\pi}^{\pi} \sin nx \, dx = \left[ -\frac{\cos nx}{n} \right]_{x=-\pi}^{x=\pi} = 0,$$

y

$$\|\cos_n\|^2 = \int_{-\pi}^{\pi} \cos^2 nx \, dx = \int_{-\pi}^{\pi} \frac{1 + \cos 2nx}{2} \, dx = \pi, \quad 1.1.4.4$$

$$\|\sin_n\|^2 = \int_{-\pi}^{\pi} \sin^2 nx \, dx = \int_{-\pi}^{\pi} \frac{1 - \cos 2nx}{2} \, dx = \pi.$$

Utilizando las identidades trigonométricas

$$\cos \alpha \cos \beta = \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)], \quad 1.1.4.4.a$$

$$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)]$$

Se llega a lo siguiente:

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\cos(n+m)x + \cos(n-m)x] \, dx = 0, \quad 1.1.4.5$$

$$\int_{-\pi}^{\pi} \operatorname{sen} nx \operatorname{sen} mx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\cos(n-m)x - \cos(n+m)x] \, dx = 0$$

para enteros  $m, n$  diferentes. Finalmente, usando la identidad

$$\operatorname{sen} \alpha \cos \beta = \frac{1}{2} [\operatorname{sen}(\alpha + \beta) + \operatorname{sen}(\alpha - \beta)],$$

se concluye que

$$\int_{-\pi}^{\pi} \operatorname{sen} nx \cos mx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\operatorname{sen}(n+m)x + \operatorname{sen}(n-m)x] \, dx = 0 \quad 1.1.4.6$$

para cualquier  $n$  y  $m$ .

De 1.1.4.3, 1.1.4.4, 1.1.4.5 y 1.1.4.6 se deriva que el sistema trigonométrico básico es un sistema de funciones ortogonales.

### 1.1.5 Series de Fourier basadas en sistemas ortogonales

Sea  $f(x)$  una función definida en un intervalo  $[a, b]$ . Suponga que  $f(x)$  puede ser representada como la suma de una serie que involucra las funciones del sistema 1.1.4.1, es decir, piense que para todo punto del intervalo  $[a, b]$

$$f(x) = c_0 \varphi_0(x) + c_1 \varphi_1(x) + \dots + c_n \varphi_n(x) + \dots, \quad 1.1.5.1$$

donde  $c_0, c_1, \dots, c_n, \dots$  son constantes a ser calculadas. Para hacer esto se necesita, primeramente, suponer que la serie

$$\begin{aligned} f(x) \varphi_n(x) &= c_0 \varphi_0(x) \varphi_n(x) + c_1 \varphi_1(x) \varphi_n(x) + \dots \\ &+ c_{n-1} \varphi_{n-1}(x) \varphi_n(x) + c_n \varphi_n^2(x) \quad (n = 0, 1, 2, \dots) \\ &+ c_{n+1} \varphi_{n+1}(x) \varphi_n(x) + \dots \end{aligned} \quad 1.1.5.2$$

construida al efectuar el producto de 1.1.5.1 por  $\varphi_n(x)$ , puede ser integrada término a término. De acuerdo a 1.1.4.2, esta integración implica

$$\int_a^b f(x)\varphi_n(x) dx = c_n \int_a^b \varphi_n^2(x) dx \quad (n = 0, 1, 2, \dots).$$

Por lo que se concluye que

$$c_n = \frac{\int_a^b f(x)\varphi_n(x) dx}{\int_a^b \varphi_n^2(x) dx} = \frac{\int_a^b f(x)\varphi_n(x) dx}{\|\varphi_n\|^2} \quad (n = 0, 1, 2, \dots). \quad 1.1.5.3$$

Asúmase ahora que se da una función  $f(x)$  definida en el intervalo  $[a, b]$ , y se desea efectuar la expansión de  $f(x)$  con respecto a las funciones del sistema 1.1.4.1, sin conocer de antemano si tal expansión es posible. Si la expansión es factible (y si la condición de integrabilidad término a término se cumple también), entonces, como se ha visto, se alcanza la expresión 1.1.5.3. Así, si se intenta obtener la expansión de  $f(x)$ , resulta natural examinar primero la serie con los coeficientes dados por 1.1.5.3, y verificar si esta serie converge a  $f(x)$ . Los coeficientes dados por 1.1.5.3 son llamados **coeficientes de Fourier de  $f(x)$  con respecto al sistema 1.1.4.1**, y la serie asociada a estos coeficientes se conoce como la **serie de Fourier de  $f(x)$  con respecto al sistema 1.1.4.1**.

Si el sistema 1.1.4.1 está normalizado, entonces la fórmula para calcular los coeficientes de Fourier toman la forma

$$c_n = \int_a^b f(x)\varphi_n(x) dx \quad (n = 0, 1, 2, \dots).$$

Mientras no se tenga certeza con respecto a la convergencia de la serie de Fourier de  $f(x)$ , se denotará

$$f(x) \sim c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots$$

Si las funciones del sistema 1.1.4.1 son continuas, y la serie en el lado derecho de 1.1.5.1 es uniformemente convergente, entonces resulta sencillo demostrar que también la serie 1.1.5.2 es convergente uniformemente y, por ende, puede ser integrada término a término<sup>1</sup>.

Debido a lo anterior, se infiere que si las funciones del sistema 1.1.4.1 son continuas y si la expansión 1.1.5.1 es convergente, entonces 1.1.5.1 es la serie de Fourier de  $f(x)$  cuyos componentes  $c_n$  están dados por 1.1.5.3.

<sup>1</sup>Los términos de convergencia uniforme de las serie se resumen en el anexo I

A partir de lo anterior, es posible, dado un sistema ortogonal cualesquiera y una función en un intervalo  $[a, b]$ , construir la serie de Fourier asociada.

A manera de ejemplo, y por las características de sencillez en su comportamiento y aplicabilidad, se encontrará la serie de Fourier de una función  $f(x)$  definida en el intervalo  $[-\pi, \pi]$  con respecto al sistema trigonométrico básico presentado en la sección anterior.

Para iniciar este proceso es necesario suponer que una función  $f(x)$  de período  $\pi$  tiene una expansión de la forma

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad 1.1.5.4$$

El objetivo es encontrar los coeficientes  $a_0, a_k, b_k$  ( $k=1, 2, \dots$ ) que hacen que la igualdad 1.1.5.4 se satisfaga.

Para hacer esto es necesario asumir que la serie 1.1.5.4 es integrable término a término, es decir, la suma de las integrales es igual a la integral de las sumas. Entonces, aplicando la fórmula 1.1.5.3,

$$c_n = \frac{\int_a^b f(x) \varphi_n(x) dx}{\|\varphi_n\|^2} \quad (n = 0, 1, 2, \dots)$$

Para  $a_0/2$

$$\frac{a_0}{2} = \frac{\int_{-\pi}^{\pi} f(x) dx}{\int_{-\pi}^{\pi} 1 dx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx,$$

lo que implica

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx.$$

Para los coeficientes  $a_k$  ( $k=1, 2, \dots$ )

$$a_n = \frac{\int_{-\pi}^{\pi} f(x) \cos nx dx}{\|\cos_n\|^2} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx$$

debido a 1.1.4.4.

Utilizando también 1.1.4.4, para  $b_k$  ( $k=0,1,2,\dots$ ) se tiene

$$b_n = \frac{\int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx}{\|\operatorname{sen}_n\|^2} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx$$

Finalmente, y ya que  $\cos(0)=1$ , los coeficientes encontrados pueden ser reescritos como:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \quad (n = 0,1,2,\dots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx \, dx \quad (n = 1,2,3,\dots).$$

1.1.5.5

De este modo, si  $f(x)$  es una función integrable en cualquier intervalo de longitud del período ( $2\pi$ ), entonces los coeficientes  $a_n$  y  $b_n$  están dados por las fórmulas presentadas en 1.1.5.5.

Los coeficientes  $a_n$  y  $b_n$  calculados por las fórmulas 1.1.5.5 se denominan **Coefficientes de Fourier de  $f(x)$  con respecto al sistema trigonométrico básico**, y la serie trigonométrica con tales coeficientes es llamada **Serie de Fourier de  $f(x)$  con respecto al sistema trigonométrico básico**. Asimismo es conveniente subrayar el hecho de que las fórmulas presentadas en 1.1.5.5 involucran la integración de una función de período  $\pi$ ; por lo tanto, el intervalo de integración  $[-\pi,\pi]$  puede ser reemplazado por cualquier otro de longitud  $\pi$  de manera que, junto con las funciones 1.1.5.5, se tendrán

$$a_n = \frac{1}{\pi} \int_a^{a+2\pi} f(x) \cos nx \, dx \quad (n = 0,1,2,\dots),$$

$$b_n = \frac{1}{\pi} \int_a^{a+2\pi} f(x) \operatorname{sen} nx \, dx \quad (n = 1,2,3,\dots).$$

Aplicando la fórmula de Euler se puede relacionar la serie de Fourier con respecto al sistema trigonométrico básico:

La fórmula de Euler es

$$e^{i\varphi} = \cos\varphi + i \operatorname{sen}\varphi$$

de la que se desprende

$$\cos\varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2}, \quad \operatorname{sen}\varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i}.$$

Entonces es posible expresar

$$\begin{aligned}\cos nx &= \frac{e^{inx} - e^{-inx}}{2}, \\ \operatorname{sen} nx &= \frac{e^{inx} - e^{-inx}}{2i} = i \frac{-e^{inx} + e^{-inx}}{2}.\end{aligned}$$

Substituyendo esas expresiones en 1.1.5.4 resulta

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( \frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx} \right), \quad 1.1.5.6$$

Haciendo

$$c_0 = \frac{a_0}{2}, \quad c_n = \frac{a_n - ib_n}{2}, \quad c_{-n} = \frac{a_n + ib_n}{2} \quad 1.1.5.7$$

Por lo tanto la serie 1.1.5.6 y, al mismo tiempo la serie 1.1.5.4 tomarán como forma alternativa

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Esta es, precisamente, **la forma compleja de la serie de Fourier de  $f(x)$** ; y los coeficientes  $c_n$  dados por 1.1.5.7 se conocen como los **coeficientes complejos de Fourier de  $f(x)$** . Estos coeficientes satisfacen la siguiente relación

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \quad (n=0, \pm 1, \pm 2, \dots).$$

En efecto, utilizando la fórmula de Euler y 1.1.5.7

$$\begin{aligned}\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx &= \frac{1}{2\pi} \left[ \int_{-\pi}^{\pi} f(x) \cos nx dx - i \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx dx \right] \\ &= \frac{1}{2} (a_n - ib_n) = c_n\end{aligned}$$

para índices positivos y

$$\begin{aligned}\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} dx &= \frac{1}{2\pi} \left[ \int_{-\pi}^{\pi} f(x) \cos(nx) dx + i \int_{-\pi}^{\pi} f(x) \operatorname{sen}(nx) dx \right] \\ &= \frac{1}{2} (a_n + ib_n) = c_{-n}\end{aligned}$$

para índices negativos.

Es importante tener en mente que si  $f(x)$  es real, entonces  $c_n$  y  $c_{-n}$  son complejos

conjugados. Esto es una consecuencia inmediata de 1.1.5.7.

### 1.1.6 Error medio cuadrático (la desigualdad de Bessel y el Teorema de Parseval)

Sea  $f(x)$  una función integrable por su cuadrado<sup>2</sup> arbitraria definida en un intervalo  $[a, b]$  y sea  $\sigma_n(x)$  una combinación lineal de las primeras  $n+1$  funciones del sistema 1.1.4.1, i.e.,

$$\sigma_n(x) = \Upsilon_0 \varphi_0(x) + \Upsilon_1 \varphi_1(x) + \dots + \Upsilon_n \varphi_n(x), \quad 1.1.6.1$$

donde  $\Upsilon_0, \Upsilon_1, \dots, \Upsilon_n$  son constantes. Por hipótesis todas las funciones del sistema ortogonal 1.1.4.1 son integrables por su cuadrado (ver 1.1.4.2). Entonces, la combinación lineal  $\sigma_n(x)$  y la diferencia  $f(x) - \sigma_n(x)$ ,  $[n = 0, 1, 2, \dots]$  son también funciones integrables por su cuadrado.

Considérese ahora la cantidad

$$\delta_n = \int_a^b [f(x) - \sigma_n(x)]^2 dx, \quad 1.1.6.2$$

que se conocerá como el **error medio cuadrático** al aproximarse a  $f(x)$  por medio de  $\sigma_n(x)$ .

Ahora se deben escoger los coeficientes  $\Upsilon_0, \Upsilon_1, \dots, \Upsilon_n$  que minimicen el error cuadrático medio, i.e.,  $\delta_n$

De 1.1.6.2 se deduce que

$$\delta_n = \int_a^b f^2(x) dx - 2 \int_a^b f(x) \sigma_n(x) dx + \int_a^b \sigma_n^2(x) dx. \quad 1.1.6.3$$

Debido a 1.1.6.1 se tiene que

$$\int_a^b f(x) \sigma_n(x) dx = \sum_{k=0}^n \Upsilon_k \int_a^b f(x) \varphi_k(x) dx.$$

Y, de acuerdo a 1.1.5.3

<sup>2</sup> Una función  $f(x)$  definida en un intervalo  $[a, b]$  es integrable por su cuadrado si  $f(x)$  y su cuadrado ( $f^2(x)$ ) son integrables en  $[a, b]$ .

$$\int_a^b f(x)\varphi_k(x) dx = c_k \|\varphi_k\|^2 \quad (k = 0, 1, 2, \dots),$$

donde los  $c_k$  son los coeficientes de Fourier de la función  $f(x)$ , por lo que

$$\int_a^b f(x)\sigma_n(x) dx = \sum_{k=0}^n c_k \|\varphi_k\|^2. \quad 1.1.6.4$$

Más aún

$$\begin{aligned} \int_a^b \sigma_n^2(x) dx &= \int_a^b \left( \sum_{k=0}^n \Upsilon_k \varphi_k(x) \right)^2 dx \\ &= \int_a^b \left( \sum_{k=0}^n \Upsilon_k^2 \varphi_k^2(x) + \sum_{p \neq q} \Upsilon_p \Upsilon_q \varphi_p(x) \varphi_q(x) \right) dx \\ &= \sum_{k=0}^n \Upsilon_k^2 \int_a^b \varphi_k^2(x) dx + \sum_{p \neq q} \Upsilon_p \Upsilon_q \int_a^b \varphi_p(x) \varphi_q(x) dx. \end{aligned}$$

La última suma, extendida sobre todos los índices  $p$  y  $q$  que no exceden a  $n$ , es igual a cero por las propiedades del sistema ortogonal 1.1.4.1 dando como resultado

$$\int_a^b \sigma_n^2(x) dx = \sum_{k=0}^n \Upsilon_k^2 \|\varphi_k\|^2. \quad 1.1.6.5$$

Sustituyendo 1.1.6.4 y 1.1.6.5 en 1.1.6.3 se obtiene

$$\begin{aligned} \delta_n &= \int_a^b f^2(x) dx - 2 \sum_{k=0}^n \Upsilon_k c_k \|\varphi_k\|^2 + \sum_{k=0}^n \Upsilon_k^2 \|\varphi_k\|^2 \\ &= \int_a^b f^2(x) dx + \sum_{k=0}^n (c_k - \Upsilon_k)^2 \|\varphi_k\|^2 - \sum_{k=0}^n c_k^2 \|\varphi_k\|^2. \end{aligned}$$

Donde

$$\int_a^b f^2(x) dx = \text{CONSTANTE}, \quad \sum_{k=0}^n c_k^2 \|\varphi_k\|^2 = \text{CONSTANTE},$$

es decir, no dependen de  $\Upsilon_0, \Upsilon_1, \dots, \Upsilon_n$  y obviamente el número  $\delta_n$  es mínimo cuando

$$\sum_{k=0}^n (c_k - \Upsilon_k)^2 \|\varphi_k\|^2 = 0,$$

que es equivalente a decir que

$$\Upsilon_k = c_k \quad (k = 0, 1, 2, \dots, n).$$

Así, el error cuadrático medio es mínimo cuando los coeficientes en la combinación lineal 1.1.6.1 son los coeficientes de Fourier.

Al mínimo error cuadrático medio se le conocerá como  $\Delta_n$  y será

$$\begin{aligned}\Delta_n &= \int_a^b \left[ f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx \\ &= \int_a^b f^2(x) dx - \sum_{k=0}^n c_k^2 \|\varphi_k\|^2\end{aligned}\quad 1.1.6.6$$

En base a su definición, el número  $\Delta_n$  es no negativo, por lo que la ecuación anterior produce la siguiente desigualdad

$$\int_a^b f^2(x) dx \geq \sum_{k=0}^n c_k^2 \|\varphi_k\|^2$$

conocida como la **desigualdad de Bessel**.

Por otra parte, se dice que un sistema 1.4.1.1 es completo si para cualquier función  $f(x)$ , integrable por su cuadrado, se cumple la igualdad

$$\int_a^b f^2(x) dx = \sum_{k=0}^{\infty} c_k^2 \|\varphi_k\|^2$$

que se conoce como igualdad de Bessel.

En particular, se puede demostrar que el sistema trigonométrico básico es completo<sup>3</sup>, por lo que la desigualdad de Bessel se convierte en

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2),$$

donde  $f(x)$  es una función arbitraria integrable por su cuadrado, y  $a_n$ ,  $b_n$  son los coeficientes de Fourier de  $f(x)$ . La fórmula anterior es conocida como el **teorema de Parseval**

<sup>3</sup>Ver anexo I

## 1.1.7 Transformada de Fourier.

Considérese ahora el problema de encontrar una expansión en serie de Fourier para una función  $f(x)$  referida al sistema trigonométrico básico.

Para hacer esto suponga que  $f(x)$  esta definida para todo número real, pero  $f(x)=0$  para  $|x| > T/2$ .

Se define una nueva función  $g(x)$  de período  $T$  que es idéntica a  $f(x)$  en el intervalo  $[-T/2, T/2]$ , y tal que,

$$\begin{cases} g(x) = f(x), & -T/2 \leq x \leq T/2 \\ g(x+jT) = g(x), & j = \pm 1, \pm 2, \dots \end{cases}$$

Retomando la transformación hecha al final de la sección 1.1.3, se puede expresar

$$g(x) = \varphi(\omega_0 x) = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega_0 x}$$

donde

$$c_k = \frac{1}{T} \int_{-T/2}^{T/2} g(x) e^{ik\omega_0 x} dx$$

para  $\omega_0 = 2\pi/T$ . Debido a que  $f(x) = g(x)$  en el intervalo de integración de  $c_k$ , entonces

$$c_k = \frac{\omega_0}{2\pi} \int_{-T/2}^{T/2} f(x) e^{-ik\omega_0 x} dx$$

$$= \frac{\omega_0}{2\pi} \int_{-T/2}^{T/2} f(u) e^{-ik\omega_0 u} du$$

y

$$\begin{aligned}
 g(x) &= \sum_{k=-\infty}^{\infty} c_k e^{ik\omega_0 x} \\
 &= \sum_{k=-\infty}^{\infty} \left[ \frac{\omega_0}{2\pi} \int_{-T/2}^{T/2} f(u) e^{-ik\omega_0 u} du \right] e^{ik\omega_0 x} \\
 &= \sum_{k=-\infty}^{\infty} \frac{\omega_0}{2\pi} \int_{-T/2}^{T/2} f(u) e^{ik\omega_0(x-u)} du \\
 &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} H(k\omega_0) \omega_0
 \end{aligned}
 \tag{1.1.7.1}$$

donde

$$H(k\omega_0) = \int_{-T/2}^{T/2} f(u) e^{ik\omega_0(x-u)} du$$

Por otro lado, cada término de la suma 1.1.7.1 representa el área de un rectángulo de altura  $H(k\omega_0)$  y base  $\omega_0$ . A medida que  $T$  tiende a infinito,  $g$  tiende a  $f$ ,  $\omega_0$  tiende a 0 y el límite de la suma 1.1.7.1 se aproxima a una integral, al grado que

$$f(x) = \lim_{T \rightarrow \infty} g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) d\omega$$

$$H(\omega) = \lim_{T \rightarrow \infty} H(k\omega_0) = \int_{-\infty}^{\infty} f(u) e^{i\omega(x-u)} du$$

y

$$\begin{aligned}
 f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u) e^{i\omega(x-u)} du d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u) e^{-i\omega u} du e^{i\omega x} d\omega \\
 &= \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega
 \end{aligned}
 \tag{1.1.7.2}$$

aquí se debe hacer la aclaración de que, en el límite, la cantidad  $k\omega_0$  se convierte en una variable continua, la cual será denotada por  $\omega$ . Esto es,

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \tag{1.1.7.3}$$

La función  $F(x)$  definida en 1.1.7.3 es llamada la **transformada de Fourier de la integral de Fourier de la función  $f(x)$** , y, al mismo tiempo,  $f(x)$  es conocida como la **transformada inversa de Fourier de  $F(x)$** .

El teorema de Parseval aplicado a las funciones encontradas muestra la siguiente relación

$$\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x)|^2 dx$$

Otro hecho que debe ser tomado en cuenta es que, para algunas funciones no periódicas, la transformada de Fourier puede no existir. Sin embargo las siguientes condiciones dadas por Dirichlet son suficientes para garantizar su existencia:

1.  $f(x)$  es absolutamente integrable, es decir

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty.$$

2.  $f(x)$  sólo tiene un número finito de discontinuidades en cualquier intervalo finito.

## 1.2 Series de Fourier para funciones discretas

### 1.2.1 Funciones ortogonales

Sean  $\phi_k(t)$  y  $\phi_j(t)$  funciones complejas definidas en el dominio  $D$ , el cual es un subconjunto de la recta real. Estas funciones del tiempo definidas en un conjunto discreto son ortogonales si

$$\sum_{t \in D} \phi_k(t) \phi_j^*(t) \begin{cases} = 0, & k \neq j \\ \neq 0, & k = j \end{cases} \quad 1.2.1.1$$

Suponga que el conjunto de funciones trigonométricas  $\text{sen}(2\pi kt/n)$  y  $\text{cos}(2\pi kt/n)$  están definidas sobre un número finito  $n$  de puntos, i.e, para  $t=1,2,\dots,n$ . Para  $k=0,1,2,\dots, [n/2]$ , donde  $[x]$  se define como el máximo entero menor o igual a  $x$ , el sistema contiene exactamente  $n$  funciones no nulas. Esto es consecuencia de que la función seno es igual a 0 para  $k=0$  y para  $k=[n/2]$  si  $n$  es par. Específicamente  $[n/2]=n/2$  cuando  $n$  es par, y el sistema consta de  $\text{cos}(2\pi 0kt/n)=1$ ,  $\text{cos}(2\pi kt/n)$  y  $\text{sen}(2\pi kt/n)$  para  $k=1,2,\dots, n/2-1$  y  $\text{cos}(2\pi(n/2)t/n)=\text{cos}(\pi t)=(-1)^t$ . Si  $n$  es impar,  $[n/2]=(n-1)/2$  y el sistema consiste de  $\text{cos}(2\pi 0kt/n)=1$ ,  $\text{cos}(2\pi kt/n)$  y  $\text{sen}(2\pi kt/n)$  para  $k=1,2,\dots, (n-1)/2$ . En cualquiera de los casos, existen exactamente  $n$  funciones en el sistema.

$$\text{sen}(2\pi kt/n), \text{cos}(2\pi kt/n) \quad (k=0,1,2,\dots,[n/2]) \quad 1.2.1.2$$

Ahora se probará que el sistema 1.2.1.2 es un sistema ortogonal. Para hacer esto se aprovechará la ecuación de Euler vista en la primera parte.

$$e^{i\varphi} = \cos\varphi + i\text{sen}\varphi$$

Y un resultado conocido del estudio de series

$$\sum_{k=1}^n y^k = \frac{y-y^{n+1}}{1-y} = y \frac{1-y^n}{1-y}, \quad |y| < 1 \quad y \in \mathbb{C}$$

Entonces si  $y=e^{i\varphi}$ , se tiene que

$$\begin{aligned} \sum_{t=1}^n e^{i\varphi t} &= e^{i\varphi} \left( \frac{1-e^{i\varphi n}}{1-e^{i\varphi}} \right) = e^{i\varphi} \left( \frac{e^{i\varphi n} - 1}{e^{i\varphi} - 1} \right) \\ &= e^{i\varphi} \left[ \frac{e^{i\varphi n/2} (e^{i\varphi n/2} - e^{-i\varphi n/2}) / 2i}{e^{i\varphi/2} (e^{i\varphi/2} - e^{-i\varphi/2}) / 2i} \right] \end{aligned}$$

$$\begin{aligned}\sum_{t=1}^n e^{i\varphi t} &= e^{i\varphi(n+1)/2} \frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)} \\ &= \cos\left(\frac{\varphi(n+1)}{2}\right) \frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)} + i \text{sen}\left(\frac{\varphi(n+1)}{2}\right) \frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)}\end{aligned}$$

Pero

$$\sum_{t=1}^n e^{i\varphi t} = \sum_{t=1}^n \cos\varphi t + i \sum_{t=1}^n \text{sen}\varphi t.$$

Así

$$\sum_{t=1}^n \cos\varphi t = \cos\left(\frac{\varphi(n+1)}{2}\right) \frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)}, \quad 1.2.1.3$$

$$\sum_{t=1}^n \text{sen}\varphi t = \text{sen}\left(\frac{\varphi(n+1)}{2}\right) \frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)}. \quad 1.2.1.4$$

Sea  $\varphi=2\pi k/n$ . Se tiene

$$\frac{\text{sen}(\varphi n/2)}{\text{sen}(\varphi/2)} = \frac{\text{sen}(\pi k)}{\text{sen}(\pi k/n)} = \begin{cases} n, & k = 0, \\ 0, & k = 1, 2, \dots, [n/2], \end{cases}$$

y  $\cos\varphi=1$  y  $\text{sen}\varphi=0$  para  $\varphi=0$  por lo que las ecuaciones 1.2.1.3 y 1.2.1.4 implican que

$$\sum_{t=1}^n \cos(2\pi kt/n) = \begin{cases} n, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad 1.2.1.5$$

$$\sum_{t=1}^n \text{sen}(2\pi kt/n) = 0, \quad (k = 0, 1, \dots, [n/2]). \quad 1.2.1.6$$

Ahora, auxiliados por las identidades trigonométricas 1.1.4.4.a,

$$\text{sen}\alpha \cos\beta = \frac{1}{2}[\text{sen}(\alpha+\beta) + \text{sen}(\alpha-\beta)]$$

y de 1.2.1.5 y 1.2.1.6 se obtiene

$$\begin{aligned}\sum_{t=1}^n \cos\left(\frac{2\pi kt}{n}\right) \cos\left(\frac{2\pi jt}{n}\right) &= \frac{1}{2} \sum_{t=1}^n \cos\left(\frac{2\pi t(k+j)}{n}\right) + \frac{1}{2} \sum_{t=1}^n \cos\left(\frac{2\pi t(k-j)}{n}\right) \\ &= \frac{1}{2} \cos\left(\frac{\pi t(k+j)(n+1)}{n}\right) \frac{\text{sen}(\pi t(k+j)/n)}{\text{sen}(\pi t(k+j)/n)} \\ &\quad + \frac{1}{2} \cos\left(\frac{\pi t(k-j)(n+1)}{n}\right) \frac{\text{sen}(\pi t(k-j)/n)}{\text{sen}(\pi t(k-j)/n)} \\ &= \frac{1}{2} \begin{cases} n & k+j=0 \\ 0 & k+j \neq 0 \end{cases} + \frac{1}{2} \begin{cases} n & k-j=0 \\ 0 & k-j \neq 0 \end{cases}\end{aligned}$$

$$\sum_{t=1}^n \cos\left(\frac{2\pi kt}{n}\right) \cos\left(\frac{2\pi jt}{n}\right) = \begin{cases} n & k=j=0 \text{ o } n/2 \text{ (n par)} \\ n/2 & k=j \neq 0 \text{ o } n/2 \text{ (n par)} \\ 0 & k \neq j, \end{cases}$$

Mediante un procedimiento análogo se llega a

$$\sum_{t=1}^n \operatorname{sen}\left(\frac{2\pi kt}{n}\right) \operatorname{sen}\left(\frac{2\pi jt}{n}\right) = \begin{cases} 0 & k=j=0 \text{ o } n/2 \text{ (n par)} \\ n/2 & k=j \neq 0 \text{ o } n/2 \text{ (n par)} \\ 0 & k \neq j \end{cases}$$

$$\sum_{t=1}^n \operatorname{sen}\left(\frac{2\pi kt}{n}\right) \cos\left(\frac{2\pi jt}{n}\right) = 0 \quad \text{para todo } j, k$$

A partir de la ecuación de Euler, es claro que el sistema de funciones trigonométricas 1.2.1.1 también puede ser representado en su forma compleja.

Esta forma resulta mucho más simple y compacta, además de útil en ciertas aplicaciones. En términos de esta representación, el sistema 1.2.1.1 contiene las siguientes exponenciales complejas

$$\left\{ e^{i2\pi kt/n} \quad -\frac{n}{2}+1 \leq k \leq \frac{n}{2} \text{ si } n \text{ es par y } -\frac{(n-1)}{2} \leq k \leq \frac{(n-1)}{2} \text{ si } n \text{ es impar,} \right\}$$

el cual, obviamente, es ortogonal.

## 1.2.2 Serie de Fourier para una sucesión finita

Sea  $Z_1, Z_2, \dots, Z_n$  una sucesión (función discreta) de  $n$  números. Esta sucesión puede ser considerada como un conjunto de coordenadas de un punto en un espacio  $n$ -dimensional. En álgebra superior generalmente se construyen vectores de estas características mediante combinaciones lineales de un conjunto de vectores, llamados base del espacio vectorial, los cuales tiene la propiedad de ser linealmente independientes y, en algunos casos, ortogonales entre sí.

Entonces, aplicando el concepto anterior y considerando a cada una de las funciones (que para el desarrollo de este trabajo serán consideradas como funciones del tiempo  $t$ ) del sistema trigonométrico 1.2.1.1 como un elemento de la base propuesta, existe una combinación lineal de funciones seno y coseno que representan la sucesión  $\{Z\}$ ; es decir,

$$Z_t = \sum_{t=1}^{[n/2]} \left[ a_k \cos(2\pi kt/n) + b_k \operatorname{sen}(2\pi kt/n) \right] \quad 1.2.2.1$$

Esta ecuación es conocida como la serie de Fourier de la sucesión  $\{Z_t\}$ ; mientras los  $a_k$  y  $b_k$  reciben el nombre de **coeficientes de Fourier**. Usando las propiedades de ortogonalidad de las funciones del sistema 1.2.1.1 se pueden calcular los  $a_k$  y  $b_k$  mediante la multiplicación de  $\cos(2\pi kt/n)$  y  $\operatorname{sen}(2\pi kt/n)$  en ambos lados de 1.2.2.1, respectivamente, y posteriormente sumando sobre  $t=1, 2, \dots, n^4$ . Así, se obtiene

$$a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos\left(\frac{2\pi kt}{n}\right), & k=0, \text{ y } k=n/2 \text{ si } n \text{ par,} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos\left(\frac{2\pi kt}{n}\right), & k=1, 2, \dots, \left[\frac{n-1}{2}\right]. \end{cases} \quad 1.2.2.2$$

$$b_k = \frac{2}{n} \sum_{t=1}^n Z_t \operatorname{sen}\left(\frac{2\pi kt}{n}\right) \quad k=1, 2, \dots, \left[\frac{n-1}{2}\right].$$

Sean las frecuencias de Fourier  $\omega = 2\pi k/n$ ,  $k=0, 1, \dots, [n/2]$ . Usando el sistema de exponenciales complejas, el equivalente a 1.2.2.1 sería

$$Z_t = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i\omega_k t} & \text{si } n \text{ es par} \\ \sum_{k=-n/2+1}^{n/2} c_k e^{i\omega_k t} & \text{si } n \text{ es impar} \end{cases} \quad 1.2.2.3$$

Donde los coeficientes  $c_k$  están dados por

$$c_k = \frac{1}{n} \sum_{t=1}^n Z_t e^{-i\omega_k t}. \quad 1.2.2.4$$

Por último, las ecuaciones 1.2.2.2 y 1.2.2.4 proporcionan la última relación, existente entre los coeficientes  $a_k$  y  $b_k$  de la serie 1.2.2.1 con los  $c_k$  de 1.2.2.3

$$\begin{aligned} c_0 &= a_0 \\ c_k &= \frac{a_k - ib_k}{2} \\ c_{-k} &= c_k^* = \frac{a_k + ib_k}{2} \end{aligned}$$

<sup>4</sup>Véase la sección 1.1.5 para un seguimiento detallado

En este punto es importante hacer un alto para considerar la convergencia de la suma 1.2.2.1. Una condición que garantiza que la suma no se indefina es que la sucesión  $\{Z_t\}$  sea absolutamente sumable. Esto es

$$\sum_{t=-\infty}^{\infty} |Z_t| < \infty.$$

El desarrollo anterior muestra que cualquier sucesión puede ser escrita como una combinación lineal de funciones seno-coseno o de exponenciales complejas.

### 1.2.3 Series de Fourier para sucesiones periódicas

En este caso es del todo válida la definición de periodicidad dada el 1.1.1. Ahora, suponga que la sucesión (o función discreta)  $Z_t$  tiene período  $n$ , i.e,

$$Z_{t+n} = Z_t$$

para todo  $t$ , entero.

A esta sucesión, por lo dicho anteriormente, le corresponde en un intervalo de longitud  $n$ , una expansión en serie de Fourier dada por 1.2.2.1, o equivalentemente por 1.2.2.3, donde  $\omega = 2\pi k/n$ .

Se observa que las funciones  $\text{sen}(\omega_k t)$  y  $\text{cos}(\omega_k t)$  en 1.2.2.1 y las exponenciales  $e^{i\omega_k t}$  tienen como período a  $n$ , lo que provoca que

$$Z_{t+jn} = Z_t,$$

para todo  $t$  y  $j$  enteros. Por lo que 1.2.2.1 y 1.2.2.3 continúan siendo válidas para todo el dominio de  $Z_t$ .

Para una sucesión  $Z_t$  de período  $n$ , se define la energía asociada a  $Z_t$  como

$$\sum_{t=1}^n Z_t^2.$$

Multiplicando  $Z_t$  en ambos lados de (1.2.2.1), sumando de  $t=1$  hasta  $t=n$ , y haciendo uso de 1.2.2.2 se tiene

$$\sum_{t=1}^n Z_t^2 = \sum_{k=0}^{[n/2]} \left[ a_k \sum_{t=1}^n Z_t \cos \omega_k t + b_k \sum_{t=1}^n Z_t \sin \omega_k t \right]$$

$$= \begin{cases} na_0^2 + \frac{n}{2} \sum_{k=1}^{(n-1)/2} (a_k^2 + b_k^2) & \text{si } n \text{ es impar} \\ na_0^2 + \frac{n}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + na_0^2 \frac{n}{2} & \text{si } n \text{ es par} \end{cases} \quad 1.2.3.1$$

La ecuación anterior es conocida como la **relación de Parseval para series de Fourier**. Equivalentemente, usando las igualdades 1.2.2.3 y 1.2.2.4, se tiene la forma correspondiente a 1.2.3.1 para exponenciales

$$\sum_{t=1}^n Z_t^2 = \begin{cases} n \sum_{k=-(n-1)/2}^{(n-1)/2} |c_k|^2 & \text{si } n \text{ es impar} \\ n \sum_{k=-n/2+1}^{n/2} |c_k|^2 & \text{si } n \text{ es par} \end{cases} \quad 1.2.3.2$$

adonde  $|c_k|^2 = c_k c_k^*$ .

Las ecuaciones 1.2.3.1 y 1.2.3.2 implican que la energía total de una sucesión sobre el horizonte  $t=0, 1, 2, \dots$  es infinita. Por eso, se considerará esta energía por unidad de tiempo, a la que se denominará la fuerza de la sucesión, dado por

$$\text{Fuerza} = \begin{cases} a_0^2 + \frac{1}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) & \text{si } n \text{ es impar} \\ a_0^2 + \frac{1}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + a_{n/2}^2 & \text{si } n \text{ es par.} \end{cases}$$

o equivalentemente

$$\text{Fuerza} = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} |c_k|^2 & \text{si } n \text{ es impar,} \\ \sum_{k=-n/2+1}^{n/2} |c_k|^2 & \text{si } n \text{ es par.} \end{cases}$$

Como se dijo antes, el  $j$ -ésimo componente armónico incluye los términos  $k=+j$  y  $k=-j$  que corresponden a la misma frecuencia  $j(2\pi/n)$ . Así, se interpreta la cantidad

$$f_0 = c_0^2 = a_0^2$$

$$f_k = |c_{-k}|^2 + |c_k|^2 = 2|c_k|^2 = \frac{a_k^2 + b_k^2}{2}$$

del término de la serie de Fourier de  $Z_t$  en la  $k$ -ésima frecuencia  $\omega = 2\pi k/n$  como la contribución a la fuerza total. La cantidad  $f_k$  dibujada como una función de  $\omega_k$  se conoce como el **espectro de fuerza** y describe como se distribuye la fuerza total sobre las diferentes frecuencias de la sucesión  $Z_t$ .

### 1.2.4 Transformada de Fourier

Considérese una sucesión o función discreta  $Z_t$  de duración finita tal que  $Z_t = 0$  para  $|t| > M$ . Sea  $n = (2M+1)$  y una función  $Y_t$

$$Y_{t+jn} = Z_t \quad -(n-1)/2 \leq t \leq (n-1)/2; \quad j=0, \pm 1, \pm 2, \dots$$

la cual tiene como período a  $n$ , por lo que puede ser expresada por 1.2.1.3 como:

$$Y_t = \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i2\pi kt/n}, \quad 1.2.4.1$$

$$c_k = \frac{1}{n} \sum_{t=-(n-1)/2}^{(n-1)/2} Y_t e^{-i2\pi kt/n}.$$

Pero en el intervalo  $-(n-1)/2 \leq t \leq (n-1)/2$ ,  $Y_t = Z_t$ , entonces

$$c_k = \frac{1}{n} \sum_{t=-(n-1)/2}^{(n-1)/2} Z_t e^{-i2\pi kt/n}$$

$$= \frac{1}{n} \sum_{t=-\infty}^{\infty} Z_t e^{-i2\pi kt/n}$$

donde se utilizó el hecho de que  $Z_t = 0$  para  $t < -(n-1)/2$  o  $t > (n-1)/2$ .

Sea

$$f(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} Z_t e^{-i\omega t}.$$

Entonces

$$c_k = \frac{2\pi}{n} f(k\Delta\omega)$$

donde  $\Delta\omega = 2\pi/n$ , combinando este resultado con 1.2.4.1 se tiene

$$Y_t = \sum_{k=-(n-1)/2}^{(n-1)/2} \frac{2\pi}{n} f(k\Delta\omega) e^{ik\Delta\omega t}$$

y ya que  $\Delta\omega = 2\pi/n$ , entonces

$$Y_t = \sum_{k=-(n-1)/2}^{(n-1)/2} f(k\Delta\omega) e^{ik\Delta\omega t} \Delta\omega$$

Cada término de la suma anterior representa el área de un rectángulo de base  $\Delta\omega$  y altura  $f(k\Delta\omega)e^{ik\Delta\omega t}$ . A medida que  $n \rightarrow \infty$  se observa que  $Y_t \rightarrow Z_t$  y  $\Delta\omega \rightarrow 0$ . Por lo que en el límite, la suma se convierte en integral. Más aún, debido a que la suma se realiza sobre  $n$  valores consecutivos de longitud  $\Delta\omega = 2\pi/n$ , el intervalo total de integración siempre será de  $2\pi$  de ancho, i.e.,

$$\begin{aligned} Z_t &= \lim_{t \rightarrow \infty} Y_t \\ &= \lim_{t \rightarrow \infty} \sum_{k=-\infty}^{\infty} f(k\Delta\omega) e^{ik\Delta\omega t} \Delta\omega \\ &= \int_{2\pi} f(\omega) e^{i\omega t} d\omega. \end{aligned}$$

En vista de que  $f(\omega)e^{i\omega t}$  es una función de  $\omega$ , tiene un período de  $2\pi$ , lo que permite que cualquier intervalo de integración de  $2\pi$  cumpla con lo establecido anteriormente.

En particular, si se considera el intervalo  $-\pi < \omega < \pi$ , entonces las ecuaciones anteriores quedarán como

$$Z_t = \int_{-\pi}^{\pi} f(\omega) e^{i\omega t} d\omega \quad t=0, \pm 1, \pm 2, \dots \quad 1.2.4.2$$

y

$$f(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} Z_t e^{-i\omega t}, \quad -\pi < \omega < \pi. \quad 1.2.4.3$$

La función  $f(\omega)$  en 1.2.4.3 se conoce como la **transformada de Fourier de  $Z_t$**  (en el caso discreto), y a  $Z_t$  en 1.2.4.2 se le denomina **transformada inversa de Fourier de  $f(\omega)$**  (para funciones discretas).

El hecho de encontrar la integral 1.2.4.2 como el límite de una suma implica que la transformada inversa de Fourier representa a  $Z_t$  como una combinación lineal de funciones que son complejas, de forma senoidal, con frecuencias muy cercanas y amplitudes dadas por  $|f(\omega)| d\omega$ . Así,  $|f(\omega)|$  se conoce comúnmente como el espectro o amplitud espectral de la sucesión, ya que provee información acerca del peso que tienen las frecuencias en la composición de  $Z_t$ .

La energía asociada con la sucesión esta dada por la relación de Parseval:

$$\begin{aligned} \sum_{l=-\infty}^{\infty} |Z_l|^2 &= \sum_{l=-\infty}^{\infty} Z_l \int_{-\pi}^{\pi} f^*(\omega) e^{-i\omega l} d\omega \\ &= \int_{-\pi}^{\pi} f^*(\omega) \sum_{l=-\infty}^{\infty} Z_l e^{-i\omega l} d\omega \\ &= 2\pi \int_{-\pi}^{\pi} f^*(\omega) f(\omega) d\omega \\ &= 2\pi \int_{-\pi}^{\pi} |f(\omega)|^2 d\omega. \end{aligned}$$

1.2.4.4

Entonces, la relación de Parseval asocia la energía en el dominio del espacio con la energía en el dominio de frecuencias. En otras palabras, la energía de una sucesión puede ser encontrada ya sea, calculando la energía por unidad de tiempo y sumando sobre el dominio del tiempo, o encontrando la energía por unidad de frecuencia e integrando sobre todas las frecuencias.

En vista de lo anterior,  $g(\omega) = 2\pi |f(\omega)|^2$  como una función de  $\omega$  es también conocida como el **espectro de energía o la función de densidad espectral**.

Es importante, para el desarrollo del trabajo, señalar las diferencias entre las propiedades del dominio de la frecuencia para funciones discretas periódicas y no periódicas.

1. El espectro de frecuencias de una función discreta periódica está armónicamente relacionado y forma un conjunto discreto, mientras que el espectro de una función no periódica forma un conjunto continuo.

2. La energía sobre todo el eje del tiempo para una sucesión periódica es infinito. Por lo tanto, sus propiedades tienen que ser estudiadas en términos del espectro de potencias sobre un conjunto finito de frecuencias armónicamente relacionadas. El espectro resultante se denomina espectro de línea. La energía sobre todo el horizonte del tiempo para una sucesión aperiódica es finita y está garantizada por la condición para la convergencia.

## Capítulo 2 Series de Tiempo

## Introducción

En la vida real existen muchos fenómenos a los que resulta necesario estudiar. La idea para realizar este estudio se basa en una serie de abstracciones de las relaciones que mantienen los componentes del fenómeno entre sí.

A esta serie de abstracciones se le conoce con el nombre de modelo matemático.

En particular, es posible derivar un modelo basado en leyes estáticas (como las físicas), lo cual permite calcular exactamente el valor de cualquier variable dependiente del tiempo en un instante dado. Si éste cálculo es posible, nos encontramos con un modelo determinístico.

Por otro lado, hay fenómenos en los que influyen una gran cantidad de factores y en los que no es posible establecer un modelo que permita determinar de forma precisa el comportamiento de dicho fenómeno. Sin embargo, en ciertos casos es posible encontrar un modelo que calcule la probabilidad de un estado futuro del fenómeno entre dos límites específicos. Tales modelos se conocen como modelos probabilísticos o estocásticos.

Las series de tiempo, a las que este trabajo está dedicado, son un caso especial de modelo estocástico.

## 2.1 Series de tiempo y procesos estocásticos

Se define una Serie de Tiempo como un conjunto de observaciones generadas secuencialmente en un espacio paramétrico (generalmente en el tiempo). Si el espacio paramétrico es continuo, la serie de tiempo es continua. Si el espacio paramétrico es discreto, se dice entonces que la serie es discreta.

Si se analiza una serie de tiempo discreta, se considera que se tienen  $N$  valores sucesivos disponibles que se escriben como  $z_1, z_2, \dots, z_N$  y denotan observaciones hechas en intervalos equidistantes en el tiempo  $t_0+h, t_0+2h, \dots, t_0+Nh$ .

Las series de tiempo surgen de dos formas:

- Haciendo un muestreo de una serie continua.
- Acumulando los resultados de una variable sobre un lapso de tiempo.

Si los valores futuros de una serie de tiempo están determinados exactamente por una función matemática, se dice que tal serie es determinística. Si los valores futuros de una serie sólo pueden ser descritos en términos de una función de probabilidades, esta serie llevará el nombre de no-determinística o simplemente serie de tiempo estocástica.

Un proceso estocástico es una familia ordenada de variables aleatorias  $Z(w, t)$ , donde  $w$  pertenece a un espacio muestral y  $t$  a un conjunto ordenado. Para un  $t$  fijo,  $Z(w, t)$  es una variable aleatoria. Para un  $w$ ,  $Z(w, t)$ , como función de  $t$ , se conoce como función muestral o realización. Usualmente La población que consiste de todas las realizaciones posibles se llama ensamble. Se hace necesario en lo subsecuente establecer la diferencia entre un proceso estocástico y una serie de tiempo. Así, una serie de tiempo  $z_1, z_2, \dots, z_N$  de  $N$  observaciones sucesivas es considerada como una realización o muestra de una población infinita de series que podrían haber sido generadas por un proceso estocástico.

Un tipo de proceso estocástico, denominado estacionario, se basa en la suposición de que el mismo se encuentra en un estado de equilibrio estadístico. Se dice que un proceso es estrictamente estacionario si sus propiedades no se afectan al cambiar el origen del tiempo; esto es, si la función de distribución conjunta de probabilidades asociada con  $N$  observaciones  $z_{t_1}, z_{t_2}, \dots, z_{t_N}$ , hechas en un conjunto  $t_1, t_2, \dots, t_N$  cualquiera, es la misma que la función de distribución asociada a  $N$  observaciones  $z_{t_1+k}, z_{t_2+k}, \dots, z_{t_N+k}$  realizadas en los tiempos  $t_1+k, t_2+k, \dots, t_N+k$ .

Formalmente considere un conjunto finito de variables aleatorias  $\{z_{t_1}, z_{t_2}, \dots, z_{t_N}\}$  de un proceso estocástico  $\{Z(w, t): t=0, \pm 1, \pm 2, \dots\}$ . La función  $N$ -dimensional de distribución esta definida por

$$F(z_{t_1}, \dots, z_{t_N}) = \rho \{ w: z(w, t_1) \leq z_{t_1}, \dots, z(w, t_N) \leq z_{t_N} \}$$

Se dice que un proceso es estacionario de primer orden si su función de distribución unidimensional es invariante en el tiempo, es decir, si  $F(z_{t_1}) = F(z_{t_1+k})$  para enteros  $t_1$  y  $k$  cualesquiera; estacionario de segundo orden si  $F(z_{t_1}, z_{t_2}) = F(z_{t_1+k}, z_{t_2+k})$  para enteros  $t_1, t_2$  y  $k$ ; y estacionario de  $N$ -ésimo orden si

$$F(z_{t_1}, \dots, z_{t_N}) = F(z_{t_1+k}, \dots, z_{t_N+k}) \quad 2.1.1$$

para toda  $N$ -tupla  $(t_1, \dots, t_N)$  y  $k$  entero. Un proceso estocástico es estrictamente estacionario si (2.1.1) se cumple para toda  $N$ .

En adelante se utilizara la notación  $Z_t$  para representar el proceso estocástico  $Z(w, t)$ .

### 2.1.1 Media y varianza de un proceso estacionario

Para un proceso estacionario  $\{Z_t; t=0, \pm 1, \pm 2, \dots\}$ , la suposición de estacionaridad implica que la función de distribución de probabilidades  $F(z_t)$  es la misma para todos los instantes  $t$  y puede ser escrita como  $F(z)$  solamente. Por consiguiente, el proceso antes citado tiene una función de media constante

$$\mu = \mu_t = E[z_t] = \int_{-\infty}^{\infty} z F(z) dz$$

que define el nivel alrededor del cual el proceso fluctua, y una función de varianza constante

$$\sigma^2 = \sigma_t^2 = E[(z-\mu)^2] = \int_{-\infty}^{\infty} (z-\mu)^2 F(z) dz$$

que mide la dispersión alrededor de este nivel. Además, la media  $\mu$  del proceso estocástico puede ser estimada por la media muestral

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$$

de la serie de tiempo, y la varianza  $\sigma^2$  del proceso estocástico se obtiene con el siguiente estimador de la serie.

$$\sigma_z^2 = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2$$

### 2.1.2 Funciones de autocovarianza y autocorrelación

De la suposición de estacionaridad se desprende que la naturaleza de la función de distribución de probabilidades puede ser referida graficando un diagrama difuso utilizando las parejas  $(z_t, z_{t+k})$ , de la serie de la serie de tiempo, separadas por  $k$  intervalos de la misma longitud (o un atraso de  $k$  unidades de tiempo). Por ejemplo, para un conjunto de datos la figura (a) representa el diagrama difuso con  $k=1$  intervalos de separación, que se obtiene de graficar  $z_{t+1}$  contra  $z_t$ ; y la figura (b) representa el diagrama de flujo de el mismo conjunto de datos pero con una diferencia de  $k=2$  intervalos (trazando  $z_{t+2}$  contra  $z_t$ ).

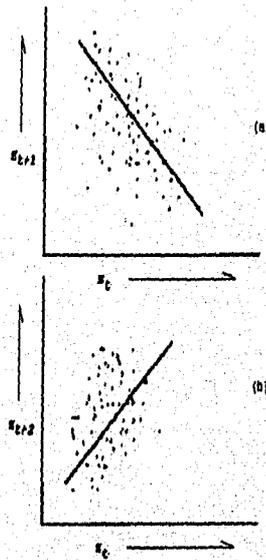


Fig. 1

Puede observarse que los valores en la vecindad de la serie de tiempo están correlacionados y que la correlación (pendiente de la recta) entre  $z_t$  y  $z_{t+1}$  aparece como negativa y la correlación de  $z_t$  y  $z_{t+2}$  es positiva. La covarianza entre  $z_t$  y  $z_{t+k}$  separados por  $k$  intervalos de tiempo

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

$$= \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2}$$

ya que para un proceso estacionario, la varianza  $\sigma_z^2 = \gamma_0$  es la misma en el instante  $t$ .

De este modo, la autocorrelación con  $k$  periodos de diferencia es

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

Los números  $\gamma_k$  y  $\rho_k$  son conocidos como los coeficientes de autocovarianza y autocorrelación.

Se define la función de autocovarianza, de una serie de tiempo, como la regla que asocia a cada unidad de atraso  $k$ , su coeficiente de autocovarianza  $\gamma_k$ . Similarmente se conoce a la función de autocorrelación (o correlograma en el caso en el que la regla se especifique gráficamente) de una serie de tiempo (ACF) a la regla que asocia a cada unidad de atraso  $k$ , su coeficiente de autocorrelación  $\rho_k$ . En la siguiente figura se puede observar una función de autocorrelación representada por una matriz y su gráfica asociada.

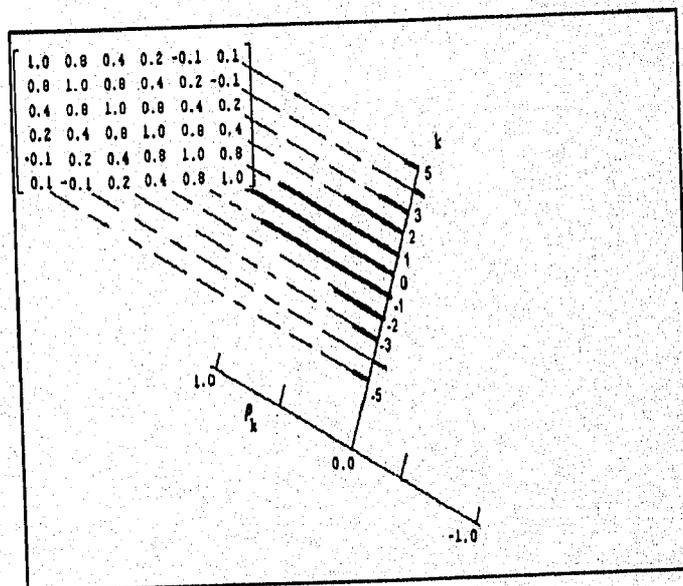


Fig. 2

Es sencillo observar que para un proceso estacionario la función de autocovarianza y la de autocorrelación tienen las siguientes propiedades:

a)  $\gamma_0 = \text{VAR}(Z_t)$ ;  $\rho_0 = 1$ .

b)  $|\gamma_k| \leq \gamma_0$ ;  $|\rho_k| \leq 1$ .

c)  $\gamma_k = \gamma_{-k}$  y  $\rho_k = \rho_{-k}$ , para toda  $k$ , es decir, las funciones de autocovarianza y autocorrelación son pares alrededor del origen del tiempo,  $k=0$ . Esto se debe a que la diferencia temporal entre  $Z_t$  y  $Z_{t+k}$  y  $Z_t$  y  $Z_{t-k}$  es la misma. Por lo tanto la función de autocorrelación normalmente se grafica sólo para los valores positivos del dominio.

Para el caso práctico, en el que se tiene un conjunto de  $N$  observaciones  $z_1, \dots, z_N$ , es necesario trabajar con un estimador de los parámetros o coeficientes de autocorrelación, en particular se ha llegado a la conclusión de que el mejor estimador para los coeficientes de correlación  $\rho_k$  con un atraso de  $k$  unidades es

$$r_k = \frac{c_k}{c_0}$$

donde

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \quad k=0,1,2,\dots,K$$

es el estimador de la autocovarianza  $\gamma_k$ , y  $\bar{z}$  es la media de la serie de tiempo.

#### - Matrices de autocorrelación

La matriz de autocovarianza asociada a un proceso estacionario para un conjunto de observaciones  $(z_1, \dots, z_n)$ , hechas en  $n$  periodos sucesivos es:

$$\Gamma_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \dots & \gamma_0 \end{bmatrix}$$

$$\Gamma_n = \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_2 & \dots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{n-3} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho_{n-3} & \rho_{n-2} & \rho_{n-1} & \dots & 1 \end{bmatrix} = \sigma_z^2 P_n$$

Una matriz de covarianzas  $\Gamma_n$  de esta forma, simétrica con elementos constantes en cualquier diagonal, se conocerá como matriz de autocovarianzas y la matriz de correlaciones correspondiente  $P_n$ , denotará la matriz de autocorrelaciones de la serie. Considere ahora una combinación lineal de variables aleatorias  $z_1, z_2, \dots, z_{t-n+1}$

$$L_t = l_1 z_t + l_2 z_{t-1} + \dots + l_n z_{t-n+1}$$

Ya que  $\text{cov}[z_t, z_j] = \gamma_{|t-j|}$  para un proceso estacionario, la varianza de  $L_t$  es

$$\text{var}[L_t] = \sum_{i=1}^n \sum_{j=1}^n l_i l_j \gamma_{|i-j|}$$

que necesariamente es mayor que cero si alguno de los  $l_i$  es distinto de cero, lo que implica que la matriz de autocovarianza y autocorrelación son positivas definidas para un proceso estacionario.

Para que las matrices mencionadas sean consideradas como definidas positivas su determinante y el de todos sus menores principales deben ser mayores que cero, lo que provoca el surgimiento de ciertas restricciones para los elementos que las componen, en particular, cuando  $n=2$  es sencillo encontrar que

$$-1 < \rho_1 < 1$$

para  $n=3$  nos encontramos las restricciones

$$-1 < \rho_1 < 1$$

$$-1 < \rho_2 < 1$$

$$-1 < \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} < 1$$

y demás, debido a que  $P_n$  debe ser positiva definida para todos los valores de  $n$ , las autocorrelaciones del proceso estacionario tienen que satisfacer un número muy grande de condiciones, pero como se verá más adelante, todas esas condiciones están contenidas en la definición del espectro.

### 2.1.3 Ejemplo

Para mostrar prácticamente todo lo considerado hasta el momento suponga que se tiene una serie de 10 datos

t	Zt	Zt+1	Zt+2	Zt+3	...	Zt-2	Zt-1
1	13	08	15	04			
2	08	15	04	04		13	
3	15	04	04	12		08	13
4	04	04	12	11		15	08
5	04	12	11	07		04	15
6	12	11	07	14		04	04
7	11	07	14	12		12	04
8	07	14	12			11	12
9	14	12				07	11
10	12					14	07

La media muestral de esta serie es

$$\bar{x} = \frac{13+8+15+4+4+12+11+7+14+12}{10} = \frac{100}{10} = 10$$

La varianza es

$$\sigma^2 = \frac{(13-10)^2 + (8-10)^2 + (15-10)^2 + (4-10)^2 + (4-10)^2 + (12-10)^2 + (11-10)^2 + (7-10)^2 + (14-10)^2 + (12-10)^2}{10} = \frac{144}{10} = 14.4$$

Por otro lado, los coeficientes de autocovarianza y autocorrelación son

$$c_0 = 14.4$$

$$\rho_0 = 1$$

$$c_1 = (13-10)(8-10) + (8-10)(15-10) + (7-10)(14-10) + (14-10)(12-10) \\ = -27$$

$$\rho_1 = -\frac{27}{144} = -0.188$$

$$c_2 = (13-10)(15-10) + (8-10)(4-10) + (11-10)(14-10) + (7-10)(12-10) \\ = -29$$

$$\rho_1 = -\frac{29}{144} = -0.201$$

$$c_3 = (13-10)(4-10) + (8-10)(4-10) + (12-10)(14-10) + (11-10)(12-10) \\ = 26$$

$$\rho_3 = \frac{26}{144} = 0.181$$

$$c_4 = -19; \quad c_5 = -17; \quad c_6 = 17$$

$$\rho_4 = -0.131; \quad \rho_5 = -0.118; \quad \rho_6 = 0.118$$

$$c_7 = -7; \quad c_8 = 8; \quad c_9 = 6$$

$$\rho_7 = -0.048; \quad \rho_8 = 0.055; \quad \rho_9 = 0.041$$

La matriz de autocovarianzas y autocorrelaciones se presentan a continuación

$$Y = \begin{bmatrix} 14.4 & -27 & -29 & 26 & -19 & -47 & 17 & 7 & 8 & 6 \\ -27 & 14.4 & -27 & -29 & 26 & -19 & -47 & 17 & 7 & 8 \\ -29 & -27 & 14.4 & -27 & -29 & 26 & -19 & -47 & 17 & 7 \\ 26 & -29 & -27 & 14.4 & -27 & -29 & 26 & -19 & -47 & 17 \\ -19 & 26 & -29 & -27 & 14.4 & -27 & -29 & 26 & -19 & -47 \\ -47 & -19 & 26 & -29 & -27 & 14.4 & -27 & -29 & 26 & -19 \\ -19 & -47 & -19 & 26 & -29 & -27 & 14.4 & -27 & -29 & 26 \\ 26 & -19 & -47 & -19 & 26 & -29 & -27 & 14.4 & -27 & -29 \\ -29 & 26 & -19 & -47 & -19 & 26 & -29 & -27 & 14.4 & -27 \\ -29 & -29 & 26 & -19 & -47 & -19 & 26 & -29 & -27 & 14.4 \end{bmatrix}$$

$$P = \begin{bmatrix} 1.000 & -.188 & -.201 & .181 & -.131 & -.326 & .118 & -.048 & .055 & .041 \\ -.188 & 1.000 & -.188 & -.201 & .181 & -.131 & -.326 & .118 & -.048 & .055 \\ -.201 & -.188 & 1.000 & -.188 & -.201 & .181 & -.131 & -.326 & .118 & -.048 \\ .181 & -.201 & -.188 & 1.000 & -.188 & -.201 & .181 & -.131 & -.326 & .118 \\ -.131 & .181 & -.201 & -.188 & 1.000 & -.188 & -.201 & .181 & -.131 & -.326 \\ -.326 & -.131 & .181 & -.201 & -.188 & 1.000 & -.188 & -.201 & .188 & -.131 \\ .118 & -.326 & -.131 & .181 & -.201 & -.188 & 1.000 & -.188 & -.201 & .188 \\ -.048 & .118 & -.326 & -.131 & .181 & -.201 & -.188 & 1.000 & -.188 & -.201 \\ .055 & -.048 & .118 & -.326 & -.131 & .181 & -.201 & -.188 & 1.000 & -.188 \\ .041 & .055 & -.048 & .118 & -.326 & -.131 & .181 & -.201 & -.188 & 1.000 \end{bmatrix}$$

Y por último la gráfica de la función de autocorrelación se muestra en seguida

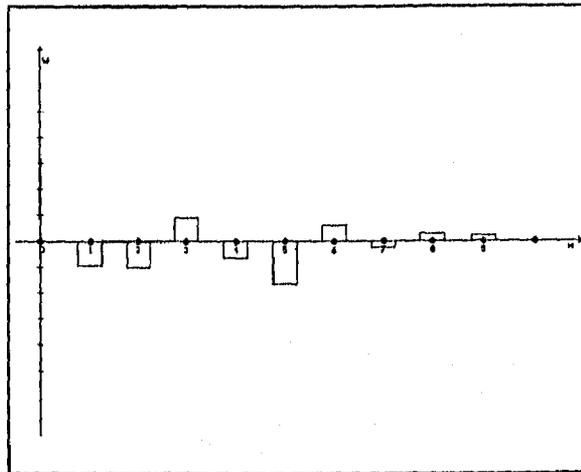


Fig. 3 Función de autocorrelación muestral

### 2.1.4 Representación de procesos estocásticos

En el análisis de series de tiempo existen dos maneras muy útiles para expresar un proceso de esta índole. Una es escribiendo el proceso  $Z_t$  como una combinación lineal de una sucesión de variables aleatorias independientes, es decir,

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

donde  $\psi_0=1$ ,  $\{a_t\}$  es una variable aleatoria con media cero y varianza  $\sigma^2$  llamada normalmente ruido blanco, y

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty. \quad 2.1.4.1$$

Introduciendo el operador de salto hacia atrás (backshift operator)  $B^j x_t = x_{t-j}$  es posible reescribir la fórmula (2.1.4.1) como

$$\tilde{Z}_t = \psi(B) a_t \quad 2.1.4.2$$

donde  $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$  y  $\tilde{Z}_t = Z_t - \mu$ .

Al aplicar la esperanza a la ecuación (2.1.4.1), dado que las variables  $a_t$  es independiente en el tiempo con media 0, se obtiene que

$$\begin{aligned} E(Z_t) &= E(\mu) + \sum_{j=0}^{\infty} E(\psi_j a_{t-j}) \\ &= \mu + \sum_{j=0}^{\infty} \psi_j E(a_{t-j}) \\ &= \mu + \sum_{j=0}^{\infty} \psi_j 0 = \mu \end{aligned}$$

$$\begin{aligned} \text{Var}(Z_t) &= \text{Var}(\mu) + \sum_{j=0}^{\infty} \text{Var}(\psi_j a_{t-j}) \\ &= 0 + \sum_{j=0}^{\infty} \psi_j^2 \text{Var}(a_{t-j}) \\ &= \sum_{j=0}^{\infty} \psi_j \sigma^2 = \sigma^2 \sum_{j=0}^{\infty} \psi_j \end{aligned}$$

Por otro lado el coeficiente de covarianza de  $\tilde{Z}_t = Z_t - \mu$  es

$$\begin{aligned} \gamma_k &= E(\tilde{Z}_t \tilde{Z}_{t+k}) \\ &= E\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j a_{t-i} a_{t+k-j}\right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j E(a_{t-i} a_{t+k-j}) \end{aligned}$$

Tomando en cuenta que la esperanza de  $a_t a_{t+j}$  es igual a 0 cuando  $j$  es diferente de cero porque la variable  $a$  es independiente en el tiempo, y por lo tanto  $a_t$  no tiene correlación con  $a_{t+j}$ , y cuando  $j$  es nula esta esperanza se convierte en la varianza  $\sigma^2$  (recuerde que la media de  $a_t$  es cero), entonces la suma anterior será nula para todos aquellos coeficientes en los que el subíndice  $t-i$  no coincida con el  $t+k-j$ , por lo que el coeficiente de covarianza se puede escribir como

$$\gamma_k = \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} \quad 2.1.4.3$$

Lo que arroja como resultado que el coeficiente de correlación sea

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

Se observa que las funciones de autocovarianza y autocorrelación del proceso 2.1.4.1 son dependientes exclusivamente del intervalo  $k$ . Sin embargo, ya que involucran sumas infinitas, para ser estacionario  $\gamma_k$  tiene que ser finita para toda  $k$ . Ahora

$$|\gamma_k| = |E(\ddot{Z}_t \ddot{Z}_{t+k})| \leq \sqrt{\text{Var}(Z_t) \text{Var}(Z_{t+k})} = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2$$

entonces es necesario que

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

para que el proceso 2.1.4.1 sea estacionario.

La forma empleada en 2.1.4.1 es conocida como representación en medias móviles (**MA** por su traducción del inglés) o representación de **Wold** para un proceso estocástico.

Para una sucesión de autocovarianzas  $\gamma_k$ , con  $k=0, \pm 1, \pm 2, \dots$  la función generadora de autocovarianzas se define como

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k,$$

donde la varianza del proceso,  $\gamma_0$ , es el coeficiente del operador  $B^0$  y la autocovarianza  $\gamma_k$  es el coeficiente de los operadores  $B^k$  y  $B^{-k}$ . Utilizando 2.1.4.3 se puede reescribir la función generadora como

$$\begin{aligned} \gamma(B) &= \sigma^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+k} B^k \\ &= \sigma^2 \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \psi_j \psi_{j+l} B^{-l} \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j B^j \sum_{l=0}^{\infty} \psi_l B^{-l} \\ &= \sigma^2 \psi(B) \psi(B^{-1}) \end{aligned}$$

Donde  $j=i+k$  y hay que destacar que  $\psi_k=0$  para  $j<0$ . Este es un método conveniente de calcular las autocovarianzas para algunos procesos. Similarmente la función generadora de autocorrelaciones será

$$\rho(B) = \sum_{k=-\infty}^{\infty} \rho_k B^k = \frac{\gamma(B)}{\gamma_0}.$$

Otra forma para escribir un proceso  $Z_t$  es utilizando una representación autorregresiva (**AR**), en la cual el valor de  $Z$  en el tiempo  $t$  está dado por una combinación lineal de sus valores anteriores más una variable de choque aleatorio, es decir,

$$\ddot{Z}_t = \pi_1 \ddot{Z}_{t-1} + \pi_2 \ddot{Z}_{t-2} + \dots + a_t$$

o equivalentemente

$$\pi(B)\ddot{Z}_t = a_t$$

en donde  $\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$ , y  $1 + \sum_{j=0}^{\infty} |\pi_j| < \infty$ .

Aunque las representaciones autoregresivas y de medias móviles son útiles, no se recomiendan como formas iniciales para construir el modelo porque contienen una suma infinita de parámetros que es imposible estimar de un número finito de observaciones disponibles. En su lugar se deben utilizar modelos con una cantidad finita de parámetros.

$$\phi_p(B)Z_t = Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t$$

En la representación autoregresiva de un proceso, si sólo un número de  $\pi$  ponderadores son distintos de cero, es decir,  $\pi_1 = \phi_1, \pi_2 = \phi_2, \dots, \pi_p = \phi$  y  $\pi_k = 0$  si  $k > p$ , se dice que el proceso resultante es autoregresivo de orden  $p$  **AR(p)** y se escribe

$$Z_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} = \theta_q(B)a_t$$

Análogamente, si únicamente existen  $q$  ponderadores no nulos en la representación por medias móviles, entonces el proceso será llamado de medias móviles de orden  $q$  **MA(q)** y estará dado por

$$Z_t - \phi_1 Z_{t-1} - \dots = a_t - \phi_1 a_{t-1} - \dots$$

Sin embargo, si se restringe la representación de un modelo a formas puramente autoregresivas o de medias móviles, la cantidad de parámetros aún puede resultar demasiado larga. Una alternativa natural son los modelos en los que se encuentran partes autoregresivas y de medias móviles mezclados (**ARMA**). Escrito en forma abreviada resulta

$$\pi(B)Z_t = \psi(B)a_t$$

Por último, un proceso ARMA que tenga únicamente  $p$  parámetros autorregresivos y  $q$  de medias móviles se le da el nombre de proceso **ARMA(p,q)**

$$\phi_p(B)Z_t = \theta_q(B)a_t$$

## 2.2 Series no estacionarias

Hasta el momento se ha estado haciendo referencia a análisis de procesos considerados como estacionarios, desafortunadamente, esta característica no siempre se cumple en el estudio de las series de tiempo (especialmente en las series que surgen de fenómenos económicos).

La no estacionaridad puede presentarse de diversas formas; hay procesos cuya media no permanece constante en el tiempo; con varianzas que cambian de intervalo a intervalo; y combinaciones de ambos aspectos.

Si un proceso carece de estabilidad en la media, este representa un problema al estimar la función de medias para un instante dado  $t$ . Afortunadamente, se pueden construir modelos a partir de una sola realización para describir el fenómeno.

Existen dos planteamientos distintos que llevan a la construcción de este modelo

### 1.- De tendencia determinística:

Este planteamiento se basa en la suposición de que la función de medias de un proceso no estacionario tiene una tendencia determinística. En tal caso, debe usarse un modelo de regresión estándar para describir el fenómeno. Por ejemplo, si la función de medias  $\mu_t$  sigue una tendencia lineal  $\mu_t = \alpha_0 + \alpha_1 t$ , se puede utilizar el modelo lineal

$$Z_t = \alpha_0 + \alpha_1 t + \dots + \alpha_k t^k + a_t$$

De manera general, si la tendencia determinística de un proceso está dada por un polinomio de orden  $k$ , el modelo asociado al fenómeno será

$$Z_t = \alpha_0 + \alpha_1 t + a_t$$

Por otro lado, si el proceso observa un comportamiento senoidal, entonces el modelo adecuado es

$$\begin{aligned} Z_t &= v_0 + v \cos(\omega t + \theta) + a_t \\ &= v_0 + \alpha \cos(\omega t) + \beta \sin(\omega t) + a_t \end{aligned}$$

donde

$$\begin{aligned} \alpha &= v \cos \theta, & \beta &= v \sin \theta, \\ v &= \sqrt{\alpha^2 + \beta^2} \quad \text{y} \quad \theta = \tan^{-1}(-\beta/\alpha). \end{aligned}$$

Más generalmente

$$Z_t = \nu_0 + \sum_{j=1}^m (\alpha_j \cos \omega_j t + \beta_j \sin \omega_j t) + a_t$$

que se conoce como modelo de periodicidades conocidas.

## 2.- De tendencia no determinística

Aunque existen muchas series no-estacionarias, debido a algunas fuerzas de equilibrio, diferentes partes de estas series se comportan de la misma forma excepto por sus diferencias en el nivel (o valor) de la media local.

A estas series se les denomina homogéneas no estacionarias y, en virtud de esta característica, su comportamiento local es independiente de su nivel (o media). Por tanto, haciendo  $\Psi(B)$  el operador autoregresivo que describe la serie, se tiene

$$\Psi(B)(Z_t + C) = \Psi(B)Z_t$$

para una constante  $C$  dada. Esto implica que  $\Psi(B)$  es de la forma

$$\Psi(B) = \phi(B)(1-B)^d,$$

para alguna  $d > 0$ , donde  $\phi(B)$  es un operador estacionario autorregresivo.

Así, una serie homogénea no estacionaria puede ser reducida a una serie estacionaria escogiendo alguna diferencia de la serie original. En otras palabras, si la serie  $\{Z_t\}$  es no estacionaria, su diferencia de orden  $d$   $\{(1-B)^d Z_t\}$ , para algún entero  $d$  mayor o igual a 1, sí lo es.

El concepto de diferenciación visto en esta sección puede añadirse en la construcción de modelos ARMA para hacerlos más versátiles.

De este modo surgen los modelos **ARIMA(p,d,q)** (que incluyen una diferencia de orden  $d$  en la serie original) cuya forma general es

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t.$$

Entonces, cuando un proceso sea no estacionario en la media, el paso a seguir será diferenciarlo (lo menos posible) hasta que este sea, al menos, estacionario en la media.

Que un proceso sea estacionario en la media no implica necesariamente que tenga

una varianza constante. Sin embargo, un fenómeno que sea no estacionario en la media, también lo será en la varianza.

Para estabilizar la varianza de un proceso cuya media es constante, es necesario utilizar una transformación adecuada.

Es común que la varianza de un proceso no estacionario cambie con la media. Así,

$$\text{Var}(Z_t) = cf(\mu_t)$$

para una constante  $c$  y una función  $f$ .

Se necesita encontrar una transformación  $T(Z_t)$  que estabilice la varianza. Para hacer esto, se usará una aproximación a  $T$  mediante la Serie de Taylor de primer orden alrededor de  $\mu_t$ .

$$T(Z_t) \approx T(\mu_t) + T'(\mu_t)(Z_t - \mu_t)$$

en donde  $T'(\mu_t)$  es la primera derivada de  $T(Z_t)$  evaluada en  $\mu_t$ . Ahora

$$\begin{aligned} \text{Var}[T(Z_t)] &= [T'(\mu_t)]^2 \text{Var}(Z_t) \\ &= c[T'(\mu_t)]^2 f(\mu_t). \end{aligned}$$

y ya que la varianza debe ser constante, entonces

$$[T'(\mu_t)]^2 = \frac{1}{f(\mu_t)} \quad \text{y} \quad T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}}$$

por lo que

$$T(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t$$

A partir de este razonamiento, Box y Cox llegaron a la siguiente transformación

$$T(Z_t) = Z_t^{(\lambda)} = \frac{Z_t^\lambda - 1}{\lambda}$$

donde  $\lambda$  representa el parámetro de la transformación.

Algunos parámetros  $\lambda$  y sus transformaciones asociadas se presentan a continuación:

$\lambda$	Transformación
-1.0	$1/Z_t$
-0.5	$1/\sqrt{Z_t}$
0.0	$\ln Z_t$
0.5	$\sqrt{Z_t}$
1.0	$Z_t$

Para terminar:

Si una serie es no estacionaria en la varianza, pero su media es constante, entonces a esta serie debe aplicarse la transformación de Box-Cox con el parámetro  $\lambda$  correspondiente.

## 2.3 Series estacionales

En los fenómenos económicos (y en muchos otros), es común encontrar patrones de comportamiento que se repiten con cierta regularidad. El menor lapso de tiempo que tarda en repetirse el patrón es conocido como el período estacional. Por ejemplo, la cantidad de artículos escolares vendidos en el país aumenta al acercarse Septiembre, en este caso se está hablando de un fenómeno con periodo estacional de 12 meses.

Actualmente existen 2 corrientes que proporcionan, desde sus diferentes puntos de vista, una alternativa para aproximar un modelo que explique el fenómeno de estacionalidad.

Por un lado, y bajo la suposición de que una serie de tiempo consiste en una mezcla de un componente de tendencia ( $P_t$ ), un componente estacional ( $S_t$ ), y un componente irregular ( $e_t$ ); y que estas tres partes son independientes y aditivas, entonces, es factible describir la serie  $Z_t$  como

$$Z_t = P_t + S_t + e_t$$

Para estimar cada uno de los componentes se han introducidos varios métodos de descomposición.

Un método de descomposición es el que se basa en la regresión, el cual se basa en que

$$\begin{aligned} Z_t &= P_t + S_t + e_t \\ &= \alpha_0 + \sum_{i=1}^m \alpha_i U_{it} + \sum_{j=1}^k \beta_j V_{jt} + e_t \end{aligned}$$

donde  $P_t = \alpha_0 + \sum_{i=1}^m \alpha_i U_{it}$ , y las  $U_{it}$  son las variables de tendencia,  $S_t = \sum_{j=1}^k \beta_j V_{jt}$ , y las  $V_{jt}$  son las variables estacionales. Por ejemplo, un componente de tendencia lineal  $P_t$  sería

$$P_t = \alpha_0 + \alpha_1 t,$$

o de manera general

$$P_t = \alpha_0 + \sum_{i=1}^m \alpha_i t^i.$$

Similarmente, el componente estacional  $S_t$  puede interpretarse como una combinación lineal de funciones seno y coseno de distintas frecuencias, por ejemplo

$$S_t = \sum_{j=1}^{\lfloor s/2 \rfloor} \left[ \beta_j \operatorname{sen} \left( \frac{2\pi j}{s} \right) + \gamma_j \operatorname{cos} \left( \frac{2\pi j}{s} \right) \right]$$

donde  $\lfloor s/2 \rfloor$  es la parte entera de la división  $s/2$ . Por lo que el modelo original queda reescrito como

$$Z_t = \alpha_0 + \sum_{i=1}^m \alpha_i t^i + \sum_{j=1}^{\lfloor s/2 \rfloor} \left[ \beta_j \operatorname{sen} \left( \frac{2\pi j}{s} \right) + \gamma_j \operatorname{cos} \left( \frac{2\pi j}{s} \right) \right] + e_t.$$

El método expuesto anteriormente pretende que la serie de tiempo permita una descomposición en 3 partes independientes y de otras características mencionadas. Sin embargo estas condiciones no siempre se satisfacen (la mayor parte de las veces), es por esto que es necesario un enfoque alternativo.

Para empezar, supóngase que se desconoce que una serie  $\{Z_t\}$  contiene variación estacional y que se ajusta un modelo **ARIMA** no estacional a la serie, i.e.,

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)b_t, \quad 2.3.1$$

Obviamente, la serie residual  $\{b_t\}$  no será ruido blanco ya que contendrá correlaciones inexplicables entre los períodos. Sea

$$\rho_{js} = \frac{E(b_{t-js} - \mu_b)(b_t - \mu_b)}{\sigma_b^2} \quad j=1,2,\dots$$

la función de autocorrelación de  $\{b_t\}$  que representa las relaciones inexplicables entre los períodos. Resulta claro que esta relación entre los períodos se ajusta también a un modelo ARIMA

$$\Phi_p(B^s)(1-B^s)^D b_t = \Theta_Q(B^s)a_t, \quad 2.3.2$$

donde

$$\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$$

y

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_p B^{ps}$$

Son polinomios de  $B^s$  con raíces distintas y  $a_t$  es ruido blanco.

Combinando 2.3.1 con 2.3.2 se obtiene un modelo multiplicativo de período  $s$  **SARIMA(p,d,q)x(P,D,Q)**,

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^s)a_t$$

donde

$$\ddot{Z}_t = \begin{cases} Z_t^{-\mu}, & \text{si } d=D=0 \\ Z_t, & \text{otro caso} \end{cases}$$

para el que  $\phi_p(B)$  y  $\theta_q(B)$  son los factores regulares autorregresivos y de medias móviles y  $\Phi_p(B)$  y  $\Theta_q(B)$  son los factores (polinomios) estacionales autorregresivos y de medias móviles respectivamente.

## Capítulo 3 Teoría Espectral

## Introducción

En este capítulo se analizará a las series de tiempo mediante el estudio del dominio de las frecuencias. En efecto, el espectro de un proceso estacionario es sencillamente la transformada de Fourier de la función de covarianzas absolutamente sumable del proceso. De forma general, a todo proceso estacionario le corresponde una función de distribución espectral.

Además de la teoría se incluyen ejemplos de espectros para algunos modelos autorregresivos ARMA.

## 3.1 El Espectro

### 3.1.1 Propiedades del espectro

Sea  $Z_t$  un proceso estacionario real con una sucesión de autocovarianzas  $\gamma_k$  sumable absolutamente. Entonces la transformada de  $\gamma_k$  existe y es

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k}, \quad -\pi \leq \omega \leq \pi \\ &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k (\cos \omega k - i \operatorname{sen} \omega k) \\ &= \frac{1}{2\pi} \left[ \sum_{k=-\infty}^{\infty} \gamma_k \cos \omega k - i \sum_{k=-\infty}^{\infty} \gamma_k \operatorname{sen} \omega k \right] \end{aligned}$$

Considerando que  $\gamma_k = \gamma_{-k}$ ,  $\operatorname{sen} \omega k = \operatorname{sen}(-\omega k)$  y  $\operatorname{sen} 0 = 0$ , entonces la segunda suma de la expresión anterior se anula.

Por otro lado,  $\cos \omega k = \cos(-\omega k)$  y  $\cos 0 = 1$ , lo que significa

$$\begin{aligned} f(\omega) &= \frac{\gamma_0}{2\pi} + \frac{1}{2\pi} \sum_{k=1}^{\infty} 2\gamma_k \cos \omega k \\ &= \frac{\gamma_0}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \gamma_k \cos \omega k, \quad -\pi \leq \omega \leq \pi \end{aligned} \quad 3.1.1.1$$

La sucesión  $\gamma_k$  puede ser recuperada de  $f(\omega)$  a través de la transformada inversa de Fourier

$$\gamma_k = \int_{-\pi}^{\pi} f(\omega) e^{i\omega k} d\omega. \quad 3.1.1.2$$

$f(\omega)$  tiene las siguientes propiedades:

1.  $f(\omega)$  es una función real, continua y no negativa, es decir,  $|f(\omega)| = f(\omega)$ . De este modo, de la ecuación (1.2.4.4) del capítulo 1,  $f(\omega)$  también se conoce como el espectro de la sucesión de autocovarianzas  $\gamma_k$  correspondiente al proceso estacionario  $Z_t$ . Resulta sencillo notar que  $f(\omega)$  es una función continua, ya que es resultado de una suma de cosenos. Para demostrar que es no negativa se utiliza el hecho de que  $\gamma_k$  es positiva definida<sup>5</sup>, o sea

<sup>5</sup> Ver capítulo anterior

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j \gamma_{(k_i - k_j)} \geq 0$$

para cualesquiera números reales  $c_i, c_j$  y enteros  $k_i$  y  $k_j$ . En particular, sea  $c_i = \cos \omega i$ ,  $c_j = \cos \omega j$ ,  $k_i = i$ , y  $k_j = j$ , entonces,

$$\sum_{i=1}^n \sum_{j=1}^n \gamma_{(i-j)} \cos \omega i \cos \omega j \geq 0$$

Igualmente

$$\sum_{i=1}^n \sum_{j=1}^n \gamma_{(i-j)} \sin \omega i \sin \omega j \geq 0$$

De este modo,

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n \gamma_{(i-j)} [\cos \omega i \cos \omega j + \sin \omega i \sin \omega j] \\ = \sum_{i=1}^n \sum_{j=1}^n \gamma_{(i-j)} \cos(i-j) \end{aligned}$$

Al desarrollar la suma anterior se observa que existen varios sumandos iguales, esto se debe a que diferentes combinaciones de  $i-j$  dan el mismo resultado, además de que las funciones  $\gamma_k$  y coseno son pares. Asociando los términos que son iguales se llega a que la suma es

$$\begin{aligned} \sum_{k=-(n-1)}^{n-1} (n-|k|) \gamma_k \cos \omega k \\ = \sum_{k=-(n-1)}^{n-1} \left[ 1 - \frac{|k|}{n} \right] \gamma_k \cos \omega k \geq 0 \end{aligned}$$

donde  $k=(i-j)$ . Por consiguiente

$$\sum_{k=-(n-1)}^{n-1} \left[ 1 - \frac{|k|}{n} \right] \gamma_k \cos \omega k \geq 0$$

cuando  $n$  tiende a infinito

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{k=-(n-1)}^{n-1} \left[ 1 - \frac{|k|}{n} \right] \gamma_k \cos \omega k \\ = \sum_{k=-\infty}^{\infty} \gamma_k \cos \omega k \geq 0 \end{aligned}$$

Si se multiplica la última expresión por  $1/2\pi$  se concluye que  $f(\omega)$  es mayor o igual a cero.

2.  $f(\omega) = f(\omega + 2\pi)$  y por lo tanto  $f(\omega)$  tiene período  $2\pi$ . Además, y puesto que  $f(\omega) = f(-\omega)$ , su gráfica normalmente se dibuja sólo para  $0 \leq \omega \leq \pi$ .

3. De 3.1.1.2 se concluye

$$\text{Var}(Z_t) = \gamma_0 = \int_{-\pi}^{\pi} f(\omega) d\omega$$

Esta relación prueba que el espectro  $f(\omega)$  puede ser interpretado como la descomposición de la varianza de un proceso. El término  $f(\omega)d\omega$  es la contribución a la varianza atribuible al componente del proceso con frecuencias en el intervalo  $(\omega, \omega + d\omega)$ . Un máximo en la curva indica una contribución importante a la varianza de los componentes a frecuencias en el intervalo correspondiente. Por ejemplo, en la figura 1 se muestra que los componentes cercanos a la frecuencia  $\omega_0$  son las de mayor importancia y los componentes de alta frecuencia próximos a  $\pi$  resultan los de menor importancia.

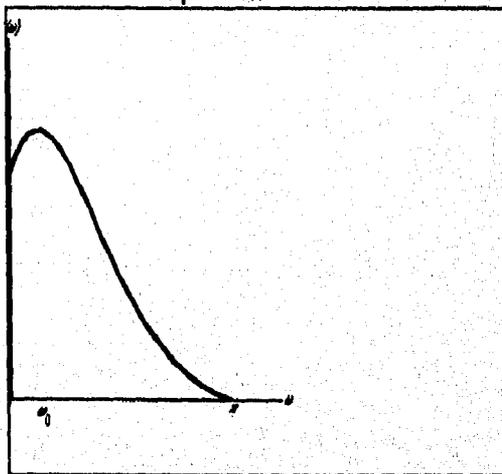


Fig. 1

4. Las ecuaciones 3.1.1.1 y 3.1.1.2 implican que el espectro  $f(\omega)$  y la sucesión de autocovarianzas  $\gamma_k$  de una transformada de Fourier par, con una siendo determinada totalmente por la otra. Por lo tanto, el acercamiento mediante el estudio del dominio del tiempo y mediante el dominio de las frecuencias son teóricamente equivalentes. La razón que justifica ambas aproximaciones es que hay ocasiones en las que un acercamiento es preferible al otro para presentación o interpretación.

### 3.1.2 Representación espectral de la función de autocovarianza

Advierta que la representación de  $\gamma_k$  dada es válida exclusivamente para una función de autocovarianza absolutamente sumable. Más generalmente, para una función de autocovarianzas dada, su representación espectral esta dada por la siguiente integral de Fourier-Stieltjes:

$$\gamma_k = \int_{-\pi}^{\pi} e^{i\omega k} dF(\omega), \quad 3.1.2.1$$

donde  $F(\omega)$  es la función de distribución espectral. La ecuación 3.1.2.1 se conoce como la representación espectral de la función de autocovarianzas  $\gamma_k$ . Al igual que toda función de distribución, la espectral es no decreciente que puede ser particionada en 3 componentes: (1) una función seccionalmente continua con un número finito o infinito numerable de discontinuidades; (2) una función continua; y (3) una función singular o puntual. El tercer componente es, para la mayoría de los cálculos, insignificante y por lo general se elimina del análisis. Así la función de distribución espectral es

$$F(\omega) = F_s(\omega) + F_c(\omega)$$

donde  $F_s(\omega)$  es la función de continua a trozos y  $F_c(\omega)$  es la función continua. La ecuación 3.1.1.2 indica que para un proceso cuya función de autocovarianzas absolutamente sumable,  $F(\omega) = F_c(\omega)$  y  $dF(\omega) = f(\omega)d\omega$ .

Para ilustrar una función de distribución espectral discontinua considere el siguiente modelo

$$Z_t = \sum_{i=1}^M A_i \text{sen}(\omega_i t + \theta_i) \quad 3.1.2.2$$

en donde los  $A_i$  son constantes y las  $\theta_i$  son variables aleatorias independientes que se ajustan a una distribución uniforme en el intervalo  $[-\pi, \pi]$ . Entonces

$$\begin{aligned} E(Z_t) &= \sum_{i=1}^M A_i E[\text{sen}(\omega_i t + \theta_i)] \\ &= \sum_{i=1}^M A_i \int_{-\pi}^{\pi} \frac{\text{sen}(\omega_i t + \theta_i)}{2\pi} d\theta_i = 0 \end{aligned}$$

y

$$\begin{aligned}
 \gamma_k &= E(Z_t Z_{t+k}) \\
 &= \sum_{i=1}^M A_i^2 E\{\text{sen}(\omega_i t + \theta_i) + \text{sen}[\omega_i(t+k) + \theta_i]\} \\
 &= \sum_{i=1}^M A_i^2 E\left\{\frac{1}{2}[\cos\omega_i k - \cos(\omega_i(2t+k) + 2\theta_i)]\right\} \\
 &= \frac{1}{2} \sum_{i=1}^M A_i^2 \cos\omega_i k, \quad k=0, \pm 1, \pm 2, \dots
 \end{aligned}$$

3.1.2.3

donde se empleó el resultado

$$E\{\cos[\omega_i(2t+k) + 2\theta_i]\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos[\omega_i(2t+k) + 2\theta_i] d\theta_i = 0.$$

La función de autocovarianzas en la ecuación 3.1.2.3 no es sumable absolutamente para todas las  $k$ , en particular para  $k=0$  es claro que la suma infinita de los absolutos de las  $A_i$  al cuadrado, es infinita. En consecuencia, no es posible representar esta función en términos de 3.1.1.2. Sin embargo, es factible escribirla en la forma 3.1.2.1. Para hacer esto hay que tomar en cuenta las siguientes observaciones:

1.  $Z_t$  es la suma de  $M$  componentes independientes.

2.  $\omega_i$  se encuentra en el intervalo  $[-\pi, \pi]$ ,  $\gamma_k$  en 3.1.2.1 es así mismo la función de autocovarianza del proceso en cuestión para frecuencias  $\omega$ .

3.

$$\gamma_0 = \frac{1}{2} \sum_{i=1}^M A_i^2,$$

y en consecuencia la varianza  $A^2/2$  del  $i$ -ésimo componente establece su contribución a la varianza total de  $Z_t$  en ambas frecuencias  $-\omega_i$  y  $\omega_i$ . En otras palabras, una mitad de la varianza del  $i$ -ésimo componente queda asociado con la frecuencia  $-\omega_i$  mientras que la otra mitad está asociada con la frecuencia  $\omega_i$ . Por lo tanto la ecuación 3.1.2.3 queda como:

$$\gamma_k = \sum_{i=-M}^M c_i \cos\omega_i k$$

donde

$$c_i = c_{-i} = \begin{cases} A_i^2/4, & \omega_i \neq 0 \\ A_i^2/2, & \omega_i = 0 \end{cases}$$

o equivalentemente

$$\begin{aligned}\gamma_k &= \int_{-\pi}^{\pi} \cos \omega k \, dF(\omega) \\ &= \int_{-\pi}^{\pi} e^{i\omega k} dF(\omega)\end{aligned}$$

en la cual  $F(\omega)$  es una función seccional monótonamente decreciente con intervalos continuos de longitud  $A^2/4$  en  $\omega=\omega_i \neq 0$  y un intervalo de tamaño  $A^2/2$  cuando  $\omega=\omega_i=0$  para  $i=0,1,\dots,M$ .

La figura 2 ilustra la función de distribución espectral  $F(\omega)$  para un proceso con  $M=2$ ,  $A_1=1$ ,  $A_2=2$ , y frecuencias  $\omega_1$  y  $\omega_2$  no negativas haciendo énfasis en las definiciones citadas se observa que es una función acumulativa definida en el intervalo  $[-\pi,\pi]$  por lo que inicialmente vale cero, en  $-\omega_2$ , tiene un incremento de 1, en  $-\omega_1$ , uno de 0.25 y así sucesivamente. El espectro con las contribuciones a la varianza para cada frecuencia es un conjunto de líneas discretas y se presenta en la figura 3.

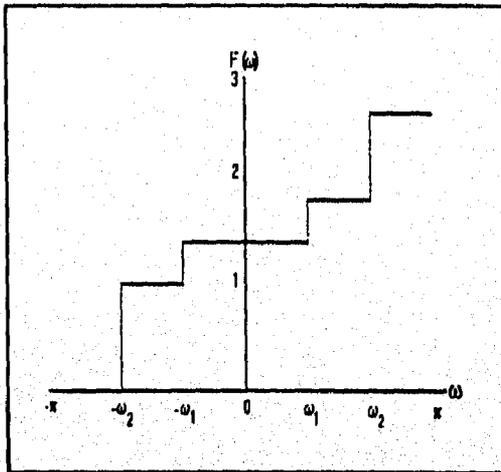


Fig 2. Distribución espectral

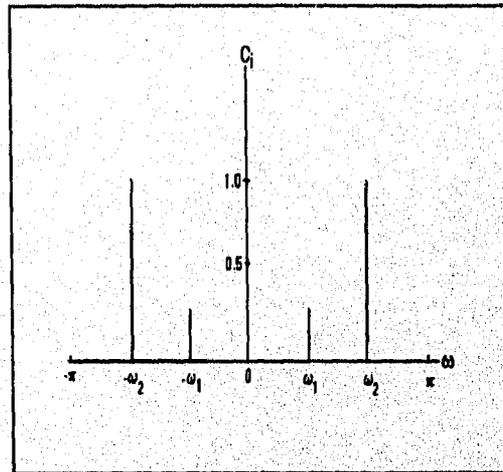


Fig 3. Espectro de líneas del proceso

Hay que considerar que aunque  $F(\omega)$  es no decreciente y positiva, aún no cumple con las todas las propiedades inherentes a una función de distribución ya que  $\int_{-\pi}^{\pi} dF(\omega) = \gamma_0$  la cual no necesariamente es 1. Sin embargo si se define

$$G(\omega) = \frac{F(\omega)}{\gamma_0},$$

entonces  $G(\omega) \geq 0$  y  $\int_{-\pi}^{\pi} dG(\omega) = 1$ . Cuando  $dF(\omega) = f(\omega)d\omega$ , se cumple que

$$p(\omega)d\omega = dG(\omega) = \frac{f(\omega)}{\gamma_0}d\omega.$$

En consecuencia, de 3.1.2.1 y 3.1.2.2 se obtiene el par de transformadas de Fourier correspondientes:

$$\begin{aligned}
 p(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k e^{-i\omega k} \\
 &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k \cos \omega k \\
 &= \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \rho_k \cos \omega k, \quad -\pi \leq \omega \leq \pi
 \end{aligned}
 \tag{3.1.2.4}$$

y

$$\rho_k = \int_{-\pi}^{\pi} p(\omega) e^{i\omega k} d\omega, \quad k=0, \pm 1, \pm 2, \dots
 \tag{3.1.2.5}$$

La función  $p(\omega)$  tiene las propiedades de una función de densidad de probabilidades sobre el rango  $[-\pi, \pi]$  y se le conoce como la función de densidad espectral.

### 3.1.3 Descomposición de Wold

En el estudio de las series de tiempo se ha hablado de series estocásticas y determinísticas por separado, pero también es posible encontrar en la realidad fenómenos cuyos modelos estén formados por una combinación de ambas, por ejemplo el modelo en donde  $A$  es una constante,  $\lambda$  es una variable aleatoria uniformemente distribuida en  $[-\pi, \pi]$  y  $[\theta(\mathbf{B})/\phi(\mathbf{B})]$  es un modelo ARMA estacionario e invertible; este modelo contiene ambos componentes (determinísticos y estocásticos). En efecto, Wold (1954) demostró que cualquier proceso estacionario puede ser escrito como

$$Z_t = Z_t^{(d)} + Z_t^{(n)}$$

en el que  $Z_t^{(d)}$  es el componente determinístico y  $Z_t^{(n)}$  es el no determinístico puro. Esta ecuación se le da el nombre de descomposición de Wold de un proceso estacionario. Esta fórmula es similar a la descomposición de la función de distribución espectral dada en el apartado anterior considerando a  $Z_t^{(d)}$  como la función de densidad seccionalmente continua y a  $Z_t^{(n)}$  como la función continua absolutamente.

## 3.2 El espectro de algunos procesos

### 3.2.1 El espectro y la función generadora de autocovarianzas

En el capítulo anterior se definió la función generadora de autocovarianzas

$$\gamma(b) = \sum_{k=-\infty}^{\infty} \gamma_k B^k,$$

en la cual la varianza del proceso es el coeficiente de  $B^0=1$ , y la autocovarianza  $\gamma_k$  con  $k$  unidades de diferencia esta asociada a  $B^k$  y  $-B^k$ . Por otro lado, si la sucesión de autocovarianzas es absolutamente sumable, entonces existe su espectro

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{i\omega k}.$$

Si comparamos las ecuaciones anteriores podemos encontrar una relación entre el espectro y la función generadora de autocovarianzas de tal forma que

$$f(\omega) = \frac{1}{2\pi} \gamma(e^{-i\omega}).$$

### 3.2.2 Espectro de modelos ARMA

Como se mencionó en el capítulo anterior, cualquier proceso estocástico  $Z_t$  puede ser descrito de la siguiente forma

$$Z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B)a_t$$

donde  $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ ,  $\psi_0 = 1$  y su función generadora de autocovarianzas es

$$\gamma(B) = \sigma^2 \psi(B)\psi(B^{-1})$$

Considere un modelo **ARMA(p,q)**

$$\phi_p(B)Z_t = \theta_q(B)a_t$$

con  $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  y  $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$  que se escribe

$$Z_t = \psi(B)a_t$$

para la que  $\psi(B) = \theta_q(B)/\phi_p(B)$ . Por tanto, la función generadora de autocovarianzas del modelo **ARMA(p,q)** viene dada por

$$\gamma(B) = \sigma^2 \frac{\theta_q(B)\theta_q(B^{-1})}{\phi_p(B)\phi_p(B^{-1})} \quad 3.2.2.1$$

Cuando el modelo es estacionario, la función de autocovarianzas es absolutamente sumable, consecuentemente, el espectro del modelo estacionario **ARMA(p,q)** es

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \gamma(e^{-i\omega}) \\ &= \frac{\sigma^2 \theta_q(e^{-i\omega})\theta_q(e^{i\omega})}{2\pi \phi_p(e^{-i\omega})\phi_p(e^{i\omega})} \\ &= \frac{\sigma^2 |\theta_q(e^{-i\omega})|^2}{2\pi |\phi_p(e^{-i\omega})|^2} \end{aligned} \quad 3.2.2.2$$

que es conocido también como espectro racional.

Si el modelo es inversible, entonces existe la inversa de  $f(\omega)$  y está dada por

$$f^{-1}(\omega) = \frac{2\pi |\phi_p(e^{-i\omega})|^2}{\sigma^2 |\theta_q(e^{-i\omega})|^2}$$

que es, en efecto, el espectro de un modelo **ARMA(q,p)**. Aplicando la transformada inversa de Fourier de  $f^{-1}(\omega)$  se obtiene

$$\gamma_k^{(f)} = \int_{-\infty}^{\infty} f^{-1}(\omega) e^{i\omega k} d\omega,$$

denominada función inversa de autocovarianzas. Obviamente

$$\rho_k^{(f)} = \frac{\gamma_k^{(f)}}{\gamma_0^{(f)}} = \frac{1}{\gamma_0^{(f)}} \int_{-\pi}^{\pi} f^{-1}(\omega) e^{i\omega k} d\omega$$

representa a la función inversa de autocorrelaciones.

La función inversa de autocorrelaciones puede usarse como una herramienta de identificación en el análisis de series de tiempo ya que presenta características similares a la función de autocorrelación parcial.

## El espectro del Ruido blanco

El proceso ruido blanco es una serie de variables aleatorias independientes

$$Z_i = a_i$$

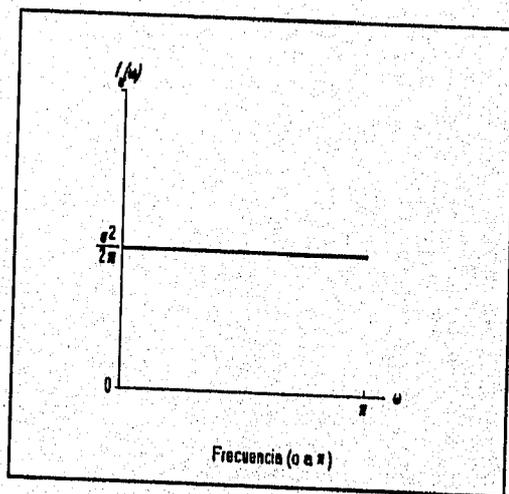
con función de autocovarianzas

$$\gamma_k = \begin{cases} \sigma^2, & k=0, \\ 0, & k \neq 0, \end{cases}$$

y función generadora de autocovarianzas  $\gamma(B) = \sigma^2$ . En consecuencia, por 3.1.1.1

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k} \\ &= \frac{\sigma^2}{2\pi}, \quad -\pi \leq \omega \leq \pi, \end{aligned}$$

cuya gráfica es una línea horizontal mostrada en la siguiente figura



**Fig 4.**

El comportamiento de esta función implica que todas las frecuencias contribuyen de igual modo a la varianza.

## El espectro de un proceso AR(1)

La función generadora de autocovarianzas para un proceso estacionario AR(1),  $(1-\phi B)Z_t = a_t$ , se deduce de 3.2.2.1 como

$$\gamma(B) = \sigma^2 \frac{1}{(1-\phi B)(1-\phi B^{-1})}$$

y por 3.2.2.2 su espectro es

$$\begin{aligned} f(\omega) &= \frac{\sigma^2}{2\pi (1-\phi e^{-i\omega})(1-\phi e^{i\omega})} \\ &= \frac{\sigma^2}{2\pi (1+\phi^2 - \phi(\cos\omega - i\sin\omega) - \phi(\cos\omega + i\sin\omega))} \\ &= \frac{\sigma^2}{2\pi (1+\phi^2 - 2\phi\cos\omega)} \end{aligned}$$

La forma de la curva depende del signo de  $\phi$ . Como se muestra en la figura 5, cuando  $\phi > 0$  y la serie esta correlacionada positivamente, es espectro está dominado por componentes de baja frecuencia (o período largo). Si  $\phi < 0$  la serie esta correlacionada negativamente, y el espectro está dominado por los componentes de alta frecuencia (o período corto). En otras palabras, un espectro en el que dominen las frecuencias bajas, indica una serie relativamente suave, es cambio, un espectro dominado por frecuencias altas implica una serie más dispereja.

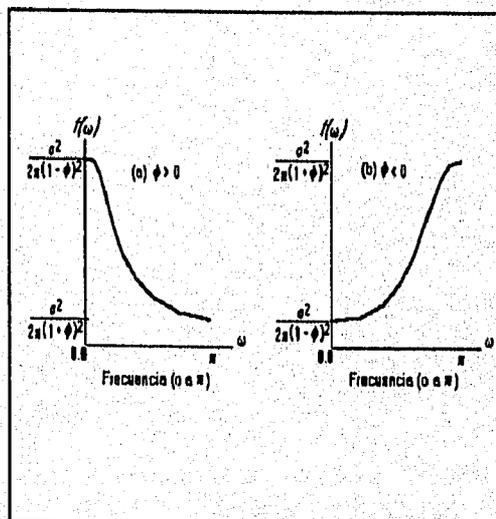


Fig 5

Conforme  $\phi$  se aproxima a 1, el proceso **AR(1)** se convierte en un modelo de caminata aleatoria, y por ende

$$f(\omega) = \frac{\sigma^2}{4\pi} \frac{1}{(1 - \cos\omega)}$$

corresponde al espectro del modelo, aunque una caminata aleatoria no tiene espectro debido a que su sucesión de autocovarianzas no es absolutamente sumable. Este efecto se refleja en la función  $f(\omega)$  cuando  $\omega$  tiende a infinito. La conclusión de este fenómeno en el análisis de series es que cuando haya un espectro muestral con una cresta cercana a cero muy alta, posiblemente exista la necesidad de diferenciar.

### El espectro de un proceso **MA(1)**

La función generadora de autocovarianzas para un proceso estacionario de media móviles **MA(1)**,  $Z_t = (1 - \theta B)a_t$ , es

$$\gamma(B) = \sigma^2(1 - \theta B)(1 - \theta B^{-1}),$$

y, nuevamente, por 3.2.2.2 su espectro es

$$\begin{aligned} f(\omega) &= \frac{\sigma^2}{2\pi} (1 - \theta e^{-i\omega})(1 - \theta e^{i\omega}) \\ &= \frac{\sigma^2}{2\pi} (1 + \theta^2 - \theta(\cos\omega - i\sin\omega) - \theta(\cos\omega + i\sin\omega)) \\ &= \frac{\sigma^2}{2\pi} (1 + \theta^2 - 2\theta\cos\omega). \end{aligned}$$

La forma de la curva depende otra vez del signo de  $\theta$ . El comportamiento del espectro de este proceso es inverso al del **AR(1)**, por lo tanto cuando  $\theta$  es positiva, la serie esta correlacionada negativamente y, por ende, dispareja, lo cual se ve reflejado en el espectro cuanto a frecuencias altas les corresponden valores altos y viceversa; cuando  $\theta$  es negativa, la serie tendrá una forma suave y en el espectro habrá valores altos en frecuencias bajas. Ambos casos están ilustrados en la figura 6.

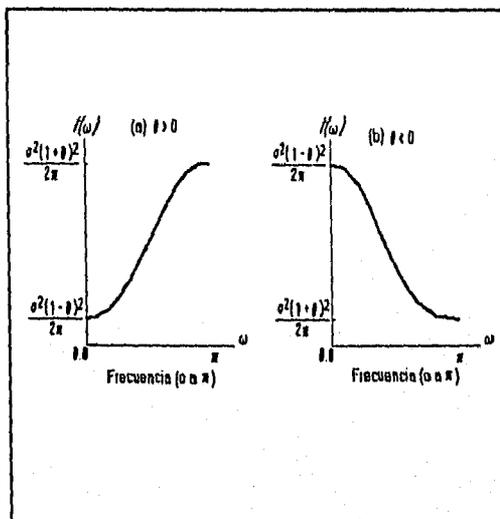


Fig 6

### El espectro de la suma de dos procesos independientes

Considere el proceso  $Z_t$ , el cual es la suma de dos procesos estacionarios  $X_t$  y  $Y_t$ ,

$$Z_t = X_t + Y_t.$$

Sean  $\gamma_z(\mathbf{B})$ ,  $\gamma_x(\mathbf{B})$ , y  $\gamma_y(\mathbf{B})$  las funciones generadoras de autocovarianzas de cada uno de los procesos, entonces,

$$\begin{aligned} \gamma_z(\mathbf{B}) &= \sum_{k=-\infty}^{\infty} \gamma_z(k) B^k \\ &= \sum_{k=-\infty}^{\infty} \text{Cov}(Z_t, Z_{t+k}) B^k \\ &= \sum_{k=-\infty}^{\infty} \text{Cov}(X_t + Y_t, X_{t+k} + Y_{t+k}) B^k \\ &= \sum_{k=-\infty}^{\infty} [\text{Cov}(X_t, X_{t+k}) + \text{Cov}(Y_t, Y_{t+k})] B^k \\ &= \sum_{k=-\infty}^{\infty} \gamma_x(k) B^k + \sum_{k=-\infty}^{\infty} \gamma_y(k) B^k \\ &= \gamma_x(\mathbf{B}) + \gamma_y(\mathbf{B}). \end{aligned}$$

Consecuentemente

$$f_z(\omega) = f_x(\omega) + f_y(\omega).$$

es decir, el espectro de la suma de dos procesos independientes es la suma de los espectros.

## El espectro de modelos estacionales

Como ya se ha visto, un modelo determinístico estacional con período  $s$  tiene un espectro discreto (lineal) en las armónicas  $\omega=2\pi ks$ ,  $k=\pm 1, \pm 2, \dots, \pm[s/2]$ . Por ejemplo el siguiente modelo

$$Z_t = \alpha_0 + \sum_{k=1}^6 \left[ \alpha_k \cos\left(\frac{2\pi kt}{12}\right) + \beta_k \text{sen}\left(\frac{2\pi kt}{12}\right) \right] + a_t$$

donde  $a_t$  representa ruido blanco.

El espectro de  $Z_t$ , considerando el apartado anterior, estará dado por la suma de dos espectros independientes. Uno de ellos representa el componente determinístico, y el otro será asociado a  $a_t$ . Debido a que el espectro del ruido blanco es una línea horizontal, el espectro resultante contendrá picos en las frecuencias armónicas estacionales  $\omega=2\pi ks$ ,  $k=\pm 1, \pm 2, \dots, \pm 6$ .

Ahora considere el siguiente modelo estacional ARMA con período 12:

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

cuya función generadora de autocovarianzas es

$$\gamma(B) = \frac{\sigma^2}{2\pi} \frac{1}{(1 - \Phi B^{12})(1 - \Phi B^{-12})}$$

Si el proceso es estacionario, existe su espectro y esta dado por

$$\begin{aligned} f(\omega) &= \frac{\sigma^2}{2\pi} \frac{1}{(1 - \Phi e^{-12i\omega})(1 - \Phi e^{12i\omega})} \\ &= \frac{\sigma^2}{2\pi} \frac{1}{1 + \Phi^2 - 2\Phi \cos(12\omega)} \end{aligned}$$

Perciba que  $\cos(12\omega)=-1$  en  $\omega=\pi(2k-1)/12$  para valores  $k=1,2,3,4,5,6$ , y  $\cos(12\omega)=1$  cuando  $\omega=0$  y  $\omega=2\pi k/12$ ,  $k=1,2,3,4,5,6$ . Así, siempre que  $\Phi>0$  además de una cresta en  $\omega=0$ , el espectro presentará máximos en las frecuencias armónicas estacionales  $2\pi k/12$  si  $k=1,2,3,4,5$  y  $6$ ; y mínimos en frecuencias  $\omega=\pi(2k-1)/12$  para los mismos valores de  $k$ , como se muestra en la figura 7.

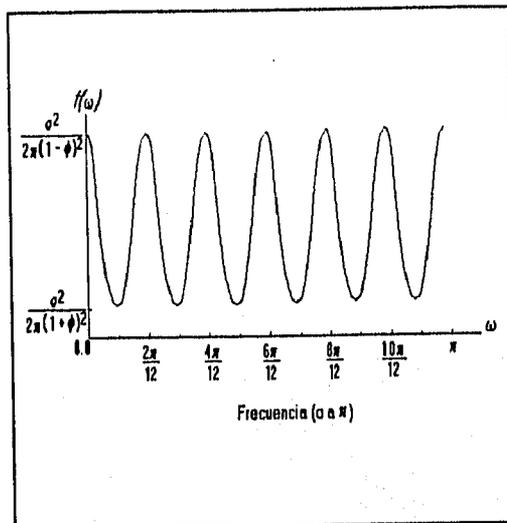


Fig 7

Como ejemplo, sea un modelo ARMA multiplicativo que contiene partes estacionales y no estacionales,

$$(1 - \Phi B^{12})(1 - \phi B)Z_t = a_t$$

a este modelo, si es estacionario, le corresponde el espectro

$$f(\omega) = \frac{\sigma^2}{(2\pi(1 + \Phi^2 - 2\Phi \cos 12\omega))(1 + \phi^2 - 2\phi \cos \omega)}$$

La gráfica cuando  $\Phi > 0$  y  $\phi > 0$  se aprecia en la figura 8.

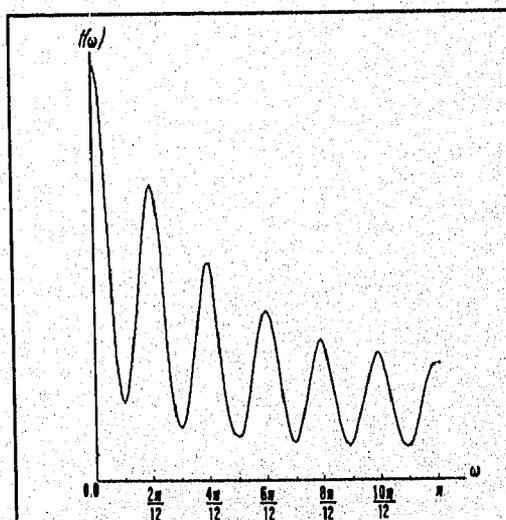


Fig 8

Aquí es conveniente aclarar que el espectro muestra del siguiente capítulo no tendrá un comportamiento tan regular como el que se presenta, sin embargo, si los datos verdaderamente siguen un patrón estacional, el espectro presentará picos o saltos en las frecuencias armónicas.

Otra aclaración que es importante es que en el análisis de la serie por medio del espectro, un pico agudo puede indicar la presencia de un componente determinístico estacional, mientras que los picos anchos usualmente señalan componentes estacionales no determinísticos similares a los procesos multiplicativos estacionales ARIMA.

Para una revisión detallada de todos los modelos posibles, se pone a disposición del lector el programa AST el cual genera el espectro de cualquier modelo SARIMA.

## Capítulo 4. Estimación del espectro

## Introducción

En este capítulo se desarrollará un método para estimar el espectro de una serie de observaciones. El espectro estimado es útil en la determinación del modelo de comportamiento de los datos ya que puede ser comparado con el espectro teórico visto en el capítulo anterior, y en base a esto encontrar el modelo que proporcione el mejor ajuste.

Por otro lado, cuando se estima el espectro, se obtiene una función cuyo comportamiento suele ser inestable y, debido a la naturaleza del espectro teórico, no resulta sencillo compararlos; por lo cual se han desarrollado métodos que permiten suavizar el espectro estimado para facilitar el proceso de identificación del modelo.

## 4.1 El periodograma

### 4.1.1. Definición

Dada una serie de  $n$  observaciones, se puede utilizar el material del capítulo 1 para representar las observaciones como una serie de Fourier.

$$Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} (a_k \cos \omega_k t + b_k \sin \omega_k t)$$

para la cual, las  $\omega_k = 2\pi k/n$ ,  $k=0, 1, \dots, \lfloor n/2 \rfloor$  son las frecuencias de Fourier y

$$a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos \omega_k t & k=0 \text{ y } k=\frac{n}{2} \text{ si } n \text{ es par,} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos \omega_k t & k=1, 2, \dots, \lfloor \frac{n-1}{2} \rfloor, \end{cases}$$

y

$$b_k = \begin{cases} \frac{2}{n} \sum_{t=1}^n Z_t \sin \omega_k t & k=1, 2, \dots, \lfloor \frac{n-1}{2} \rfloor, \end{cases}$$

son los coeficientes de Fourier. Cabe mencionar la relación que existe entre la representación anterior y el análisis de regresión visto en el capítulo 2. En efecto, los coeficientes de Fourier son, en general los coeficientes de regresión estándar al ajustar el siguiente modelo de regresión:

$$Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} (a_k \cos \omega_k t + b_k \sin \omega_k t) + e_t,$$

donde las  $\omega_k$  son las mismas frecuencias de Fourier y los  $e_t$  son variables aleatorias independientes idénticamente distribuidas (v.a.i.i.d.) con una ley de probabilidades normal con media 0 y varianza  $\sigma^2 \sim (N(0, \sigma^2))$ .

Entonces, es posible hacer uso de la noción de regresión para considerar a la relación de Parseval.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

$$\sum_{i=1}^n Z_i^2 = \begin{cases} na_0^2 + \frac{n}{2} \sum_{k=1}^{[n/2]} (a_k^2 + b_k^2) & \text{si } n \text{ es par} \\ na_0^2 + \frac{n}{2} \sum_{k=1}^{[n/2]} (a_k^2 + b_k^2) + na_{n/2}^2 & \text{si } n \text{ es impar} \end{cases}$$

como un análisis de varianza presentado a continuación:

Fuente	Grados de libertad	Suma de Cuadrados
Frecuencia $\omega_0 = \text{Media}$	1	$na_0^2$
Frecuencia $\omega_1 = 2\pi/n$	2	$(a_1^2 + b_1^2)n/2$
Frecuencia $\omega_2 = 4\pi/n$	2	$(a_2^2 + b_2^2)n/2$
⋮	⋮	⋮
Frecuencia $\omega_{(n-1)/2} = (n-1)\pi/n$	2	$(a_{(n-1)/2}^2 + b_{(n-1)/2}^2)n/2$
Frecuencia $\omega_{n/2} = \pi/n$ si $n$ par	1	$na_{n/2}^2$
Total	$n$	$\sum_{i=1}^n Z_i^2$

Donde los grados de libertad son el número de elementos (o variables aleatorias que están inmersos en cada término de la suma de cuadrados).

La cantidad  $I(\omega_k)$  definida por

$$I(\omega_k) = \begin{cases} na_0^2 & k=0 \\ \frac{n}{2}(a_k^2 + b_k^2) & k=1,2,\dots,(n-1)/2 \\ na_{n/2}^2 & k=n/2 \text{ cuando } n \text{ es par} \end{cases}$$

es conocida como el **periodograma**, el cual fue introducido en 1898 por Schuster para buscar componentes periódicos en una serie.

### 4.1.2 Propiedades muestrales del espectro

En esta sección se asumirá que el número  $n$  de variables aleatorias contenidas en la muestra es impar, ésto sin efecto alguno en el análisis cuando este número sea par.

Considérese un conjunto de  $n$  v.a.i.i.d  $Z_1, \dots, Z_n$  con distribución  $N(0, \sigma^2)$ , Entonces

$$E(a_k) = \frac{2}{n} \sum_{t=1}^n E(Z_t) \cos \omega_k t = 0$$

y

$$\begin{aligned} \text{Var}(Z) &= \text{Var} \left[ \frac{2}{n} \sum_{t=1}^n Z_t \cos \omega_k t \right] \\ &= \frac{4}{n^2} \sum_{t=1}^n \text{Var}(Z_t) \cos^2 \omega_k t \\ &= \frac{4}{n^2} \sum_{t=1}^n \sigma^2 \cos^2 \omega_k t \\ &= \frac{4\sigma^2}{n^2} \sum_{t=1}^n \cos^2 \omega_k t \\ &= \frac{4\sigma^2}{n^2} \frac{n}{2} = \frac{2\sigma^2}{n} \end{aligned}$$

Por tanto, los coeficientes  $a_k$  para  $k=1, 2, \dots, (n-1)/2$  son v.a.i.i.d.  $\sim N(0, \sigma^2)$  y, en consecuencia, las  $na_k^2/2\sigma^2$  siguen una distribución  $\chi^2$  (ji-cuadrada)<sup>6</sup> con 1 grado de libertad, haciendo un razonamiento análogo se llega misma conclusión para las  $nb_k^2/2\sigma^2$ ; más aún,  $na_k^2/2\sigma^2$  y  $nb_k^2/2\sigma^2$  son independientes para  $k=1, 2, \dots, (n-1)/2$  debido a la propiedad ortogonal del sistema seno-coseno; en otras palabras

$$\begin{aligned} \text{Cov}(a_k, b_j) &= \frac{4}{n^2} E \left( \sum_{t=1}^n Z_t \cos \omega_k t \sum_{u=1}^n Z_u \sin \omega_j u \right) \\ &= \frac{4}{n^2} \sum_{t=1}^n [E(Z_t^2) \cos \omega_k t \sin \omega_j t] \\ &= \frac{4\sigma^2}{n^2} \sum_{t=1}^n \cos \omega_k t \sin \omega_j t \\ &= 0 \quad \text{para toda } k \neq j \end{aligned} \quad 4.1.2.1$$

<sup>6</sup>Sea un conjunto de  $n$  variables aleatorias independientes e idénticamente distribuidas bajo una ley de probabilidades normal con media cero, entonces la suma de los cuadrados de éstas seguirán una distribución ji-cuadrada con  $n$  grados de libertad.

Esto hace que las ordenadas del periodograma

$$\frac{I(\omega_k)}{\sigma^2} = \frac{n}{2\sigma^2}(a_k^2 + b_k^2)$$

para  $k=1,2,\dots,(n-1)$  sean v.a.i.d. con distribución ji-cuadrada y 2 grados de libertad, y por argumentos semejantes  $I(0)/\sigma^2$  y  $I(\pi)/\sigma^2$  se ajustan a la misma ley de probabilidades pero con 1 grado de libertad.

Analícese ahora una serie cuyo comportamiento es estacional, supóngase que dicha serie esta dada por

$$Z_t = a_0 + a_k \cos \omega_k t + b_k \sin \omega_k t + e_t$$

donde  $\omega_k = 2\pi k/n$ ,  $k$  es diferente de 0 y  $e_t$  es un proceso de ruido blanco. Para examinar el modelo, se procederá a armar la siguiente prueba de hipótesis

$$H_0: a_k = b_k = 0 \quad \text{contra} \quad H_1: a_k \neq 0 \text{ o } b_k \neq 0$$

bajo el estadístico de prueba (o prueba estadística):

$$\begin{aligned} F &= \frac{\frac{n}{2}(a_k^2 + b_k^2)/2}{\frac{n}{2} \sum_{j=1}^{\lfloor n/2 \rfloor} (a_j^2 + b_j^2)/(n-3)} \\ &= \frac{(n-3)(a_k^2 + b_k^2)}{2 \sum_{j=1}^{\lfloor n/2 \rfloor} (a_j^2 + b_j^2)} \end{aligned}$$

Cabe aclarar que se está buscando comprobar probabilísticamente si la serie contiene componentes periódicos; para ello se propone la hipótesis nula  $H_0$  como: "La serie no es periódica", contra la hipótesis alternativa; "La serie es periódica"; utilizando para esto la función de distribución  $F$  con 2 y  $(n-3)$  grados de libertad ( $F(2, n-3)$ ). Además, y en vista de que el modelo propuesto consta de un nivel medio  $a_0$ , el numerador y el denominador no incluyen el término  $na_0^2/2$ . En efecto, debido a que el periodograma evaluado en la frecuencia 0 arroja la media muestral y no la periodicidad de la serie, ésta se excluye del análisis.

Más generalmente, es posible construir una prueba de hipótesis para determinar si una serie contiene varios componentes periódicos postulando el modelo

$$Z_t = a_0 + \sum_{i=1}^m (a_{k_i} \cos \omega_{k_i} t + b_{k_i} \sin \omega_{k_i} t) + e_t$$

en el que  $e_t$  es ruido blanco  $\omega_k = 2\pi k/n$  y el conjunto  $I = \{k; i=1, \dots, m\}$  es un subconjunto de  $\{k: k=1, 2, \dots, [n/2]\}$ .

El estadístico correspondiente al modelo bajo las hipótesis mencionadas es:

$$F = \frac{(n-2m-1) \sum_{i=1}^m (a_{k_i}^2 + b_{k_i}^2)}{2m \sum_{j=1}^{[n/2]} (a_j^2 + b_j^2)}$$

el cual sigue una distribución  $F$  con  $2m$  y  $(n-2m-1)$  grados de libertad.

#### 4.1.3 Pruebas para componentes periódicos

En la práctica, si una serie de tiempo contiene un componente periódico no se conoce la frecuencia que lo produce.

Por ejemplo, si se quiere probar

$$H_0: \alpha = \beta = 0 \quad \text{contra} \quad H_1: \alpha \neq 0 \quad \text{o} \quad \beta \neq 0$$

para el modelo

$$Z_t = \alpha \cos \omega t + \beta \sin \omega t + e_t \quad 4.1.3.1$$

donde  $e_t$  es ruido blanco, y la frecuencia  $\omega$  es desconocida, entonces, en vista de la incertidumbre acerca del valor de la frecuencia, la distribución  $F$  y el estadístico de prueba no se aplican directamente. Sin embargo, el análisis del periodograma aún puede ser llevado a cabo. Efectivamente, si el modelo en cuestión encierra un componente periódico en la frecuencia  $\omega$ , se espera que el periodograma  $I(\omega_k)$  en la frecuencia de Fourier  $\omega_k$  más cercana a  $\omega$  llegue al máximo.

Así, es posible efectuar una búsqueda de la mayor ordenada del periodograma y establecer de nuevo, mediante una prueba de hipótesis, si esta ordenada puede ser considerada como el máximo en una muestra de  $[n/2]$  v.a.i.i.d.  $(\omega)$  que siguen una distribución  $\chi^2$  con 2 grados de libertad.

En este caso, una prueba estadística sería

$$I^{(1)}(\omega_{(1)}) = \max I(\omega_k)$$

para la cual  $\omega_{(1)}$  señala la frecuencia de Fourier a la que le corresponde el máximo en el periodograma.

Ahora, bajo la hipótesis nula  $H_0$ : "El proceso  $\{Z_t\}$  es ruido blanco" y considerando que las  $I(\omega_k)$  se distribuyen como una  $\chi^2$  con 2 grados de libertad cuya función de probabilidades esta dada por

$$p(x) = \frac{1}{2} e^{-\frac{x}{2}} \quad 0 \leq x < \infty$$

Entonces, para cualquier constante  $g$  mayor o igual a 0, se tiene

$$\begin{aligned} P \left[ \frac{I^{(1)}(\omega_{(1)})}{\sigma^2} > g \right] &= 1 - P \left[ \frac{I^{(1)}(\omega_{(1)})}{\sigma^2} \leq g \right] \\ &= 1 - P \left[ \frac{I(\omega_k)}{\sigma^2} \leq g, k=1, 2, \dots, [n/2] \right] \\ &= 1 - \left( \int_0^g \frac{1}{2} e^{-x/2} \right)^{[n/2]} \\ &= 1 - (1 - e^{-g/2})^{[n/2]} \end{aligned}$$

donde se consideró que la probabilidad de que  $I^{(1)}(\omega_{(1)})$  sea menor que  $g$  es igual a la probabilidad de que todas las ordenadas del espectro lo sean, ya que  $I^{(1)}(\omega_{(1)})$  es la máxima coordenada; además, como todas estas variables son independientes, la probabilidad de que todas ellas sean menores que  $g$  al mismo tiempo es igual al producto de la probabilidad de que cada una de ellas lo sea.

Por otro lado, si la varianza  $\sigma^2$  estuviera determinada, se podría usar el desarrollo previo para construir una prueba estadística exacta para la ordenada mayor.

Sin embargo, en la práctica,  $\sigma^2$  es desconocido en la mayoría de los casos por lo que hay que estimarla. Para encontrar un estimador insesgado de  $\sigma^2$  tómesese en consideración que

$$E \left[ \sum_{k=1}^{[n/2]} I(\omega_k) \right] = \left[ \frac{n}{2} \right] 2\sigma^2$$

por lo que

$$\hat{\sigma}^2 = \frac{1}{2[n/2]} \sum_{k=1}^{[n/2]} I(\omega_k)$$

es un estimador insesgado de  $\sigma^2$  y, en consecuencia, el estadístico de prueba es

$$V = \frac{I^{(1)}(\omega_{(1)})}{2[n/2] \sum_{k=1}^{[n/2]} I(\omega_k)}$$

Considerando que los  $I(\omega_k)$ , son independientes, entonces

$$\text{Var}(\hat{\sigma}^2) \rightarrow 0 \quad \text{cuando } n \rightarrow \infty$$

Esto se deriva de que  $\hat{\sigma}^2$  es un estimador consistente de  $\sigma^2$ . Por consiguiente, para muestras grandes,  $V$  tiene aproximadamente la misma distribución que  $I^{(1)}(\omega_{(1)})/\sigma^2$ , es decir, para todo  $g$  mayor o igual a cero,

$$P(V > g) \approx 1 - (1 - e^{-(g/2)})^{[n/2]}$$

Una prueba exacta para  $I(\omega_k)$ , fue dado por Fischer en 1929; ésta se basa en el siguiente estadístico

$$T = \frac{I^{(1)}(\omega^{(1)})}{\sum_{k=1}^{[n/2]} I(\omega_k)}$$

Bajo la hipótesis nula " $Z_t$  es un proceso Gaussiano de ruido blanco con distribución  $N(0, \sigma^2)$ ". Fischer mostró que

$$P(T > g) = \sum_{j=1}^m (-1)^{j-1} \binom{N}{j} (1-jg)^{N-1} \quad 4.1.3.2$$

donde  $N=[n/2]$ ,  $g > 0$  y  $m=[1/g]$ . Así, para cualquier nivel de significación  $\alpha$  es factible utilizar la ecuación anterior para encontrar el valor crítico de  $g_\alpha$  tal que

$$P(T > g_\alpha) = \alpha.$$

Si el valor de  $T$  calculado de la serie es mayor que  $g_\alpha$ , entonces se rechaza la hipótesis nula (el proceso es estacional). Este procedimiento es conocido como la prueba de Fischer.

La distribución  $T$  para el nivel de significación  $\alpha=0.05$  como fue dada por Fischer se presenta a continuación junto con una aproximación obtenida a partir de primer término de 4.1.3.2

$$P(T > g) \approx N(1-g)^{N-1} \quad 4.1.3.3$$

$N=[n/2]$	$g_\alpha$ (fórmula exacta)	$g_\alpha$ (1er término)
05	0.68377	0.68377
10	0.44495	0.44495
15	0.33465	0.33463
20	0.27040	0.27046
25	0.22805	0.22813
30	0.19784	0.19794
35	0.17513	0.17525
40	0.15738	0.15752
45	0.14310	0.14324
50	0.13135	0.13149

Como puede apreciarse, la aproximación es muy cercana al resultado exacto, por lo que, en la mayoría de las aplicaciones se utiliza la ecuación 4.1.3.3 para encontrar el valor crítico  $g_\alpha$  de la prueba.

Cabe aclarar que un valor significativo de  $I^{(1)}(\omega_{(1)})$  implica rechazar la hipótesis nula de 4.1.3.1 o, en otras palabras, existe un componente periódico en la serie en alguna frecuencia  $\omega$ .

Por otro lado, la frecuencia  $\omega$  no necesariamente es igual a  $\omega_{(1)}$ , ya que se escogió de las frecuencias de Fourier y no del intervalo entre 0 y  $\pi$ . Sin embargo, en 1949, Hartley demostró que la  $\omega$  desconocida asociada al máximo componente periódico puede estimarse por  $\omega_{(1)}$ , y la probabilidad de que  $|\omega - \omega_{(1)}| > 2\pi/n$  es menor que el nivel de significación de la prueba.

Sea  $I^{(2)}(\omega_{(2)})$  la segunda ordenada más grande del periodograma correspondiente a la frecuencia de Fourier  $\omega_{(2)}$ . Entonces existe una extensión a la prueba de Fischer para la segunda mayor ordenada del periodograma basada en el estadístico

$$T_2 = \frac{I^{(2)}(\omega_{(2)})}{\left\{ \sum_{k=1}^{[n/2]} I(\omega_k) \right\} - I^{(1)}(\omega_{(1)})}$$

donde se utiliza la distribución 4.1.3.2 reemplazando a  $N$  con  $N-1$ . Este proceso continua buscando componentes periódicos hasta que se obtiene un valor insignificante.

## 4.2 El espectro muestral

Para estimar el espectro de una serie con autocovarianzas absolutamente sumables, recuérdese la definición del espectro

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k} \\ &= \frac{1}{2\pi} \left( \gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos \omega k \right), \quad -\pi \leq \omega \leq \pi \end{aligned}$$

Con base en una muestra, es natural estimar  $f(\omega)$  reemplazando las covarianzas teóricas  $\gamma_k$  por las autocovarianzas muestrales  $\hat{\gamma}_k$ . Sin embargo, para una serie de  $n$  observaciones, sólo se pueden calcular  $\hat{\gamma}_k$  para  $k=0, 1, \dots, (n-1)$ . Por lo que el estimador del espectro o espectro muestral es

$$\begin{aligned} \hat{f}(\omega) &= \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \hat{\gamma}_k e^{-i\omega k} \\ &= \frac{1}{2\pi} \left( \hat{\gamma}_0 + 2 \sum_{k=1}^{n-1} \hat{\gamma}_k \cos \omega k \right), \quad -\pi \leq \omega \leq \pi. \end{aligned} \tag{4.2.1}$$

Como  $\hat{\gamma}_k$  es insesgado asintóticamente, entonces

$$\lim_{n \rightarrow \infty} E(\hat{f}(\omega)) = f(\omega)$$

lo cual implica que  $\hat{f}(\omega)$  es asintóticamente insesgado, por lo cual puede ser un estimador adecuado de  $f(\omega)$ .

Para examinar más a fondo las propiedades del espectro muestral, considérese  $\hat{f}(\omega_k)$  en las frecuencias de Fourier  $\omega_k = 2\pi k/n$   $k=0, 1, \dots, [n/2]$ . En esas frecuencias el espectro muestral y el periodograma están relacionados de la siguiente manera:

$$\begin{aligned} I(\omega_k) &= \frac{n}{2} (a_k^2 + b_k^2) \\ &= \frac{n}{2} (a_k + ib_k)(a_k - ib_k) \\ &= \frac{n}{2} \left[ \sum_{t=1}^n Z_t (\cos \omega_k t - i \sin \omega_k t) \right] \left[ \sum_{t=1}^n Z_t (\cos \omega_k t + i \sin \omega_k t) \right] \\ &= \frac{2}{n} \left[ \sum_{t=1}^n Z_t e^{-i\omega_k t} \right] \left[ \sum_{t=1}^n Z_t e^{i\omega_k t} \right] \end{aligned}$$

$$\begin{aligned}
 I(\omega_k) &= \frac{2}{n} \left[ \sum_{t=1}^n (Z_t - \bar{Z}) e^{-i\omega_k t} \right] \left[ \sum_{t=1}^n (Z_t - \bar{Z}) e^{i\omega_k t} \right] \\
 &= \frac{2}{n} \sum_{t=1}^n \sum_{s=1}^n (Z_t - \bar{Z})(Z_s - \bar{Z}) e^{-i\omega_k(t-s)}
 \end{aligned} \tag{4.2.2}$$

ya que

$$\sum_{t=1}^n e^{-i\omega_k t} = \sum_{t=1}^n e^{i\omega_k t} = 0.$$

Considerando

$$\hat{\gamma}_j = \sum_{t=1}^{n-j} (Z_t - \bar{Z})(Z_{t+j} - \bar{Z}),$$

entonces, substituyendo  $j=t-s$  en 4.2.2 se obtiene

$$\begin{aligned}
 I(\omega_k) &= 2 \sum_{j=-(n-1)}^{n-1} \hat{\gamma}_j e^{-i\omega_k j} \\
 &= 2 \left\{ \hat{\gamma}_0 + 2 \sum_{j=1}^{n-1} \hat{\gamma}_j \cos \omega_k j \right\}.
 \end{aligned}$$

Entonces, debido a 4.2.1

$$\hat{f}(\omega_k) = \frac{1}{4\pi} I(\omega_k), \quad k=1, 2, \dots, [n/2], \tag{4.2.3}$$

donde hay que aclarar que  $\hat{f}(\omega_{n/2}) = I(\omega_{n/2})/2\pi$  si  $n$  es par.

Todo lo anterior lleva a que  $Z_t$  es un proceso Gaussiano de ruido blanco con media 0 y varianza constante  $\sigma^2$  entonces los  $\hat{f}(\omega_k)$ , para  $k=1, 2, \dots, (n-1)/2$  están independiente e idénticamente distribuidos como  $(\sigma^2/4\pi)\chi^2(2) = (\sigma^2/2\pi)\chi^2(2)$ , en otras palabras,

$$\hat{f}(\omega_k) \sim \frac{\sigma^2 \chi^2(2)}{2\pi \cdot 2} \tag{4.2.4}$$

En donde  $\chi^2(2)$  es la distribución ji-cuadrada con 2 grados de libertad. Aquí  $\sigma^2/2\pi$  de 4.2.4 es el espectro muestral de  $Z_t$ . En general, siguiendo los mismos argumentos, se puede demostrar que si  $Z_t$  es un proceso Gaussiano con espectro  $f(\omega_k)$ , entonces

$$\hat{f}(\omega_k) \sim f(\omega_k) \frac{\chi^2(2)}{2}$$

<sup>7</sup>Esto se debe a que las  $\hat{f}(\omega_k)$  son el resultado de la suma de 2 variables normales elevadas al cuadrado

Aplicando la esperanza en ambos lados de la ecuación anterior

$$E(\hat{f}(\omega_k)) = E \left[ f(\omega_k) \frac{\chi^2(2)}{2} \right] = f(\omega_k),$$

y

$$\text{Var}(\hat{f}(\omega_k)) = \text{Var} \left[ f(\omega_k) \frac{\chi^2(2)}{2} \right] = [f(\omega_k)]^2, \quad 4.2.5$$

la cual no depende del tamaño de la muestra  $n$ . En consecuencia, aunque el espectro muestral, calculado en las frecuencias de Fourier, es insesgado, no es un buen estimador ya que no es consistente debido a que la varianza de  $\hat{f}(\omega_k)$  no tiende a 0 cuando  $n$  tiende a infinito. Aún más, dada 4.2.3 y el hecho de que las ordenadas del periodograma son independientes, como se mostró en 4.1.2.1, se tiene

$$\text{Cov}[\hat{f}(\omega_k), \hat{f}(\omega_j)] = 0, \quad K \neq j$$

para 2 frecuencias de Fourier  $\omega_k, \omega_j$  diferentes.

Cuando el tamaño  $n$  de la muestra aumenta y el intervalo entre cada frecuencia se vuelve más fino, la covarianza entre  $\hat{f}(\omega)$  y  $\hat{f}(\lambda)$  aún tiende a 0 cuando  $n$  tiende a infinito. Como resultado,  $\hat{f}(\omega)$  es inestable. Para corregir esta inestabilidad, se han desarrollado varios métodos que suavizan o estabilizan el comportamiento del espectro. Dichos métodos se analizan en el siguiente capítulo.

## 4.3 El espectro suavizado

### 4.3.1 La ventana espectral (suavizando el dominio de las frecuencias)

Una forma de reducir la varianza del espectro muestral es suavizándolo localmente en la vecindad de una frecuencia escogida. En otras palabras, el estimador espectral es el espectro suavizado obtenido de la suma de los valores ponderados de  $m$  valores a la derecha e izquierda de la frecuencia  $\omega_j$  elegida, es decir,

$$\hat{f}(\omega_k) = \sum_{j=-m_n}^{m_n} W_n(\omega_j) \hat{f}(\omega_k - \omega_j)$$

donde  $\omega_k = 2\pi k/n$ ,  $n$  es el tamaño de la muestra,  $k=0, \pm 1, \pm 2, \dots, \pm [n/2]$ , y  $W_n(\omega_j)$  es una función de ponderación que cumple

$$\sum_{j=-m_n}^{m_n} W_n(\omega_j) = 1,$$

$$W_n(\omega_j) = W_n(-\omega_j)$$

y

$$\lim_{n \rightarrow \infty} \sum_{j=-m_n}^{m_n} W_n^2(\omega_j) = 0. \quad 4.3.1.1$$

La función  $W_n(\omega_j)$  es conocida como la ventana espectral porque solamente se utilizan algunas de las ordenadas espectrales en el procedimiento. Si  $f(\omega)$  es plano y aproximadamente constante dentro de la ventana, entonces

$$\begin{aligned} E[\hat{f}_w(\omega_k)] &= \sum_{j=-m_n}^{m_n} W_n(\omega_j) E[\hat{f}(\omega_k - \omega_j)] \\ &\approx f(\omega_k) \sum_{j=-m_n}^{m_n} W_n(\omega_j) = f(\omega_k) \end{aligned}$$

y por 4.2.5

$$\begin{aligned} \text{Var}[\hat{f}_w(\omega_k)] &\approx \sum_{j=-m_n}^{m_n} W_n^2(\omega_j) [f(\omega_k)]^2 \\ &\approx [f(\omega_k)]^2 \sum_{j=-m_n}^{m_n} W_n^2(\omega_j), \end{aligned}$$

La propiedad de la ventana espectral mencionada en 4.3.1.1 implica que la varianza del espectro suavizado disminuye cuando  $m_n$  aumenta. El valor de  $n$  representa el número de frecuencias utilizadas para suavizar el espectro. Esto está directamente relacionado con el ancho de la ventana, también conocido como el ancho de banda de la ventana.

Cuando el ancho de banda se incrementa, se pondera un número mayor de ordenadas espectrales, y por tanto el estimador se vuelve más suave, más estable y tiene una varianza más reducida. Sin embargo, a menos que  $f(\omega)$  sea realmente plano, el sesgo se incrementa junto con el ancho de banda de la ventana, por lo que el problema será escoger entre menor varianza y mayor sesgo y lo contrario.

Aún más, ya que el espectro  $\hat{f}(\omega)$  se definió también para cualquier valor de  $\omega$  entre  $-\pi$  y  $\pi$ , entonces existe un espectro suavizado general y debe estar dado en términos de

$$\begin{aligned}\hat{f}(\omega) &= \int_{-\pi}^{\pi} W_n(\lambda) \hat{f}(\omega - \lambda) d\lambda, \\ &= \int_{-\pi}^{\pi} W_n(\omega - \lambda) \hat{f}(\lambda) d\lambda\end{aligned}\tag{4.3.1.2}$$

donde  $W_n(\omega)$  es la ventana espectral que satisface:

$$\begin{aligned}\int_{-\pi}^{\pi} W_n(\lambda) d\lambda &= 1, \\ W_n(\lambda) &= W_n(-\lambda) \\ \lim_{n \rightarrow \infty} \frac{1}{n} \int_{-\pi}^{\pi} W_n^2(\lambda) d\lambda &= 0.\end{aligned}$$

La ventana espectral se conoce como Kernel.

Nuevamente, si el espectro es más o menos constante dentro de la banda de la ventana espectral, la condición 4.3.1.2 implica que

$$E(\hat{f}_w(\omega)) = f(\omega).\tag{4.3.1.3}$$

Para la varianza, como se vio en la sección 1.2.4, es factible aproximar la igualdad 4.3.1.2 mediante una suma

$$\hat{f}(\omega) \approx \frac{2\pi}{n} \sum_{k=-[n/2]}^{[n/2]} W_n(\omega - \omega_k) \hat{f}(\omega_k)$$

donde  $\omega_k = 2\pi k/n$  Así

$$\begin{aligned} \text{Var}(\hat{f}_W(\omega)) &= \left(\frac{2\pi}{n}\right)^2 f^2(\omega) \sum_{k=-[n/2]}^{[n/2]} W_n^2(\omega - \omega_k) \\ &= \frac{2\pi}{n} f^2(\omega) \sum_{k=-[n/2]}^{[n/2]} W_n^2(\omega - \omega_k) \frac{2\pi}{n} \\ &= \frac{2\pi}{n} f^2(\omega) \int_{-\pi}^{\pi} W_n^2(\lambda) d\lambda. \end{aligned} \quad 4.3.1.2$$

Aquí se aprecia que  $\text{Var}(\hat{f}_W(\omega)) \rightarrow 0$  cuando  $n \rightarrow \infty$ , y por lo tanto,  $\hat{f}_W(\omega)$  es un estimador consistente de  $f(\omega)$ .

### 4.3.2 La ventana de retardo (suavizando el dominio del tiempo)

El espectro  $f(\omega)$  es la transformada de Fourier de la función de autocovarianzas  $\gamma_k$ . Entonces, una alternativa para suavizar el espectro es aplicar una función ponderadora  $W(k)$  a las autocovarianzas muestrales, es decir,

$$\hat{f}_W(\omega) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} W(k) \gamma_k e^{-i\omega k}.$$

Como la función de autocovarianzas muestrales  $\gamma_k$  es simétrica y pierde confiabilidad cuando  $k$  aumenta, la función  $W(k)$  debe ser elegida simétrica, y los pesos de ésta deben ser inversamente proporcionales a la magnitud de  $k$ . Entonces

$$\hat{f}_W(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) \gamma_k e^{-i\omega k},$$

donde  $W_n(k)$  es la sucesión absolutamente sumable la cual se deriva de una función continua  $W(x)$  acotada, par y que satisface

$$\begin{aligned} |W(x)| &\leq 1, \\ W(0) &= 1 \\ W(x) &= W(-x) \\ W(x) &= 0, \quad |x| > 1 \end{aligned}$$

El valor de  $M$  representa el punto donde la función se trunca, este punto depende del tamaño de la muestra  $n$ .  $W_n(k)$  se denomina generalmente la ventana de retardo.

La ventana de retardo y la espectral están muy relacionadas, como consecuencia del hecho de que la autocovarianza es la transformada inversa de Fourier del espectro.

La transformada inversa de Fourier de  $f(\lambda)$  es:

$$\hat{\gamma}_k = \int_{-\pi}^{\pi} \hat{f}(\lambda) e^{i\lambda k} d\lambda \quad \text{para } k=0, \pm 1, \pm 2, \dots, \pm(n-1)$$

Entonces,

$$\begin{aligned} \hat{f}_W(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) \hat{\gamma}_k e^{i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) \int_{-\pi}^{\pi} \hat{f}(\lambda) e^{i\lambda k} e^{-i\omega k} d\lambda \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) \hat{f}(\lambda) e^{-i(\omega-\lambda)k} d\lambda \end{aligned}$$

donde

$$W_n(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) e^{-i\omega k} \quad 4.3.2.1$$

es la ventana espectral. Aquí es claro que, de la ecuación 4.3.2.1, la ventana espectral es la transformada de Fourier de la ventana de retraso y, por lo tanto, la ventana de retraso es la transformada inversa de Fourier de la ventana espectral,

$$\begin{aligned} W_n(k) &= \int_{-\pi}^{\pi} W_n(\omega) e^{i\omega k} d\omega, & k=0, \pm 1, \pm 2, \dots, \pm M. \\ &= \int_{-\pi}^{\pi} W_n(\omega - \lambda) \hat{f}(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} W_n(\lambda) \hat{f}(\omega - \lambda) d\lambda \end{aligned}$$

Finalmente, la ventana espectral y de retraso forman una pareja de transformadas de Fourier, con una determinada unívocamente por la otra.

### 4.3.3 Ventanas más utilizadas

-La ventana rectangular

Se define la ventana rectangular o de retraso truncado como

$$W_n^R(k) = \begin{cases} 1, & |k| \leq M, \\ 0, & |k| > M, \end{cases}$$

donde  $M$ , menor que 1, es el punto donde ocurre el truncamiento de la ventana. Esta ventana se deriva de la siguiente función continua

$$W(x) = \begin{cases} 1, & |x| \leq 1, \\ 0, & |x| > 1. \end{cases}$$

De la ecuación (4.3.2.1), se deduce que la ventana espectral correspondiente es

$$\begin{aligned} W_n^R(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^M W_n^R(k) e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^M e^{-i\omega k} \\ &= \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^M \cos \omega k \right) \\ &= \frac{1}{2\pi} \left[ 1 + 2 \cos(\omega(M+1)/2) \operatorname{sen}(\omega) \right] \\ &= \frac{1}{2\pi} \frac{\operatorname{sen}(\omega/2) + [\operatorname{sen}(\omega(M+1/2)) - \operatorname{sen}(\omega/2)]}{\operatorname{sen}(\omega/2)} \\ &= \frac{1}{2\pi} \frac{\operatorname{sen}(\omega(M+1/2))}{\operatorname{sen}(\omega/2)}. \end{aligned} \tag{4.3.3.1}$$

En el desarrollo anterior se utilizaron las herramientas que se explicaron en la sección 1.2.1 y 1.2.2.

Las gráficas de éstas ventanas se presentan a continuación

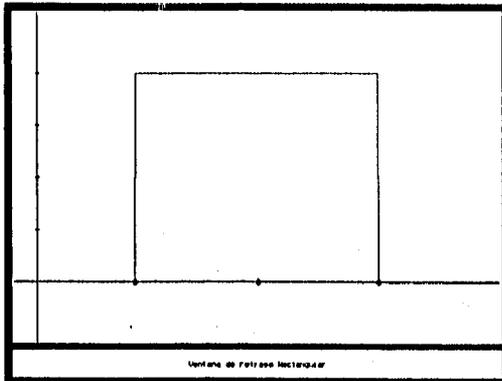


Fig. 1

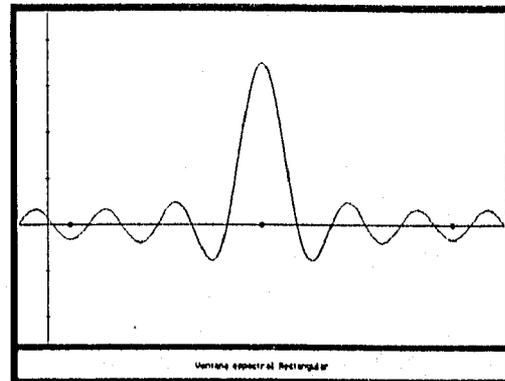


Fig. 2

En la ventana espectral anterior se encuentra un pico principal en  $\omega=0$ , con altura  $(2M+1)/2\pi$ , ceros en  $\pm 2j\pi/(2M+1)$  y picos menores de magnitud decreciente en  $\omega=\pm(4j+1)\pi/(2M+1)$  para  $j=1,2,\dots$ . Como resultado, el estimador suavizado del espectro que usa esta ventana puede arrojar resultados negativos para algunas frecuencias, y ya que el espectro es una función positiva, este resultado es indeseable.

Se define el ancho de banda de una ventana espectral  $W_n(\omega_j)$  que tiene su máximo en  $\omega=0$  como la distancia entre los puntos cuyas coordenadas están a la mitad de la distancia entre el origen y el máximo valor (ver Fig), Esto es,

$$\text{Ancho de banda} = 2\omega_p$$

donde  $\omega_p$  cumple con la condición

$$W_n(\pm\omega_p) = \frac{1}{2} W_n(0).$$

En general, el ancho de banda es inversamente proporcional al punto de truncamiento  $M$  que se usa en la ventana de retraso. Por lo tanto, cuando  $M$  crece, el ancho de banda disminuye, y consecuentemente la varianza del espectro suavizado aumenta. Por otro lado, cuando  $M$  decrece, el ancho de banda crece, y la varianza disminuye mientras que el sesgo es mayor.

-La ventana de Bartlett

La ventana que Bartlett propuso en 1950

$$W_n^R(k) = \begin{cases} 1 - |k|/M, & |k| \leq M, \\ 0, & |k| > M, \end{cases}$$

está basada en la función triangular

$$W_n^R(k) = \begin{cases} 1 - |x|, & |x| \leq 1, \\ 0 & |x| > 1. \end{cases}$$

Esta ventana también se conoce como ventana triangular por razones obvias.

La ventana espectral correspondiente es:

$$\begin{aligned} W_n^R(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^M W_n^R(k) e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^M \left(1 - \frac{|k|}{M}\right) e^{-i\omega k} \\ &= \frac{1}{2\pi M} \sum_{k=1}^M (M - |k|) e^{-i\omega k} \\ &= \frac{1}{2\pi M} \sum_{j=0}^{M-1} \sum_{k=j}^j e^{-i\omega k} \\ &= \frac{1}{2\pi M} \sum_{j=0}^{M-1} \frac{\text{sen}(\omega(j+1/2))}{\text{sen}(\omega/2)} \\ &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) + \sum_{j=1}^{M-1} \text{sen}(\omega(j+1/2)) \right\} \\ &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) \right. \\ &\quad \left. + \sum_{j=1}^{M-1} [\text{sen}(\omega j) \cos(\omega/2) + \text{sen}(\omega/2) \cos(\omega j)] \right\} \\ &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) + \cos(\omega/2) \sum_{j=1}^{M-1} \text{sen}(\omega j) \right. \\ &\quad \left. + \text{sen}(\omega/2) \sum_{j=1}^{M-1} \cos(\omega j) \right\} \\ &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) \frac{\cos(\omega/2) \text{sen}(\omega M/2) \text{sen}(\omega(M-1)/2)}{\text{sen}(\omega/2)} \right. \\ &\quad \left. + \frac{\text{sen}(\omega/2) \cos(\omega M/2) \text{sen}(\omega(M-1)/2)}{\text{sen}(\omega/2)} \right\} \end{aligned}$$

$$\begin{aligned}
 W_n^R(\omega) &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) + \frac{\text{sen}(\omega(M-1)/2)}{\text{sen}(\omega/2)} [\cos(\omega/2) \text{sen}(\omega M/2) \right. \\
 &\quad \left. \text{sen}(\omega/2) \cos(\omega M/2)] \right\} \\
 &= \frac{1}{2\pi M \text{sen}(\omega/2)} \left\{ \text{sen}(\omega/2) + \frac{\text{sen}(\omega(M-1)/2)}{\text{sen}(\omega/2)} \text{sen}(\omega(M+1)/2) \right\} \\
 &= \frac{1}{2\pi M [\text{sen}(\omega/2)]^2} \left\{ [\text{sen}(\omega/2)]^2 + \text{sen}(\omega(M-1)/2) \text{sen}(\omega(M+1)/2) \right\} \\
 &= \frac{1}{2\pi M [\text{sen}(\omega/2)]^2} \left\{ \frac{1}{2}(1 - \cos\omega) + \frac{1}{2}(\cos\omega - \cos\omega M) \right\} \\
 &= \frac{1}{2\pi M [\text{sen}(\omega/2)]^2} \left\{ \frac{1}{2}(2 - \cos\omega M) \right\} \\
 &= \frac{1}{2\pi M} \left[ \frac{\text{sen}(\omega M/2)^2}{\text{sen}(\omega/2)} \right]
 \end{aligned}$$

4.3.3.2

La ventana de retraso triangular de Bartlett y su ventana espectral correspondiente se presentan en las figuras 3 y 4.

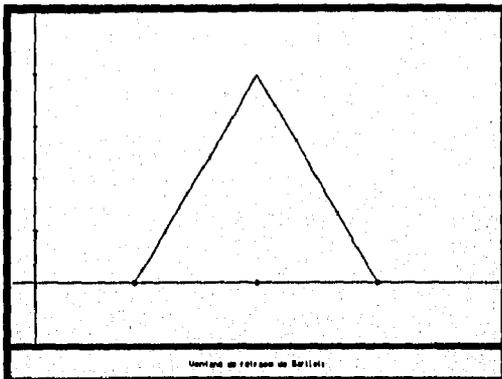


Fig. 3

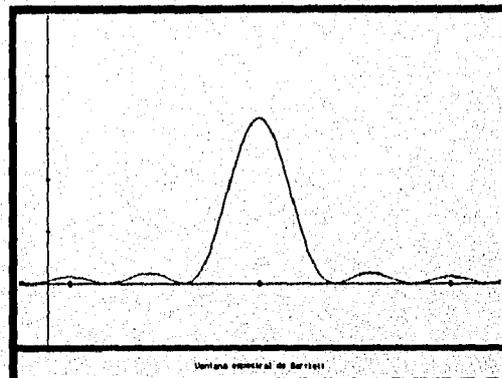


Fig. 4

Ya que la ventana espectral  $W_n^B(\omega)$  es no negativa, el estimador del espectro de Bartlett es no negativo. Más aún, si se compara 4.3.3.1 con 4.3.3.2, se observa que los picos laterales de la ventana de Bartlett son más pequeños que los producidos por la ventana triangular.

Por otro lado, se puede observar que si se utilizan ventanas con picos laterales grandes, el espectro suavizado resultante  $\hat{f}_W(\lambda)$  hará mayores contribuciones en frecuencias diferentes de  $\omega$ . Entonces, el estimador del espectro  $\hat{f}_W(\lambda)$  puede presentar el fenómeno de filtración, es decir, puede reflejar componentes espectrales significativos

en frecuencias diferentes de  $\omega$ .

-La ventana de Blackman Tuckey.

Blackman y Tuckey sugirieron la siguiente ventana de atraso:

$$W_n^T(k) = \begin{cases} 1-2a+2a\cos(\pi k/M), & |k| \leq M, \\ 0, & |k| > M, \end{cases}$$

basada en la función ponderadora continua

$$W_n^T(k) = \begin{cases} 1-2a+2a\cos(\pi k), & |x| \leq 1, \\ 0, & |x| > 1, \end{cases}$$

De nuevo,  $M$  es el punto de truncamiento de la función muestral de autocovarianzas, y la constante  $a$  se encuentra entre 0 y 0.25 para asegurar que  $W_n^T(k) \geq 0$  para toda  $k$ . La ventana espectral correspondiente se obtiene de la siguiente manera:

$$\begin{aligned} W_n^T &= \frac{1}{2\pi} \sum_{k=-M}^M W_n^T(k) e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^M [1-2a+2a\cos(\pi k/M)] e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^M [1-2a+a(e^{i\pi k/M} + e^{-i\pi k/M})] e^{-i\omega k} \\ &= a \frac{1}{2\pi} \sum_{k=-M}^M e^{-i(\omega-\pi/M)k} + (1-2a) \frac{1}{2\pi} \sum_{k=-M}^M e^{-i\omega k} \\ &\quad + a \frac{1}{2\pi} \sum_{k=-M}^M e^{-i(\omega+\pi/M)k} \\ &= \frac{a \operatorname{sen}[(\omega-\pi/M)(M+1/2)]}{2\pi \operatorname{sen}[(\omega-\pi/M)/2]} + \frac{(1-2a) \operatorname{sen}[\omega(M+1/2)]}{2\pi \operatorname{sen}(\omega/2)} \\ &\quad + \frac{a \operatorname{sen}[(\omega+\pi/M)(M+1/2)]}{2\pi \operatorname{sen}[(\omega+\pi/M)/2]} \end{aligned}$$

Como puede verse, la ventana espectral de Blackman-Tuckey es una combinación lineal de ventanas espectrales producidas por la función de retraso rectangular en las frecuencias  $(\omega-\pi/M)$ ,  $\omega$  y  $(\omega+\pi/M)$  es decir,

$$W_n^T(\omega) = a W_n^R(\omega-\pi/M) + (1-2a) W_n^R(\omega) + a W_n^R(\omega+\pi/M)$$

donde  $W_n^R(k)$  está dado en 4.3.3.1. Como consecuencia, el estimador del espectro de Blackman-Tuckey puede tomar valores negativos al ser evaluado en algunas frecuencias  $\omega$ .

Cuando la constante  $a$  de la ventana anterior toma el valor de 0.25, se le conoce como ventana de Hamming, o de Tuckey-Hamming.

A continuación se muestran las ventanas espectral y de retraso de Blackman-Tuckey con  $a=0.23$  para dar una idea del comportamiento de las mismas

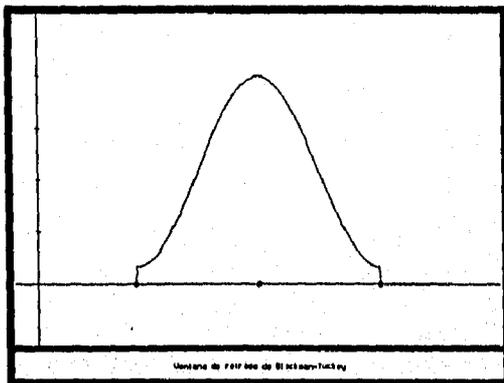


Fig. 5

-Ventana de Parzen

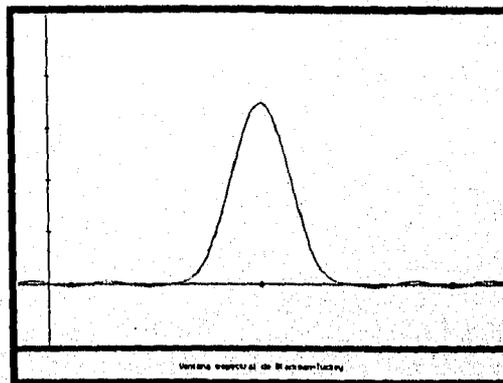


Fig. 6

Parzen propuso la siguiente ventana de retraso:

$$W_n^P(k) = \begin{cases} 1 - 6(k-M)^2 + 6(|k|/M)^3, & |k| \leq M/2, \\ 2(1 - |k|/M)^3, & M/2 < |k| \leq M \\ 0, & |k| > M, \end{cases}$$

la cual parte de la siguiente función ponderadora continua

$$W_n^P(k) = \begin{cases} 1 - 6x^2 + 6|x|^3, & |x| \leq 1/2, \\ 2(1 - |x|)^3, & 1/2 < |x| \leq 1 \\ 0, & |x| > 1. \end{cases}$$

La ventana espectral correspondiente para un valor de  $M$  par está dada por

$$\begin{aligned}
 W_n^P(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^M W_n^P(K) \cos(\omega k) \\
 &= \frac{1}{2\pi} \left\{ \sum_{k=-M/2}^{M/2} [1 - 6(k-M)^2 + 6(|k|/M)^3] \cos \omega k \right. \\
 &\quad \left. + 2 \sum_{M/2 < |k| \leq M} (1 - |k|/M) 6^3 \cos \omega k \right\} \\
 &= \frac{3}{8\pi M^3} \left[ \frac{\text{sen}(\omega M/4)}{1/2 \text{sen}(\omega/2)} \right]^4 \left\{ 1 - 2/3 [\text{sen}(\omega/2)]^2 \right\},
 \end{aligned}$$

4.3.3.4

cuando **M** es grande, la ecuación 4.3.3.4 se aproxima a

$$W_n^P(\omega) \approx \frac{3}{8\pi M^3} \left[ \frac{\text{sen}(\omega M/4)}{1/2 \text{sen}(\omega/2)} \right]^4$$

Las ventanas de Parzen se presentan a continuación

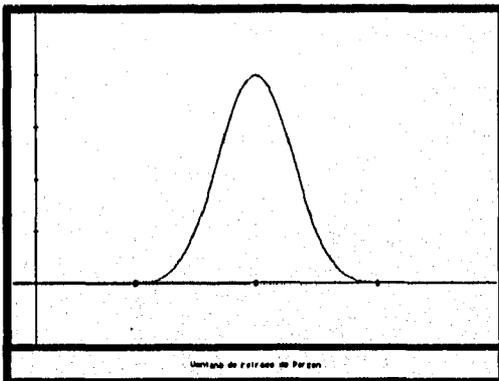


Fig. 7

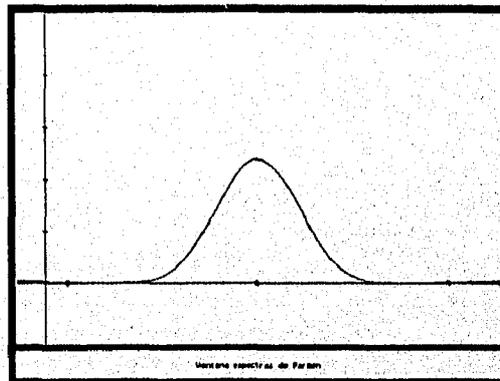


Fig. 8

La calidad de un espectro suavizado está determinado por la forma de la ventana espectral (es decir, por la forma de la función ponderadora) y por el ancho de banda (punto de truncamiento) de la misma. En el mismo sentido, el estimador del espectro obtenido por la misma ventana pero con diferentes anchos de banda varía. Así, en el suavizado del espectro se debe tener en cuenta no sólo la figura de las ventanas espectrales, sino que también se debe considerar el ancho de banda de la ventana.

Este último aspecto es a menudo un problema más importante y difícil de resolver ya que no existe un criterio único para escoger el ancho de banda óptimo. Además, anchos de banda mayores producen espectros más suaves, es decir, con varianza menor; por otro lado, si el ancho de banda es menor, en el espectro estimado existe un sesgo menor y consecuentemente, tiene una mejor resolución.

Para lograr un ancho de banda que realice una conjugación balanceada de los dos aspectos anteriores, se sugiere seguir estos pasos.

Primero escoger una ventana espectral con una forma deseable, después, calcular el estimador del espectro utilizando un ancho de banda grande y, entonces, recalculer el estimador usando anchos de banda cada vez más pequeños hasta que se alcancen una resolución y estabilidad aceptables.

Debido a que el ancho de banda de una ventana espectral está inversamente relacionada con el punto de truncamiento  $M$  usado en la ventana de retraso, se sugiere, como camino alternativo al anterior, 3 valores de  $M$ ,  $M_1, M_2, M_3$ ; con  $M_3 = 4M_1$ . Se calculan los espectros para estos valores y se utilizan para determinar si el valor de  $M$  que se está buscando queda dentro, antes o después del intervalo de  $M_1$  a  $M_3$ .

El ancho de banda espectral también puede determinarse escogiendo un valor de  $M$  tal que  $\hat{\gamma}(k)$  sea insignificante cuando  $k > M$ . Este criterio de elección de  $M$  ha resultado ser adecuado en muchos casos prácticos.

Alternativamente, es posible elegir el punto de truncamiento en función del tamaño  $n$  de la muestra. Para  $n$  lo suficientemente grande,  $M = \lfloor n/10 \rfloor$ .

En principio, si se desea ser capaz de distinguir varios picos de  $f(\omega)$  en frecuencias  $\lambda_1, \lambda_2, \dots$ , el ancho de banda de la ventana espectral no debe exceder el intervalo mínimo entre esos picos adyacentes. De otra forma, el espectro suavizado no los respetará.

#### 4.3.4 Intervalos de confianza de las ordenadas espectrales

En la sección 4.2 se mencionó que para una muestra  $Z_1, Z_2, \dots, Z_n$  de un proceso cuyo espectro es  $f(\omega)$ , las ordenadas espectrales  $\hat{f}(\omega)$  en las frecuencias de Fourier  $\omega_k = 2\pi k/n$  con  $\omega_k \neq 0$  y  $\pi$  están independiente e idénticamente distribuidas como

$$\hat{f}(\omega) \sim f(\omega_k) \frac{\chi^2(2)}{2}$$

Entonces, si se suaviza el espectro muestral con el término de medias móviles simple  $(2m+1)$

$$\hat{f}_W(\omega_k) = \frac{1}{2m+1} \sum_{j=-m}^m \hat{f}(\omega_k - \omega_j),$$

la ordenada del espectro suavizado se distribuirá como

$$\hat{f}_W(\omega_k) \sim f(\omega_k) \frac{\chi^2(DF)}{DF},$$

donde los grados de libertad  $DF=(4m+2)$  son la suma de los grados de libertad de las variables aleatorias. Sin embargo esta propiedad no se cumple si  $\omega$  no es una frecuencia de Fourier o si el espectro no está suavizado por una ventana espectral uniforme. Por lo tanto, sólo se puede aproximar a un espectro suavizado por medio de

$$\hat{f}_W(\omega) = c\chi^2(u)$$

en donde  $c$  y  $u$  cumplen:

$$E[\hat{f}_W(\omega)] = E(c\chi^2(u)) = cu$$

y

$$Var[\hat{f}_W(\omega)] = Var(c\chi^2(u)) = 2c^2u.$$

Ahora, para una ventana espectral dada  $W_n(\omega)$  de 4.3.1.3 y 4.3.1.4 se obtiene

$$E[\hat{f}_W(\omega)] = f(\omega)$$

y

$$Var[\hat{f}_W(\omega)] = \frac{2\pi}{n} [f(\omega)]^2 \int_{-\pi}^{\pi} W_n^2(\omega) d\omega.$$

Entonces

$$cu = f(\omega),$$

$$2c^2u = \frac{2\pi}{n} f^2(\omega) \int_{-\pi}^{\pi} W_n^2(\omega) d\omega,$$

de lo que se deriva

$$u = \frac{n}{\pi \int_{-\pi}^{\pi} W_n^2(\omega) d\omega},$$

$$c = \frac{f(\omega)}{u}$$

4.3.4.1

En consecuencia

$$\hat{f}_W(\omega_k) \sim f(\omega_k) \frac{\chi^2(u)}{u}$$

para la que  $u$  representa el equivalente a los grados de libertad del espectro suavizado.

De lo anterior se concluye que es posible construir un intervalo de confianza para el espectro con un nivel  $\alpha$  de significación utilizando

$$P \left\{ f(\omega) \frac{\chi^2_{1-\alpha/2}(u)}{u} \leq \hat{f}_W(\omega) \leq \frac{\chi^2_{\alpha/2}(u)}{u} \right\} = 1 - \alpha$$

en la que  $\chi^2(u)$  es el extremo superior de la distribución ji-cuadrada donde se encuentra el  $\alpha\%$  del área bajo la curva. Entonces el intervalo de confianza de  $f(\omega)$  con un nivel  $\alpha$  de significación está dado por

$$\frac{u \hat{f}_W(\omega)}{\chi^2_{\alpha/2}(u)} \leq f(\omega) \leq \frac{v \hat{f}_W(\omega)}{\chi^2_{1-\alpha/2}(u)},$$

donde  $u$  es el equivalente de los grados de libertad calculados por 4.3.4.1.

## 4.4 Estimación espectral ARMA

Basándose en una serie de tiempo dada  $Z_1, Z_2, \dots, Z_n$ , es posible aproximar el modelo que describe la serie mediante un modelo **AR(p)**,

$$(1 - \phi_1 B - \dots - \phi_p B^p) Z_t = a_t$$

para algún  $p$ . Sean  $\hat{\phi}_1, \dots, \hat{\phi}_p$  y  $\hat{\sigma}_a^2$  los estimadores de  $\phi_1, \dots, \phi_p$  y  $\sigma_a^2$ . Una alternativa razonable para estimar el espectro es sustituir esos parámetros estimados en la expresión teórica del espectro del modelo **AR(p)** que se desarrolló en el capítulo anterior, es decir,

$$\hat{f}_A(\omega) = \frac{\hat{\sigma}_a^2}{2\pi \hat{\phi}_p(e^{-i\omega}) \hat{\phi}_p(e^{i\omega})}$$

donde  $\hat{\phi}_p(e^{-i\omega}) = (1 - \hat{\phi}_1 e^{-i\omega} - \dots - \hat{\phi}_p e^{-ip\omega})$ . Este método se conoce como estimación espectral autorregresiva.

Para una  $n$  lo suficientemente grande Parzen demostró que

$$\text{Var}[\hat{f}_A(\omega)] \approx \frac{2pf^2(\omega)}{n}$$

Entonces, para controlar la varianza, el orden de  $p$  escogido en la estimación del proceso no debe ser muy grande, pero tampoco puede ser muy pequeño, ya que esto provocaría una aproximación muy pobre al modelo y, al mismo tiempo, incrementaría el sesgo del espectro estimado.

Para escoger  $p$ , existen muchos criterios de selección, pero se recomienda utilizar el criterio **CAT** de Parzen.

Más generalmente, se puede aproximar en modelo **ARMA(p,q)** a un proceso desconocido

$$(1 - \phi_1 B - \dots - \phi_p B^p)(Z_t - \mu) = (1 - \theta_1 B - \dots - \theta_q B^q) a_t$$

para algunos  $p$  y  $q$ . Sean  $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q$  y  $\hat{\sigma}_a^2$  los estimadores de  $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  y  $\sigma_a^2$ .

El espectro del proceso en cuestión puede ser estimado por:

$$\hat{f}_W(\omega) = \hat{\sigma}_a^2 \frac{\hat{\theta}_q(e^{-i\omega})\hat{\theta}_q(e^{i\omega})}{\hat{\phi}_p(e^{-i\omega})\hat{\phi}_p(e^{i\omega})} \quad 4.4.1$$

donde

$$\hat{\phi}_p(e^{-i\omega}) = (1 - \hat{\phi}_1 e^{-i\omega} - \dots - \hat{\phi}_p e^{-ip\omega}),$$

y

$$\hat{\theta}_q(e^{-i\omega}) = (1 - \hat{\theta}_1 e^{-i\omega} - \dots - \hat{\theta}_q e^{-iq\omega}),$$

La ecuación 4.4.1 es conocida como la estimación espectral ARMA. Al igual que la estimación espectral **AR**, la calidad del proceso anterior está dada por la elección apropiada de  $p$  y  $q$  en la aproximación.

## Capítulo 5 Presentación del programa AST

## Introducción

Este capítulo presenta un programa cómputo llamado "Análisis de Series de Tiempo" o AST por sus iniciales, el cual resume prácticamente lo que se ha expuesto hasta el momento.

El objetivo de tal programa es permitir al lector probar, bajo diferentes escenarios, los conceptos explicados en este documento de tal forma que resulten más claros.

Por tal motivo se analizará un ejemplo utilizando los conceptos revisados y la metodología de Box y Jenkins, por lo que se sugiere al lector que no este familiarizado con dicha metodología consulte algún texto al respecto<sup>8</sup>.

Este capítulo va a hacer referencia en varias ocasiones a las opciones que contiene el programa AST, por lo que también se recomienda revisar el Anexo II, ya que en él se encuentran explicaciones detalladas del manejo del sistema.

---

<sup>8</sup>Se sugiere GONZALEZ VIDEGARAY, MARIA DEL CARMEN. *Modelos de Decisión con Procesos Estocásticos II*. UNAM 1990.

## 5.1 Ejemplo

Para explicar el funcionamiento y las características del programa así como ilustrar la aplicación del mismo en el estudio de las series de tiempo, se va a desarrollar un ejemplo basado en el Turismo Receptivo Vía Aérea en el país entre 1979 y 1987 utilizando la metodología de Box-Jenkins e introduciendo los conceptos del análisis de Fourier donde sea conveniente. Para este propósito será necesario también el programa estadístico conocido como Statgraphics V 4.2.

Es necesario recalcar que el análisis de series de tiempo mediante series de Fourier presentada en esta tesis es una metodología alterna a la de Box-Jenkins por lo que tienen puntos en común y se complementan.

### Presentación de la Serie

A través de su historia, México se ha caracterizado por ser un país de enormes riquezas culturales así como de bellos lugares; es por eso que muchos extranjeros gustan de viajar al país para disfrutar de tales cosas.

Esto ha propiciado que el turismo sea considerado como una importante fuente de ingresos para el país; y por ende, que el sector turismo sea uno de los más fuertes actualmente.

Estos hechos hacen que el pronóstico de la afluencia de turistas sea objeto de interés y se realicen enormes esfuerzos para determinar el comportamiento de este fenómeno y poder estar mejor preparados para atender esta demanda.

La tabla que se muestra en la siguiente página contiene los datos concernientes al número de turistas recibidos vía aérea mensualmente entre 1979 y 1987<sup>9</sup>.

Esta serie se encuentra almacenada en un archivo llamado TURISMO.AST y se encuentra en el disco que se proporciona con este documento.

---

<sup>9</sup> Fuente Original: Revista anual del Instituto Nacional de Estadística Geografía e Informática.

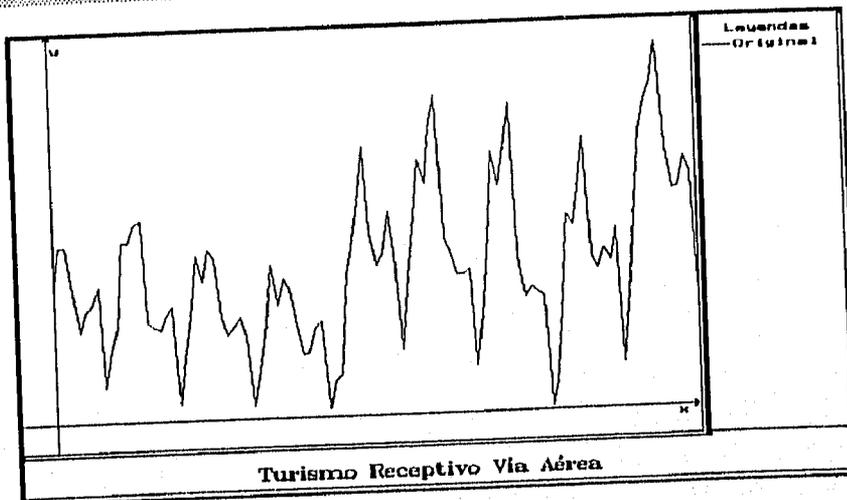
TURISMO RECEPTIVO VIA AEREA (Número de turistas en millares)									
	1979	1980	1981	1982	1983	1984	1985	1986	1987
Ene	221	249	220	202	241	286	283	253	341
Feb	247	262	242	220	274	325	310	278	354
Mar	247	265	235	213	313	349	341	315	382
Abr	215	182	194	186	251	246	231	229	305
May	186	188	181	166	228	234	202	221	277
Jun	200	186	187	167	237	220	209	235	279
Jul	205	196	193	185	266	220	206	227	299
Ago	218	202	180	189	237	223	203	249	288
Sep	146	132	130	126	167	154	123	153	198
Oct	172	165	155	146	220	190	140	202	
Nov	189	185	188	150	253	246	190	269	
Dec	249	238	230	222	303	307	260	318	

El primer paso en el análisis es determinar si la serie es estacionaria, es decir, si la media y la varianza permanecen constantes a través de la realización; así, es necesario graficar los datos y observar si éstos observan dicho comportamiento.

En este sentido, y siguiendo el ejemplo, el primer paso es graficar la los datos de turismo.

En la imagen presentada en la siguiente página se observa claramente que los datos no observan un comportamiento estacionario; se percibe una tendencia y varianza creciente además de un comportamiento estacional (anual).

En consecuencia, se deben tomar las medidas adecuadas para hacer que la serie se comporte de manera estacionaria.

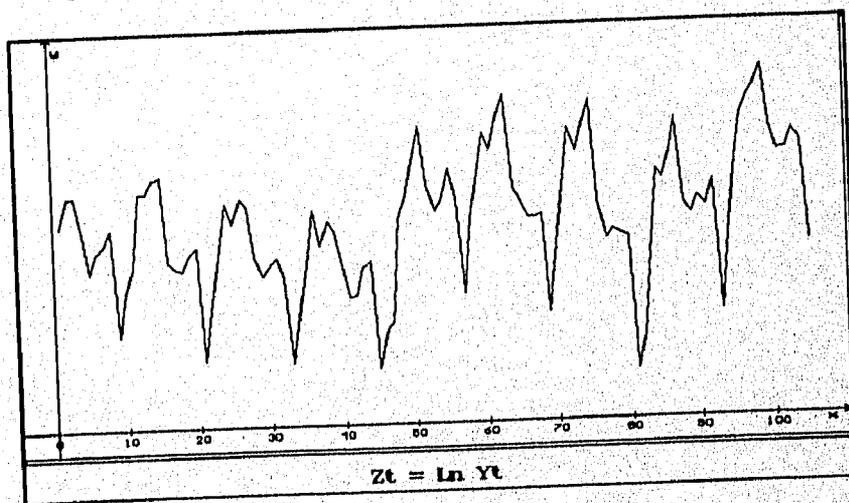


El primer paso será estabilizar la varianza mediante una transformación<sup>10</sup>.

En el ejemplo, se llegó a la conclusión de que la transformación que mejor eliminó la varianza de la serie fué **Logaritmo natural**, que se obtiene dando valores de 0 a 11 y 12. Esta transformación produce una serie que se va a denominar

$$Z_t = \text{Ln } Y_t$$

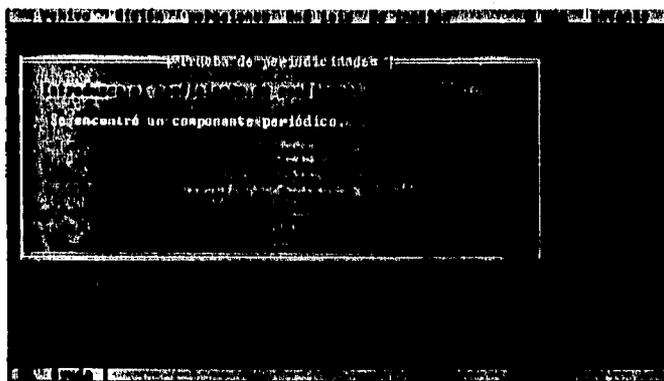
donde  $Y_t$  representa el número de turistas en millares recibidos en el mes  $t$ . A continuación se presenta la gráfica de la serie generada.



<sup>10</sup> Metodología de Box y Jenkins.

Como puede observarse en la imagen anterior, aún no se ha logrado la estacionaridad de la serie ya que se sigue observando tendencia creciente y variación estacional.

AST permite hacer una prueba de estacionalidad a la serie para determinar el período de la serie, en ese sentido, el resultado de dicha prueba es:



Esta prueba indica que la serie observa un comportamiento estacional con un período de 11.67 ~ 12 meses (anual).

Lo anterior hace proponer las siguientes series con el fin de lograr la estabilidad.

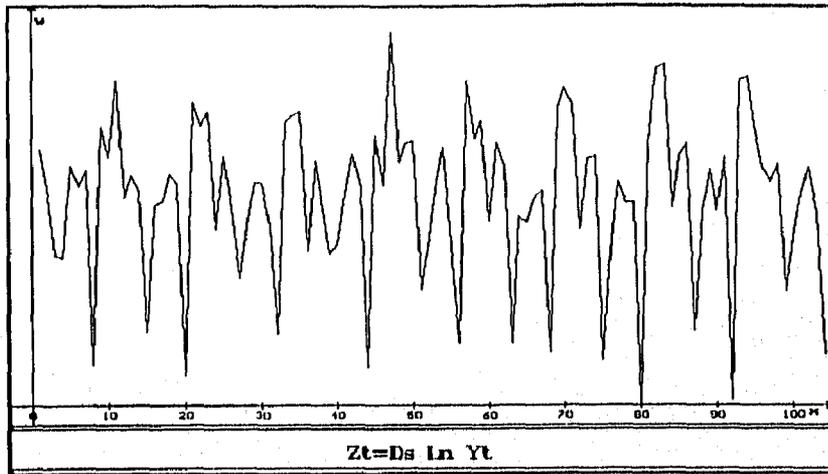
La primera serie es resultado de aplicar una diferencia simple a la serie  $Z_t = L_n Y_t$  para eliminar la tendencia. Esta serie será denominada  $Z_t = D_s L_n Y_t$ .

La segunda serie se genera a partir de una diferencia estacional con un factor estacional 12 y será conocida como  $Z_t = D_{e12} L_n Y_t$ .

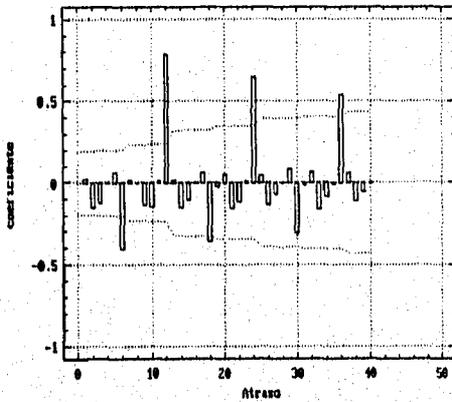
La tercera serie surge de aplicar una diferencia estacional de grado 12 y una diferencia simple a  $Z_t = L_n Y_t$  y se denominará  $Z_t = D_s D_{e12} L_n Y_t$ .

A continuación se presentan las gráficas de las series así como sus respectivas funciones de autocorrelación muestrales, autocorrelaciones parciales, espectro muestral y espectro muestral suavizado.

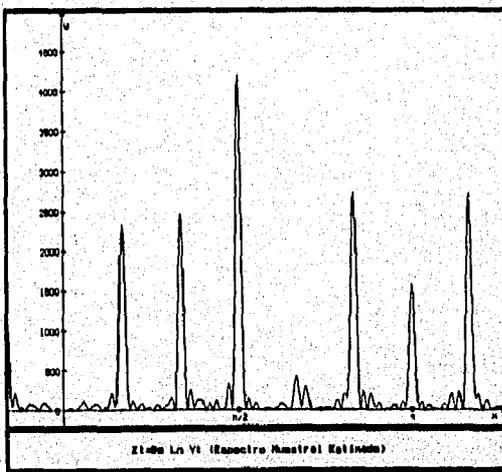
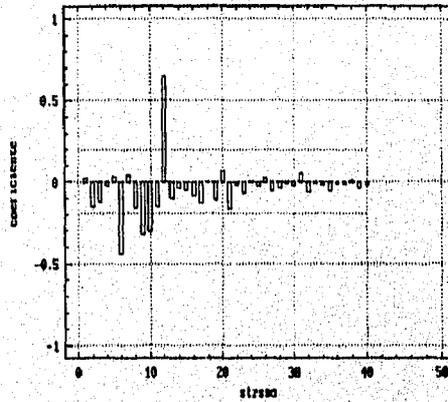
$Z_t = D_s \ln Y_t$



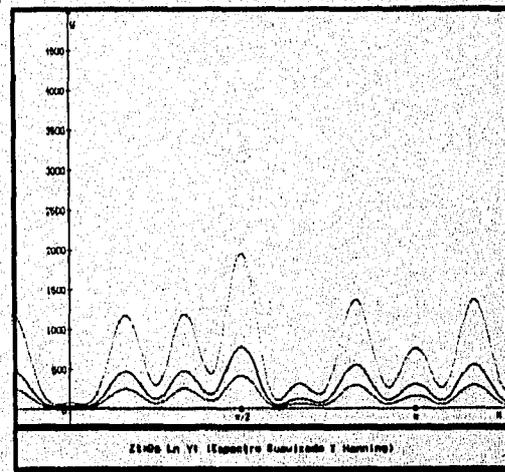
Autocorrelaciones  
De  $\ln Y_t$



Autocorrelaciones Parciales  
De  $\ln Y_t$



$Z_t = D_s \ln Y_t$  (Espectro Muestral Estimado)



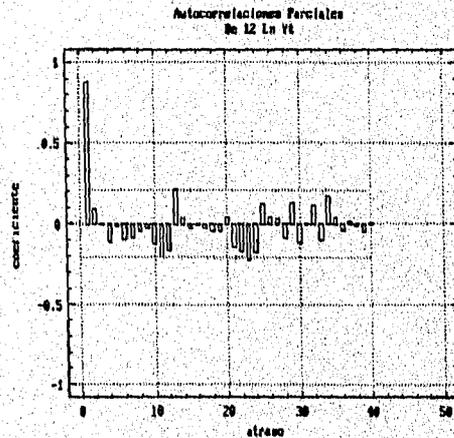
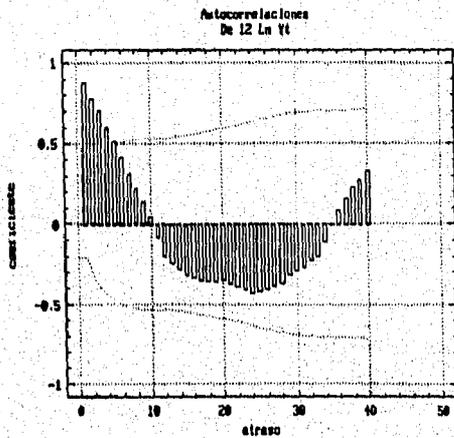
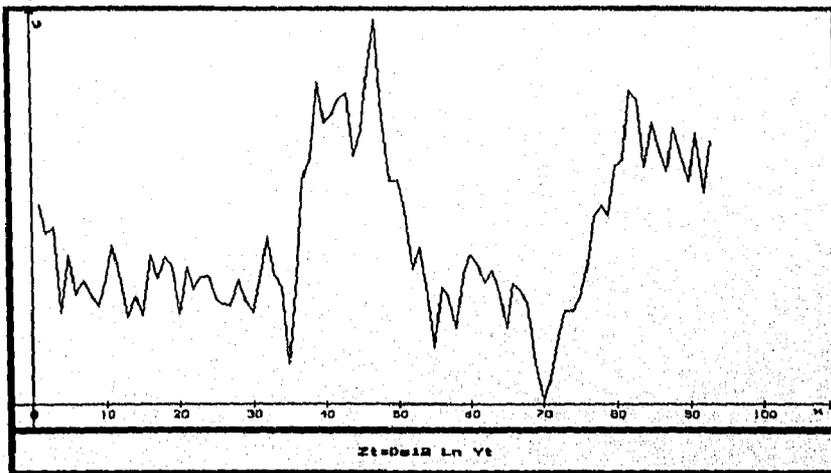
$Z_t = D_s \ln Y_t$  (Espectro Suavizado y Homino)

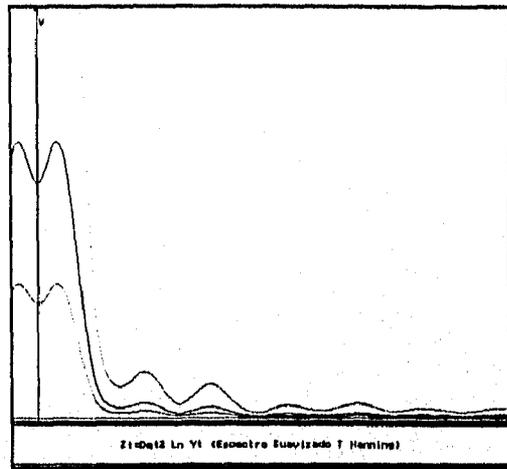
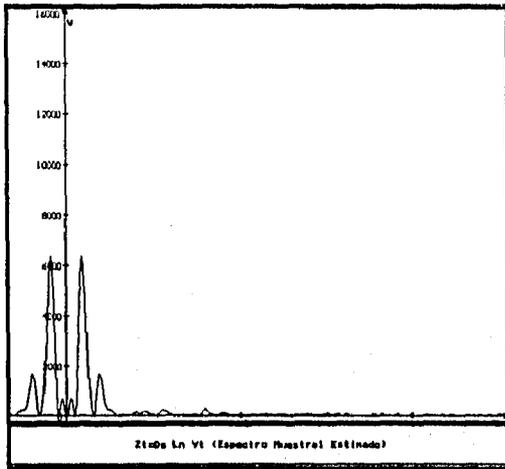
La gráfica de la serie refleja varianza y media constante, pero aún conserva estacionalidad. Por otro lado las autocorrelaciones muestrales decrecen lentamente mostrando valores significativos en 6, 12, 24 y 36 mientras que las autocorrelaciones parciales se truncan en 12 mostrando valores significativos en 6, 9 y 12.

Finalmente, el espectro suavizado observa también un comportamiento estacional del tipo AR ya que los picos de la gráfica son agudos.

Dado lo anterior se propone que la serie  $Z_t = D_{s12} \ln Y_t$  sigue un modelo SAR(1).

$Z_t = D_{s12} \ln Y_t$



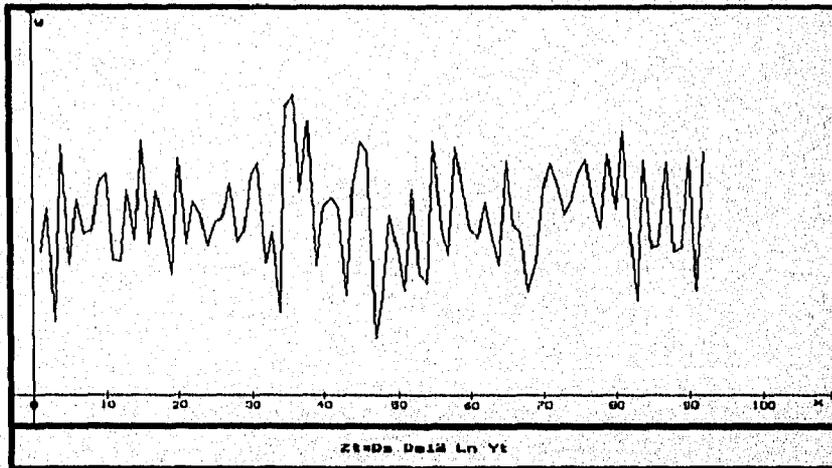


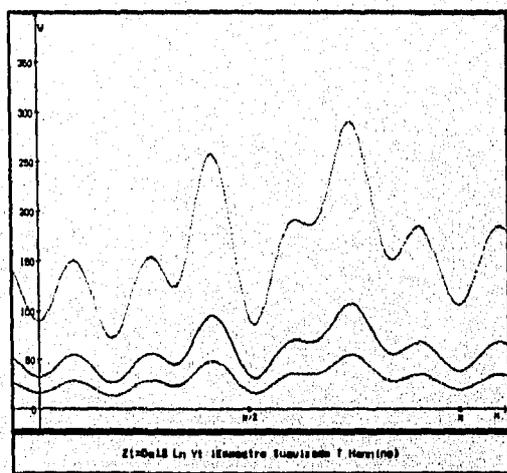
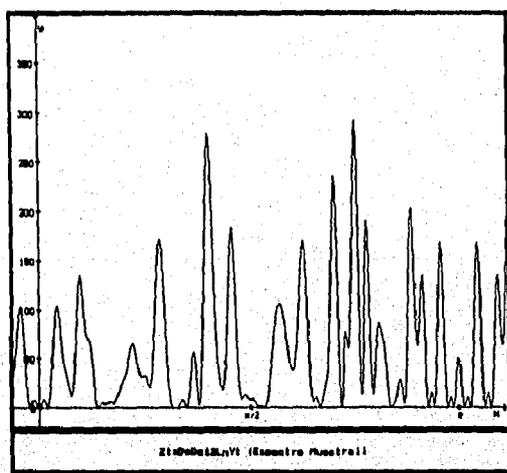
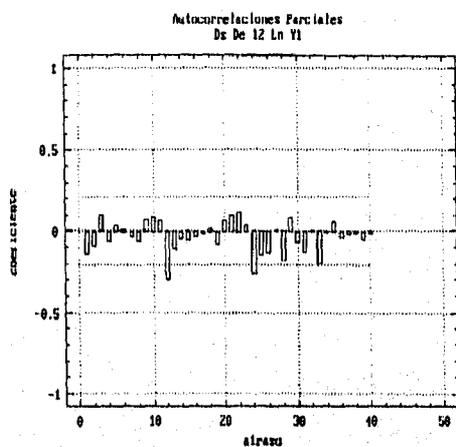
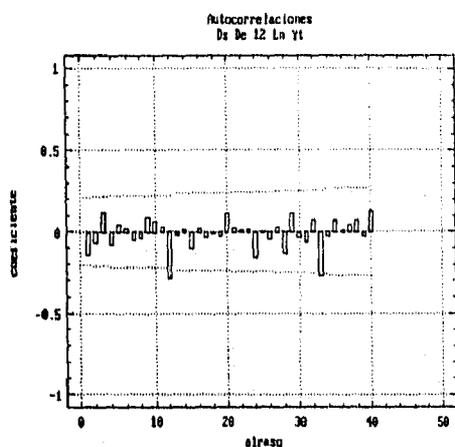
La gráfica de la serie refleja varianza y media constante y carece de fluctuación estacional. Por otro lado las autocorrelaciones decrecen muy lentamente mientras que las autocorrelaciones parciales muestra un valor significativo en 1.

Finalmente, el espectro suavizado observa un comportamiento senoidal suave decreciente cuyo valor más alto se encuentra en 0, el comportamiento decreciente señala un modelo AR, pero la tendencia senoidal suave indica que esta serie también podría seguir un modelo ARxSMA

Dado lo anterior, y por el principio de parsimonia, se propone que la serie  $Z_t = D_s \text{Ln } Y_t$  sigue un modelo AR(1) y posteriormente se puede analizar un modelo AR(1)xSMA(1) si es necesario.

$Z_t = D_s \text{Ln } Y_t$





La gráfica de la serie refleja varianza y media constante y carece de fluctuación estacional. Por otro lado las autocorrelaciones presentan un valor significativo en 12 mientras que las autocorrelaciones parciales muestran valores significativos en 12 y 24.

Finalmente, el espectro suavizado observa un comportamiento senoidal suave, lo que indica un modelo SMA.

Dado lo anterior se propone que la serie  $Z_t = DS De 12 Ln Y_t$  sigue un modelo SMA(1).

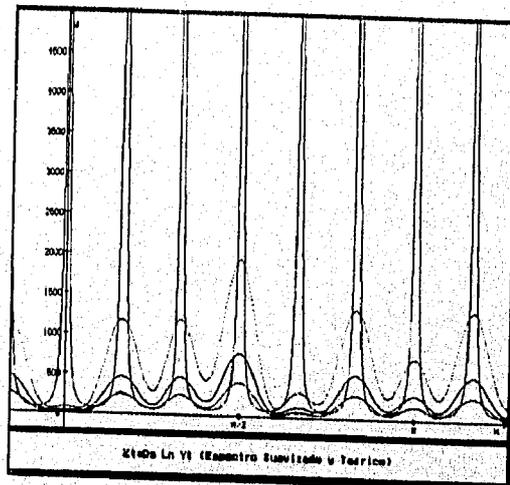
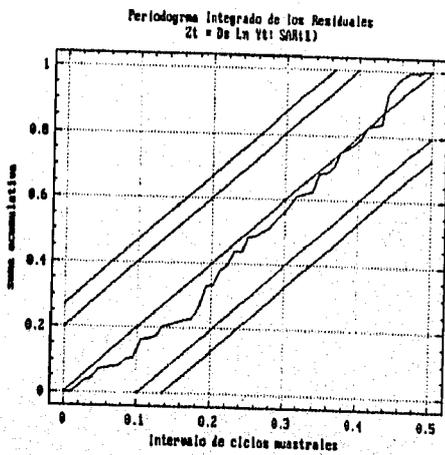
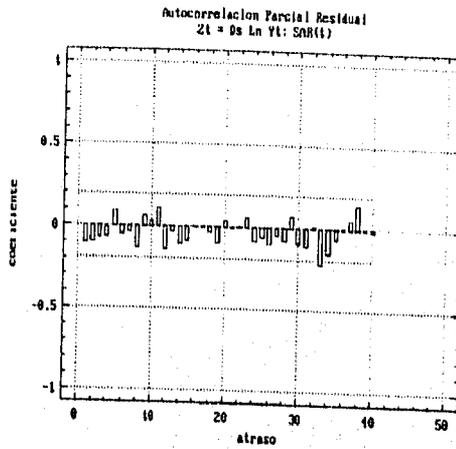
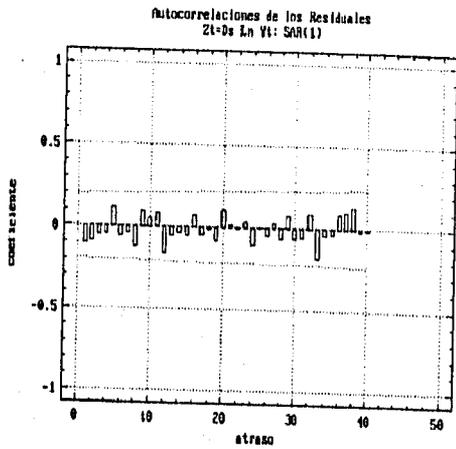
En resumen, se tienen 3 series de tiempo con los siguientes modelos propuestos:

Serie	Modelos Propuestos
Ds Ln Yt	SAR(1)
De 12 Ln Yt	AR(1) o AR(1)xSMA(1)
Ds De 12 Ln Yt	SMA(1)

A continuación se va realizar el análisis de Box y Jenkins para determinar los coeficientes de cada modelo y comparar el espectro teórico de cada serie con sus espectros.

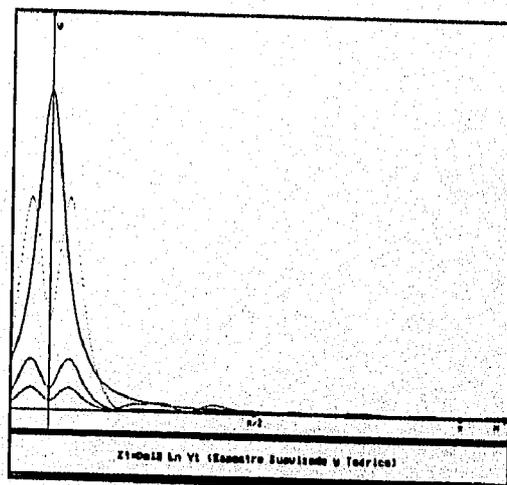
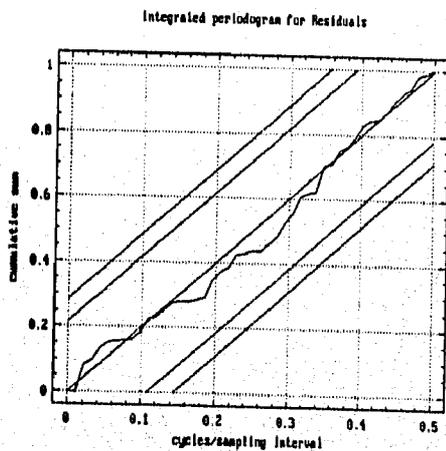
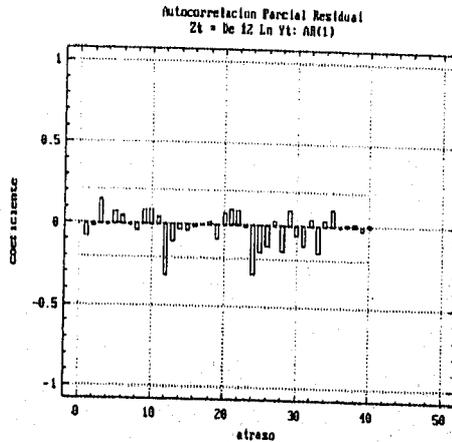
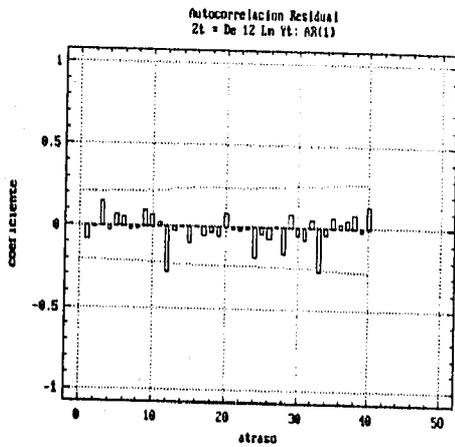
Aquí es conveniente destacar que los espectros teóricos se calcularon utilizando la varianza de la serie y los coeficientes calculados por el programa STATGRAPHICS V4 del cual también provienen las gráficas presentadas.

1.  $Z_t = D_s \ln Y_t$ : SAR(1)



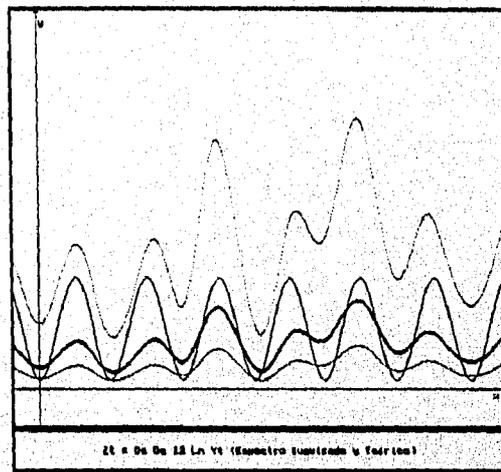
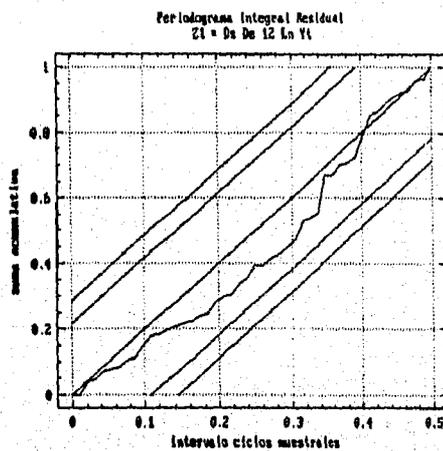
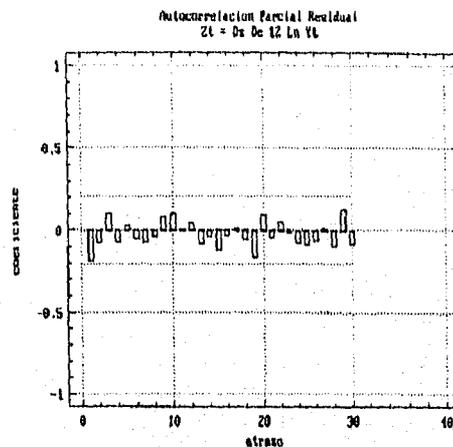
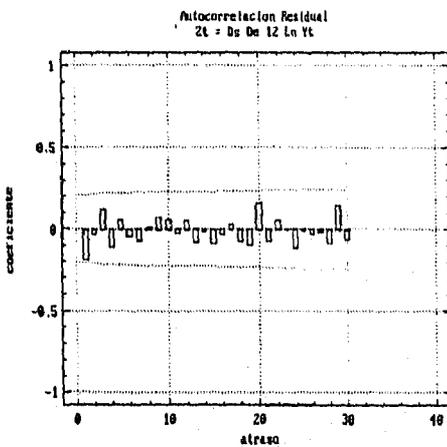
El espectro suavizado muestra ajuste aceptable. Se observa que las autocorrelaciones de los residuales sugieren ruido blanco y que el periodograma integral es bueno, por lo tanto, se puede decir que este es un buen modelo.

## 2. $Z_t = De 12 Ln Y_t: AR(1)$



En las autocorrelaciones se pueden observar valores significativos, por lo que este modelo no tiene un buen ajuste; de igual forma, el espectro suavizado difiere del espectro teórico por lo que es necesario crear otro modelo que de un mejor ajuste.

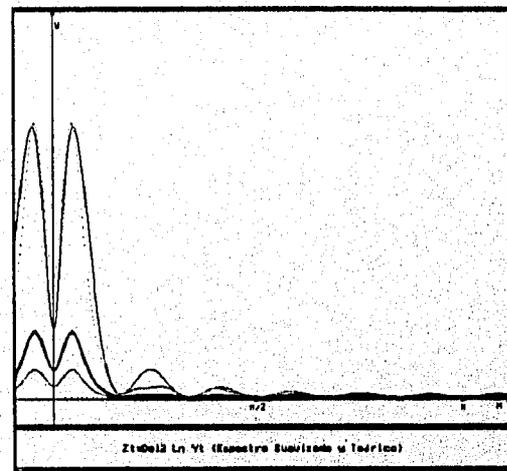
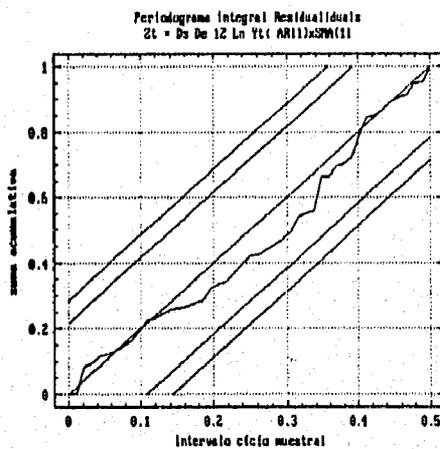
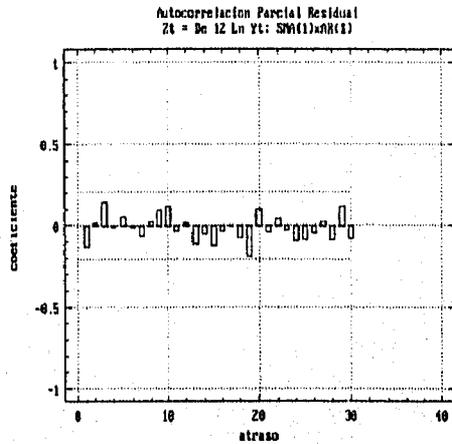
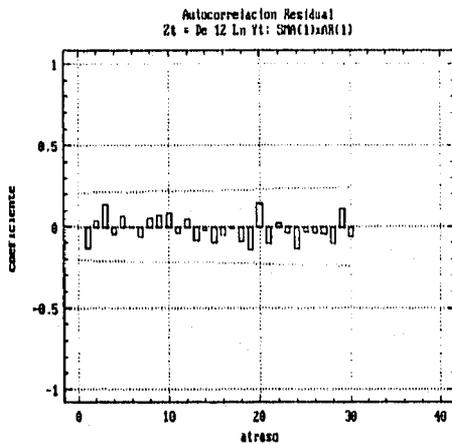
El comportamiento de las autocorrelaciones y del espectro suavizado presentado anteriormente sugieren crear otro modelo  $AR(1) \times SMA(1)$ .

3.  $Z_t = D_s D_e 12 L_n Y_t$ : SMA(1)

Las autocorrelaciones se ajustan al ruido blanco, sin embargo el periodograma integral no presenta una forma satisfactoria y el espectro suavizado indica cierta tendencia por lo que será necesario tratar con un modelo que se apegue más a la serie presentada.

En este sentido, y debido al comportamiento presentado por las autocorrelaciones y el espectro suavizado, se sugiere el modelo  $AR(1) \times SMA(1)$ .

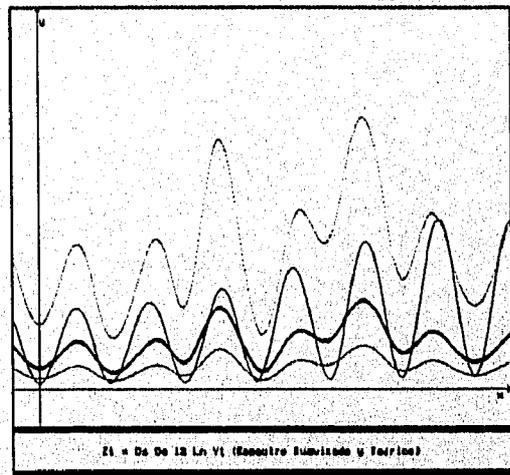
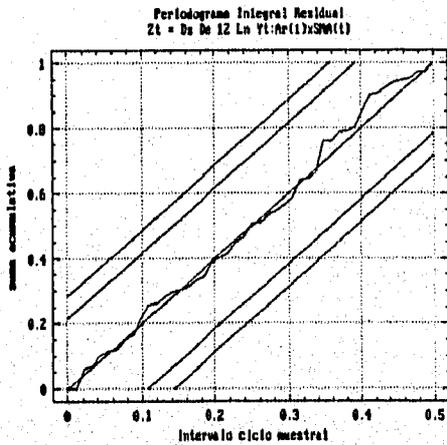
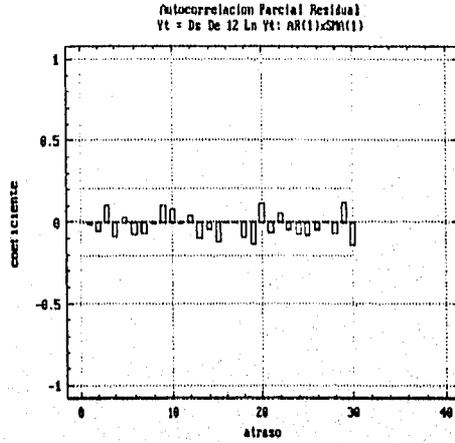
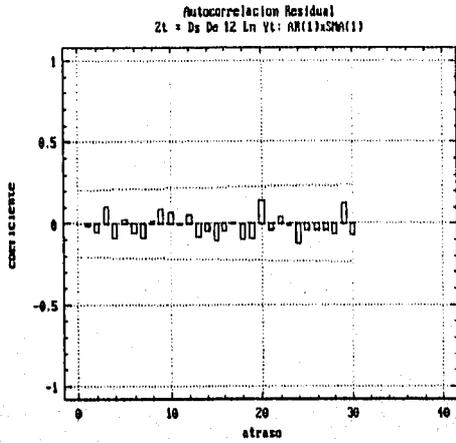
4.  $Z_t = De 12 Ln Y_t : SMA(1) \times AR(1)$



Se ve que las funciones de autocorrelación describen ruido blanco y que el periodograma integral tiene un comportamiento aceptable.

De igual manera se observa un excelente ajuste en el espectro suavizado por lo que se considera que este es un buen modelo.

5.  $Z_t = D_s D_e L_n Y_t: AR(1) \times SMA(1)$ .



Se observa que las funciones de autocorrelación se comportan como ruido blanco y que el periodograma integral y el espectro suavizado tienen un excelente ajuste, por lo que podemos afirmar que este es un modelo que describe a la serie de la mejor manera.

A continuación se presenta un cuadro comparativo donde se muestra la información correspondiente a cada modelo.

	Modelo				
	1	2	3	4	5
# Dif. Ord.	1	0	1	0	1
# Dif. Est.	0	1	1	1	1
# Parámetros	1	1	1	2	2
$\phi(1)$ (est. T)	--	0.87972 (17.32)	--	0.87477 (16.50)	-0.19068 (-1.82)
$\Phi(1)$ (est. T)	0.92751 (16.58)	--	--	--	--
$\Theta(1)$ (est. T)	--	--	0.56069 (5.82)	0.57386 (05.97)	0.57888 (6.01)
Box-Pierce	13.4118	14.2551	17.6765	12.3225	9.8521
ACF significativas	--	12	--	--	--
PACF significativas	--	12,24	--	--	--

En este cuadro y en las gráficas previas se puede ver que el modelo que mostró el mejor ajuste fue el número 5.

De esta manera se ha visto como el análisis de Fourier puede ser utilizado como un mecanismo adicional en el análisis de las series de tiempo.

## Conclusiones

En este trabajo quedó demostrado que existe una relación entre la función de autocovarianzas y el espectro de una serie. Efectivamente, el espectro de una serie es la transformada de Fourier de la función de autocovarianzas absolutamente sumable del proceso.

Esta conclusión es muy importante ya que permite establecer un vínculo entre las series de Fourier y las series de tiempo; abriendo la posibilidad de estudiar una serie de tiempo con base en el peso que cada frecuencia tiene sobre el espectro y de esta forma encontrar periodicidades (estacionalidad).

Por otro lado, se ha alcanzado el objetivo planteado al inicio. Se ha desarrollado una metodología complementaria a la de Box y Jenkins para determinar el comportamiento de una serie de tiempo y se ha elaborado un programa que permite aplicar esta metodología a un conjunto de datos.

En términos generales, esta nueva forma de analizar se basa en el cálculo del espectro de un proceso a través de las autocovarianzas de la serie; a continuación se grafica el espectro suavizado utilizando una ventana espectral y, por último, se proponen varios espectros teóricos correspondientes a modelos diferentes, los cuales se comparan con el espectro suavizado para obtener el que mejor se ajuste.

Además, usando el espectro de potencias, es posible realizar pruebas de periodicidad a los datos para encontrar componentes estacionales en la serie. Todo lo anterior utilizando pruebas de hipótesis estadísticas que garantizan la confiabilidad de los resultados.

Finalmente se ha presentado un ejemplo para ilustrar lo anterior, en él los resultados obtenidos por la metodología de Box y Jenkins han sido corroborados por los generados por el programa.

## ANEXO I (Pruebas Matemáticas)

## Sistemas Completos. Convergencia en la media

El sistema de ecuaciones definido en 1.1.4.1 se dice completo si para cualquier función integrable por su cuadrado se cumple

$$\int_a^b f^2(x) dx = \sum_{k=0}^{\infty} c_k^2 \|\varphi_k\|^2 \quad 1.1$$

en lugar de la desigualdad de Bessel. En esta ecuación los  $c_k$  ( $k=0,1,2,\dots$ ) Son los coeficientes de Fourier de la función  $f(x)$ .

**Teorema 1.** Sean  $f(x)$  y  $F(x)$  funciones integrables por su cuadrado para las cuales

$$\begin{aligned} f(x) &\sim c_0 \varphi_0(x) + c_1 \varphi_1(x) + \dots, \\ F(x) &\sim C_0 \varphi_0(x) + C_1 \varphi_1(x) + \dots, \end{aligned}$$

y sea además el sistema 1.4.1.1 completo. Entonces

$$\int_a^b f(x)F(x) dx = \sum_{k=0}^{\infty} c_k C_k \|\varphi_k\|^2.$$

**Prueba.** La suma  $f(x)+F(x)$  y la diferencia  $f(x)-F(x)$  son funciones integrables por su cuadrado. Más aún, los coeficientes de Fourier de la suma y de la resta son  $c_k+C_k$  y  $c_k-C_k$  respectivamente. Debido a la definición de una función completa se tiene

$$\begin{aligned} \int_a^b [f(x)+F(x)]^2 dx &= \sum_{k=0}^{\infty} (c_k+C_k)^2 \|\varphi_k\|^2, \\ \int_a^b [f(x)-F(x)]^2 dx &= \sum_{k=0}^{\infty} (c_k-C_k)^2 \|\varphi_k\|^2, \end{aligned}$$

y la resta de ambas ecuaciones produce

$$4 \int_a^b f(x)F(x) dx = \sum_{k=0}^{\infty} 4c_k C_k \|\varphi_k\|^2.$$

El siguiente resultado tiene importantes consecuencias:

**Teorema 2.** Una condición necesaria y suficiente para que el sistema 1.1.4.1 sea completo es que la relación

$$\lim_{n \rightarrow \infty} \int_a^b \left[ f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx = 0. \quad 1.2$$

se cumpla para cualquier función  $f(x)$  integrable por su cuadrado  $f(x)$ , donde  $c_k$  ( $k=0,1,2,\dots$ ) son los coeficientes de Fourier de  $f(x)$  con respecto al sistema 1.1.4.1.

Prueba. Utilizando la ecuación 1.1.6.6 y el hecho de que la definición de función completa es equivalente a

$$\lim_{n \rightarrow \infty} \int_a^b \left[ f(x)^2 - \sum_{k=0}^n c_k^2 \|\varphi_k(x)\|^2 \right] dx = 0.$$

Si la relación 1.2 se satisface, se dice que la serie de Fourier converge a  $f(x)$  en la media. Así el teorema 2 puede ser formulado también como:

Una condición necesaria y suficiente para que el sistema 1.1.4.1 sea completo es que la serie de Fourier de cualquier función integrable por su cuadrado  $f(x)$  converge a  $f(x)$  en la media.

Debe ser señalado que la convergencia ordinaria de una serie de Fourier a la función de la cual fue formada no siempre ocurre, aún si el sistema 1.1.4.1 es completo. Sin embargo, como se acaba de demostrar, la convergencia en la media siempre se cumple para sistemas completos (se entiende que se está hablando de funciones integrables por su cuadrado). En particular, estas consideraciones aplican al sistema trigonométrico.

Lo mostrado anteriormente muestra la importancia del concepto de convergencia en la media y sugiere que este tipo de convergencia puede ser considerada como una generalización de la convergencia ordinaria. Para justificar completamente este acercamiento, se mostrará que una serie de Fourier puede converger en la media a una función solamente. Así, suponga que en adición a 1.2 la relación

$$\lim_{n \rightarrow \infty} \int_a^b \left[ F(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx = 0, \quad 1.3$$

se cumple también. Entonces, utilizando la desigualdad

$$(a + b)^2 \leq 2(a^2 + b^2),$$

se deduce que

$$\begin{aligned}
0 &\leq \int_a^b [F(x) - f(x)]^2 dx \\
&= \int_a^b \left[ \left( F(x) - \sum_{k=0}^n c_k \varphi_k(x) \right) + \left( \sum_{k=0}^n c_k \varphi_k(x) - f(x) \right) \right]^2 dx \\
&\leq 2 \int_a^b \left[ F(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx + 2 \int_a^b \left[ f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx.
\end{aligned}$$

Debido a 1.2 y 1.3, esto implica que

$$\int_a^b [F(x) - f(x)]^2 dx = 0,$$

ya que el integrando es no negativo, se deduce que

$$F(x) = f(x)$$

en los intervalos continuos del integrando. Pero la integral tiene sólo un número finito de puntos de discontinuidad. En consecuencia, las funciones  $F(x)$  y  $f(x)$  conciden donde sea excepto en un número finito de puntos. Dos funciones que tienen las características anteriores difícilmente se pueden considerar diferentes en la teoría de las series de Fourier (ya que sus coeficientes de Fourier están expresados en términos de integrales y las integrales no dependen de los valores del integrando en un número finito de puntos). Lo anterior permite formular la siguiente conclusión:

**Teorema 3.** Si el sistema 1.1.4.1 es completo, entonces cualquier función integrable por su cuadrado  $f(x)$  está completamente determinada (excepto por sus valores en un conjunto finito de puntos) por su serie de Fourier, sin importar si la serie converge.

## Propiedades importantes de los sistemas completos

A continuación se establecerán algunas propiedades muy importantes de los sistemas completos.

**Teorema 1.** Si el sistema 1.1.4.1 es completo, entonces cualquier función continua  $f(x)$  ortogonal a todas las funciones del sistema debe ser 0 (nula).

Prueba. Si  $f(x)$  es ortogonal a todas las funciones del sistema, entonces todos los coeficientes de Fourier de  $f(x)$  son nulos. Entonces la definición de sistema completo implica que

$$\int_a^b f^2(x) dx = 0,$$

en consecuencia

$$f(x) \equiv 0,$$

ya que  $f(x)$  es continua.

**Teorema 2.** Si el sistema 1.1.4.1 es completo, si las funciones de sistema son continuas, y si la serie de Fourier de la función continua  $f(x)$  es convergente uniformemente, entonces la suma de la serie es igual a  $f(x)$ .

Prueba. Sea

$$f(x) \sim c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots,$$

y

$$s(x) \equiv c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots. \quad 1.4$$

Ya que las funciones del sistema 1.4.1.1 son continuas, y ya que la serie 1.4 es convergente uniformemente, la suma de las series es continua. Esto se infiere del capítulo 1 que la serie es la serie de Fourier de  $s(x)$ . Por lo tanto las funciones continuas  $f(x)$  y  $s(x)$  tienen la misma serie de Fourier. Pero el teorema 3 de la sección anterior implica que

$$f(x) \equiv s(x).$$

y por 1.4

$$f(x) \equiv c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots.$$

**Teorema 3.** Si el sistema 1.1.4.1 es completo, entonces la serie de Fourier de toda función  $f(x)$  integrable por su cuadrado puede ser integrada término a término, sin importar la convergencia de la serie.

En otras palabras, si

$$f(x) \sim c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots,$$

entonces

$$\int_{x_1}^{x_2} f(x) dx = c_0 \int_{x_1}^{x_2} \varphi_0(x) dx + c_1 \int_{x_1}^{x_2} \varphi_1(x) dx + \dots + c_n \int_{x_1}^{x_2} \varphi_n(x) dx + \dots, \quad 1.5$$

donde  $x_1$  y  $x_2$  son dos puntos cualesquiera en el intervalo  $[a, b]$ .

Prueba. Asumiendo por definición que  $x_1 < x_2$ , se tiene

$$\begin{aligned} \left| \int_{x_1}^{x_2} f(x) dx - \sum_{k=0}^n c_k \int_{x_1}^{x_2} \varphi_k(x) dx \right| &\leq \int_{x_1}^{x_2} \left| f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right| dx, \\ &\leq \int_a^b \left| f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right| dx \quad 1.6 \\ &\leq \sqrt{\int_a^b \left[ f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx} \int_a^b 1 dx \end{aligned}$$

Donde se ha empleado la desigualdad de Schwartz. Por el teorema 2 de la sección precedente, el último término en 1.6 tiende a 0 cuando  $n$  tiende a infinito. Entonces,

$$\lim_{n \rightarrow \infty} \left[ \int_{x_1}^{x_2} f(x) dx - \sum_{k=0}^n c_k \int_{x_1}^{x_2} \varphi_k(x) dx \right] = 0,$$

el cual es equivalente a 1.5.

## **ANEXO II (Guía del usuario de AST)**

## A. Comandos de edición

Para editar una serie es necesario cargar el programa en la memoria tecleando

**AST**

desde el prompt del sistema operativo DOS en la ruta donde se encuentre el archivo ejecutable.

Cuando se realiza lo anterior, el programa presenta la siguiente pantalla



En ella aparece un menú de posibles acciones, este menú es dinámico, es decir, sólo es posible acceder las opciones que tienen significado en el contexto actual.

Aquí es importante resaltar ciertos puntos:

El sistema cuenta con ayuda en línea sensitiva al contexto que se activa al presionar la tecla **F1**.

Así mismo, presionando **Ctrl-K** se tiene acceso a un listado de todas las funciones que se pueden realizar por medio de combinaciones de teclas. En la siguiente imagen se presenta la pantalla cuando se presiona **Ctrl-K**

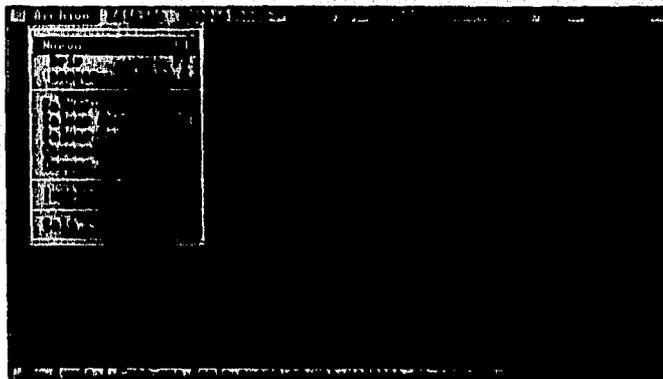


Cada vez que se inicia una nueva sesión en **AST** únicamente las opciones de edición y de ayuda están disponibles.

El proceso de edición involucra varios pasos:

En primer lugar se escoge la fuente de la serie. Es posible crear una serie capturando los datos en la pantalla de edición; recuperando los datos de un archivo en disco o generándolos a partir de un modelo dado.

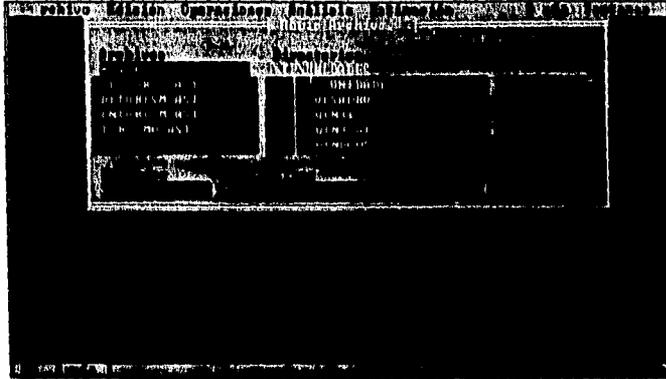
Todas estas opciones están disponibles desde el menú de **Archivo** que se muestra a continuación:



Para crear una serie capturando los datos de la pantalla de edición, escoja la opción **Nuevo** del menú de archivo o presione **F4**.

Para recuperar un conjunto de datos de un archivo en disco, escoja la opción **Abrir** o **F3**.

Al escoger abrir una serie de disco AST presenta la pantalla "Abrir archivo" que se presenta en seguida



Esta ventana tiene 4 secciones entre las que se puede pasar utilizando **Tab** o **Cambio-Tab**.

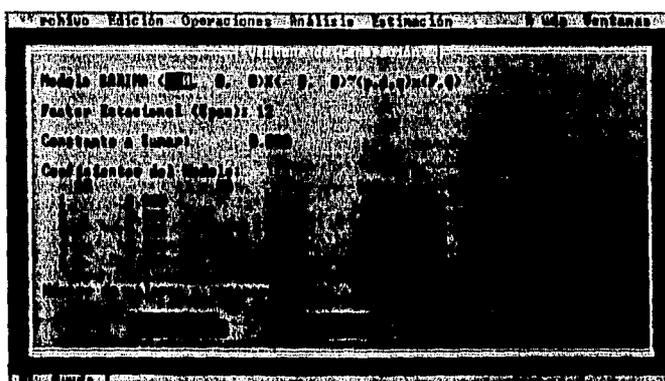
En la primera sección usted puede escoger un archivo de la lista o introducir uno nuevo incluyendo opcionalmente la ruta completa.

La segunda muestra el árbol de directorios en la unidad de disco actual y permite cambiar la ruta en la que se encuentra el archivo que se desea abrir. Aquí es importante mencionar que el cualquier rama del árbol puede expandirse o cerrarse utilizando **Ctrl-A** para abrir la rama y **Ctrl-C** para cerrarla.

En la tercera sección se encuentran las extensiones de los archivos que se desean abrir (La extensión por defecto de un archivo creado por **AST** es **.AST**) y en la cuarta las unidades de disco disponibles para hacer la selección.

Cuando termine de introducir el nombre de la serie a recuperar, presione **Retorno** o el botón **Aceptar** para efectuar la acción; o presione **Escape** o el botón **Cancelar** para abortar Abrir Archivo.

Por último, si lo que desea es construir una serie que siga un modelo teórico, entonces escoja la opción **Generar** o presione **Ctrl-G**, haciendo que el programa despliegue la siguiente "Ventana de Generación" de series.



En esta ventana se captura primeramente el tipo de modelo al cual se van a ajustar los datos y, dependiendo de éste, los coeficientes apropiados:

El factor estacional es necesario sólo si se desea utilizar modelos estacionales; la constante a sumar es opcional para cualquier modelo; los coeficientes Ar, Ma, Sar y Sma son necesarios si su grado correspondiente fué introducido previamente. Finalmente, el número de valores a generar es necesario siempre.

Al igual que en la ventana de abrir archivo el usuario puede escoger si desea aceptar, abortar o pedir ayuda acerca de la acción que está por realizar.

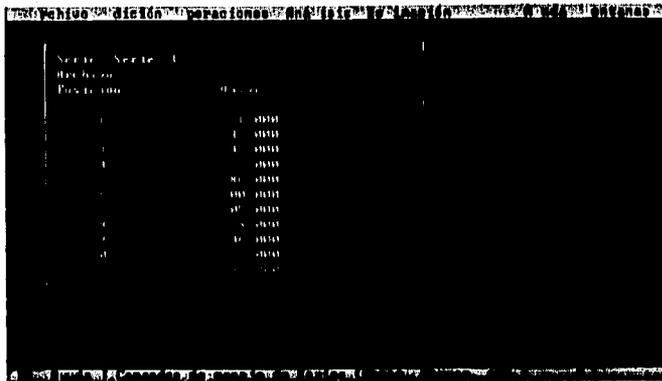
Adicionalmente, el usuario puede escoger en cualquier momento salvar o cerrar la serie eligiendo las opciones de **Salvar**, **Salvar Como**, **Salvar Todo**, **Cerrar** o **Cerrar Todo** para terminar la edición y análisis de la serie.

Cuando escoja las opciones de **Salvar**, la pantalla que el sistema presenta es análoga a la que presenta al abrir el archivo, por lo cual se omite su explicación.

Es importante mencionar que **AST** está diseñado para soportar el análisis de hasta 10 series de tiempo, y que el número máximo de datos en cada serie depende únicamente a la cantidad de memoria RAM con la que cuente la computadora.

Una vez que ha sido creada la serie, el sistema presenta la pantalla de edición de la serie, la cual permite la visualización y modificación de la información.

La pantalla de edición se compone de dos secciones y se muestra en la siguiente figura.



La sección superior contiene el nombre de la Serie de Tiempo que se está editando y el nombre del archivo en el que se almacenan los datos. El nombre de la serie por defecto es "Serie XX" donde "XX" representa el número de series que se han abierto hasta el momento (se puede trabajar hasta con 10 series al mismo tiempo si es que la memoria lo permite); y no existe nombre de archivo hasta que la serie es salvada en disco.

La segunda sección presenta un desplegado en donde pueden verse 11 datos al mismo tiempo, en esta sección es donde se lleva a cabo la captura los datos de la serie de tiempo.

Una vez en esta ventana, se procederá a introducir los datos que se van a analizar. Estos datos pueden darse como números reales (2, 3.4, 7.902, etc) o como fracciones (23/4, 1/3, etc).

Para aceptar la entrada de un dato oprima una de estas teclas: **Retorno**, **Flecha arriba** o **abajo**, **Cambio-Tab** o **Tab**. Si no desea modificar la información y quiere recuperar el valor anterior, presione **Escape** si no ha aceptado la modificación o **Ctrl-U** si ya lo ha hecho.

Una vez que los datos han sido introducidos, si es necesario se procederá a modificarlos utilizando los siguientes pasos.

En primer lugar se debe navegar a la posición donde se va a comenzar a modificar la información, después se selecciona el conjunto de datos que se van a cambiar; y por último se efectúa la operación.

Por tal motivo se expondrán a continuación los comandos relacionados a cada uno de los pasos antes citados, y además se darán un conjunto de comandos misceláneos que han sido añadidos para un mejor manejo de la información.

- **Comandos de navegación:**

Para navegar se tienen las siguientes funciones y teclas:

Tecla	Función
<b>Flecha Arriba</b>	Posiciona el cursor el dato previo.
<b>Flecha Abajo</b>	Posiciona el cursor en el dato siguiente (Si este no existe, lo crea)
<b>Inicio (Home)</b>	Pone el cursor en el primer dato de la serie
<b>Fin (End)</b>	Coloca el cursor en el último dato de la serie
<b>Pág. Arriba</b>	Mueve el cursor 10 posiciones antes de la actual
<b>Pág. Abajo</b>	Lleva el cursor 10 posiciones después de la actual
<b>Ctrl-I</b>	Abre una ventana en la que se solicita la posición a la que se ha de llevar el cursor
<b>Tab</b>	Pasa el control de una serie a la siguiente (si es que se está editando más de una serie en el momento)

Además, existen opciones de menú para ir a una posición determinada dentro de una serie (**Edición/Ir a**) y para navegar entre diferentes series (**Ventanas**).

- **Comandos de Selección**

Para escoger el conjunto de datos a los que se aplicará una operación, se cuenta con las siguientes funciones:

Tecla	Función
<b>Cambio Arriba</b>	Selecciona el dato previo.
<b>Cambio Abajo</b>	Selecciona el dato siguiente.
<b>Cambio Inicio</b>	Selecciona los datos que se encuentran desde el inicio hasta la posición actual.
<b>Cambio Fin</b>	Selecciona los datos desde la posición actual hasta el final de la serie.
<b>Cambio Pág. Arr.</b>	Selecciona los 10 valores anteriores.

Tecla	Función
<b>Cambio Pág. Ab.</b>	Selecciona los 10 valores siguientes
<b>Ctrl-T</b>	Selecciona todos los datos en la serie

Cabe aclarar que si no se hace alguna operación de selección, la acción que se realice aplicará sobre el dato en el cual se encuentre el cursor.

- **Comandos de edición**

Los comandos de edición permiten modificar un conjunto de datos de diversas formas.

En la siguiente tabla se presenta un compendio de comandos de edición disponibles al usuario del programa:

Tecla	Función
<b>F10</b>	Edita el dato actual.
<b>Inserción</b>	Inserta un dato en la posición actual (Este comando no funciona si se ha seleccionado más de un dato).
<b>Borrar</b>	Borra el conjunto de datos seleccionado.
<b>Ctrl-X</b>	Corta los datos seleccionados y los pone en un área temporal.
<b>Ctrl-C</b>	Copia el conjunto de datos seleccionado y los pone en un área temporal.
<b>Ctrl-V</b>	Copia el contenido del área temporal en la posición actual del cursor.
<b>Ctrl-U</b>	Deshacer el último cambio.
<b>Ctrl-E</b>	Almacena el conjunto de datos seleccionado en el disco.
<b>Ctrl-L</b>	Inserta en la posición actual los datos almacenados en un archivo.

Cabe destacar que cuando el usuario copia o borra un conjunto de datos en una serie, este conjunto queda almacenado en un área común, por lo cual es posible

que las series que se editan en un momento determinado compartan información.

Además de las combinaciones aquí mencionadas, es posible realizar las mismas operaciones a través de opciones de menú.

- **Comandos Varios**

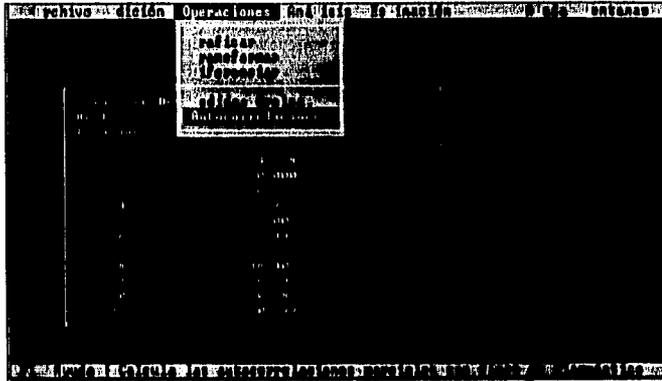
Los comandos varios permiten un mejor manejo y visualización de la información contenida en la serie.

A continuación se presenta una lista de estos comandos:

<b>Tecla</b>	<b>Función</b>
<b>Ctrl-F</b>	Cambia el formato utilizado para presentar los datos.
<b>Ctrl-N</b>	Cambia el nombre a la serie.
<b>Ctrl-G</b>	Genera una nueva serie.
<b>Ctrl-F5</b>	Mueve la ventana de edición.
<b>Ctrl-K</b>	Muestra las teclas de función disponibles.
<b>F1</b>	Trae la ayuda.

## B. Operaciones generales

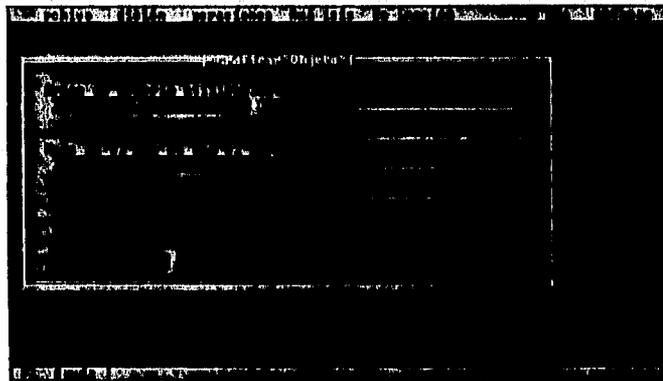
El menú de **Operaciones** se presenta a continuación



- **Graficación.**

Para graficar una serie es necesario escoger la opción **Graficar** del menú de **Operaciones**.

Al escoger dicha opción, el sistema presenta la siguiente pantalla de opciones de graficación:



la cual se divide en 7 secciones que determinan las características de la gráfica.

La primera sección permite elegir un área de memoria para almacenar la salida. Es importante destacar que el número de áreas disponibles está determinado por la memoria disponible en la computadora.

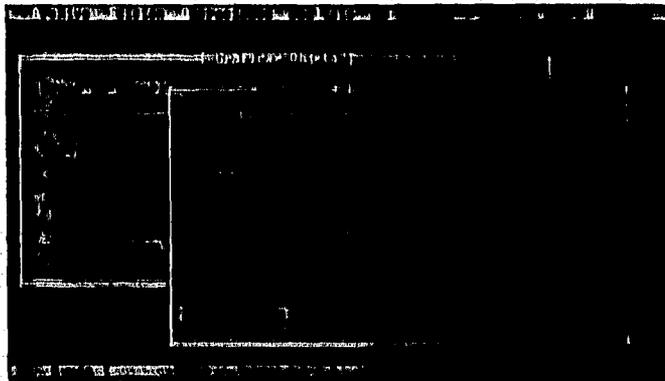
El hecho de salvar una gráfica permite hacer comparaciones con lo que se puede lograr un mejor análisis.

La segunda sección permite escoger el modo el que se usará el buffer, en otras palabras, si el contenido de la gráfica salvada anteriormente va ser destruido o utilizado como base para la siguiente salida.

Aquí hay que considerar que si se escoge la opción de utilizar el buffer, la sección de límites y divisiones no va a estar disponible, de igual forma, no será posible cambiar algunas de las opciones al definir las leyendas.

La sección de leyendas permite definir el título y la leyenda de la gráfica, así mismo el usuario puede escoger el tipo y tamaño de letra para ambos.

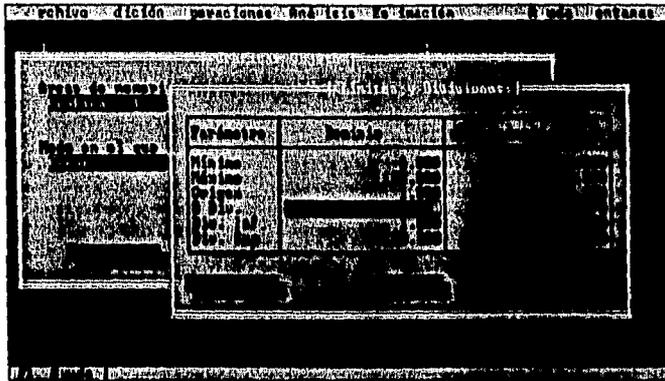
La pantalla de leyendas se presenta en seguida:



En este punto hay que señalar que las opciones de utilizar título y etiquetas, así como el tipo y tamaño de letra para la gráfica están disponibles sólo si se está generando una gráfica nueva.

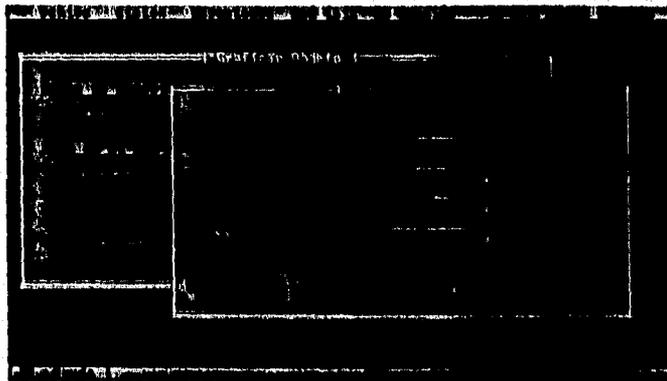
En la sección de límites y divisiones se pueden modificar las características de los ejes, los límites de los dominios y el número de divisiones que los ulteriores observarán.

Esta pantalla se muestra a continuación



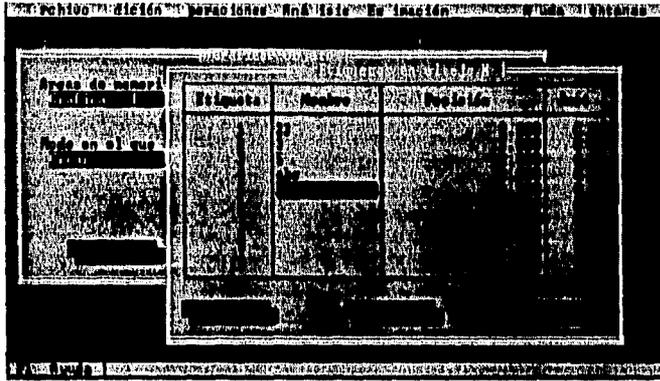
En la sección de estilos se definen los tipos de gráfica (Barras, Líneas, Discreta y de Puntos para las discretas, mientras que para las continuas este campo no es disponible); El tipo, color y ancho de línea; y el patrón de relleno si se habla de una gráfica de barras.

La pantalla de estilos se presenta enseguida.



La sección de etiquetas permite introducir al usuario hasta diez etiquetas que se imprimirán en el eje X; es posible definir la posición, texto y justificación de la etiqueta.

La siguiente figura contiene dicha pantalla



La pantalla de sombreado sólo está disponible si se está procesando una función continua; en ella se definen hasta 5 intervalos de sombreado que definirán áreas bajo la curva con respecto al origen que se haya definido previamente.

A continuación se presenta esta pantalla

- **Transformaciones**

Para transformar una serie se utiliza la opción **Transformar** del Menú de **Operaciones** la cual permite aplicar la transformación de Box-Cox a un conjunto de datos.

La pantalla de transformación de Box-Cox se presenta a continuación.

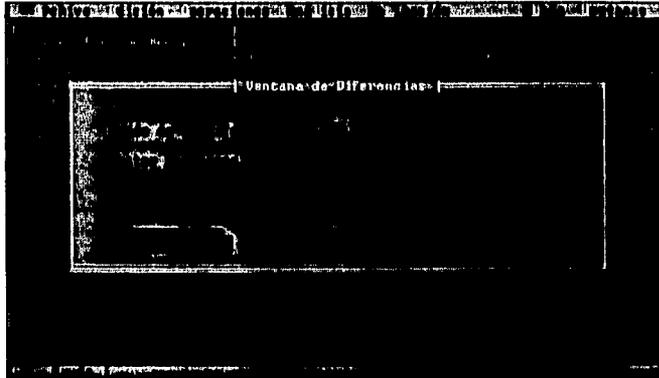


Cuando se aplica esta operación a la serie, el sistema genera una nueva pantalla de edición donde almacena los datos resultantes los cuales formarán parte de una nueva serie independiente de la original, y por tanto podrán ser editados, salvados y analizados separadamente.

- **Diferencias**

Para diferenciar una serie es necesario escoger la opción **Diferenciar** del menú de **Operaciones** e introducir el grado de la diferencia así como el factor estacional.

La ventana de diferencias se muestra en seguida.



Al igual que al aplicar transformaciones, cada vez que se acepta una diferencia se crea una nueva serie independiente de la anterior.

- **Medidas Generales**

Como información adicional, AST permite visualizar algunos de los estadísticos de las series que se están editando. Para obtener una tabla que muestra la media, varianza, desviación estándar y rango de cada una de las series activas, elija la opción **Medidas Generales** del Menú de **Operaciones**.

La salida de este comando es una tabla con el siguiente formato:

Serie	Media	Varianza	Desviación Estándar	Rango	
				Mínimo	Máximo
Serie 1	100	169	13	55	231
Serie 2	50	64	8	-45	321

- **Autocorrelaciones.**

Para obtener las autocorrelaciones de una serie escoja la opción **Autocorrelaciones** del menú de **Operaciones**.

Al escoger esta opción el sistema genera una pantalla de edición en la que coloca los resultados. Estos datos pueden ser modificados, graficados, etc. ya que son independientes de la serie original.

## C. Análisis

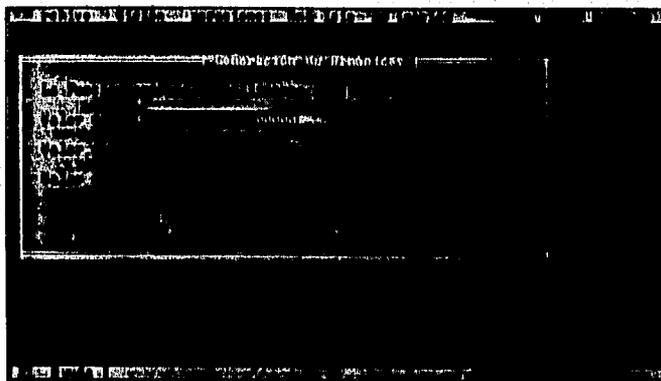
El menú de Análisis contiene 5 opciones las cuales se muestran en la siguiente figura.



- **Armónicas.**

Esta opción sirve para graficar la función  $f(x) = A \text{ sen } (\mu x + p)$  donde  $A$  es la amplitud,  $\mu$  la frecuencia y  $p$  la fase de la función armónica.

Al escoger esta opción el sistema despliega esta pantalla



en la que el usuario puede introducir valores para las constantes antes mencionadas.

Cuando el usuario acepta graficar la armónica, el sistema invoca al procedimiento de opciones de graficación en el cual puede elegir la manera en que la función será dibujada.

- **Coefficientes de Fourier.**

Esta opción le indica al sistema calcular los coeficientes de Fourier asociados a la serie que se encuentre activa en ese momento.

Los coeficientes se generan y se presentan al usuario en una tabla que tiene el formato que se describe a continuación.

Tabla de Coeficientes de Fourier

K	Frecuencia	ak	bx	Amplitud	Fase
0	0.000	9.555	0.000	9.555	1.571
1	0.105	-2.370	-1.733	2.960	6.435

- **Periodograma**

Esta opción grafica el periodograma (suma al cuadrado de los coeficientes de Fourier) o espectro de potencias asociado a la serie que se este editando en esos momentos.

Al escoger esta opción el sistema calcula, si es necesario, los coeficientes de Fourier y a continuación presenta la pantalla de graficación donde el usuario definirá las características de la salida.

- **Polinomio trigonométrico**

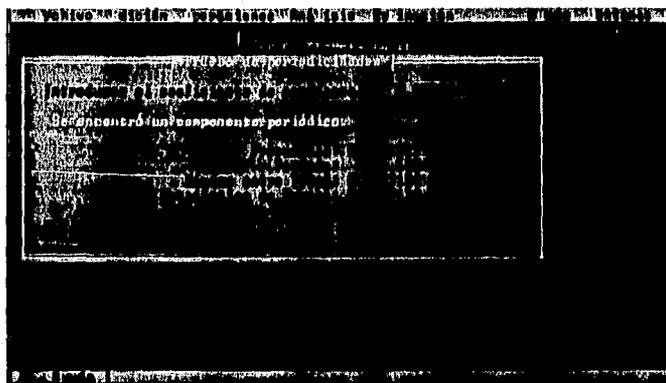
Esta opción grafica el polinomio trigonométrico o serie de Fourier asociado a la serie que se esté editando en esos momentos.

Al escoger esta opción el sistema calcula, si es necesario, los coeficientes de Fourier y a continuación presenta la pantalla de graficación donde el usuario definirá las características de la salida.

- **Prueba de periodicidades**

Esta opción realiza una búsqueda de componentes periódicos en la serie que se esté editando.

Cuando el usuario escoge esta opción, el sistema presenta la siguiente figura donde se captura el coeficiente de confianza e inmediatamente se emite un reporte donde se muestra en primer lugar si se encontró o no un componente periódico y la información concerniente a este hecho.



## D. Estimación

El menú de estimación cuenta con 4 opciones que se muestran en la siguiente figura.



- **Espectro estimado**

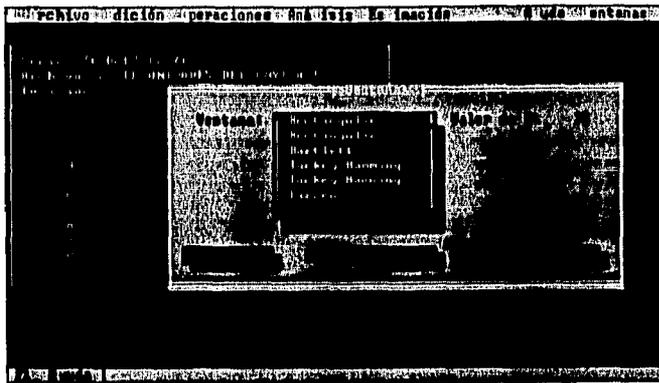
Esta opción grafica el espectro estimado asociado a la serie que se edita en esos momentos.

Al escoger esta opción el sistema calcula, si es necesario, las covarianzas y a continuación presenta la pantalla de graficación donde el usuario definirá las características de la salida.

- **Ventanas**

Permite graficar las ventanas espectrales.

Al elegir esta opción, el sistema despliega la siguiente pantalla donde el usuario puede decidir que ventana quiere revisar y el valor de  $M$  que utilizará.

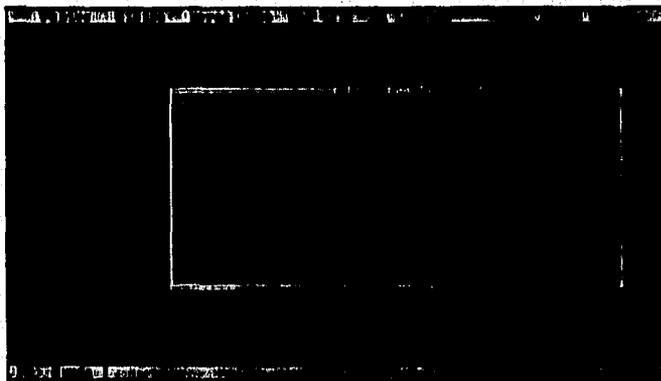


Al aceptar graficar la ventana el sistema presenta la pantalla de opciones de graficación donde el usuario elegirá las características de la salida.

- **Espectro Suavizado**

Esta opción grafica el espectro suavizado mediante alguna ventana espectral correspondiente a la serie que se encuentre activa.

Para este fin el sistema muestra la siguiente pantalla

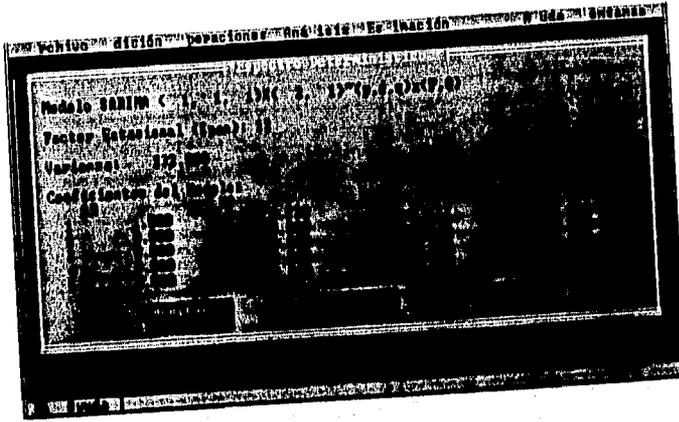


en la cual se introduce la ventana espectral que se va a utilizar, el valor de  $M$  que define el ancho de banda y el coeficiente de confianza.

- **Espectro Determinístico**

Al elegir esta opción del menú, el usuario es capaz de introducir un modelo SARIMA para que el sistema grafique el espectro teórico correspondiente.

La pantalla en la que el usuario definirá este modelo se muestra en seguida



En esta pantalla el usuario introduce en primer lugar los parámetros que definen el modelo en términos generales; después el factor estacional si introdujo un valor diferente de 0 para los componentes Sar o Sma; el valor de la varianza del modelo; y por último los coeficientes del modelo.

Al término de la captura, el sistema presenta la pantalla de graficación donde se introducen las características de la salida del espectro determinístico.

## D. Misceláneos

En esta sección del anexo se incluyen los menús de **Ayuda** y **Ventanas** que son parte complementaria del sistema.

En primer lugar, el menú de ayuda contiene tres opciones, **Contexto**, **Teclas** y **Acerca de**.

La opción **Contexto** presenta una pantalla que contiene una breve explicación de la acción que se está llevando a cabo en esos momentos.

La opción **Teclas** presenta una lista con todas las teclas que tienen asociada una función y que pueden utilizarse en el instante.

La opción **Acerca de** presenta la siguiente pantalla con los datos del autor.



Finalmente el menú de ventanas contiene 3 opciones **Siguiente**, **Anterior** y **Cerrar todo**.

La opción **Siguiente** pasa el control de edición a la serie que continúe en el orden en el que fueron creadas.

La opción **anterior** pasa el control de edición a la serie previa en el orden en el que fueron creadas.

La opción **Cerrar todo** termina con la edición de todas las series que se encuentran activas actualmente. Para cada una de las series que han sido modificadas, pregunta si se desea actualizar la información en el disco con estos cambios.

Adicionalmente, el menú de ventanas mantiene al final una lista con los nombres de las series que se están editando. Si el usuario escoge alguno de estos nombres, entonces el control de la edición pasará inmediatamente a la serie elegida.

## Bibliografía

**"PC Programmer's Bible"**

Peter Norton.

Microsoft Press, 1992.

**"TURBO PASCAL 6. The complete reference"**

Stephen O'brien.

McGraw Hill, 1991.

**"Object-oriented modeling and design"**

Rumbaugh, James.

Prentice Hall 1991.

**"Pascal applications for the sciences"**

Crandall, Richard E.

Wiley, 1984

**"Fourier Series"**

Tolstov, Georgii Pavlovich.

Dover, 1976.

**"Fourier Series"**

Gowar, N. W.

Chatto & Windus, 1974.

**"Fourier Series"**

Hardy, Godfrey Harold.

University of Cambridge, 1950.

**"Modelos de Decisión con Procesos Estocásticos II  
(Metodología de Box-Jenkins)"**

González Videgaray, María del Carmen.

UNAM, 1990.

**"Time Series Analysis: Forecasting and control"**

Box, George E. P.

Prentice Hall, 1994.