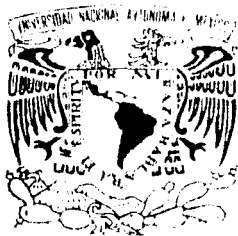


24021
26



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

**ESCUELA NACIONAL DE ESTUDIOS
PROFESIONALES "ACATLÁN"**

**"ANÁLISIS DE SERIES DE TIEMPO
NO ESTACIONARIAS MEDIANTE
MÉTODOS DE VENTANAS"**

TESIS

QUE PARA OBTENER EL TITULO DE

**LICENCIADO EN MATEMÁTICAS
APLICADAS Y COMPUTACIÓN**

PRESENTA

ERIK LÓPEZ MARROQUÍN

ASESOR: VÍCTOR JOSÉ PALENCIA GÓMEZ

FECHA: ENERO DE 2003

TESIS CON
FALLA DE ORIGEN



A



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.

Prólogo

El presente trabajo se enfoca al tema del análisis de las series de tiempo no estacionarias, partiendo de las limitaciones que presenta estudiar estas series mediante los métodos convencionales que por lo general suponen que la serie de tiempo es estacionaria. Aunque se incluyen diferentes métodos de análisis en el trabajo, se enfatiza el uso de aquellos que utilizan ventanas para el análisis, es decir, utilizando intervalos de tiempo para observar los cambios que presentan a través del tiempo la media, varianza, autocovarianza, frecuencia y fluctuación estacional (conocidas como "propiedades espectrales" de la serie de tiempo).

Dado que la no estacionaridad en las series de tiempo puede presentarse de manera individual o conjunta en las diferentes propiedades espectrales, el trabajo se encuentra organizado en tres capítulos, desarrollándose por separado el análisis de la no estacionaridad en cada propiedad espectral.

El capítulo primero de este trabajo trata entonces el tema del análisis de series de tiempo no estacionarias en la media, se definen ahí los conceptos básicos del tema, comparando diferentes métodos para analizar series de tiempo no estacionarias en la media, mostrando las ventajas y desventajas de cada método, así como la forma en como pueden complementarse. Se enfatiza en éste capítulo particularmente el uso de las ventanas (medias móviles) para el análisis de la tendencia.

En el capítulo segundo se extiende el uso de las ventanas para analizar la varianza y muy especialmente se toca el tema del análisis de la autocovarianza no estacionaria. Se muestra aquí una importante aplicación del uso de ventanas para afrontar las limitantes que se tienen al analizar series de tiempo no estacionarias en la autocovarianza, pues este tipo de no estacionaridad presenta particular dificultad para estudiarse.

TESIS CON
FALLA DE ORIGEN

El capítulo tercero toca el tema de la no estacionaridad en la frecuencia o periodicidad de las series de tiempo. Se muestran aquí principalmente dos métodos utilizando el concepto de ventaneo: El análisis de Fourier con ventanas, y el análisis Wavelet. Este último método de aplicación relativamente reciente.

En general, además de mostrar el uso de ventanas como una alternativa viable para estudiar series no estacionarias, el trabajo busca observar las virtudes y defectos que presenta utilizar los métodos de ventanas.

Considero importante manifestar mi agradecimiento por el apoyo y orientación recibido durante la planeación y desarrollo del trabajo por parte de mis profesores **Víctor José Palencia Gómez** y **Elvira Beatriz Clavel Díaz**, cuyos consejos, sugerencias, correcciones, etc. ayudaron a llevar este proyecto a buen término. De la misma manera quiero expresar mi gratitud con la profesora **Maricarmen González Videgaray**, a quien debo el gusto por el tema de las series de tiempo. Finalmente, agradezco a los miembros de mi jurado: **Jorge Jiménez Zamudio**, **Manuel Valadéz Rodríguez** y **Carmen Villar Patiño**, por las revisiones, comentarios y sugerencias que se sirvieron prestar.

TESIS CON
FALLA DE ORIGEN

Introducción

El creciente desarrollo que ha tenido durante las últimas dos décadas la tecnología de las computadoras ha ampliado significativamente el uso de algoritmos numéricos para resolver problemas analíticos que, por su complejidad resultarían muy difíciles y laboriosos de resolver analíticamente, y en algunos casos, prácticamente imposibles de resolver sin el apoyo de una computadora. El uso de las computadoras ha permitido también el desarrollo y la aplicación de ciertos métodos y técnicas que nos permiten analizar eficientemente y con rapidez grandes volúmenes de datos, tales como indicadores financieros, cuentas de banco, registros académicos, etc.

De la misma manera, aunque han existido métodos para analizar series de tiempo desde antes de la expansión de los microprocesadores, estos métodos requieren realizar una cantidad grande de cálculos que los convierten en herramientas laboriosas y extremadamente difíciles de utilizar para quien no dispone de una herramienta tan potente como lo es el ordenador. Se entiende entonces que muy pocas personas podían realizar este tipo de análisis, pues no sólo requería del conocimiento necesario, sino también de mucha paciencia y dedicación.

Sin embargo, el desarrollo de la tecnología no sólo expandió y perfeccionó los métodos existentes, sino que además proporcionó las herramientas necesarias para el surgimiento de nuevas técnicas de análisis que permitieran obtener resultados aún superiores, pues la rapidez y desarrollo creciente de los análisis efectuados por las computadoras sacó a relucir también los defectos de estas técnicas, que por lo general sólo pueden aplicarse a un tipo específico de datos, y surge entonces la necesidad de nuevos métodos de análisis para casos especiales como lo es, por ejemplo, el de las series de tiempo no estacionarias.

Las series de tiempo, definidas éstas como secuencias de datos (observaciones) ordenados cronológicamente, se pueden utilizar en una gran cantidad de casos prácticos, y para casi

TESIS CON
FALLA DE ORIGEN

cualquier disciplina que requiera del conocimiento sistemático de un proceso o sistema a medida que el tiempo cambia. Sus aplicaciones pueden abarcar procesos económicos, sociales, administrativos, físicos, biológicos, químicos, etc.

La importancia fundamental del estudio de las series de tiempo consiste en comprender el comportamiento de la serie, de manera que podamos apreciar cómo el tiempo puede afectarla y detectar patrones sistemáticos de comportamiento tales como: tendencia, periodicidad, heteroscedasticidad, etc. Mediante el estudio de las series de tiempo se puede predecir, analizar las relaciones entre causas y efectos, identificar el comportamiento estocástico de la serie, así como establecer esquemas de control que permitan detectar desviaciones significativas de una meta deseada.

George E. P. Box y Gwilym M. Jenkins propusieron en 1976 una metodología de análisis de series de tiempo que permitió obtener modelos predictivos de éstas, partiendo del análisis de la dependencia de un dato con los anteriores. Sin embargo, la metodología de Box y Jenkins se enfoca particularmente a las series de tiempo estacionarias, por lo que el análisis y pronóstico que se realizan suponen que el comportamiento general de la serie no cambia con el tiempo. En realidad, en la vida cotidiana existen una gran cantidad de series con comportamiento no estacionario, por lo que las aplicaciones de la metodología se encuentran algo limitadas. No obstante, la metodología de Box y Jenkins permite obtener modelos de predicción para ciertos tipos de series no estacionarias, mediante diferencias y transformaciones que se aplican a la serie original, aunque las predicciones obtenidas son por lo general fiables a corto plazo para series no estacionarias.

Aunque el presente trabajo se enfoca a mostrar una alternativa de análisis para el caso de las series de tiempo no estacionarias, es importante señalar que esta alternativa se limita únicamente a realizar el análisis de la serie, no se ofrece aquí una alternativa para obtener un modelo de predicción como el que se obtiene de la metodología de Box y Jenkins. No obstante, cuando sea

TESIS CON
FALLA DE ORIGEN

posible se mostrarán algunas alternativas viables para realizar predicciones en una serie de tiempo no estacionaria, como es el caso de los modelos de regresión.

El estudio de las series de tiempo no estacionarias involucra detectar cambios en el comportamiento del sistema durante ciertos intervalos de tiempo, de esta manera, las propiedades espectrales de una serie de tiempo (media, varianza, autocovarianza, fluctuación estacional, frecuencia y variación aleatoria) deben analizarse de manera que podamos detectar los cambios que presentan a través de tiempo. Una forma de realizar el análisis de la serie a través del tiempo consiste en utilizar intervalos de tiempo fijos, denominados "ventanas", para los cuales se calculan las propiedades espectrales de los datos, de manera que el intervalo de tiempo (ventana) se recorre para volver a calcular las propiedades espectrales en el siguiente subconjunto de datos de la serie. Este concepto de ventaneo se tratará con mayor detalle en el primer capítulo de éste trabajo.

El presente trabajo muestra el uso de las ventanas para analizar cada una de las propiedades espectrales de las series de tiempo, así como las ventajas y desventajas de aplicación que proporcionan, y en algunos casos se presentan comparaciones con otros métodos alternativos o complementarios.

Cabe señalar que las gráficas incluidas en el primer capítulo de este trabajo se generaron utilizando el paquete Excel 2000, y posteriormente, debido a la complejidad de los cálculos requeridos, se desarrolló un programa utilizando Matlab 6.0 para generar gráficas requeridas a lo largo de los capítulos restantes.

TESIS CON
FALLA DE ORIGEN

Capítulo I

Análisis de la tendencia

La función de media es sin duda la primera propiedad espectral que se analiza en una serie de tiempo, pues proporciona parte de la información más general del comportamiento del sistema. Dicha función nos ayuda a determinar si en general los valores de la serie tienden a crecer, decrecer o mantener constante su valor. No obstante, la utilidad principal del análisis de la tendencia consiste fundamentalmente en el estudio de la forma en como la serie tiende a crecer o decrecer. Ésta puede, por ejemplo, crecer en forma lineal, exponencial, potencial, logarítmica, etc. El análisis de la tendencia nos permite, por un lado, analizar los efectos de diversos factores que intervienen en el comportamiento de la serie, y por el otro, obtener modelos matemáticos que puedan permitir pronosticar los valores promedio de la serie.

Al analizar la función de media de series de tiempo no estacionarias en la media (con tendencia), es importante tener en cuenta la diferencia entre analizar y predecir, puesto que el análisis consiste en el estudio sistemático de la relación entre las causas y efectos, pero no necesariamente implica la obtención de un modelo de predicción. La predicción de la serie implica por lo general utilizar sólo los rasgos generales del comportamiento de la serie y mirar hacia adelante en el tiempo, y no tanto estudiar los datos como puede ser el efecto de una crisis económica.

En este capítulo se exponen tres métodos para realizar el análisis de una serie con media no estacionaria: El uso de ventanas o subintervalos de tiempo para realizar un análisis relativo al tamaño del subintervalo; el uso de modelos de regresión lineal o no lineal, para obtener un modelo que refleje el comportamiento general de la media y permita predecir; y el uso de diferencias ordinarias, particularmente útiles en los casos de comportamiento estocástico de la tendencia. Las ventajas y desventajas que presentan los métodos que se exponen en este capítulo, así como sus casos específicos de aplicación son expuestos con la finalidad de ilustrar la capacidad, el alcance y los límites de cada método.

TESIS CON
FALLA DE ORIGEN

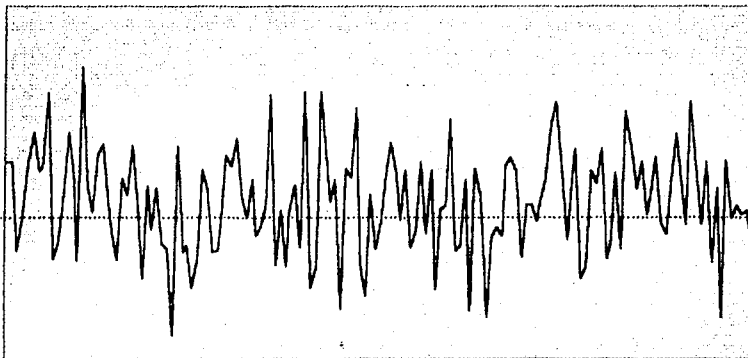
1.1 Medias Móviles

Este primer método utiliza la fórmula tradicional para obtener el valor promedio de los datos. Sin embargo, considera que los valores promedio de la serie varían con el tiempo, por lo que el método recibe comúnmente el nombre de "medias móviles". Teniendo en cuenta que la función de media representa el valor promedio de las observaciones durante ciertos subintervalos de tiempo, es decir, el valor promedio de los datos a través del tiempo, un estimador muestral para la media μ en un proceso estacionario está dado por:

$$\hat{\mu} = \bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$$

en donde Z_i representa los valores de la serie para cada valor de tiempo i .

Sin embargo, realizar el cálculo de la media de esta forma bastaría cuando el comportamiento de la serie presentara una media constante (ver gráfica 1), pero en el caso de una función de media cuyos valores promedio varían con el tiempo, el cálculo de la función de media se puede realizar utilizando promedios parciales, es decir, obteniendo el valor promedio de la serie para ciertos subintervalos de tiempo.



Gráfica 1 Función de media constante (Ruido Blanco).

Para obtener valores apropiados de los promedios parciales es necesario, primero que nada, determinar la longitud de subintervalo de acuerdo al tipo de cambios que queramos analizar en la media de la serie. Si elegimos un subintervalo de tiempo pequeño podremos visualizar los cambios pequeños en la media de la serie, en tanto que si elegimos un subintervalo de tiempo grande sólo se reflejarán los cambios drásticos en la media. Esto se debe a que cuando se calcula la media para un subconjunto pequeño de datos, los valores de la media cambian con mucha facilidad, pues siendo pocos datos hacen que la media sea más sensible a cambios debidos al último dato que se ha considerado. Por otro lado, cuando el subconjunto de datos es grande, por ejemplo de unos 100 datos, los últimos datos por lo general no son muy significativos, pues la media de la serie representa el comportamiento de una cantidad grande de datos, y por tanto un cambio pequeño en los últimos datos de la serie no se reflejará fácilmente en la función de media resultante.

En lo sucesivo nos referiremos al subintervalo de tiempo que se utiliza en un análisis como "ventana"¹, en tanto que al tamaño de la ventana le llamaremos "escala"². El análisis de series de tiempo por medio de ventanas con múltiples tamaños (escalas), se aplica específicamente a series con comportamiento no estacionario.

Serie original	Media con ventana de tamaño 3
9.6208	9.3488
9.1321	8.8317
9.2934	8.4053
8.0696	7.9663
7.8528	8.1743
7.9765	
8.6937	

Tabla 1

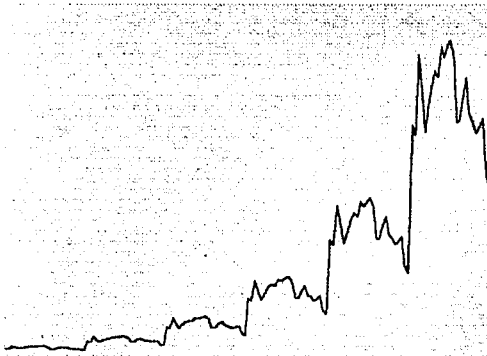
¹ En general, a los intervalos de tiempo, independientemente de su longitud, se les puede denominar ventanas, y por tanto en éste documento se adopta ésta práctica.

² Se utiliza en este trabajo la denominación de "escala" en el sentido de que se realiza un microanálisis o macroanálisis de los datos, es decir, se contemplan los cambios en las propiedades espectrales ya sea a corto o a largo plazo, dependiendo del tamaño de la ventana, por lo que podemos hablar de un "análisis a escala".

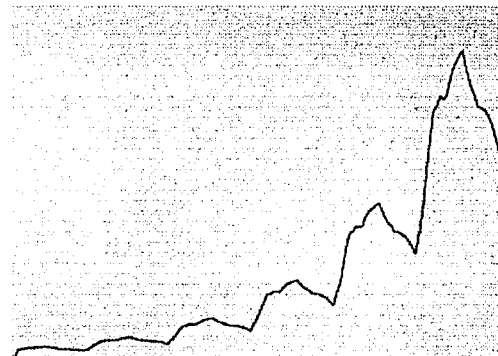
Consideremos la serie de tan solo seis datos que se muestra en la Tabla 1. Suponiendo que deseamos obtener una función de media utilizando una ventana de tamaño 3, la primera media se obtendrá utilizando los primeros tres datos, la segunda se obtendrá con los datos 2, 3 y 4; y así sucesivamente. Es importante hacer notar que al obtener las medias se pierden 2 datos en este caso particular.

En general cuando manejamos n datos en una serie de tiempo cualquiera y utilizamos una ventana de tamaño s , se pierden $s - 1$ datos, quedándonos finalmente un total de $n - s + 1$ datos. Se entiende entonces que el límite en el tamaño de una ventana está dado en función del tamaño de la serie que se analiza. Por lo general es preferible no utilizar ventanas muy grandes, sino sólo lo que sea necesario para apreciar el comportamiento general de la tendencia. Por otro lado, una ventana muy chica no proporcionará mayor información sobre la media que la gráfica de la serie original.

La selección de un tamaño de ventana apropiado para visualizar la serie no es entonces tan sencilla y en ocasiones es preferible obtener la función de media con diferentes escalas (tamaños de ventana). Consideremos ahora un segundo ejemplo de una serie simulada de 150 datos cuya representación se muestra en la Gráfica 2.

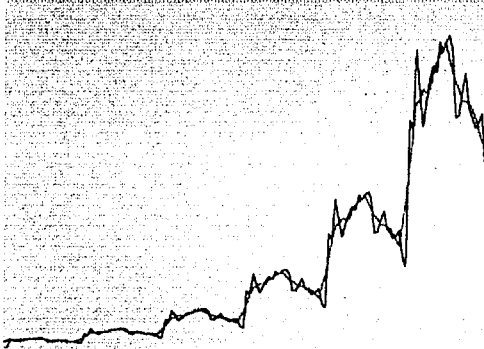


Gráfica 2 Serie de Tiempo

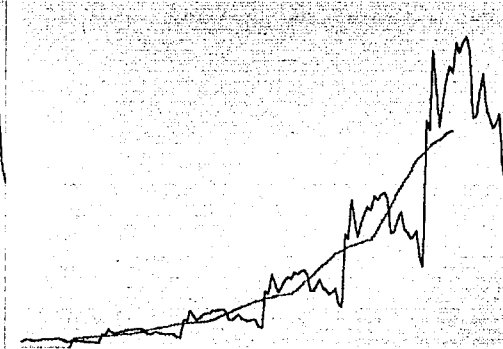


Gráfica 3 Función de Media con ventana de tamaño 5

La gráfica 2 a simple vista presenta en general una tendencia creciente, en tanto que la gráfica 3 se puede decir que es una versión suavizada de la gráfica original, que entre algunas ventajas puede tener la propiedad de "suavizar" los datos atípicos, y de esta manera entre más grande sea la ventana en la que se calculan los promedios, los datos atípicos influirán menos en la función de media resultante.



Gráfica 4 Función de Media con ventana de tamaño 5



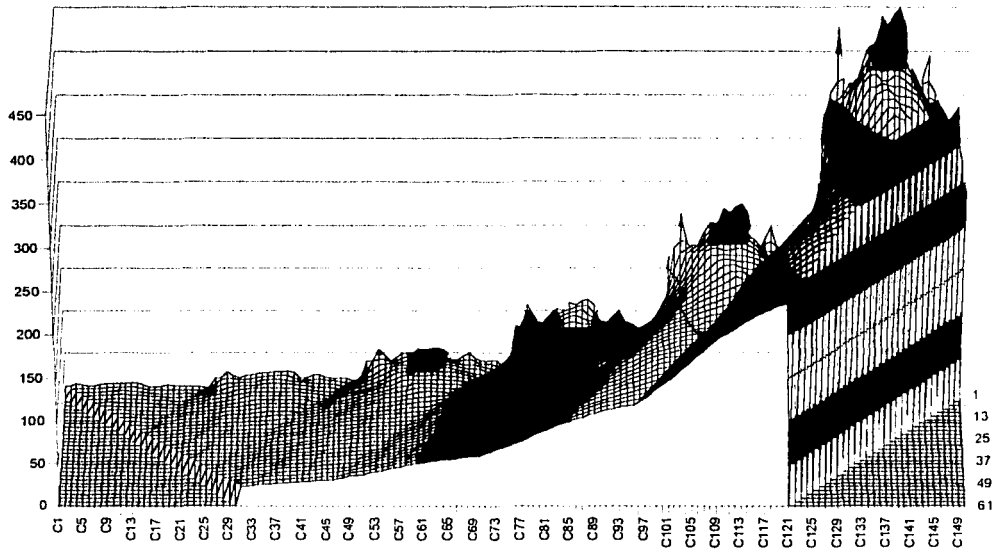
Gráfica 5 Función de Media con ventana de tamaño 33

En la gráfica 4 se muestran simultáneamente la *función de media*³ (en rojo) y la serie original, corroborándose que se trata de una versión suavizada, en tanto que en la gráfica 5 tenemos una función de media con ventana más grande, de tamaño 33, en donde se puede visualizar la tendencia exponencial creciente con marcados escalamientos periódicos. En el caso particular de esta serie, el comportamiento periódico se puede apreciar desde la gráfica original, pero no siempre es posible hacer esto. Por otro lado, en el caso de las series con periodicidad el comportamiento suavizado se reflejará mejor cuando el tamaño de la ventana sea igual al periodo de la serie.

Como sea, en la mayoría de los casos nos interesa obtener sólo los rasgos generales de la tendencia, y por lo tanto deseamos obtener la función de media que mejor refleje el

³ A diferencia de cuando se tiene una media constante, en este caso la media es una función del tiempo (y de la escala), por lo que se utiliza el concepto de "función de media".

comportamiento general de la tendencia. Se pueden obtener simultáneamente las funciones de media en diferentes tamaños de ventana, como se muestra en la gráfica 6.



Gráfica 6 Función de media con ventanas de tamaños 2 a 61

La gráfica 6 muestra simultáneamente las ventanas de tamaños 2 al 61, donde se aprecia un comportamiento gradualmente suavizado. La forma de trapecio que presenta la base se debe a la pérdida de datos cuando se obtienen los promedios de medias móviles. Con el objeto de poder graficar los datos sin problemas se han utilizado ceros en las zonas externas del trapecio. Esta gráfica tridimensional también se puede mostrar en el plano como si fuera vista desde arriba. La gráfica 7 muestra la función de media vista desde arriba.



Gráfica 7 Función de Media con múltiples tamaños de ventana en el plano

Esta gráfica presenta en las escalas pequeñas (1 al 9) un ascenso y descenso constante, nótese que los ascensos y descensos son periódicos. Por otro lado, a escalas más grandes podemos observar una tendencia creciente, y que por añadidura es exponencial, esto lo podemos concluir dado que las áreas de un cierto color son cada vez más chicas de izquierda a derecha, lo cual significa que cada vez crecen más rápido los valores de la serie.

En general, podríamos pensar que realizar un análisis de la tendencia mediante este método puede resultar excesivamente tedioso, sin embargo, obtener estas gráficas y analizarlas no presenta demasiada dificultad con un software apropiado como Matemática, Matlab o Statgraphics, o bien en un software desarrollado por nosotros para ese fin.

Aunque estrictamente hablando, la función de media se calcula en la forma descrita anteriormente, cuando deseamos obtener predicciones de los valores de la función de media, puede resultar más práctico utilizar la regresión lineal o polinomial, pues tiene mayor utilidad para realizar predicciones sobre la serie, y aunque no estamos calculando precisamente promedios de la serie, son aproximaciones bastante útiles.

1.2 Tendencia lineal

Aunque el método de ventanas puede ser utilizado para analizar la tendencia en cualquier tipo de serie, si se trata de obtener un modelo predictivo cuando la serie muestra tendencia determinística, en este caso conviene utilizar un modelo de regresión, especialmente si el comportamiento de la tendencia se puede representar mediante un modelo simple. El caso del comportamiento lineal creciente o decreciente se aproxima mediante un modelo de regresión lineal. Dicho de otra manera, cuando suponemos a partir de la gráfica original que se tiene una tendencia de tipo lineal, podemos utilizar la regresión lineal como herramienta para obtener una recta que aproxime la tendencia. Desde luego que si la serie no tiene tendencia lineal, la recta resultante no se aproximará a la tendencia real de la serie.

Al utilizar la regresión lineal obtendremos una ecuación de predicción conocida como "ecuación de regresión". El término **regresión lineal** implica que la media de la serie de tiempo μ_{z_t} está linealmente relacionada con t por la ecuación de regresión lineal poblacional:

$$\mu_{z_t} = \alpha + \beta t$$

donde los coeficientes de regresión α y β son parámetros que deben estimarse a partir de los datos muestrales (serie de tiempo). Si a y b representan estas estimaciones, respectivamente, se puede entonces estimar μ_{z_t} por :

$$\bar{Z}_t = a + bt$$

Para estimar los valores de a y b , podemos utilizar el método de mínimos cuadrados. Este método se basa en el hecho de que cada observación de la serie de tiempo Z_t en el i -ésimo valor de tiempo satisface la siguiente relación:

$$Z_t = a + bt_t + e_t$$

en donde e_t se le llama "residuo" o "error", y representa la diferencia entre el valor real observado y el valor obtenido por la ecuación de regresión, es decir: $e_t = Z_t - \bar{Z}_t$.

Para obtener los valores de los parámetros de la recta de regresión, se requiere obtener aquella recta que minimice la suma de los cuadrados de los residuos, representada como SSE. Entonces calculamos SSE:

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Z_i - \bar{Z}_i)^2 = \sum_{i=1}^n (Z_i - a - bt_i)^2$$

Ahora bien, como deseamos obtener el valor mínimo de la SSE, se necesita que las derivadas parciales con respecto a los parámetros a y b sean iguales a cero. Cabe aclarar que la función SSE no tiene máximos, puesto que los valores de a y b pueden ser tan grandes como se quiera, por lo que al obtener el punto en que las derivadas parciales son cero se tratará forzosamente de un mínimo.

TESIS CON
FALLA DE ORIGEN

Diferenciando SSE con respecto a a y b , se obtiene:

$$\frac{\partial(SSE)}{\partial a} = -2 \sum_{i=1}^n (Z_i - a - bt_i)$$

$$\frac{\partial(SSE)}{\partial b} = -2 \sum_{i=1}^n (Z_i - a - bt_i) \cdot t_i$$

Las ecuaciones obtenidas se igualan ahora a cero, y reacomodando los términos obtenemos el siguiente sistema de ecuaciones:

$$an + b \sum_{i=1}^n t_i = \sum_{i=1}^n Z_i$$

$$a \sum_{i=1}^n t_i + b \sum_{i=1}^n t_i^2 = \sum_{i=1}^n t_i Z_i$$

Este sistema de dos ecuaciones con dos incógnitas se resuelve para obtener los valores de a y b . La solución del sistema de ecuaciones anterior da como resultado las siguientes fórmulas para obtener los valores de a y b :

$$b = \frac{n \sum_{i=1}^n t_i Z_i - \left(\sum_{i=1}^n t_i \right) \left(\sum_{i=1}^n Z_i \right)}{n \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i \right)^2}$$

$$a = \frac{\sum_{i=1}^n Z_i - b \sum_{i=1}^n t_i}{n}$$

Consideremos la serie de seis datos de la tabla 2, la cual representa el valor de un cierto inmobiliario usado en los años de 1987 a 1992. Los valores de esta tabla a simple vista nos muestran una tendencia decreciente y podemos suponer que es lineal, por lo que obtendremos la ecuación de regresión con las fórmulas descritas anteriormente.

i	t_i	Z_i
1	1987	6350
2	1988	5695
3	1989	5750
4	1990	5395
5	1991	4985
6	1992	4895

Tabla 2

TESIS CON
FALLA DE ORIGEN

Al realizar los cálculos de las sumatorias obtenemos los siguientes resultados:

$$\sum_{i=1}^6 t_i = 11,937 \quad \sum_{i=1}^6 Z_i = 33,070 \quad \sum_{i=1}^6 Z_i t_i = 65,787,885 \quad \sum_{i=1}^6 t_i^2 = 23,748,679$$

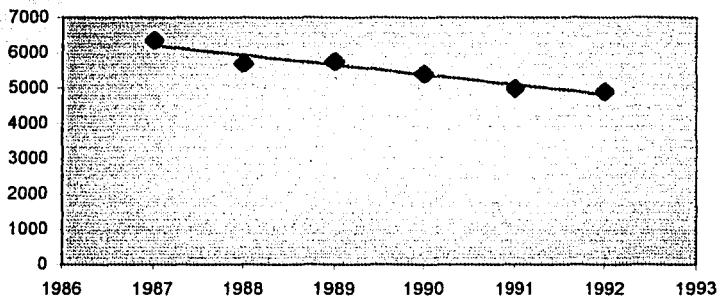
$$b = \frac{(6)(65,787,885) - (11,937)(33,070)}{(6)(23,748,679) - (11,937)^2} = -278.857$$

$$a = \frac{(33,070) - (-278.857)(11,937)}{6} = 560,297.952$$

Por lo tanto la ecuación de regresión lineal para esta serie de tiempo está dada por:

$$\bar{Z}_i = 560,297.952 - 278.857 \cdot i$$

La gráfica 8 muestra la línea de regresión para esta pequeña serie de tiempo:

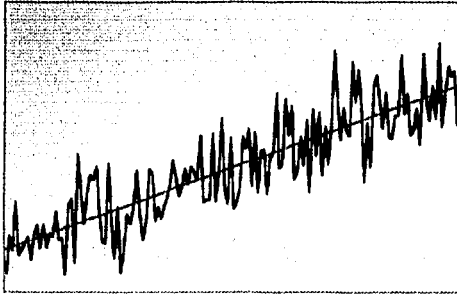


Gráfica 8 Recta de regresión

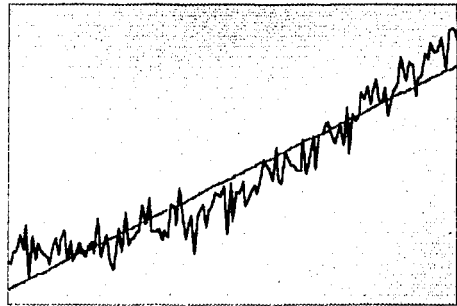
Cabe mencionar que en este ejemplo se han usado como valores de t_i los años correspondientes (1987, 1988,...). Esto se debe a que el modelo de regresión puede calcular la recta de regresión sin importar si la distancia entre las observaciones es uniforme, de hecho si se hubiera omitido por alguna razón el valor del inmobiliario en algún año, esto no habría afectado mucho dado que el modelo considera los valores del tiempo. A diferencia del modelo de medias móviles, éste no se ve afectado por la omisión de datos. Sin embargo, generalmente cuando manejamos una serie de tiempo, en donde las observaciones se hacen en intervalos uniformes de tiempo, y sobre todo cuando no se han omitido datos, el vector de valores del tiempo (1987, 1988,...) se puede sustituir por otro vector de números sucesivos (1,2,3,...), es decir, en lugar de 1987 utilizamos 1, 1988 lo

sustituimos por 2 y así sucesivamente. El efecto será un valor más pequeño en los parámetros de regresión, pero con el mismo resultado. No obstante, si por algún motivo el analista desea obtener los resultados con los valores de tiempo reales, puede hacerlo sin mucha dificultad mientras no tenga planeado realizar ningún cálculo manual.

En las series que se muestran a continuación se puede observar que cuando tenemos una serie con tendencia lineal o cuasilineal, la recta de regresión se ajustará bien al comportamiento de la serie, en tanto que si tenemos otro tipo de tendencia debemos ser cuidadosos para elegir el modelo de regresión apropiado, pues como se muestra en la gráfica 10 (datos simulados), una tendencia exponencial será pobremente representada por una recta de regresión.



Gráfica 9 Recta de regresión bien ajustada



Gráfica 10 Recta de regresión en una tendencia exponencial

Para los casos en que el comportamiento de la tendencia sea no lineal, es conveniente utilizar la regresión polinomial, o bien, un modelo exponencial o logarítmico, con el fin de aproximar mejor el comportamiento de la serie.

1.3 Tendencia Polinomial, Exponencial y Logarítmica

El método de mínimos cuadrados que se mostró anteriormente puede ser aplicado no sólo en la regresión lineal, sino también en la regresión polinomial. Supongamos que tenemos una serie cuya tendencia es no lineal, en ese caso podemos aplicar el método de mínimos cuadrados para

obtener un polinomio de grado N , que aproxime con mayor precisión la tendencia de la serie.

Entonces, se busca un polinomio de la forma

$$Z_i = \alpha_0 + \alpha_1 t_i + \alpha_2 t_i^2 + \dots + \alpha_N t_i^N + e_i$$

en donde $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N$ son los estimadores de los parámetros. Al igual que en el caso lineal,

la suma de los cuadrados de los residuos queda expresada como:

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Z_i - \alpha_0 - \alpha_1 t_i - \alpha_2 t_i^2 - \dots - \alpha_N t_i^N)^2$$

Y calculamos las derivadas parciales con respecto a $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N$:

$$\frac{\partial(SSE)}{\partial \alpha_0} = -2 \sum_{i=1}^n (Z_i - \alpha_0 - \alpha_1 t_i - \alpha_2 t_i^2 - \dots - \alpha_N t_i^N)$$

$$\frac{\partial(SSE)}{\partial \alpha_1} = -2 \sum_{i=1}^n (Z_i - \alpha_0 - \alpha_1 t_i - \alpha_2 t_i^2 - \dots - \alpha_N t_i^N) \cdot t_i$$

$$\dots\dots\dots$$

$$\frac{\partial(SSE)}{\partial \alpha_N} = -2 \sum_{i=1}^n (Z_i - \alpha_0 - \alpha_1 t_i - \alpha_2 t_i^2 - \dots - \alpha_N t_i^N) \cdot t_i^N$$

Posteriormente, igualamos a cero las derivadas parciales y reacomodando los términos obtenemos el sistema de ecuaciones de $N+1$ ecuaciones con $N+1$ incógnitas:

$$\alpha_0 n + \alpha_1 \sum_{i=1}^n t_i + \alpha_2 \sum_{i=1}^n t_i^2 + \dots + \alpha_N \sum_{i=1}^n t_i^N = \sum_{i=1}^n Z_i$$

$$\alpha_0 \sum_{i=1}^n t_i + \alpha_1 \sum_{i=1}^n t_i^2 + \alpha_2 \sum_{i=1}^n t_i^3 + \dots + \alpha_N \sum_{i=1}^n t_i^{N+1} = \sum_{i=1}^n Z_i t_i$$

$$\dots\dots\dots$$

$$\alpha_0 \sum_{i=1}^n t_i^N + \alpha_1 \sum_{i=1}^n t_i^{N+1} + \alpha_2 \sum_{i=1}^n t_i^{N+2} + \dots + \alpha_N \sum_{i=1}^n t_i^{2N} = \sum_{i=1}^n Z_i t_i^N$$

A diferencia del caso lineal, en donde debido a que es un sistema sencillo de resolver en estos términos se da la solución al sistema de ecuaciones de 2×2 , obtenemos primero todas las sumatorias necesarias y las sustituimos en el sistema de ecuaciones anterior, para después resolver el sistema de ecuaciones lineal mediante algún método conveniente, como el de Gauss-Jordan. La solución al sistema de ecuaciones son los valores de $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N$ que se sustituirán finalmente en el polinomio de grado N para obtener el modelo resultante.

La selección del grado de nuestro polinomio de regresión depende en buena medida de la complejidad de nuestra serie de tiempo. Conviene probar con diferentes grados de polinomio para elegir finalmente aquél que, siendo el más sencillo, sea a la vez el más representativo.

Adicionalmente a los modelos anteriores, podemos tener diversos tipos de modelos no lineales que tal vez se ajusten mucho mejor a nuestros datos. Sin embargo, obtener los valores de los parámetros requiere en este caso de resolver un sistema de ecuaciones no lineal, lo que nos obligaría en la mayoría de los casos a aplicar un método numérico como el de Newton-Raphson, siempre y cuando el modelo no lineal sea derivable.

No obstante, aunque los modelos no lineales generan un sistema de ecuaciones no lineal mediante el método de mínimos cuadrados, algunos casos específicos pueden linealizarse. Esto se logra aplicando una transformación a los datos originales. En la tabla 3 se muestran algunos modelos útiles de regresión y su transformación apropiada para obtener los parámetros mediante regresión lineal simple (Datos de Walpole & Myers, 1992).

Tipo de regresión	Modelo	Transformación Apropiada	Forma de regresión lineal simple
Exponencial	$y = \alpha e^{\beta x}$	$y^* = \ln y$	Regresión de y^* contra x
Exponencial (2)	$y = \alpha + \beta e^x$	$x^* = e^x$	Regresión de y contra x^*
Logarítmica	$y = \alpha + \beta \ln x$	$x^* = \ln x$	Regresión de y contra x^*
Potencial	$y = \alpha x^\beta$	$y^* = \ln y$ $x^* = \ln x$	Regresión de y^* contra x^*
Recíproca	$y = \alpha + \beta \left(\frac{1}{x}\right)$	$x^* = \frac{1}{x}$	Regresión de y contra x^*
Hiperbólica	$y = \frac{x}{\alpha + \beta x}$	$x^* = \frac{1}{x}; y^* = \frac{1}{y}$	Regresión de y^* contra x^*

Tabla 3 Transformaciones para modelos de regresión no lineales

TESIS CON
FALLA DE ORIGEN

En esta tabla, los valores de tiempo vienen representados en la variable x en lugar de t , en tanto que los valores de la serie están representados por " y " en lugar de Z . Cabe mencionar de nueva cuenta que en el caso de las series de tiempo no es usual realizar operaciones con los valores de tiempo, dado que las observaciones se hacen en intervalos fijos de tiempo. En realidad, el uso principal de los modelos de regresión se enfoca a observar la relación que existe entre dos o más conjuntos de variables, aunque no necesariamente implica una relación con el tiempo.

Al aplicar un modelo de regresión a una serie de tiempo, trabajamos bajo el supuesto de que el comportamiento de la serie puede explicarse únicamente en función del tiempo, y por esta misma razón, la ecuación de regresión obtenida será una función sólo del tiempo. Por estas razones, la ecuación de regresión obtenida no refleja la relación de un dato con los anteriores, y características como comportamientos periódicos o comportamientos especiales de la varianza se verán omitidas en la función resultante.

Por otro lado, los modelos de regresión que se presentan aquí sólo reflejan rasgos simples de la serie, pues con excepción de la regresión polinomial presentan trayectorias simples, sin puntos críticos (máximos o mínimos) o con cuando mucho uno. Por esta razón, las ecuaciones de regresión proporcionan casi exclusivamente datos sobre la tendencia. Así, una ecuación de regresión lineal con pendiente positiva, nos hablará de una tendencia creciente, en tanto que el nivel de ajuste que tenga cada modelo nos dirá si se trata de una tendencia lineal, exponencial, logarítmica, etc. De la misma manera, una recta de regresión lineal sin tendencia, o dicho de otra manera, con el parámetro β muy cercano a cero, implicará que la serie de tiempo es estacionaria en la media.

Para poder evaluar el nivel de ajuste de un modelo de regresión específico, obtenemos el coeficiente de correlación entre los datos de la serie y los valores generados por la función de regresión. Una aproximación apropiada del coeficiente de correlación ρ se obtiene con el

TESIS CON
FALLA DE ORIGEN

estimador r utilizando la siguiente fórmula, que se explicará más adelante en el análisis de la función de autocorrelación.

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

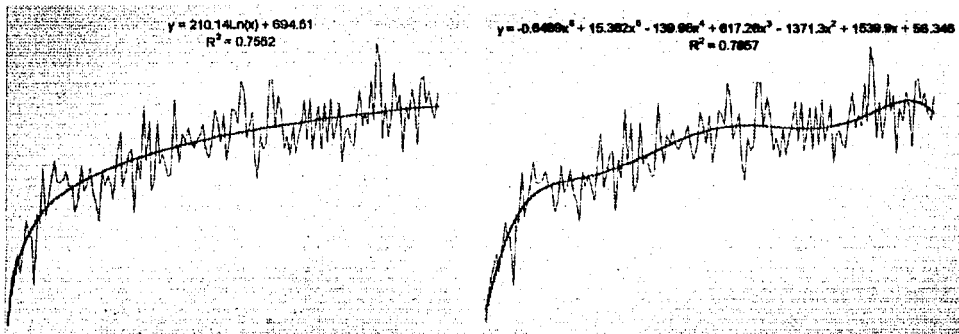
En este caso, una de las variables x ó y representará a la serie original, y la otra los valores de la función de regresión para los mismos tiempos. El coeficiente de correlación r representa la fuerza y tipo de correlación que existe ente los dos conjuntos de datos. Cuando el signo de la correlación es positivo, esto indica que a valores grandes de x se tienen valores grandes de y , y a valores pequeños de x se tienen valores pequeños de y . Por otro lado, cuando el signo de la correlación es negativo, esto indica que a valores grandes de x se tienen valores pequeños de y , y viceversa.

No obstante, aunque podemos utilizar el coeficiente de correlación para evaluar el ajuste del modelo de regresión, se suele utilizar con más frecuencia r^2 , que se le llama usualmente "*coeficiente de determinación muestral*", y este valor nos dirá el porcentaje de variaciones en un conjunto de datos que están linealmente correlacionados con el segundo conjunto de datos. De esta manera, para determinar si un modelo de regresión es apropiado, calculamos el coeficiente de determinación muestral y podemos decir que la correlación será:

- 1) **Débil:** Si $0 \leq |r| \leq .5$, o bien si $r^2 \leq .25$
- 2) **Fuerte:** Si $.8 \leq |r| \leq 1$, o bien si $r^2 \geq .64$
- 3) **Moderada:** en otro caso.

y seleccionaremos aquel modelo con el mayor coeficiente de determinación que a la vez sea el más sencillo. Es probable por ejemplo, que una regresión polinomial de orden alto supere un poco en la correlación a un modelo exponencial, en este caso siempre será preferible utilizar el modelo exponencial, a menos de que el nivel de ajuste sea muy pobre.

TESIS CON
FALLA DE ORIGEN



Gráfica 11 Modelo de regresión logarítmico

Gráfica 12 Modelo de regresión polinomial de orden 6.

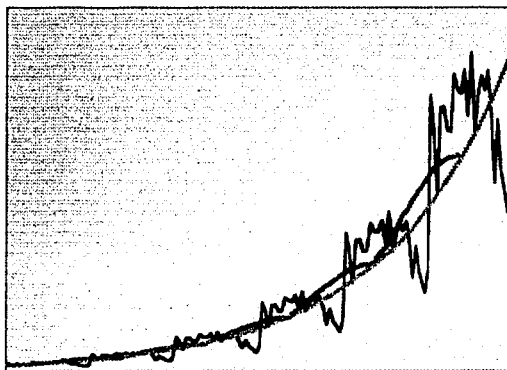
Las gráficas 11 y 12 (datos simulados) muestran la comparación de un modelo logarítmico con un modelo polinomial de orden 6. Nótese que las curvas que presenta el modelo polinomial aunque reflejan bien el comportamiento de la serie, hacen al modelo muy complejo en comparación con el modelo logarítmico, el cual se calcula mediante regresión lineal simple y tiene mayor valor predictivo. Por otro lado, podemos ver que $r^2 = 0.7552$, lo cual significa que tenemos un nivel de determinación muestral bastante aceptable, pues 75% de los valores de la ecuación de regresión están linealmente correlacionados con 75% de los datos de la serie original.

En la mayoría de los casos debe evitarse utilizar modelos de regresión polinomial debido a que por un lado pueden no tener mucho valor predictivo dado que se basan sólo en el tiempo y no consideran la relación de un dato con los anteriores, y por otro lado un comportamiento muy complejo de nuestro modelo, aunque se ajustará más a los datos, reflejará curvas que pueden deberse a datos atípicos o a razones de otro tipo. El uso de un modelo polinomial se justifica cuando ningún modelo se ajusta con suficiente precisión a la serie, pero en este caso debemos tener cuidado, porque se puede tratar de una tendencia estocástica, sin un patrón de crecimiento o decrecimiento específico o con saltos aleatorios ocasionales, sobre todo cuando el grado del polinomio requerido es grande para poder obtener un nivel aceptable de ajuste.

Cabe la posibilidad también, de aplicar un modelo de regresión sobre un conjunto de valores de medias móviles, sobre todo cuando el tamaño de las ventanas utilizadas es grande, pues de otra

manera se puede tener un sesgo en la función de media. Es importante asegurarnos en este caso de que la función de media obtenida por la regresión refleja bien la curva de medias móviles cuando el tamaño de la ventana tiende a valores grandes, puesto que un modelo de regresión bien ajustado a la media además de presentar curvas simples bien suavizadas, permite realizar una extrapolación con la que obtenemos valores estimados de media para los datos que se han perdido en el proceso de la obtención de las medias móviles.

Desafortunadamente, el cálculo de la función de media utilizando la regresión sobre un modelo de medias móviles tiene el serio defecto de que, aunque suaviza el efecto de los datos atípicos, la correlación que existe entre un dato y los anteriores provoca que la ecuación de regresión que se obtenga sea una estimación sesgada de la media, pues refleja pequeñas variaciones debidas a elementos estacionales o estocásticos. En particular en el caso de las fluctuaciones estacionales, la función de media sólo se reflejará bien cuando el tamaño de la ventana sea igual al periodo de la fluctuación estacional, o bien, a un múltiplo de éste. La gráfica 13 muestra una comparación entre un modelo exponencial y un modelo de medias móviles, usando la serie de la gráfica 2.



Gráfica 13 Modelo Exponencial vs Medias Móviles

En esta gráfica se puede apreciar que existe una estacionalidad marcada, la línea en rojo representa el modelo de medias móviles, en tanto que la línea azul representa el modelo de regresión exponencial. Nótese que ambos modelos concuerdan en los puntos intermedios de cada

periodo de la fluctuación estacional. Lo que significa que, el promedio de los valores de cada periodo se ajusta muy bien al modelo de crecimiento exponencial, el cual no se ve afectado por curvas debidas a la variación estacional. La ventana utilizada aquí es igual al periodo de la fluctuación estacional.

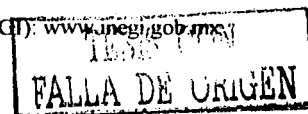
De la misma manera en que podemos utilizar funciones exponenciales, recíprocas, logarítmicas, etc., también se pueden utilizar las funciones seno y coseno para aproximar series cíclicas. El uso de los senos y cosenos para representar funciones se detallará más adelante en el análisis de Fourier.

1.4 Modelos de tendencia estocástica

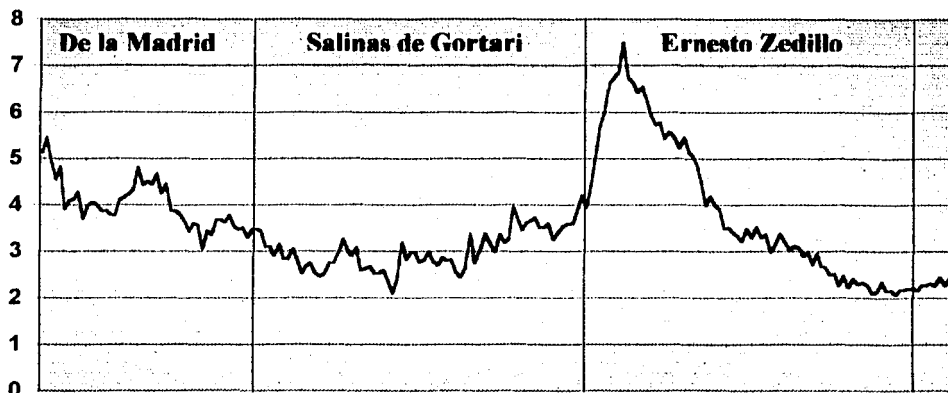
Los modelos de regresión mostrados anteriormente parten del supuesto de que la serie tiene una tendencia con comportamiento determinístico. Si este supuesto es verdadero, nuestro modelo de regresión no sólo aproximará la serie, sino que además podrá realizar predicciones muy confiables. Sin embargo, existen diversas series en las que el comportamiento no es tan sencillo de modelar, y en los casos en que debemos recurrir a una regresión polinomial, existe una gran probabilidad de que el comportamiento de la tendencia no sea determinístico, sobre todo si el grado del polinomio de regresión es alto. En este caso, los cambios en la tendencia pueden deberse a eventos aleatorios que no forman parte del comportamiento sistemático en general.

La gráfica 14, obtenida con datos del INEGI⁴, representa la tasa de desempleo mensual desde 1985 hasta el año 2001. En esta gráfica las líneas azules separan los sexenios de los presidentes correspondientes. Observamos al final del sexenio de Miguel de la Madrid una tendencia lineal decreciente, posteriormente el sexenio de Salinas de Gortari presenta un ligero ascenso en forma de curva y, en el punto final de este sexenio un evento aleatorio mejor conocido como "error de diciembre" provoca una crisis generando desempleo, lo cual rompe totalmente con el patrón de

⁴ Página Web del Instituto Nacional de Estadística, Geografía e Informática (INEGI): www.inegi.gob.mx



comportamiento anterior. Se podría asumir que el comportamiento en el resto del sexenio de Zedillo se encuentra íntimamente ligado al error de diciembre.



Gráfica 14 Tasa de desempleo mensual de 1985 a 2001

Este tipo de eventos aleatorios en diversas series de tiempo pueden resultar muy significativos, pues un simple cambio atípico cambia totalmente la forma en como se comporta la serie. El ejemplo anterior es un caso muy sensible a variaciones aleatorias, sobre todo porque existe una dependencia muy alta entre un dato y el que le precede. No olvidemos que el desempleo en un mes dado se da a partir del número de desempleados que había el mes anterior.

Este tipo de tendencia se le conoce como "tendencia estocástica", y se caracteriza por ser resultado de una acumulación de fluctuaciones aleatorias en el tiempo. Cabe señalar que el ejemplo anterior además de tener factores de tendencia estocástica tiene también comportamientos sistemáticos durante ciertos subintervalos de tiempo, debidos a políticas de empleo o a diversos factores que se aplican en cada sexenio. Un ejemplo de este tipo de políticas puede ser el tratado de libre comercio, que aunque parte del sexenio de Salinas es un factor para disminuir el desempleo durante el sexenio de Zedillo, pues a muchas empresas norteamericanas les conviene más el pago de la mano de obra en pesos que en dólares, lo cual propicia que muchas empresas norteamericanas establezcan centros de producción en México.

Cuando se tiene una serie con tendencia estocástica, es obvio que no puede modelarse mediante un modelo determinístico, puesto que los cambios en la tendencia son acumulaciones de fluctuaciones aleatorias y no precisamente de un patrón sistemático. En este caso, un método para eliminar esta dependencia de las fluctuaciones aleatorias anteriores, es realizar diferencias a la serie original. Es decir, restarle al valor de una observación, el valor de la que le precede. Así, Nuestra nueva serie diferenciada queda:

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

De esta manera obtenemos una nueva serie que representa los incrementos entre una observación y otra. Si los cambios en la tendencia son realmente consecuencia de una acumulación de valores aleatorios, aplicando diferencias eliminaremos el comportamiento aleatorio de la tendencia. Cuando los incrementos entre un dato y otro sean en general constantes, una diferencia bastará para obtener una serie estacionaria en la media, en tanto que cuando los incrementos crecen o decrecen tendremos que realizar una segunda diferencia como se muestra a continuación.

$$W_t = \nabla^2 Z_t = \nabla Z_t - \nabla Z_{t-1} = (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2})$$

La primera diferencia representa una "velocidad" de crecimiento o decrecimiento (incremento por unidad de tiempo), mientras que la segunda diferencia representa la "aceleración" de estos incrementos. En general, para eliminar la tendencia se puede aplicar el número de diferencias que se necesiten.

$$W_t = \nabla^d Z_t$$

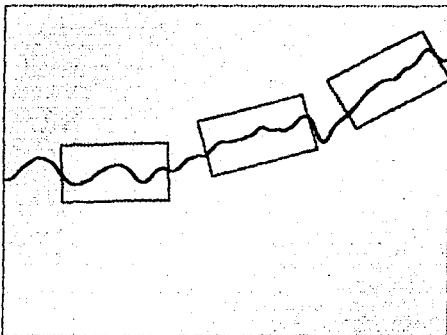
Ahora bien, existen muchas series de tiempo no estacionarias que, debido a diversos componentes de equilibrio, tienen una gran similitud entre sus diferentes partes, excepto por la diferencia en sus medias locales. Box y Jenkins se refieren a este tipo de series como "homogéneas no estacionarias". Básicamente el uso de las diferencias se enfoca a las series homogéneas no estacionarias, aunque bien puede aplicarse en una serie con comportamiento determinístico.

Un ejemplo de serie homogénea no estacionaria es cuando la d -ésima diferencia da como resultado un ruido blanco, así tenemos que

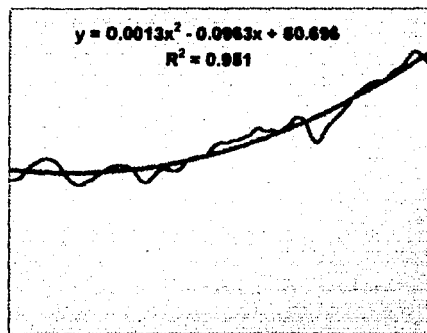
$$\nabla^d Z_t = a_t$$

en donde a_t es ruido blanco. En este ejemplo podemos ver que cuando $d = 1$, es el modelo de caminata aleatoria $Z_t = Z_{t-1} + a_t$, en donde el nivel de la serie en el tiempo t es $\mu_t = Z_{t-1}$, el cual está sujeto a la variación aleatoria en el tiempo $(t-1)$. Es decir, la media de la serie Z_t cambia a través del tiempo aleatoriamente, y decimos que la serie tiene una tendencia estocástica.

El uso de las diferencias no siempre resulta eficiente, y debemos cuidarnos de no abusar de su uso. No se trata sólo de obtener una serie más bonita que elimine la tendencia, sino que debemos tener en cuenta que, en primer lugar lo que se está analizando ya no es el comportamiento real de la serie, sino el comportamiento de los incrementos de ésta. Las propiedades espectrales de la nueva serie son en realidad las de los incrementos. Por otro lado no debemos caer en el error de pensar que por un acto de magia "hacemos" estacionaria la serie, sino que tan sólo estamos aprovechando que, aunque la serie no es estacionaria, sus incrementos sí lo son. De esta manera, para que las diferencias sean útiles, es necesario que la serie sea homogénea, es decir, que aunque la tendencia no sea estacionaria, la serie presente una gran semejanza entre diferentes partes, excepto por la diferencia entre sus medias locales.

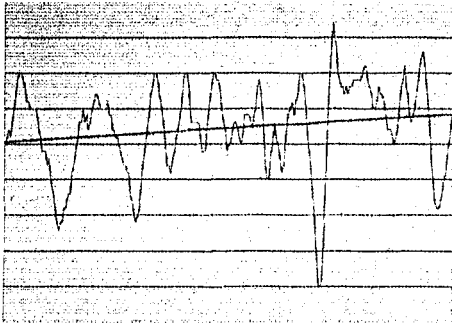


Gráfica 15 Actividad Industrial 1980-2001

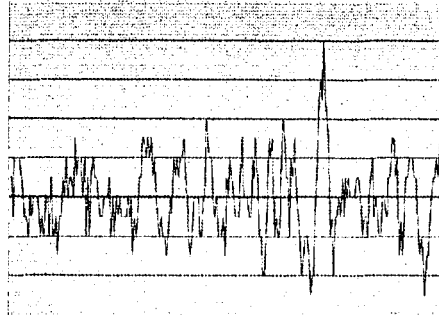


Gráfica 16 Regresión Polinomial de orden 2

TESIS CON
FALLA DE ORIGEN



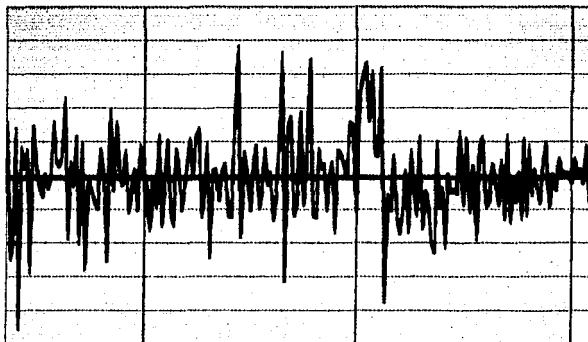
Gráfica 17 Primera diferencia Ordinaria



Gráfica 18 Segunda diferencia ordinaria

La gráfica 15 representa los niveles de actividad industrial desde 1980 (Datos del INEGI). Los cuadros que se señalan muestran una cierta estabilidad en la media durante algunos subintervalos de tiempo. Al observar la inclinación de los cuadros observamos un crecimiento en la pendiente. La gráfica 16 muestra el modelo de regresión polinomial de orden 2, que fue el que mejor se ajustó a los datos, pues tiene un coeficiente $r^2 = 0.951$ contra la serie original, y un coeficiente $r^2 = 0.98$ contra las medias móviles. Al aplicar una diferencia ordinaria se obtiene la serie que aparece en la gráfica 17, mostrando una tendencia lineal creciente en los incrementos de actividad industrial. Por otro lado, al aplicar una segunda diferencia ordinaria tenemos una serie estacionaria en la media, lo cual representa que la aceleración del crecimiento de la actividad industrial es constante.

En general, una serie que se ajusta bien a un modelo Polinomial de orden N , al aplicarle diferencias tendrá una N -ésima diferencia estacionaria en la media. Sin embargo, en casos con tendencia estocástica, como el planteado en el ejemplo de desempleo (Gráfica 14), aunque para ajustar la serie se requeriría un polinomio de grado alto, pueden necesitarse tan sólo un par de diferencias, o inclusive una. La gráfica 19 muestra una diferencia ordinaria para el ejemplo de desempleo.



Gráfica 19 Una diferencia Ordinaria, desempleo

Al igual que el ejemplo de desempleo, una serie que presente comportamiento de caminata aleatoria se verá reducida a ruido blanco con tan sólo una diferencia, aunque inicialmente pueda parecer que requerirá muchas diferencias. En general, las series más comúnmente analizadas requerirán pocas diferencias. Sin embargo, es importante no sobre-diferenciar la serie, sobre todo si se va a obtener un modelo para predicción a partir de la serie diferenciada, pues el modelo puede complicarse entre más diferencias se realicen.

Capítulo II

Análisis de la Varianza y Autocovarianza

En el capítulo anterior se mostraron algunos métodos para analizar la tendencia de una serie de tiempo. Particularmente es importante señalar que cualquier serie de tiempo se puede analizar con métodos de ventanas, pero como se desprende de la comparación entre estos métodos, la obtención de un modelo predictivo para la tendencia requiere de otros métodos como la regresión, es decir, aunque el análisis mediante ventanas se aplica a cualquier serie de tiempo, su utilidad principal es la de analizar y no predecir. No obstante es posible combinar los métodos de medias móviles y de regresión en algunos casos específicos.

En este segundo capítulo se plantea ahora la cuestión sobre la utilidad del método de ventanas para realizar el análisis y la predicción de valores en la varianza y la autocovarianza. Cabe mencionar que en este capítulo, a diferencia del anterior, se tratarán casi exclusivamente métodos de ventaneo. De esta manera, aunque existen otros métodos para el análisis de la autocovarianza, como la metodología de Box-Jenkins, no se han incluido aquí. Esto se debe a que la metodología se aplica específicamente a una autocovarianza estacionaria, no obstante se mencionan aquí las características en comparación con este método, y la forma en como pueden complementarse estos métodos. En lo que se refiere al análisis de la varianza, se muestra también el tema de la estabilización de la varianza.

2.1 Función de Varianza Muestral

Una vez que obtengamos un modelo apropiado para representar la función de media, podemos proceder al análisis de la varianza de la serie. Esto se debe a que el cálculo de la varianza se basa en el estudio de las variaciones de los datos con respecto a la media. De esta manera, cuando obtenemos una función de media, ya sea por medio de diferencias ordinarias (media constante) o mediante una ecuación de regresión, el cálculo de la función de varianza puede hacerse a partir de

TESIS CON
FALLA DE ORIGEN

los valores de media obtenidos. Cuando una serie presenta una varianza constante, el cálculo de la varianza se realiza mediante la fórmula:⁵

$$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{n}$$

es decir, se trata de un promedio de las desviaciones de la media al cuadrado.

Obtener una varianza constante no presenta mayor dificultad, sin embargo es muy común obtener una serie con comportamiento creciente en la varianza y en algunas veces decreciente. Cuando se tiene un comportamiento sistemático de la varianza (no constante), se le conoce como "*heteroscedasticidad*", la cual puede ser analizada mediante el método de ventanas mostrado en el análisis de la tendencia.

2.1.1 Varianza heteroscedástica

El procedimiento para calcular la función de varianza muestral puede realizarse utilizando tanto los valores de medias móviles como los valores de regresión, o en su caso el valor de media constante. El método que se describe a continuación utiliza prácticamente el mismo concepto de ventana y escala utilizada para las medias móviles, de esta manera para obtener la función de varianza necesitamos escoger un tamaño de ventana apropiado, es decir, suficientemente grande para apreciar el comportamiento general de la varianza.

Aunque en principio estamos aplicando el mismo método que en las medias móviles, no necesariamente implica que la función de media se tenga que obtener por este mismo método, pero es importante aclarar que si se trata de una media obtenida mediante un modelo de regresión

⁵ A diferencia de una muestra estadística común, los valores de una serie de tiempo por lo general son dependientes entre sí, además de no presentar necesariamente una distribución normal, por lo que el estimador más apropiado es el que se describe aquí.

o una media constante, debemos estar muy seguros de que el modelo de regresión o la media constante son apropiados, esto significa que debe asemejarse a una función de medias móviles a escalas grandes.

Para proceder con el análisis de la varianza conviene obtener primero una versión de la serie con la tendencia estabilizada, para esto diferenciamos la serie con respecto a los valores de media en los mismos tiempos

$$w_i = z_i - \bar{Z}_i$$

Aún para valores de media constantes, se puede hacer también la diferencia con el objeto de obtener directamente la función de varianza y facilitar los cálculos posteriores con una media de cero, puesto que al diferenciar se obtienen las desviaciones de la media. Cabe señalar que en el caso de las medias que se obtuvieron mediante el método de medias móviles, sólo podrán obtenerse las diferencias para los datos en los cuales se dispone de una media, recordando que se perdieron $s - 1$ datos. Por otro lado, con el fin de facilitar las diferencias es recomendable utilizar escalas impares para las medias móviles.

Básicamente el procedimiento consiste en calcular la varianza, con la fórmula descrita anteriormente, para subintervalos de s datos. Entonces, podemos elegir el tamaño del subintervalo de acuerdo al tipo de variaciones que se quieren analizar. Aunque en la mayoría de los casos sólo nos interesa el comportamiento general a grandes rasgos, en ocasiones puede ser útil conocer los cambios de la varianza en intervalos de tiempo pequeños, que bien pueden estar relacionados con el nivel de la media o presentar incrementos o decrementos periódicos.

Retomando el ejemplo de la serie de seis datos que se dio en la sección de medias móviles, podemos calcular los valores de las varianzas para los intervalos de tiempo de tamaño 3. Podemos ver que la media ha sido calculada mediante medias móviles con ventana de tamaño 3. Posteriormente realizamos las diferencias para obtener la nueva serie cuya media es cero. De esta manera, se calcula el promedio de estas diferencias al cuadrado.

ESTO CON
FALLA DE ORIGEN

$$\frac{\sum_{i=1}^n (Z_i - \bar{Z}_i)^2}{n}$$

en donde $\bar{Z}_i = 0$. Los valores obtenidos se muestran en la tabla 4, en este caso particular se utilizó el mismo tamaño de ventana para la media y la varianza, aunque no necesariamente deben ser iguales.

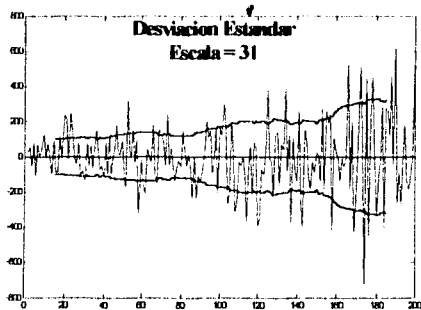
Serie original Z_i	Media con ventana de tamaño 3 \bar{Z}_i	$Z_i - \bar{Z}_i$	Varianza Escala = 3 $Var(Z_i)$
9.6208	9.3488	-0.2167	0.1243
9.1321	8.8317	0.4617	0.1129
9.2934	8.4053	-0.3357	0.0549
8.0696	7.9663	-0.1135	
7.8528	8.1743	-0.1978	
7.9765			
8.6937			

Tabla 4

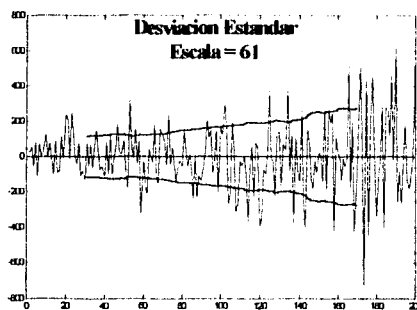
En esta tabla, se muestran primero los valores de las medias móviles en los que se han perdido dos datos, posteriormente se realiza la diferencia para estabilizar la tendencia, y al calcular la varianza se vuelven a perder otros 2 datos. Así, la primera varianza se calcula:

$$s_1^2 = \frac{(-0.2167)^2 + (0.4617)^2 + (-0.3357)^2}{3} = 0.1243$$

puede observarse que como también en este caso se pierden datos, al aplicar el método de ventanas a las medias y después a la varianza puede resultar en la pérdida de una cantidad considerable de datos. Las gráficas 20 y 21 muestran la desviación estándar obtenida por este método para una serie de 200 datos, se observa en este caso un crecimiento en el valor de la varianza a través del tiempo.



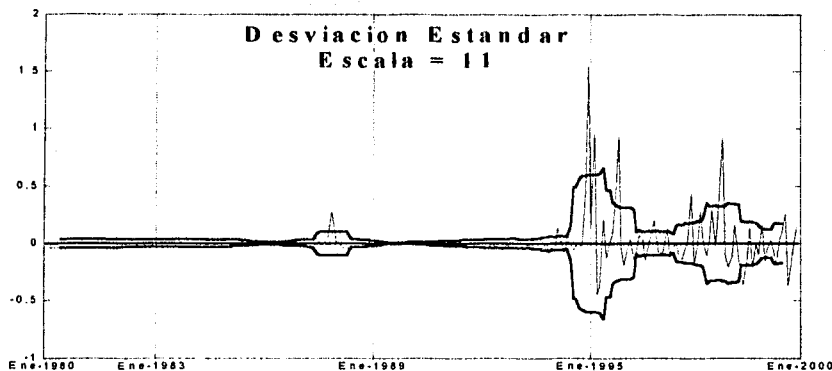
Gráfica 20 Desviación estándar por medio de ventanas, Escala = 31



Gráfica 21 Desviación estándar por medio de ventanas, Escala = 61

La gráfica 20 (datos simulados) utiliza una escala 31, y refleja las disminuciones de la desviación estándar en algunos intervalos de tiempo, en tanto que la gráfica 21, realizada en escala 61 nos muestra que en realidad el comportamiento creciente de la desviación estándar tiende a ser lineal.

Aunque puede utilizarse también en este caso un modelo de regresión, por lo general muchas series de tiempo en la práctica no presentan un comportamiento totalmente sistemático en la varianza, sino que puede presentar patrones de comportamiento que sólo duran un intervalo específico de tiempo. Un ejemplo de este tipo de comportamiento se muestra en la gráfica 22, que representa las variaciones en el valor del dólar con respecto al mes anterior, desde 1980 al 2000 (Datos del INEGI).



Gráfica 22 Variaciones en el tipo de cambio mensual (1980-2000)

Esta serie tiene cambios en el patrón de varianza cada sexenio, la primera parte muestra un patrón ligeramente decreciente, después una devaluación al inicio del sexenio de Salinas, posteriormente una varianza ligeramente creciente, luego otra devaluación sexenal (error de diciembre), y luego una varianza decreciente, pero mucho más alta que en los dos sexenios anteriores.

De nueva cuenta el análisis por medio de ventanas nos permite analizar series con diferentes patrones de comportamiento en diferentes subintervalos de tiempo. En este tipo de series el análisis por medio de ventanas supera por mucho a la información que puede proporcionarnos un modelo de regresión con respecto a la media o la varianza, aunque desde luego no es posible realizar predicciones de la serie por este método. De hecho, realizar predicciones de una serie cuyos patrones de comportamiento cambian constantemente resulta extremadamente complicado, salvo que se puedan identificar los periodos de cambio de estos patrones, como en el ejemplo en que se mostró un cambio sexenal.

2.1.2 Transformaciones a la Varianza

En el caso de la función de media, observamos que era posible pseudoeliminar⁶ la tendencia, ya sea mediante un modelo de regresión y luego diferenciando los valores de la serie con los de regresión, o bien, realizando diferencias ordinarias. Al igual que en la tendencia, la varianza también puede estabilizarse, con la finalidad de facilitar el análisis de una serie de tiempo mediante la función de autocovarianza, esto se logra realizando una transformación a los datos originales, de acuerdo al comportamiento que presente la varianza. En la mayoría de los casos no es sencillo apreciar a simple vista cual es la más apropiada, pero pueden probarse diferentes transformaciones para determinar cual es la más adecuada.

⁶ Su utiliza aquí el término "pseudoeliminar" en el sentido de que no se está realmente descartando, modificando u "olvidando" que existe una tendencia, sino que en lugar de trabajar con la serie original, utilizamos la característica de que las desviaciones de los datos con respecto al modelo de regresión tienen una media constante, o bien, en el caso de las diferencias, que los incrementos tienden a ser constantes.

En la siguiente tabla se muestran algunas transformaciones que pueden aplicarse a los datos con el fin de pseudoeliminar la varianza de la serie:

Tipo	Transformación
Inversa	$W = \frac{1}{Z}$
Logaritmo	$W = \ln Z$
Raíz Cuadrada	$W = \sqrt{Z}$
Inversa del Logaritmo	$W = \frac{1}{\ln Z}$
Inversa de la Raíz Cuadrada	$W = \frac{1}{\sqrt{Z}}$

Existen otras transformaciones que se pueden realizar a las series de tiempo con el fin de pseudoeliminar la varianza, podemos notar que estas transformaciones se asemejan a las que se aplicaron para obtener modelos de regresión no lineal. De hecho, el uso de las transformaciones en este caso adolece de las mismas limitaciones que se mencionaron para la regresión no lineal, es decir, sólo puede aplicarse en aquellos casos donde la propiedad espectral (media o varianza) se puede linealizar. Dicho de otra manera, la transformación sólo se aplica cuando el nivel de la varianza es proporcional al nivel de media de la serie. Como ejemplo podemos mencionar el caso de la transformación logaritmo, que puede linealizar una tendencia exponencial, y a la vez servirnos para estabilizar la varianza.

Si se desea estabilizar tanto la media como la varianza, debe realizarse primero la transformación necesaria a la varianza, para posteriormente estabilizar la tendencia, esto se debe a la dependencia que existe entre el nivel de media y la varianza de la serie. Consideremos que una transformación a los datos en la mayoría de los casos no sólo afecta a la varianza de la serie. La estabilización de la varianza sólo es útil cuando existe esta dependencia, en otro caso, probablemente no nos servirá de mucho estabilizar la varianza.

TESIS CON
FALLA DE ORIGEN

2.2 Función de Autocovarianza

Sin duda la propiedad espectral más útil para realizar predicciones en una serie de tiempo es la función de autocovarianza, puesto que ésta representa la relación que tiene un dato con los anteriores. La función de autocovarianza es utilizada para obtener los modelos autoregresivos y de medias móviles del tipo

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

conocidos como modelos ARMA. Los parámetros necesarios para obtener una buena estimación se pueden determinar utilizando la función de autocovarianza. De hecho, la metodología de Box y Jenkins es prácticamente un análisis de la autocovarianza.

La función de autocovarianza es una medida de la dependencia entre un dato y los que lo preceden, de esta manera, la autocovarianza entre los valores de la serie que se encuentran separados por un intervalo de tiempo k está dada por

$$\hat{\gamma}_k = \text{Autocov}(Z_t, Z_{t+k}) = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z}_t)(Z_{t+k} - \bar{Z}_{t+k})}{n}$$

Entonces, el producto de las desviaciones con respecto a la media local de cada dato se promedia y se obtiene un indicador de la dependencia entre los datos separados por un intervalo de tiempo k . Nótese que se trata de un estimador sesgado de la autocovarianza, dado que aunque se realizan $n-k$ productos, se promedian como si se tratara de n grados de libertad⁷. Esto se debe a que, cuando el parámetro μ es desconocido, tanto $n-k$ como n generan estimadores sesgados de γ_k , pues básicamente la estimación de la autocovarianza se realiza utilizando las desviaciones de las medias locales, y en este caso el estimador descrito en la ecuación anterior es, en realidad, el "menos sesgado". Dicho de otra manera, si consideramos que entre más grande

⁷ Se entiende por "grados de libertad" el número de observaciones independientes en una muestra, en este caso se tiene una muestra de $n-k$ productos.

sea el valor de k , tendremos una menor cantidad de productos, y por tanto, la estimación de la autocovarianza será cada vez menos confiable, pues considera una menor muestra de datos. De esta manera, dividiendo entre n en lugar de $n-k$, subestimamos los valores de la autocovarianza para valores de k grandes, lo que genera un estimador asintóticamente insesgado⁸.

El signo de los valores de la autocovarianza nos indica el tipo de dependencia que existe entre los datos separados por un intervalo de tiempo k . Cuando el valor de la autocovarianza es positivo, significa que a valores grandes de Z_t corresponden valores grandes de Z_{t+k} , y que a valores pequeños de Z_t corresponden valores pequeños de Z_{t+k} . Por otro lado, cuando la autocovarianza es negativa significa que a valores grandes de Z_t corresponden valores pequeños de Z_{t+k} , y viceversa.

Sin embargo, aunque el signo de la autocovarianza nos indica la forma en que se relacionan los datos, la función de autocovarianza no proporciona información respecto a la fuerza de esta relación, por lo que para obtener un indicador de ésta se utiliza la función de autocorrelación (ACF)

$$\hat{\rho}_k = r_k = \frac{\hat{\gamma}_k}{\hat{\sigma}^2} = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

Dividiendo la autocovarianza entre la varianza de la serie se obtiene el indicador de la fuerza de correlación entre las observaciones separadas por un intervalo de tiempo k . Desafortunadamente, este indicador supone que la varianza de la serie es constante, por lo que la función de autocorrelación puede no ser muy eficiente para analizar series con varianza no estacionaria. Sin embargo, es aquí donde podemos utilizar una vez más el análisis por medio de ventanas para observar cambios en la función de varianza y autocovarianza a través del tiempo.

⁸ Puede consultarse respecto al tema de la estimación de la autocovarianza el artículo de Donald B. Percival, "Three curious properties of the sample variance and autocovariance for stationary processes with unknown mean", 1999.

2.2.1 Función de Autocorrelación

Como es de suponerse, cuando una serie de tiempo cambia sus patrones de comportamiento a través del tiempo, el problema de analizar la serie se complica, especialmente porque puede hacer que el comportamiento de la serie sea casi imposible de predecir. De hecho, la metodología de Box y Jenkins aunque puede analizar y predecir ciertas series de tiempo no estacionarias en la media y la varianza, es prácticamente inútil si la autocovarianza es no estacionaria.

De esta manera, si trabajamos bajo el supuesto de que el patrón de comportamiento de la serie permanece constante, podemos utilizar la metodología de Box-Jenkins para obtener un modelo de predicción, aun cuando se tenga una tendencia o heteroscedasticidad, puesto que éstas pueden pseudoeliminarse. Sin embargo, cuando el comportamiento sea no estacionario y no determinístico, las predicciones serán fiables a muy corto plazo.

Una serie de tiempo no estacionaria en la autocovarianza no nos deja entonces la opción de predecir, salvo en el caso poco probable y complicado de que el cambio en el patrón de comportamiento sea sistemático, pero aún en este caso obtener un modelo que represente las variaciones en el tiempo de los coeficientes de correlación resulta extremadamente complicado. Sin embargo, aunque no podemos predecir los valores de una serie de tiempo no estacionaria en la autocovarianza, es posible analizar los cambios de patrones y dependencias entre datos a través del tiempo, y de esta manera, podemos utilizar ventanas para realizar estos análisis.

Sin embargo, el procedimiento no es tan sencillo como parecería a primera vista: Recordemos que el análisis de la tendencia mediante ventanas requería del uso de tres dimensiones (escala, tiempo y media), en tanto que el de la varianza requería también tres dimensiones (escala, tiempo y varianza), en estos casos para evitar utilizar gráficas tridimensionales (como la que se mostró en las medias móviles en el capítulo primero), podemos realizar la gráfica bidimensional para una escala específica, y apreciar de esta manera en dicha escala el comportamiento de la propiedad

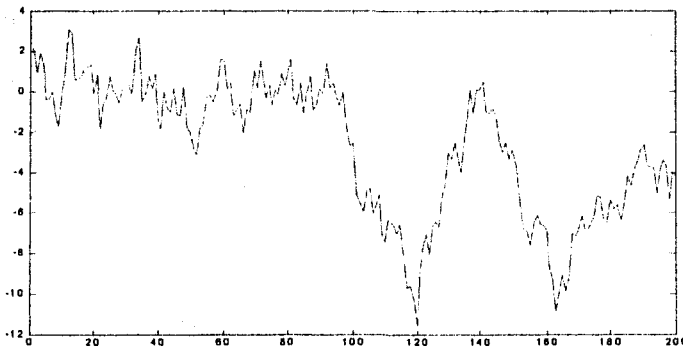
TESIS CON
FALLA DE ORIGEN

espectral a través del tiempo. Pero al momento de realizar un análisis mediante ventanas de la autocovarianza requerimos de cuatro dimensiones (escala, tiempo, intervalo que separa a dos observaciones, y valor de autocorrelación). De esta manera, como se trata de un conjunto de datos en cuatro dimensiones, debemos ser muy cuidadosos con las representaciones gráficas que utilizamos.

Consideremos una serie de tiempo de 200 datos con dos patrones de comportamiento diferentes, definidos de la siguiente manera:

$$Z_t = \begin{cases} 0.4Z_{t-1} + e_t, & t = \overline{1,100} \\ Z_{t-1} + e_t, & t = \overline{101,200} \end{cases}$$

La gráfica de una simulación de estos datos se muestra en la figura 23, en donde aunque tenemos dos patrones diferentes de comportamiento en el tiempo, si nosotros no supiéramos de antemano esto, sería un poco complicado darnos cuenta cuando la autocovarianza es no estacionaria. De hecho, es importante notar que esta gráfica se asemeja mucho a las gráficas económicas, en donde los eventos estocásticos (devaluaciones, crisis, atentados, etc.) provocan un cambio en la función de autocovarianza.



Gráfica 23 Serie de tiempo no estacionaria en la autocovarianza

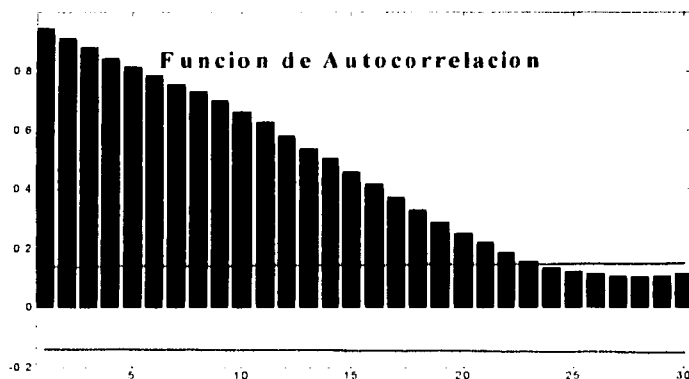
Los primeros cien datos de la serie muestran en general un comportamiento sin tendencia y con varianza estable, en tanto que la segunda parte de los datos es el modelo de caminata aleatoria.

Si suponemos que la media, la varianza y la autocovarianza de la serie son constantes (lo cual no se aplica en la serie anterior), calculamos la función de autocorrelación utilizando la fórmula

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \text{ en donde}$$

$$\hat{\gamma}_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{n}$$

La gráfica 24 muestra la función de autocorrelación que se obtiene en la serie de la gráfica 23, en donde obtenemos una función de autocorrelación decreciente infinita, lo cual era de esperarse dado que los dos modelos utilizados para generar la serie son del tipo autoregresivo AR(1).

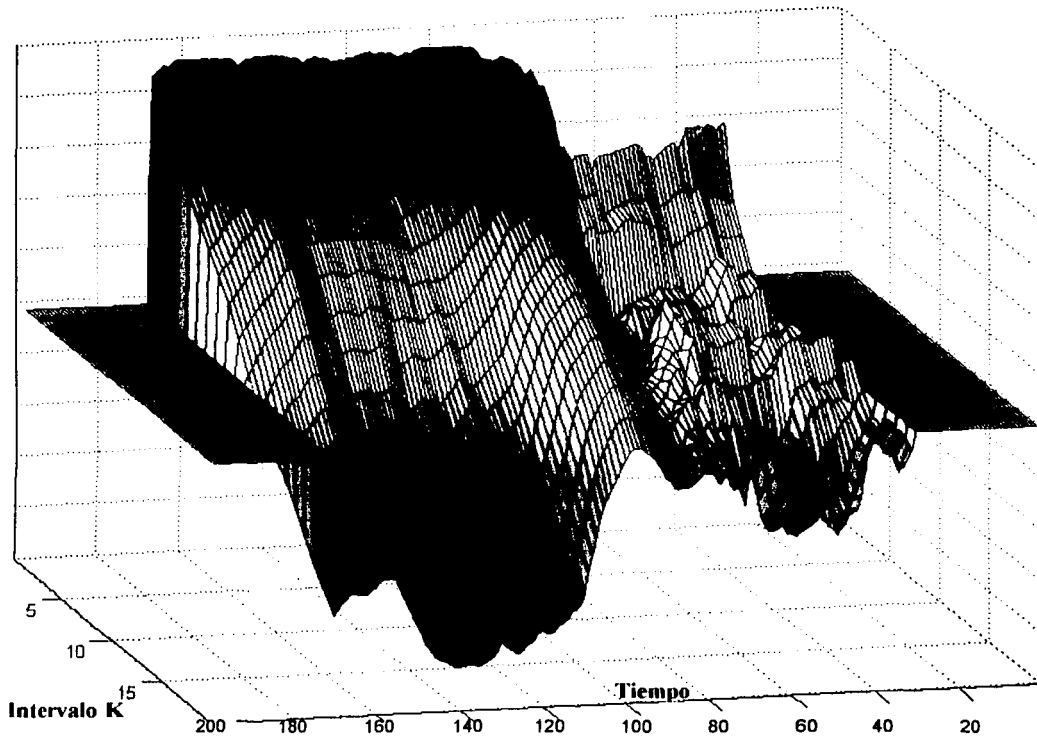


Gráfica 24 Función de Autocorrelación

La línea roja representa el intervalo de confianza de 95% para aceptar la hipótesis de que el valor de una cierta correlación sea cero. De esta manera, cualquier correlación estimada cuyo valor quede dentro del intervalo de confianza se considerará cero, y los valores que sobresalen ligeramente serán considerados a criterio del analista.

Como podemos ver en la gráfica anterior no se aprecian los dos patrones de comportamiento de la serie debido a que el cálculo realizado de esta manera supone que la autocovarianza es estacionaria. En consecuencia, para apreciar cambios en la autocovarianza, podemos obtener la

ACF para ciertos subintervalos de tiempo. Entonces, si utilizamos un tamaño de ventana de 60 datos, por ejemplo, y calculamos la ACF para cada 60 datos recorriendo en el tiempo, cada vez que calculamos la ACF tendremos el valor de cada una de las autocorrelaciones para esos 60 datos. El resultado de este procedimiento se puede mostrar en una gráfica tridimensional para cada escala, cuando se quiera ver el cambio general en la función de autocovarianza a través del tiempo. La gráfica 25 muestra las autocorrelaciones con escala de 60 y calculado para las primeras 20 autocorrelaciones. Debido a la complejidad de la gráfica no se han agregado en ésta los intervalos de confianza para la hipótesis de que los valores sean cero.

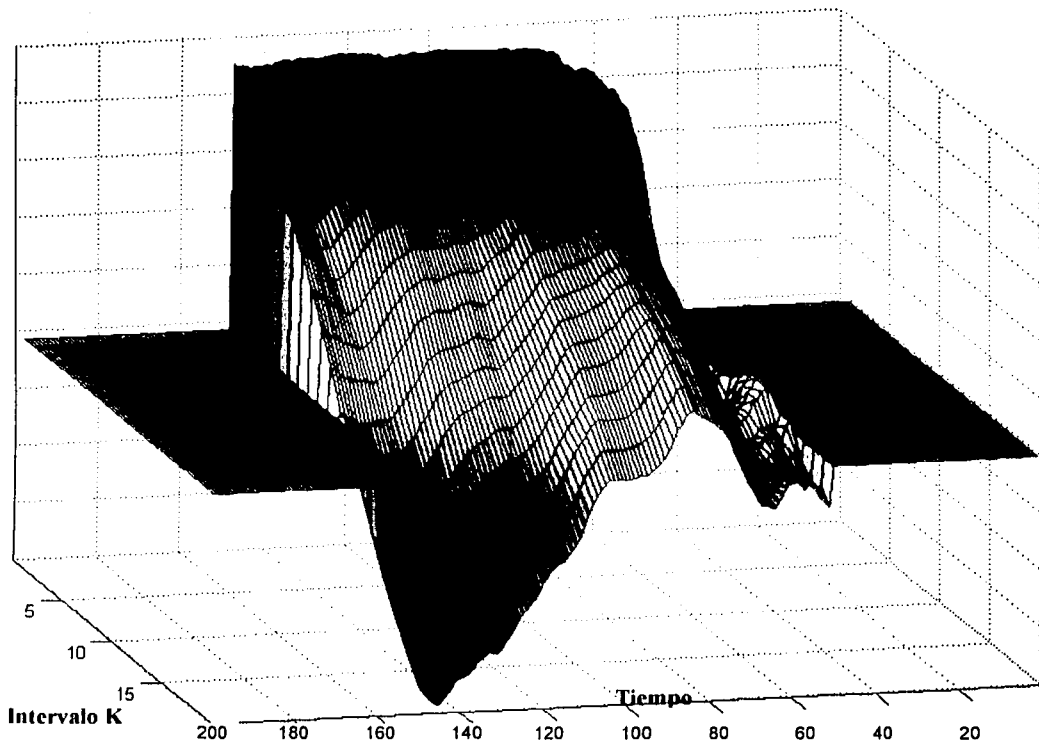


Gráfica 25 Función de Autocorrelación con Escala = 60 para 20 autocorrelaciones

En esta gráfica se pueden ya ver claramente los dos diferentes patrones de comportamiento, el eje de derecha a izquierda representa el tiempo (200 datos), en tanto que el eje del fondo hacia adelante representa los valores de K (intervalo que separa a dos variables). Nótese que aunque

ambos patrones de comportamiento son decrecientes, los valores de la autocorrelación para los primeros 100 datos son mucho mas pequeños que los valores de los 100 datos finales.

La gráfica anterior tiene la utilidad de ayudarnos a detectar cambios significativos en la función de autocovarianza, y en caso de que se tengan diferentes patrones de comportamiento, es posible utilizar el último patrón de comportamiento para predecir datos. Si este ejemplo se tratara de una serie real, nos convendría utilizar únicamente los últimos 100 datos para obtener un modelo de predicción mediante la metodología de Box y Jenkins.



Gráfica 26 Función de Autocorrelación con Escala = 100 para 20 autocorrelaciones

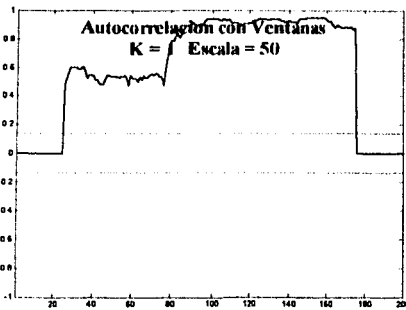
También debemos tener en cuenta que, al igual que en la media y la varianza, el tamaño de la escala influye en los resultados que obtenemos, y por lo tanto las escalas pequeñas reflejarán mejor los cambios en el tiempo, en tanto que las escalas grandes sólo reflejarán los cambios

TESIS CON
FALLA DE ORIGEN

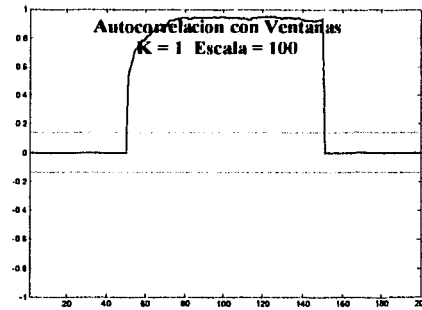
drásticos en la autocorrelación a través del tiempo. La gráfica 26 muestra la ACF del mismo ejemplo para una ventana de tamaño 100. A diferencia de la ACF obtenida con ventana de tamaño 60, en esta gráfica no se nota un cambio drástico entre los dos patrones de comportamiento, sino que aparenta un cambio continuo en la función de autocovarianza, que como se ve en la autocorrelación $K = 20$, aparece como decreciente a través del tiempo. Puede apreciarse que la ACF en los primeros tiempos decrece más lentamente que en los tiempos posteriores.

Es evidente que la representación tridimensional de la ACF con ventanas deja mucho que desear, por otro lado, como se mencionó anteriormente, no se aprecian en esa gráfica los límites del intervalo de confianza, lo cual sería deseable en estos casos. No obstante, se pueden realizar otras representaciones gráficas bidimensionales para un valor específico de K o del tiempo, en donde podemos utilizar el intervalo de confianza.

Consideremos las escalas de 50 y 100 con el fin de obtener información tanto específica como general. Para estas dos escalas podemos elegir ya sea un valor específico de K , o bien un valor de tiempo para calcular la ACF en ese tiempo. En el primer caso, supongamos que deseamos analizar el comportamiento en el tiempo de las autocorrelaciones con $K=1$ y $K=20$ ($\hat{\rho}_1$ y $\hat{\rho}_{20}$), las gráficas obtenidas se muestran a continuación:

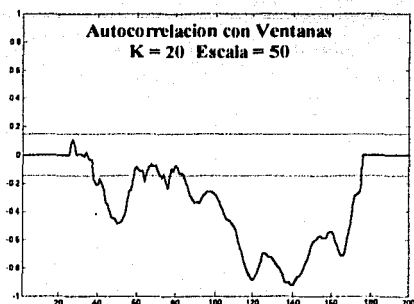


Gráfica 27 Autocorrelación con escala 50 para $K = 1$

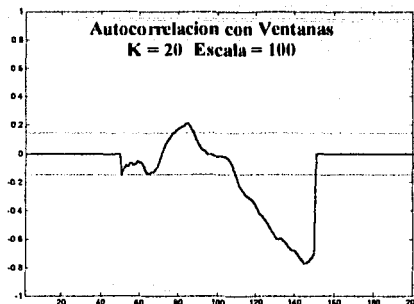


Gráfica 28 Autocorrelación con escala 100 para $K = 1$

En la gráfica 27 es muy clara la diferencia entre los dos valores de la autocorrelación para $K=1$, no así en la gráfica 28, donde el uso de una escala mayor (100) afecta la visualización de los dos patrones, que sólo se distingue por una curva inicial ascendente que indica que la autocorrelación para $K=1$ es más fuerte en los datos posteriores, pero no se refleja que el valor bajo de la autocorrelación en los primeros datos es en realidad constante durante 100 datos.



Gráfica 29 ACF con escala 50 para $K = 20$



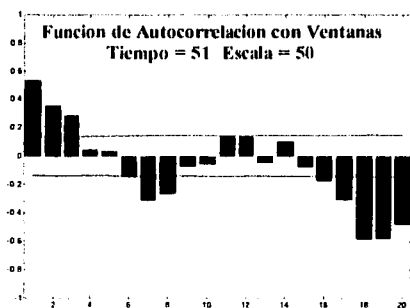
Gráfica 30 ACF con escala 100 para $K = 20$

Las gráficas 29 y 30 muestran una información similar para el caso de la autocorrelación en $K=20$, aunque en este caso no es tan clara la diferencia entre los dos patrones como en el caso anterior en la escala 50, en tanto que en la escala de 100 se observa una clara diferencia entre las dos partes de la serie. En la primera mitad tiende a tener una correlación $\hat{\rho}_{20} = 0$, mientras que en la segunda mitad esta correlación es claramente negativa, nótese que el punto en donde empiezan los valores negativos de autocorrelación es el valor de tiempo 100. Cabe mencionar que las gráficas son continuas debido a que se ubican en el dominio del tiempo, además de que debido a los datos que se pierden con el uso de escalas, a los valores de los extremos se les ha asignado el valor de cero, con objeto de facilitar su representación gráfica.

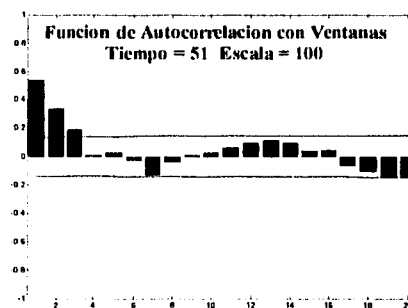
La media de la ACF permanece constante durante la primera parte de la gráfica 27, luego asciende y vuelve a permanecer constante, lo cual sugiere que la función de autocovarianza es estacionaria en esos subintervalos, o si se puede expresar así, podemos decir que es estacionaria por partes, aunque esta definición puede no ser muy apropiada. En general, podemos decir que una serie de

tiempo tiene una autocovarianza estacionaria cuando los valores de la ACF tienen una media constante a través del tiempo.

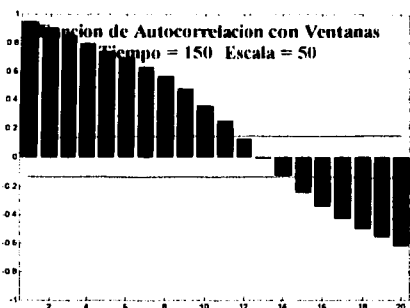
Las gráficas anteriores nos permiten apreciar hasta cierto punto el valor de una correlación muestral $\hat{\rho}_k$ a través del tiempo, quedando siempre a nuestro criterio juzgar si el tamaño de la escala es apropiada para esto. También podemos desde luego obtener la ACF para ciertos periodos específicos de tiempo. Supongamos ahora que nos interesa visualizar la función de autocorrelación en los periodos específicos de tiempo 51 y 150, que corresponden a los puntos intermedios de nuestros dos periodos con patrones diferentes. Las siguientes gráficas muestran los resultados obtenidos.



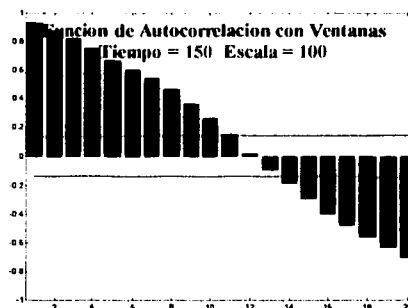
Gráfica 31 ACF con escala 50 y tiempo = 51



Gráfica 32 ACF con escala 100 y tiempo = 51



Gráfica 33 ACF con escala 50 y tiempo = 150



Gráfica 34 ACF con escala 100 y tiempo = 150

TESIS CON
FALLA DE ORIGEN

Al analizar estas gráficas hay que tener en cuenta que los valores de tiempo mencionados son en realidad el punto medio de los datos de los que se obtuvo la ACF, por ejemplo, el valor de Tiempo 51 con una ventana de tamaño 50 indica que esa gráfica se calculó con los datos del Tiempo 26 al 75, en tanto que si la ventana fuera de tamaño 100, se han considerado en esa gráfica los datos del Tiempo 1 al 100.

En esta forma de representación de datos siempre es preferible usar una escala grande, a diferencia de las gráficas 27 al 30, donde son necesarias escalas pequeñas. Entre más chica sea la escala, más imprecisas serán las últimas autocorrelaciones debido a que, si utilizamos una ventana de tamaño 20 por ejemplo, la autocorrelación $\hat{\rho}_{20}$ se obtendrá utilizando tan sólo la relación entre el primer y último dato de la ventana. De esta manera, el número de datos de los cuales se obtiene la autocorrelación es más grande entre mayor sea el tamaño de la ventana, y a escalas chicas sólo podemos considerar las primeras autocorrelaciones.

La gráfica 31 muestra un comportamiento decreciente en las primeras autocorrelaciones, para después sobresalir en los valores de $K=7$ y $K=8$, las autocorrelaciones que sobresalen posteriormente en $K=17$ pueden no ser muy significativas por el tamaño de la ventana. En el caso de las autocorrelaciones $K=7$ y $K=8$ reflejan una característica particular de los 50 datos en cuestión, en tanto que las últimas autocorrelaciones no deben ser tomadas muy en serio, puesto que pueden verse afectadas por un número insuficiente de datos. La gráfica 32 utiliza una ventana de 100, por lo que es más significativa. Ahí podemos observar bien que en general se trata de un comportamiento de la ACF decreciente infinita.

Las gráficas 33 y 34 son muy similares, aun cuando la diferencia en el tamaño de la escala es grande, y muestran un comportamiento decreciente, en donde en escalas grandes sobresalen los valores negativos en las autocorrelaciones grandes.

2.2.2 Función de Autocorrelación Parcial

Aunque la función de autocorrelación nos proporciona información importante respecto a la dependencia de un dato con los anteriores, existe otra herramienta para apoyarnos en el análisis de estas dependencias, esta herramienta es la función de autocorrelación parcial (PACF).

La PACF, representada por $\hat{\rho}_{kk} = r_{kk}$ se define como la contribución de la variable autoregresiva Z_{t-k} al modelo AR(k), dadas las variables anteriores $Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots, Z_{t-k+1}$. Esta fuerza de aportación se estima mediante las ecuaciones de Yule-Walker:

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 + \dots + \phi_k \rho_{1-k} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1 + \dots + \phi_k \rho_{2-k} \\ &\dots\dots\dots \\ \rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + \dots + \phi_k\end{aligned}$$

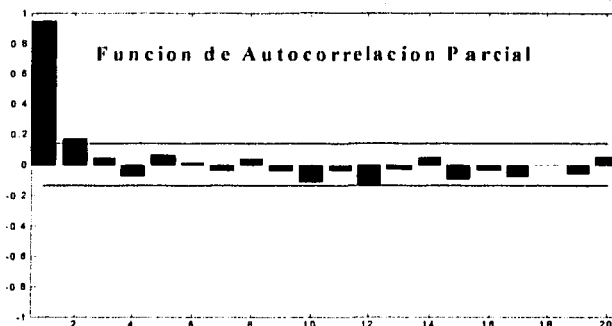
Sustituyendo los valores de autocorrelación en las ecuaciones de Yule-Walker obtenemos un sistema de ecuaciones de K ecuaciones con K incógnitas, en donde sólo nos interesa obtener el valor de $\phi_k = r_{kk}$, de esta manera para cada K debemos resolver el sistema de ecuaciones de KxK. Se puede obtener la forma escalonada por renglones (no reducida) para un cierto número k de autocorrelaciones que se desea obtener, o bien utilizar la solución general:

$$r_{kk} = \begin{cases} r_1 & k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & k > 1 \end{cases}$$

en donde

$$r_{k,j} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad j = 1, 2, \dots, k-1$$

La función de autocorrelación parcial (PACF) del modelo utilizado en la sección anterior se muestra en la gráfica 35, donde se supone estacionaria la función de autocorrelación parcial.

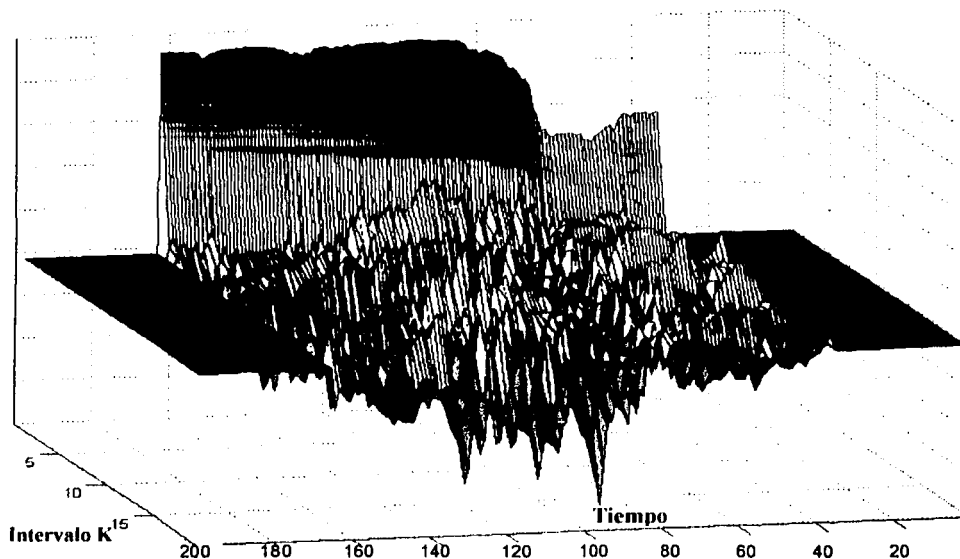


Gráfica 35 Función de Autocorrelación Parcial

La PACF se trunca en 2, aunque el valor de la segunda autocorrelación parcial sobresale muy poco del intervalo de confianza, por lo que pudiera considerarse un cero. De esto y de la información obtenida en la función de autocorrelación (ACF) se infiere que probablemente el modelo más apropiado para representar la serie si fuera estacionaria en la autocovarianza sería un AR(1) o AR(2), dado que la ACF es decreciente infinita, en tanto que la PACF se trunca en 2. En este caso al obtener un modelo con la metodología de Box y Jenkins se pueden obtener predicciones relativamente precisas, pero se pueden obtener predicciones mucho más útiles si se utilizan únicamente los datos más convenientes (el último patrón de comportamiento de los datos), en este caso los últimos 100 datos, aunque no siempre es fácil determinar donde empieza el nuevo patrón, pero independientemente de donde se trunquen los datos, descartar datos con patrones que ya no se encuentran vigentes puede mejorar mucho el modelo de predicción.

Al igual que en el caso de la ACF con ventanas, la PACF se puede calcular utilizando una escala fija y una representación tridimensional de los datos, o bien una escala y un parámetro fijos (tiempo o intervalo que separa a dos observaciones). La gráfica 36 representa la PACF para una escala de 70 con 20 autocorrelaciones parciales.

TESIS CON
FALLA DE ORIGEN

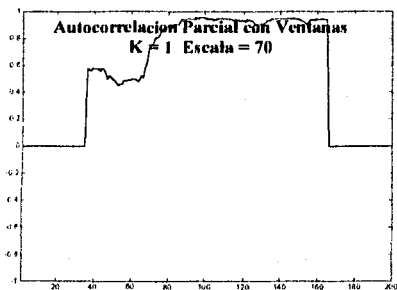


Gráfica 36 Función de Autocorrelación Parcial con Escala = 70

Las dos partes de la gráfica se truncan en uno, aunque aquí, al igual que con la ACF, observamos una diferencia significativa en la fuerza de correlación. En general podemos decir que tanto la PACF como la ACF pueden proporcionarnos la misma información con respecto al comportamiento de la autocovarianza a través del tiempo, aunque pueden servirnos para determinar simultáneamente que tipo de modelo representaría bien el comportamiento de la autocovarianza en diferentes periodos de tiempo. Por supuesto que si se desea estimar el modelo específico de un periodo de tiempo, se requerirá utilizar la metodología de Box-Jenkins en forma normal utilizando sólo los datos de interés.

También son posibles las representaciones gráficas en el plano para la PACF, en este caso sólo nos interesan los valores de $K=1$ y $K=2$, dado que de las gráficas 35 y 36 se concluye que para los demás valores de K las autocorrelaciones parciales son cero.

TESIS CON
FALLA DE ORIGEN



Gráfica 37 PACF con Escala = 70 para K=1

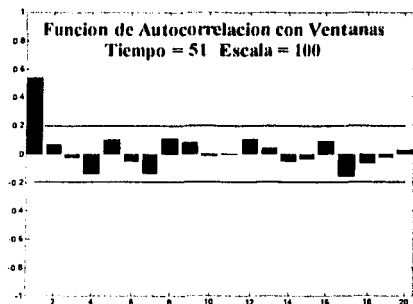


Gráfica 38 PACF con Escala = 70 para K=2

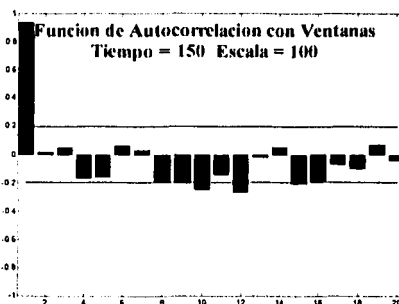
La gráfica 37 proporciona la misma información que la representación tridimensional, con dos diferentes valores de autocorrelación parcial para $K=1$, es decir, que en la segunda parte de los datos el valor de la autocorrelación es más grande. El caso interesante es cuando $K=2$, que en la PACF de toda la serie (sin escalas) sobresale ligeramente del intervalo de confianza, en tanto que cuando se calcula por ventanas ninguna de las dos partes refleja este mismo valor, sólo aparece en medio de los dos grupos de datos en la gráfica 38. Se infiere de esto que si se tienen dos patrones en una serie, cuando la función general refleje una característica que no se encuentre en ninguno de los grupos de datos, se reflejará en el punto medio de estos datos. Este puede ser de gran ayuda si se desea trincar los datos para tomar el grupo con el patrón de comportamiento más apropiado para realizar predicciones.

Las gráficas 39 y 40 son las PACF para los dos grupos de datos representativos, nótese que en ninguna de las dos gráficas es significativo el valor de $K=2$. La gráfica 40 tiene sin embargo, ciertos valores fuera del intervalo de confianza que deben ser analizados si se desea obtener un modelo de estos datos (101-200).

TESIS CON
FALLA DE ORIGEN



Gráfica 39 PACF con Escala = 100, para Tiempo = 51



Gráfica 40 PACF con Escala = 100 para Tiempo = 150

En general, la función de autocovarianza proporciona información bastante útil para obtener modelos que nos permitan predecir. Sin embargo, aunque tanto la media como la varianza se pueden pseudo-eliminar con el fin de facilitar el análisis de la serie de tiempo, una función de autocovarianza no estacionaria es prácticamente imposible de modelar, especialmente porque en la inmensa mayoría de los casos de la vida real los cambios en la función de autocovarianza son aleatorios y prácticamente impredecibles. De esta manera, lo más que podemos realizar cuando no es posible obtener un modelo que represente los cambios en la función de autocovarianza a través del tiempo es reducir nuestro análisis a un subconjunto de datos con un comportamiento que se considere representativo, así el modelo que se obtiene se comportará de acuerdo al patrón actual, sujeto desde luego a errores aleatorios.

TESIS CON
FALLA DE ORIGEN

Capítulo III

Análisis de la Frecuencia y Fluctuación Estacional

En este capítulo se tratará ahora el tema de la periodicidad, que aunque en muchos casos suele permanecer constante el patrón periódico de una serie (temporadas anuales, por ejemplo), existen otros casos en los que se pueden tener ciertos tipos de periodicidad que tienen una duración de tiempo finita. Cuando esto sucede, se hace necesario el uso de otros métodos que nos permitan ubicar los diferentes tipos de periodicidad a través del tiempo.

Cabe señalar que existe una fuerte relación entre la frecuencia o periodicidad de una serie y la función de autocovarianza, puesto que si se tienen diferentes comportamientos en la frecuencia a través del tiempo, esto implica que existen diferentes patrones de comportamiento de la serie y por tanto una autocovarianza no estacionaria.

3.1 Diferencias Estacionales

Hemos visto anteriormente que al analizar la tendencia y la varianza, se pueden presentar algunas dificultades cuando se tiene una variación estacional. De hecho, como se mencionó en el caso de las diferencias ordinarias, cuando tenemos fluctuaciones estacionales, se cumple la condición de homogeneidad, pues los movimientos o comportamientos cíclicos ocasionan que las diferentes partes de la serie tengan semejanzas entre si, con la diferencia del nivel de media en casos con tendencia. Generalmente después de eliminar la tendencia de una serie, la variación estacional puede resultar evidente, salvo en algunos casos en que la varianza tenga cambios muy significativos.

Aunque en general una serie con fluctuación estacional puede ser analizada en la forma original, es preferible en algunos casos intentar "eliminar" la variación estacional, es decir, trabajar con otra

TESIS CON
FALLA DE ORIGEN

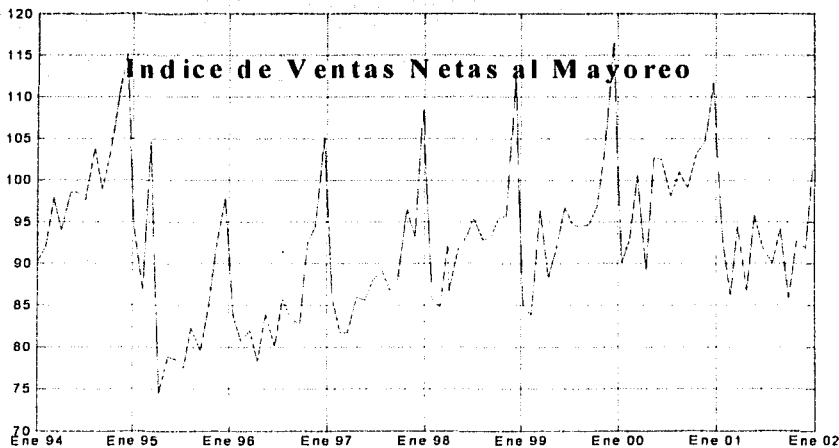
serie equivalente que represente las diferencias entre los valores de un periodo y los del siguiente. Por ejemplo cuando las observaciones se realizan cada mes, en cuyo caso es muy común tener una periodicidad anual y se suelen utilizar los valores con respecto al mismo mes del año anterior (diferencias estacionales).

Si sabemos que la longitud del periodo de la fluctuación estacional es de s , entonces la nueva serie diferenciada estacionalmente se obtiene como:

$$\nabla_s Z_t = Z_t - Z_{t-s}$$

Lo cual nos da como resultado valores de comparación con el último periodo, y en donde se tienen en total $n - s$ diferencias estacionales, lo que significa que tenemos s datos menos que la serie original.

Consideremos la serie de tiempo del índice de ventas nacionales netas al mayoreo mensualmente (Base 1994=100, Datos del INEGI), cuya gráfica se muestra a continuación.

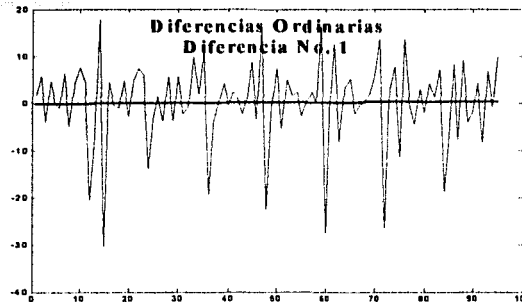


Gráfica 41 Índice de Ventas Netas al Mayoreo mensual (Base 1994=100)

Se observa un comportamiento periódico con una temporada alta de ventas en diciembre, aunque con diferentes niveles de media cada año. Aunque en este caso es obvia la fluctuación estacional,

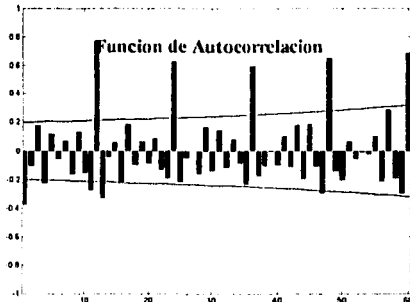
TESIS CON
FALLA DE ORIGEN

conviene estabilizar primero la tendencia siempre que sea posible. Como en este ejemplo se trata de una tendencia estocástica, realizando una diferencia ordinaria se estabiliza la tendencia,

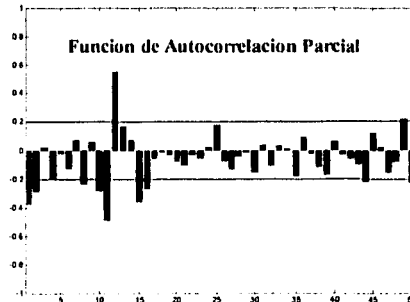


Gráfica 42 Serie estacional con tendencia estabilizada

Después de estabilizar la tendencia ya no es tan obvia la variación estacional, aunque en esta forma podemos utilizar la ACF y la PACF a fin de detectar autocorrelaciones estacionales, como se muestra a continuación:



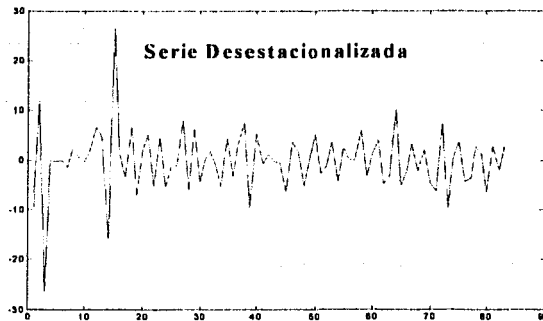
Gráfica 43 ACF de la serie con fluctuación estacional.



Gráfica 44 PACF de la serie con fluctuación estacional.

LA ACF tiene en este caso autocorrelaciones significativas en múltiplos de 12 y en forma decreciente, lo cual es un indicio de que se trata de una variación estacional con periodicidad 12. Esta puede modelarse mediante la metodología de Box & Jenkins, o bien pseudoeliminarse mediante diferencias estacionales.

TESIS CON
FALLA DE ORIGEN



Gráfica 45 Serie diferenciada estacionalmente

La gráfica 45 muestra la serie de ventas después de realizar la diferencia estacional para un periodo de 12. La interpretación de esta nueva serie a la que se han aplicado ya dos diferencias (una ordinaria y una estacional) no es tan simple, no obstante podemos concluir que los incrementos en las ventas de un año a otro tienen un comportamiento estocástico.

Aunque las diferencias estacionales pueden resultar muy útiles para analizar series, se requiere que este patrón de comportamiento estacional se encuentre presente en toda la serie, lo cual suele ser lo más común en series de tiempo con datos mensuales y periodicidad anual, por ejemplo, si el comportamiento cíclico se presenta sólo durante un intervalo de tiempo específico, en ese caso no servirá de nada aplicar la diferencia estacional, a menos que se desee analizar sólo ese subintervalo.

3.2 Análisis de Fourier

Al igual que en el caso de la media, la varianza y la autocovarianza, donde se utilizaron métodos de ventanas para analizar las propiedades espectrales, de la misma manera, el análisis de Fourier puede aplicarse mediante el método de ventanas para obtener una mejor apreciación de la frecuencia para series de tiempo no estacionarias. Pues al igual que en el caso de las propiedades

espectrales anteriores, una serie puede tener diferentes tipos de periodicidad o cambios en el patrón de los datos que afectan a la frecuencia de la serie.

Aunque las series de Fourier son particularmente eficientes para analizar series periódicas, cuando se analizan secuencias aperiódicas puede resultar incongruente o poco útil aproximar la serie mediante funciones periódicas. No obstante, en este caso existen dos alternativas viables: La transformada de Fourier con ventanas y la transformada Wavelet. En este apartado se presenta el uso de la transformada de Fourier utilizando ventanas para analizar diferentes tipos de secuencias aperiódicas o periódicas por partes.

Básicamente el procedimiento consiste en aproximar la serie de datos mediante una superposición o combinación lineal de senos y cosenos de diferentes frecuencias. En donde los coeficientes de estas sinusoides representan la fuerza de aportación de las frecuencias respectivas.

3.2.1 Transformada de Fourier Discreta

Dado que se desea aproximar la serie de tiempo a través de una combinación lineal de senos y cosenos de diferentes frecuencias, utilizamos para este fin una base ortogonal de funciones sinusoidales, de esta manera, se puede demostrar que si n es el total de datos de la serie que se van a aproximar, entonces las funciones

$$\{1, \cos w_0 t, \cos 2w_0 t, \dots, \cos(n-1)w_0 t, \sin w_0 t, \sin 2w_0 t, \dots, \sin kw_0 t, \dots, \sin(n-1)w_0 t\}$$

constituyen una base de funciones ortogonales en el intervalo de $1 < t < n$, en donde a

$w_0 = \frac{2\pi}{n}$ se le conoce como "frecuencia angular fundamental" y $w_k = kw_0$ es la k-ésima

frecuencia angular, dada en radianes. A la frecuencia $\frac{k}{n}$, se le conoce como la "k-ésima

armónica".

Se observa que como la duración de un periodo de una senoide es de 2π , al dividir $\frac{2\pi}{n}$ se obtiene una versión de la senoide expandida en el tiempo que tiene una duración de n . Por otro lado, $\frac{2\pi k}{n}$ indica que la senoide recorrerá k periodos en n valores de tiempo.

La serie de Fourier se escribe entonces, como una combinación lineal de los elementos de la base ortogonal:

$$Z_t = \sum_{k=0}^{n-1} (a_k \cos w_0 kt + b_k \sin w_0 kt)$$

Esta combinación lineal de sinusoides se acostumbra expresar con más frecuencia en su forma exponencial compleja:

$$Z_t = \sum_{k=0}^{n-1} c_k e^{iw_0 kt}$$

que significa que la serie de Fourier se puede escribir como una suma de exponenciales complejas, en donde los coeficientes c_k son complejos. Aunque no lo parezca, ésta suma sigue siendo una combinación lineal de senos y cosenos que genera a la serie de valores reales Z_t , considerando que

$$e^{iw_0 kt} = \cos w_0 kt + i \sin w_0 kt$$

De esta manera, la forma compleja de las series de Fourier nos permite representar con un solo coeficiente complejo la fuerza de aportación de los senos y cosenos que tienen la misma frecuencia. Así, la base de n funciones ortogonales que generan al espacio de series de n datos se puede escribir como:

$$B = \{1, e^{iw_0 t}, e^{iw_0 2t}, e^{iw_0 3t}, \dots, e^{iw_0 kt}, \dots, e^{iw_0 (n-1)t}\}$$

Es muy importante aclarar que las funciones que constituyen la base están definidas para valores discretos de tiempo, y que por tanto se trata en realidad de una base de series ortogonales, en donde el k -ésimo elemento de la base está dado por

$$e^{iw_0 kt} = \begin{bmatrix} e^{iw_0 k} \\ e^{iw_0 k(2)} \\ \dots \\ e^{iw_0 kt} \\ \dots \\ e^{iw_0 kn} \end{bmatrix}$$

La demostración de la ortogonalidad de estas series, o funciones discretas, es sencilla, puesto que el producto interno entre dos de estas funciones genera una serie geométrica. De esta manera, si $u \neq v$, entonces el producto interno de dos elementos de la base se escribe como

$$\sum_{t=1}^n e^{iw_0 ut} e^{-iw_0 vt} = \sum_{t=1}^n e^{iw_0 (u-v)t}$$

Considerando que el n -ésimo elemento de esta suma es siempre igual a 1, podemos por conveniencia sacarlo de la sumatoria, lo que equivale también a ajustar los índices de la sumatoria desde cero hasta $n-1$. Así, ajustando los índices y aplicando la fórmula para la suma de series geométricas, obtenemos:

$$\sum_{t=0}^{n-1} e^{iw_0 (u-v)t} = \frac{e^{iw_0 (u-v)n} - 1}{e^{iw_0 (u-v)} - 1} = \frac{e^{2\pi i (u-v)} - 1}{e^{2\pi i (u-v)/n} - 1} = \frac{1 - 1}{\alpha - 1} = 0$$

Así, el numerador siempre es cero dado que el coseno de un múltiplo de 2π es uno y el seno siempre es cero en ese mismo caso. El denominador no puede ser cero debido a que $u \neq v$.

De esta manera, la base de n vectores ortogonales nos permite representar cualquier serie de tiempo de n datos como combinación lineal de los elementos de la base ortogonal. Dicho de otra manera, cualquier serie de tiempo Z_t se puede escribir como la suma de las proyecciones ortogonales de Z_t sobre los elementos de la base B. Esto es:

$$Z_t = \text{proy}_{B_0} Z_t + \text{proy}_{B_1} Z_t + \dots + \text{proy}_{B_{n-1}} Z_t$$

en donde β_k es el k-ésimo elemento de la base B, en tanto que la k-ésima proyección ortogonal esta dada por:

$$\text{proy}_{\beta_k} Z_i = \frac{\langle Z_i, e^{i\omega_0 kt} \rangle}{\|e^{i\omega_0 kt}\|^2} e^{i\omega_0 kt} = \left(\frac{\sum_{j=1}^n Z_j e^{-i\omega_0 kj}}{n} \right) e^{i\omega_0 kt}$$

Cabe señalar que al dividir entre la norma equivale a utilizar una base ortonormal. Podemos ver que n representa la norma al cuadrado de los elementos de la base B. La serie de Fourier se puede escribir finalmente como la suma de las proyecciones

$$Z_i = \sum_{k=0}^{n-1} c_k e^{i\omega_0 kt}$$

en donde los coeficientes de proyección (coeficientes de Fourier) se calculan como:

$$c_k = \frac{1}{n} \sum_{j=1}^n Z_j e^{-i\omega_0 kj}$$

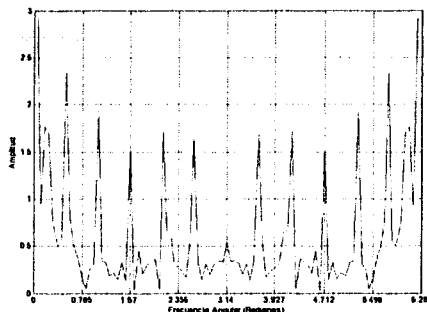
De esta manera, podemos definir la *transformada discreta de Fourier* como el producto interno de la serie con el k-ésimo elemento de la base B, esto es:

$$\hat{Z}_k = \sum_{j=1}^n Z_j e^{-i\omega_0 kj}$$

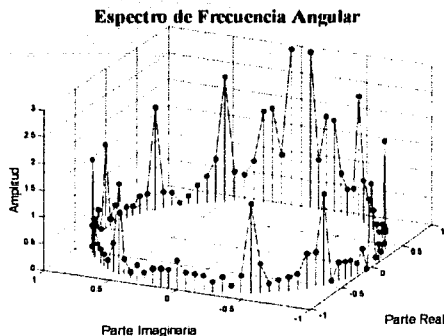
que es en realidad n veces el coeficiente de proyección. Como la transformada discreta de Fourier representa el comportamiento de la serie en función de k (frecuencia) en lugar del tiempo, podemos decir que se trata de una representación de la serie trasladada del dominio del tiempo al dominio de la frecuencia.

La gráfica de las frecuencias (armónicas) contra las amplitudes (magnitud de los coeficientes de Fourier) constituye el espectro lineal de Fourier. Cabe señalar la diferencia entre el uso de la frecuencia angular y las armónicas de Fourier, puesto que los valores de la frecuencia angular van desde cero hasta 2π . Como se ve en la gráfica 46, la gráfica es simétrica con respecto al valor de

π . Esto se debe a que la transformada de Fourier de cualquier serie de valores reales genera valores conjugados complejos para las frecuencias w_0k y $w_0(n-k)$, como se ve en la gráfica 47 (Datos de la gráfica 41, serie de ventas netas al mayoreo), donde se tiene la simetría con respecto a la parte imaginaria del dominio. Nótese que $w_0(n-k) = w_0(-k)$



Gráfica 46 Espectro de frecuencia angular



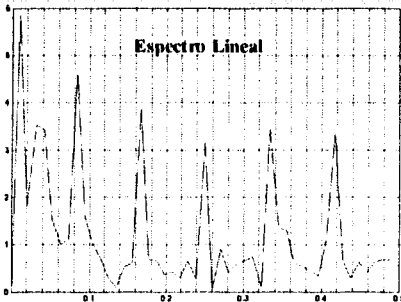
Gráfica 47 Espectro de frecuencia sobre el dominio de $\cos w_0k + i \sin w_0k$

De esta manera, considerando que los coeficientes de Fourier generan por separado valores para las frecuencias angulares positivas y negativas (aun cuando se refieren a la misma armónica de Fourier), el espectro lineal para las frecuencias de 0 a $\frac{1}{2}$ se calcula sumando las amplitudes de frecuencias equivalentes pero con diferente signo, esto equivale a multiplicar los valores de las amplitudes por dos (desde $k = 1$ hasta el máximo valor de k que satisface $\frac{k}{n} < \frac{1}{2}$), dado que las

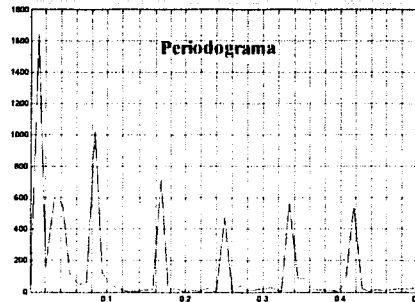
amplitudes para las frecuencias $\frac{k}{n}$ y $-\frac{k}{n}$ son iguales puesto que sus coeficientes son conjugados complejos.

La gráfica 48 muestra el espectro lineal para la serie de ventas al mayoreo utilizada en la sección anterior. En ella se aprecia en primer lugar la frecuencia de $1/96$ como la de mayor amplitud, siendo 96 datos en la muestra. Cuando la frecuencia predominante es de $\frac{1}{n}$, entonces existen indicios para suponer que no hay periodicidad, sobre todo si es prácticamente la única frecuencia

predominante. Sin embargo, cuando la diferencia es pequeña en relación con la segunda frecuencia predominante, en ese caso cabe la posibilidad de que exista periodicidad.



Gráfica 48 Espectro Lineal



Gráfica 49 Periodograma

La gráfica 49 muestra el periodograma, el cual se obtiene como la magnitud al cuadrado de los valores de la transformada de Fourier:

$$I(k/n) = \left| \hat{Z}_k \right|^2$$

De nueva cuenta, para las frecuencias de 0 a $1/2$, sumamos los valores para k y $-k$. Entonces el periodograma queda como:

$$I(k/n) = \begin{cases} 2 \left| \hat{Z}_k \right|^2 & 0 < k < n/2 \\ \left| \hat{Z}_k \right|^2 & k = 0; k = n/2 \end{cases}$$

Podríamos decir que el periodograma es una versión del espectro lineal en donde se han exagerado los valores de las amplitudes al elevarlos al cuadrado, lo que resulta en una gráfica donde se aprecian mejor las frecuencias que predominan. Esta representación del espectro es llamada comúnmente "espectro de potencia".

3.2.2 Espectro Teórico y Niveles de Significancia

El periodograma refleja con mayor claridad las frecuencias predominantes, aunque no siempre es claro qué frecuencias se deben considerar como significativas, por lo que conviene obtener un

TESIS CON
FALLA DE ORIGEN

intervalo de confianza para determinar las frecuencias estadísticamente significativas. Así, para obtener el intervalo de confianza, se requiere primero normalizar el espectro, de manera que éste tenga un valor promedio de uno. Dicho de otra manera, la serie debe normalizarse para tener una media de cero y varianza de uno.

Cabe señalar la relación que existe entre la varianza de la serie y los valores de los coeficientes de Fourier, puesto que, mientras con la frecuencia $k = 0$, el valor de c_k refleja la media de la serie, con las frecuencias distintas de cero el coeficiente c_k proporciona información sobre la varianza de la serie para cada componente de frecuencia k . De hecho, utilizando la identidad de Parseval:

$$\|Z_t\|^2 = \frac{\|\hat{Z}_k\|^2}{n}$$

podemos obtener la varianza de la serie como:

$$\sigma^2 = \frac{\sum_{t=1}^n |Z_t - \bar{Z}|^2}{n} = \frac{\sum_{k=1}^{n-1} |\hat{Z}_k|^2}{n^2} = \sum_{k=1}^{n-1} |c_k|^2$$

Que significa que la suma de la magnitud al cuadrado de los coeficientes de Fourier es igual a la varianza de la serie. Nótese que no se incluye el coeficiente c_0 en la sumatoria. Esto se debe a que, para normalizar la media de la serie en el dominio de la frecuencia, basta con hacer $c_0 = 0$, lo que en el dominio del tiempo equivaldría a restarle a cada elemento la media de la serie ($Z_t - \bar{Z}$). Dividiendo ambos lados de la identidad de Parseval por n , obtenemos la fórmula para calcular la varianza de la serie en el dominio de la frecuencia.

Conociendo la relación existente entre la varianza de la serie y el espectro de potencia, podemos ver que normalizar la varianza de la serie equivale a tener una media de uno en el espectro de potencia. Esta normalización se realiza entonces dividiendo los valores del espectro de potencia entre la media de éste:

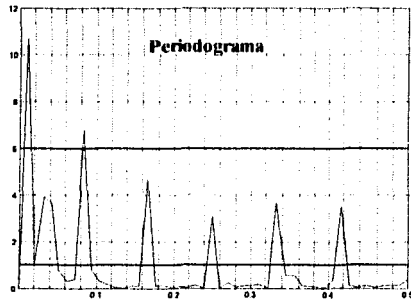
$$N(k) = \frac{|\hat{Z}_k|^2}{\frac{1}{n} \sum_{j=1}^{n-1} |\hat{Z}_j|^2} = \frac{|\hat{Z}_k|^2}{n\sigma^2} = \frac{|c_k|^2}{\sigma^2/n}$$

Esta forma normalizada del espectro nos permite comparar directamente la serie que se analiza con un espectro teórico, es decir, un modelo de comportamiento que nos sirva de base para comparar si la serie sigue ese comportamiento hipotético o si existen ciertas frecuencias que no se ajustan a éste. Por ejemplo, cuando comparamos una serie de tiempo con el espectro de ruido blanco, podemos determinar que valores del espectro de la serie pueden deberse a fluctuaciones aleatorias, o a causas ajenas a la aleatoriedad.

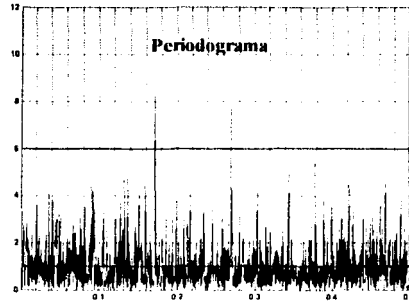
Sin embargo, para que dos espectros sean directamente comparables, es necesario realizar la normalización antes mencionada. El espectro teórico nos permite obtener el intervalo de confianza con un nivel de significancia α de que los valores del espectro de la serie se ajusten al espectro teórico. Para obtener el intervalo de confianza partimos del supuesto de que Z_t se encuentra normalmente distribuida, y por tanto las partes real e imaginaria de la transformada de Fourier están también normalmente distribuidas (Chatfield, 1989). Dado que el cuadrado de una variable normalmente distribuida, se distribuye como una chi-cuadrada con un grado de libertad, entonces $|\hat{Z}_k|^2$ se distribuye con dos grados de libertad (parte real y compleja). De esta manera, si asumimos un espectro teórico, posiblemente ruido blanco, la distribución del espectro de potencia normalizado es:

$$\frac{|\hat{Z}_k|^2}{n\sigma^2} \Rightarrow P_k \chi_2^2$$

en donde P_k es la función de media del espectro teórico que se utiliza, y " \Rightarrow " indica "se distribuye como". En el caso del ruido blanco la función de media tiene un valor constante de uno.



Gráfica 50 Ventas Netas al Mayoreo



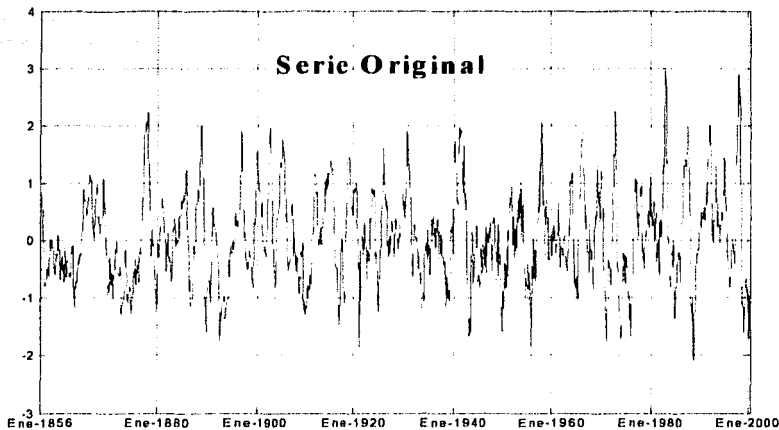
Gráfica 51 Ruido Blanco

La gráfica 50 muestra el intervalo de confianza del 95% para la serie de tiempo de ventas netas al mayoreo (en rojo), en tanto que la línea azul representa el valor promedio del espectro teórico (en este caso ruido blanco). La gráfica 51 muestra el espectro de una serie de ruido blanco, donde podemos ver que el 95 % de los datos queda bajo la gráfica. Otro aspecto a notar es que, debido a la normalización, las dos gráficas son comparables directamente. Entonces, los intervalos de confianza del espectro de ruido blanco muestran las frecuencias que pueden deberse a la aleatoriedad, si suponemos que existe independencia entre las observaciones.

El intervalo de confianza nos permite de esta manera descartar ciertas frecuencias que no sean estadísticamente significativas aunque se distingan de las demás. Es decir, que pueden deberse a factores aleatorios.

Aunque el espectro teórico de ruido blanco es útil para analizar una gran cantidad de series, hay casos en que se requiere comparar la serie con modelos de comportamiento más complejos. Consideremos la serie de tiempo que se muestra en la gráfica 52, que representa la temperatura de la superficie del mar meridional (SST, Sea Surface Temperature) cada mes desde 1856 al 2002⁴, estos niveles de temperatura se relacionan con el fenómeno de "El niño".

⁴ Valores mensuales para el NINO3.4 SST (5N-5S, 170-120W). Los datos de 1856-1949 se obtuvieron de la página "Kaplan's OS SST" en la URL: <http://ingrid.ldgo.columbia.edu/SOURCES/KAPLAN/Indices/> Los datos de 1950-2002 se obtuvieron del "Climate Diagnostics Center OISST" en la URL: <http://ftp.ncep.noaa.gov/data/cdddb/>.



Gráfica 52 Temperatura de la superficie del mar meridional (SST) mensual, 1856-2002

Como es de suponerse, la temperatura de la superficie del mar cada mes se encuentra relacionada con la temperatura del mes anterior. De hecho, utilizando la ACF y PACF se obtiene que la autocorrelación parcial de la serie se trunca en uno con un valor aproximado de 0.9 para la primera correlación. Este tipo de comportamiento es muy común en la vida real y se puede representar mediante el modelo autoregresivo AR(1):

$$Z_t = \rho_1 Z_{t-1} + e_t$$

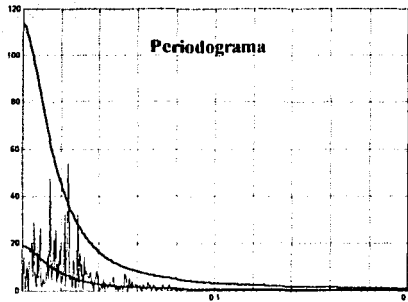
El espectro generado por este tipo de modelos para valores de ρ_1 menores que 1, se conoce como "espectro de ruido rojo", llamado así debido a que en estos modelos las frecuencias predominantes son las más pequeñas (periodos largos), en relación al espectro de la luz.

De acuerdo con Gilman (1963), la función del promedio del espectro del modelo autoregresivo de ruido rojo está dada por:

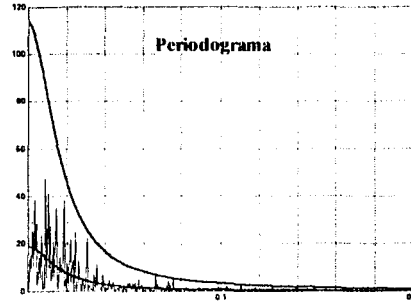
$$P_k = \frac{1 - \rho_1^2}{1 + \rho_1^2 - 2\rho_1 \cos(2\pi k/n)}$$

TESIS CON
 FALLA DE ORIGEN

Esta función puede calcularse obteniendo la transformada de Fourier cuando se supone que el valor de las aleatoriedades e_t es cero. Nótese que cuando $\rho_1 = 0$, la función da como resultado el espectro de ruido blanco (media constante de 1).



Gráfica 53 Espectro de la SST

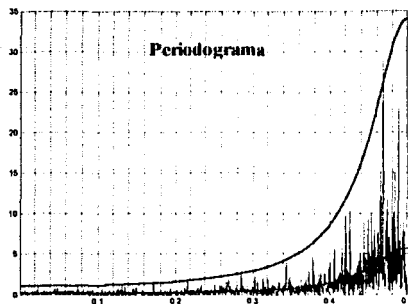


Gráfica 54 Espectro de ruido rojo .

La gráfica 53 presenta el espectro de la serie de tiempo de la SST, donde la línea azul representa el valor medio P_k del espectro teórico de ruido rojo, con $\rho_1 = 0.9$. La gráfica 54 es el espectro del modelo autoregresivo que se ha utilizado para comparar, es decir, es una serie de ruido rojo con $\rho_1 = 0.9$. Nótese que fuera del comportamiento normal de un ruido rojo con esa autocorrelación, se alcanza a tener dos frecuencias significativas. De esta manera, podemos apreciar las frecuencias que se distinguen de nuestro comportamiento teórico.

El uso del intervalo de confianza utilizando un espectro teórico nos permite distinguir las frecuencias que se salen del comportamiento normal de la serie (en comparación con el espectro teórico que se ha supuesto). El espectro de ruido blanco es útil cuando no existe dependencia entre un dato y el que le precede, o bien, cuando la correlación es muy pequeña y puede suponerse que es nula. Por otro lado, el espectro de ruido rojo es útil en muchas series con comportamiento no estacionario. Existen también otros modelos de espectro teórico que pueden utilizarse, tales como el espectro de ruido azul caracterizado por presentar especialmente frecuencias altas.

TESIS CON
FALLA DE ORIGEN



Gráfica 55 Espectro de ruido azul, $\rho_1 = -0.7$.

La gráfica 55 muestra el espectro de una serie de ruido azul, en donde la función de media del espectro teórico se obtiene utilizando la misma fórmula que en el ruido rojo, pero con valores negativos de ρ_1 , esto significa que cuando se tiene una autocorrelación negativa en ρ_1 , es posible suponer un espectro teórico de ruido azul.

Aunque la elección de ρ_1 para obtener el intervalo de confianza del espectro teórico es un tanto arbitraria, puede utilizarse el valor estimado de la primera autocorrelación de la serie utilizada, aunque no siempre es conveniente. En muchos casos puede ser preferible realizar diferencias ordinarias a la serie para comparar su espectro con el de ruido blanco, aunque, al aplicar el análisis mediante ventanas es poco usual diferenciar la serie, especialmente porque en algunos casos el comportamiento de ruido rojo o azul puede ser significativo sólo durante un intervalo de tiempo. El uso del intervalo de confianza mediante un espectro teórico, es particularmente útil en el análisis de Fourier con ventanas como se planteará mas adelante.

3.2.3 Funciones Ponderadoras

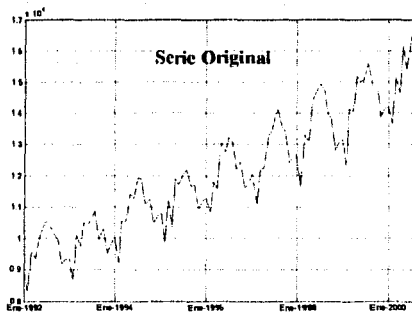
Aunque el periodograma de una serie de tiempo proporciona información muy útil respecto a la frecuencia, tiene algunos defectos que deben tenerse en consideración. En primer lugar, el periodograma no detecta fácilmente la periodicidad de una serie no estacionaria, por lo que se

requiere que la serie sea estabilizada previamente mediante diferencias ordinarias, modelos de regresión o transformaciones apropiadas, y aunque se tiene la alternativa de utilizar un espectro teórico apropiado no siempre es claro qué espectro teórico utilizar. Por otro lado, la información del periodograma es insuficiente cuando existen diferentes comportamientos de frecuencia a través del tiempo.

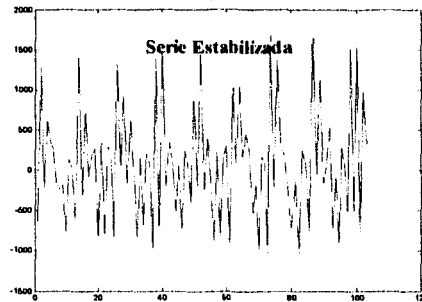
Cabe señalar la relación que existe entre la frecuencia no estacionaria y la función de autocovarianza, dado que la función de autocovarianza estudia el comportamiento de la serie en el dominio del tiempo, mostrándonos la dependencia entre las variables a través del tiempo; en tanto que el periodograma refleja el comportamiento de la serie en el dominio de la frecuencia. De esta manera, el periodograma y la ACF son dos formas distintas de representar el comportamiento de la serie, pero lo más importante que debe tenerse en consideración es que si el comportamiento de la serie cambia a través del tiempo, esto se reflejará tanto en la función de autocovarianza como en el análisis de la frecuencia con ventanas. De esto se concluye que una no estacionaridad en la autocovarianza implica necesariamente una frecuencia no estacionaria, y viceversa. Sin embargo, una serie de tiempo no estacionaria en la frecuencia puede ser analizada también por métodos de ventaneo, utilizando una escala apropiada de análisis y analizando la frecuencia para esos subintervalos de tiempo.

Existe también otro defecto que se presenta particularmente en muestras pequeñas de datos, pues cuando se toman observaciones durante un cierto intervalo de tiempo, es posible que algunos periodos de la serie queden truncados o a medias al principio y al final de las observaciones. Estos periodos incompletos afectan a la interpretación de las frecuencias significativas en el intervalo de tiempo analizado, especialmente cuando se trata de pocos datos. Por el contrario, cuando se trata de una cantidad grande de datos los periodos truncados no afectarán mucho al periodograma resultante.

TESIS CON
FALLA DE ORIGEN



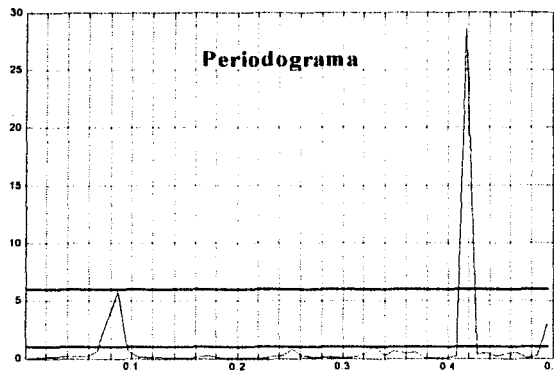
Gráfica 56 Generación de energía eléctrica



Gráfica 57 Serie Estabilizada (1 diferencia)

Consideremos la serie de tiempo representada en la gráfica 56, en donde se muestra la generación neta de energía eléctrica mensual en miles de millones de Watt/Hora, desde Enero de 1992 hasta la mitad del año 2000 (Datos del INEGI). Esta serie presenta un periodo incompleto en el año 2000, que se encuentra truncado a la mitad. La gráfica 57 muestra la serie estabilizada al aplicar una diferencia ordinaria, para facilitar el análisis de la frecuencia (estacionaria). Aunque esta serie presenta una clara periodicidad anual, en el periodograma (gráfica 58), la frecuencia de $\frac{1}{12} \approx 0.083$ queda ligeramente abajo del límite del intervalo de confianza para considerarla significativa. Los periodos truncados al final de la serie pueden afectar en ocasiones ligeramente o en algunos casos drásticamente a la frecuencia de la serie. Cuando se trata de observaciones mensuales es recomendable por tanto analizar años completos para evitar truncar el periodo anual. Aunque en este caso la periodicidad de la serie es bastante obvia, cuando existen otros patrones de frecuencia cuyo periodo desconocemos, en ese caso es inevitable llegar ocasionalmente a truncarlos.

TESIS CON
FALLA DE ORIGEN



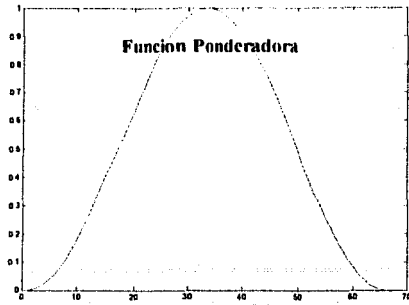
Gráfica 58 Periodograma de la serie estabilizada

En la gráfica 58 se observan dos frecuencias predominantes, la primera es la periodicidad de 12 (0.083), en tanto que la segunda es la frecuencia de aproximadamente 0.42 (periodo de 2.4). La primera, como se mencionó anteriormente, aparece como poco significativa a pesar de ser la frecuencia más evidente, la segunda frecuencia es la predominante, ésta sólo refleja el constante ascenso y descenso de la serie cada 2.4 datos, es decir, en general se tiene un ascenso seguido de un descenso, y ocasionalmente algunos ascensos o descensos seguidos en el valor de la serie estabilizada.

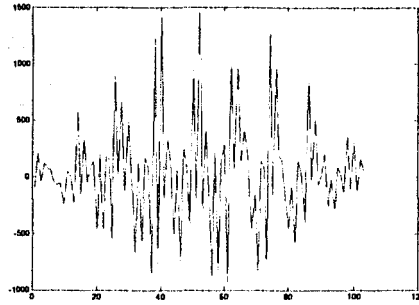
Para evitar que los periodos truncados al principio y al final de la serie afecten demasiado al análisis de la frecuencia, es conveniente utilizar una función ponderadora, es decir, una función que respete los valores centrales de la serie y gradualmente minimice los valores de las esquinas, para reducir el efecto de los periodos incompletos en el espectro resultante. La gráfica 59 muestra un ejemplo de función ponderadora que tiene el mismo tamaño que la serie estabilizada (gráfica 57). La versión ponderada de la serie que suaviza el efecto de los datos en las esquinas de la gráfica se obtiene mediante el producto

$$Z_i^* = Z_i * W(i) \quad \text{para } i = \overline{1, n}$$

en donde $W(i)$ es el valor de la función ponderadora en el tiempo i .



Gráfica 59 Función Ponderadora

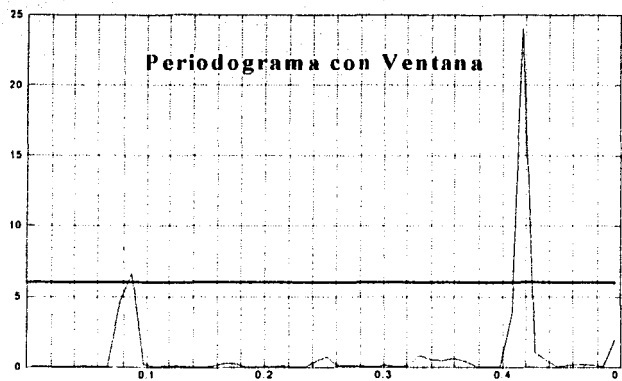


Gráfica 60 Versión Ponderada de la serie

La gráfica 60 muestra la serie ponderada en donde se han minimizado los datos de las esquinas. De esta manera, para ponderar la serie debemos obtener primero una función ponderadora (por lo general en forma de campana) cuya duración de tiempo sea igual a la duración de la serie analizada. Posteriormente, se realiza el producto de los valores de la función ponderadora y la serie, para finalmente calcular la transformada de Fourier

$$\hat{Z}_k = \sum_{j=1}^n Z_j \cdot W(t) \cdot e^{-i\omega_j k}$$

El periodograma de la serie de energía eléctrica utilizando la función ponderadora de la gráfica 59 se muestra a continuación:



Gráfica 61 Periodograma de la serie ponderada

TESIS CON
FALLA DE ORIGEN

En esta gráfica ya se aprecia la frecuencia de $\frac{1}{12}$ como significativa, y en general, al ponderar una serie los valores del periodograma representan principalmente el comportamiento de los datos centrales de la serie o del subconjunto de datos que se está analizando. Podemos ver entonces que el uso de funciones ponderadoras es útil para realizar el análisis de la frecuencia mediante ventanas, puesto que al analizar la frecuencia para los subintervalos de la ventana se truncan frecuentemente comportamientos periódicos. Por otro lado, al ponderar la serie podemos ver que las frecuencias cercanas a $1/12$ han crecido un poco, es decir, el pico se ve más grueso.

La función ponderadora que se utiliza en el análisis de la serie se elige de manera arbitraria. Existen diferentes tipos de funciones ponderadoras, entre las que destacan las funciones: Hann, Hamming, Blackman, Bartlett, Triangular y Rectangular o de Caja. Dado que estas funciones generan ceros fuera del área de interés, pueden servir también como filtros cuando se desea realizar el análisis de ventanas, por esta razón, las funciones ponderadoras son generalmente llamadas "funciones de ventana", dado que pueden utilizarse funciones ponderadoras más pequeñas que el tamaño de la serie para analizar un subintervalo específico de tiempo (ventana), en donde el producto entre los valores de la función ponderadora y la serie genera ceros fuera de la ventana.

A continuación se presentan las descripciones de algunas de las funciones de ventana más comunes (Fórmulas tomadas de la documentación de Matlab 6.0 R12, 2000):

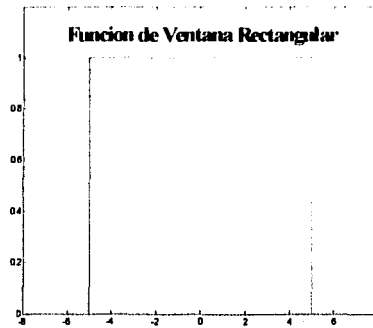
a) Función de Ventana Rectangular:

También llamada Función de Caja, el uso de esta función de ventana equivale a no ponderar la serie, puesto que sólo filtra los valores que están dentro de la ventana y el resto los vuelve ceros, de esta manera, no suaviza el efecto de los periodos truncados al considerar sólo segmentos de la serie (ventanas). La función de caja se define como:

TESIS CON
FALLA DE ORIGEN

$$W(t) = \begin{cases} 1, & 1 \leq t \leq n \\ 0, & \text{otro caso} \end{cases}$$

en donde n es el tamaño de la ventana que se desea analizar. En esta fórmula para la función de ventana se asume que se utilizarán valores discretos de t . Esto se debe a que los valores de tiempo de las observaciones de la serie son discretos; y de esta manera para ponderar una serie de n datos, se requiere obtener n valores de la función de ventana. Las fórmulas para las funciones de ventana dadas en este apartado están enfocadas a generar valores discretos, por lo que pueden diferir de las definiciones generalmente usadas de éstas. La gráfica 62 muestra la gráfica de la función de ventana rectangular.



Gráfica 62 Función de Ventana Rectangular

El uso de la función de ventana rectangular se justifica cuando por alguna razón sólo se desea filtrar la serie en un subintervalo de tiempo específico, sin ponderar la serie, por lo que principalmente es usada en el análisis con ventanas.

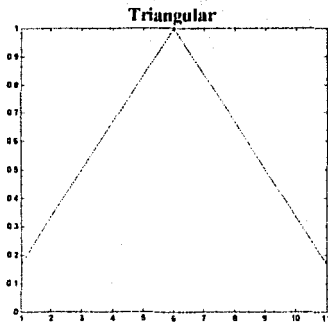
b) Función de Ventana Triangular:

Como su nombre lo indica, esta función de ventana presenta forma de pico triangular. La fórmula para construir una función de ventana triangular de tamaño n es como sigue:

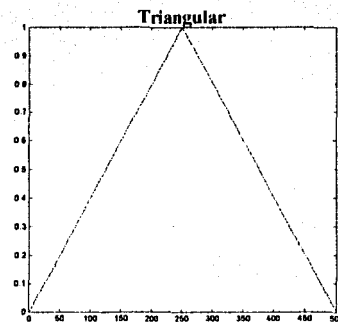
$$\text{Para } N \text{ non: } W(t) = \begin{cases} \frac{2t}{n+1} & 1 \leq t \leq (n+1)/2 \\ \frac{2(n-t+1)}{n+1} & (n+1)/2 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

$$\text{Para } N \text{ par: } W(t) = \begin{cases} \frac{2t-1}{n} & 1 \leq t \leq (n/2) \\ \frac{2(n-t+1)}{n} & (n/2)+1 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

En las gráficas 63 y 64 se muestra la función de ventana triangular para dos tamaños de ventana diferentes. En la gráfica 63 se utiliza una ventana de tamaño 11, en tanto que en la gráfica 64 la ventana es de tamaño 500.



Gráfica 63 Función Triangular de tamaño 11.



Gráfica 64 Función Triangular de tamaño 500

Comparando estas dos escalas, podemos observar que con escalas chicas, los valores de la función en las esquinas de la ventana son diferentes de cero, en tanto que a escalas grandes las esquinas toman valores muy cercanos a cero. Esto se debe a que en general las esquinas tienen un valor de $\frac{2}{n+1}$, de manera que cuando n tiende a valores grandes, el valor de la función tiende a cero. La utilidad principal de esta propiedad de la función triangular consiste en que cuando se analizan subconjuntos muy pequeños de datos, no se minimiza demasiado el efecto de los valores

en los extremos. En general, la característica de la función triangular es precisamente que no minimiza demasiado los valores en los extremos al grado de volverlos ceros.

c) Función de Ventana Bartlett:

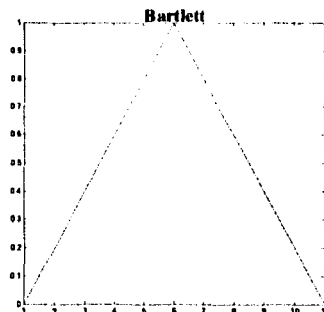
Esta función es muy semejante a la función triangular, aunque con la diferencia esencial de que la función de Bartlett genera ceros en las esquinas de la ventana para cualquier tamaño de ventana.

La función de Bartlett se calcula como:

$$\text{Para } N \text{ non: } W(t) = \begin{cases} \frac{2(t-1)}{n-1} & 1 \leq t \leq (n+1)/2 \\ \frac{2(n-t)}{n-1} & (n+1)/2 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

$$\text{Para } N \text{ par: } W(t) = \begin{cases} \frac{2(t-1)}{n-1} & 1 \leq t \leq (n/2) \\ \frac{2(n-t)}{n-1} & (n/2)+1 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

La gráfica 65 muestra la función de Bartlett de tamaño 11, en donde se aprecia la diferencia con la función triangular en las esquinas de la ventana.



Gráfica 65 Función de Ventana Bartlett de tamaño 11.

La función de Bartlett se utiliza cuando se desea que, al tratarse de ventanas de tiempo pequeñas, los valores de las esquinas también sean minimizados, puede ser particularmente útil cuando existen comportamientos periódicos pequeños truncados.

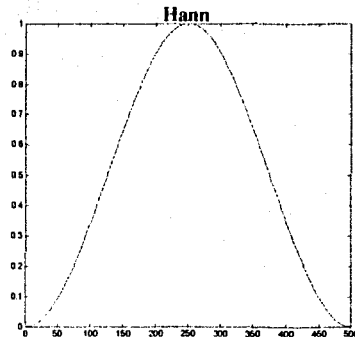
d) Función de Ventana Hann:

Una de las funciones de ventana más utilizadas, tiene forma de campana y a diferencia de las funciones triangular y de Bartlett reduce más los valores en las esquinas de la ventana de análisis.

Su fórmula es:

$$W(t) = \begin{cases} 0.5 \left[1 - \cos\left(\frac{2\pi(t-1)}{n-1}\right) \right] & 1 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

La gráfica 66 muestra el aspecto de la función de ventana de Hann:



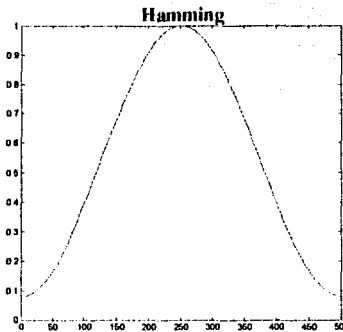
Gráfica 66 Función de Ventana de Hann.

e) Función de Ventana Hamming:

La función de ventana Hamming es semejante a la ventana de Hann, aunque aquí los valores de la función en las esquinas de la ventana son distintos de cero y su valor mínimo es de 0.08. La fórmula de la función de ventana de Hamming es:

$$W(t) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi(t-1)}{n-1}\right) & 1 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

Como se ve, la fórmula es prácticamente la misma que la función de Hann, con diferencia de los valores de 0.54 y 0.46, que en la función de Hann son de 0.5. La función de Hamming es conveniente cuando se desea minimizar los valores en los extremos de la ventana sin llegar a eliminarlos. La gráfica 67 muestra la función de ventana de Hamming.



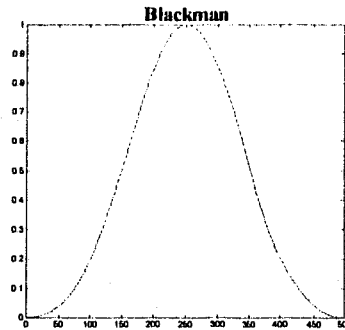
Gráfica 67 Función de Ventana de Hamming.

f) Función de Ventana Blackman:

La función de ventana de Blackman se caracteriza por disminuir drásticamente los valores en los extremos de la serie analizada. La curva en forma de campana es mucho más delgada que las funciones de Hann y Hamming. La fórmula para la función de Blackman es:

$$W(t) = \begin{cases} 0.42 - 0.5 \cos\left(\frac{2\pi(t-1)}{n-1}\right) + 0.08 \cos\left(\frac{4\pi(t-1)}{n-1}\right) & 1 \leq t \leq n \\ 0 & \text{otro_caso} \end{cases}$$

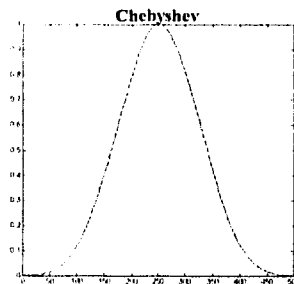
La función de ventana de Blackman es de hecho una forma más general de la fórmula utilizada en las funciones de Hann y Hamming, y puede utilizarse cuando prácticamente se desea descartar la información que aporten los extremos de la serie analizada. La gráfica 68 muestra la función de Blackman.



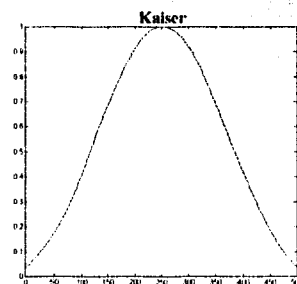
Gráfica 68 Función de Ventana de Blackman.

La elección de la función de ventana apropiada para el análisis de una serie utilizando la transformada de Fourier es arbitraria, y como no es tan sencillo en algunos casos determinar cuál es la mejor a utilizar, puede ser conveniente probar con diferentes ventanas, y comparar la información que proporcionan.

Las funciones descritas anteriormente son sólo algunas de las más utilizadas, existen desde luego muchas otras funciones (algunas muy sofisticadas) que pueden ser más convenientes en algunos casos. Como ejemplo podemos citar las funciones de Chebyshev y Kaiser (gráficas 69 y 70), las cuales utilizan ciertos parámetros que permiten ajustar la forma de la campana de acuerdo a nuestras necesidades (por ejemplo, una función que no reduzca a más del 50% los valores de la serie analizada). Desde luego, los criterios para la elección de dichos parámetros son un poco complicados, y en algunos casos se eligen de forma arbitraria.



Gráfica 69 Función de Chebyshev



Gráfica 70 Función de Kaiser

3.2.4 Transformada Discreta de Fourier con Ventanas

Debido a las características del método de ventaneo, donde se calculan los valores de las propiedades espectrales para subintervalos de tiempo de la serie, frecuentemente se truncan periodos al analizar la frecuencia mediante este método. El uso de las funciones ponderadoras soluciona el problema que se genera al truncar los periodos, permitiendo obtener una mejor estimación de la frecuencia de la serie a través del tiempo.

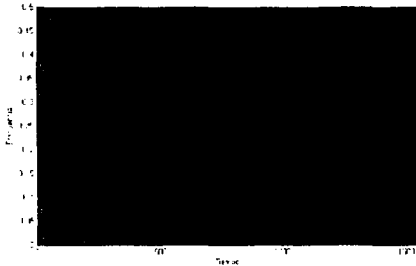
Para calcular la transformada de Fourier con ventanas, supongamos que deseamos utilizar un tamaño de ventana de S para realizar el análisis de la función. El primer paso es obtener una función de ventana ponderadora de tamaño S (seleccionada arbitrariamente), y posteriormente utilizar esta función de ventana para ponderar la serie para cada subconjunto de S datos, aplicando finalmente la transformada de Fourier para cada subconjunto de datos ponderado. Dicho de otra manera, tomamos primero los datos Z_1, Z_2, \dots, Z_S , los ponderamos con la función de ventana $W(1), W(2), \dots, W(S)$ y calculamos su transformada de Fourier; después procedemos a tomar los valores Z_2, Z_3, \dots, Z_{S+1} y los ponderamos también con $W(1), W(2), \dots, W(S)$, calculando su transformada de Fourier, y así sucesivamente.

Sin embargo, debe notarse que como la función de ventana esta definida para los tiempos 1 al S , los índices de ésta función deben ajustarse con el fin de ponderar el t -ésimo subconjunto de datos. Lo que equivale a recorrer virtualmente en el tiempo la función de ventana. De esta manera, podemos definir la transformada de Fourier con ventanas como:

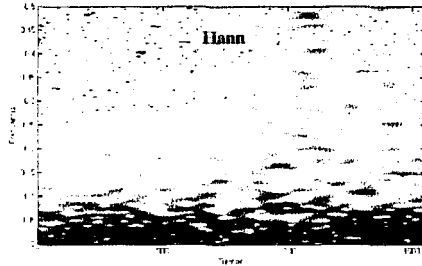
$$\tilde{Z}(k, t) = \sum_{u=1}^n Z_u \cdot W(u-t) \cdot e^{-iw_k u} \quad \begin{array}{l} t = 0, 1, 2, \dots, (n-S) \\ k = 0, 1, \dots, (n-1) \end{array}$$

en donde t representa el t -ésimo subconjunto de datos, para el cual se calculan los valores de la transformada en cada frecuencia k . Nótese que la función de ventana se recorre en el tiempo para

cada subconjunto de datos I . Las gráficas 71 y 72 muestran el espectro de Fourier con ventanas para la serie de tiempo SST, utilizando las funciones de ventana Rectangular y Hann, respectivamente. El tamaño de la ventana es de 180 datos (15 años), y al igual que las gráficas tridimensionales que se han presentado en este trabajo, los tonos más azules representan los valores más bajos del espectro, en tanto que los tonos más rojos representan los valores más altos.

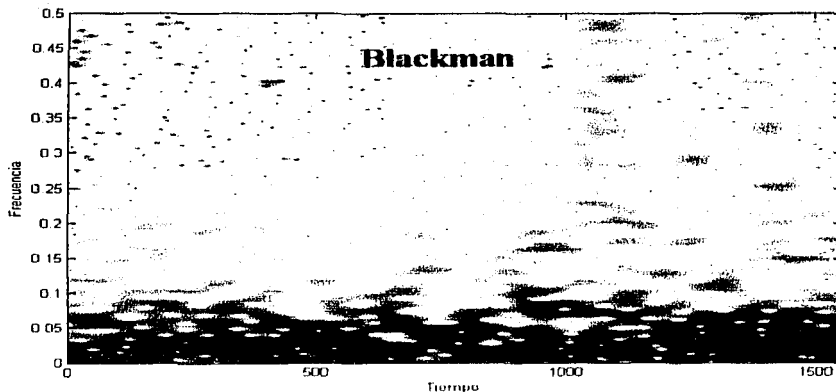


Gráfica 71 Espectro con ventana rectangular de tamaño 180 (15 años).



Gráfica 72 Espectro con ventana Hann de tamaño 180 (15 años).

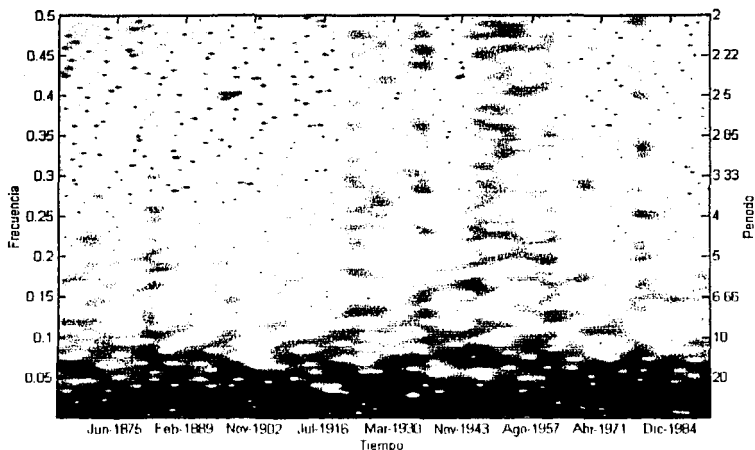
Salta a la vista la diferencia entre los espectros provocada por el uso de la función ponderadora, pues al utilizar la función de ventana rectangular, que equivale a no ponderar la serie, se tiene un espectro un poco difícil de interpretar, puesto que prácticamente toda la gráfica está en rojo. La función de ventana de Hann nos permite visualizar más claramente que las frecuencias más pequeñas son las que predominan, y sobre todo, detectar con mayor precisión las frecuencias relevantes durante ciertos subintervalos de tiempo.



Gráfica 73 Espectro con ventana de Blackman de tamaño 180 (15 años).

La gráfica 73 muestra el espectro obtenido utilizando la función de ventana de Blackman, el cual se asemeja mucho al espectro de Hann. Podemos notar que en la primera parte de la gráfica es más notoria la ausencia de frecuencias altas mayores a 0.3, más aún, las frecuencias más significativas se encuentran entre 0.05 y 0, esto significa que los periodos son mayores a 2 años.

Aunque la gráfica del espectro de Fourier con ventanas nos da una idea del comportamiento de la serie a través del tiempo, es útil determinar los intervalos de confianza asumiendo un espectro teórico apropiado a un nivel de significancia deseado. Para esto basta comparar cada subintervalo de S datos con el espectro teórico. En la gráfica 74 se observa el intervalo de confianza del 95% (línea negra) con la función de Blackman y utilizando un espectro teórico de Ruido Blanco

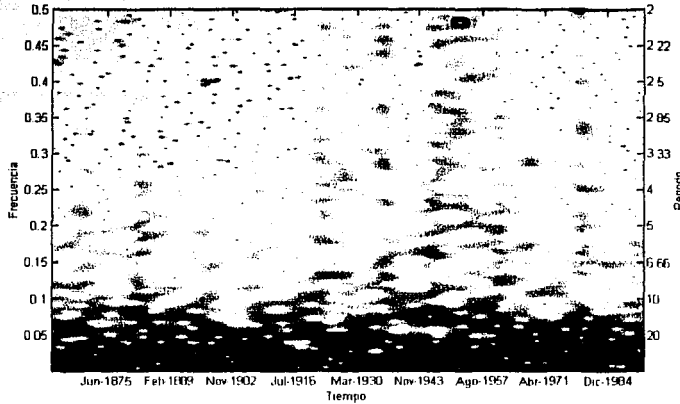


Gráfica 74 Espectrograma con espectro teórico de ruido blanco.

Se han agregado aquí las fechas intermedias de las ventanas en cuestión, por lo que por ejemplo, la fecha "Junio-1875" representa los valores de la transformada de Fourier utilizando los datos desde enero de 1868 a diciembre de 1882 (15 años). Así, podemos ver en la gráfica que el periodo caracterizado por la ausencia de frecuencias altas va de 1856 a 1923 aproximadamente. No obstante, es visible también que cuando se supone un espectro de ruido blanco, en general no hay frecuencias significativas mayores a 0.1.

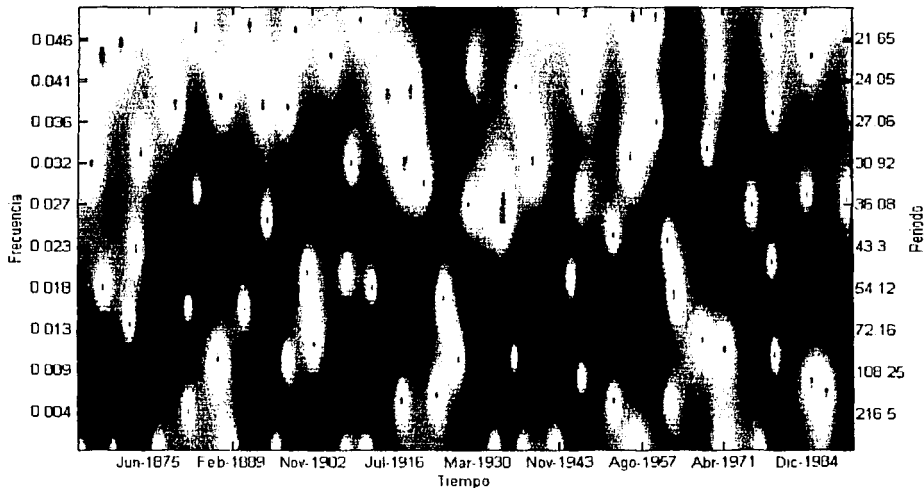
TESIS CON
FALLA DE ORIGEN

Sin embargo, como se había mencionado anteriormente para esta serie SST, el espectro teórico más apropiado es el de ruido rojo. Así pues, al aplicar un espectro teórico de ruido rojo con $\rho_1 = 0.7$ se obtiene el siguiente espectro:



Gráfica 75 Espectrograma con espectro teórico de ruido rojo.

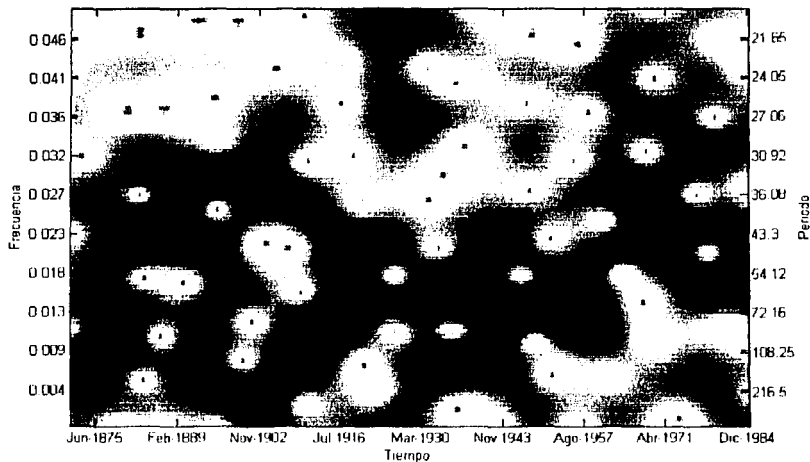
En esta gráfica el intervalo de confianza ha descartado ya ciertas frecuencias que se ajustan al comportamiento del ruido rojo. No obstante, con el fin de apreciar mejor las frecuencias predominantes, conviene analizar con mayor detalle el intervalo de frecuencia de 0 a 0.5.



Gráfica 76 Espectrograma con espectro teórico de ruido rojo (frec=0,0,5).

En la gráfica 76 podemos observar que los periodos de las frecuencias predominantes se encuentran entre 22 y 108, es decir que el periodo de los datos va de los 2 a los 9 años. No obstante, el periodo de 2 años es característico particularmente del periodo de 1921 a 1934. Muchos periodos específicos de tiempo pueden analizarse de esta manera, como por ejemplo, el último periodo predominante con valores de 4 a 5 años.

Naturalmente, cada función de ventana y cada escala proporcionarán información variada sobre la serie, considerando las características de las funciones que se han descrito anteriormente. La gráfica siguiente es el espectro de la misma serie SST, utilizando la función de ventana de Hann, con un tamaño de ventana de 360 datos (30 años).



Gráfica 77 Espectrograma con ventana de 30 años, (Hann, Espectro teórico de ruido rojo)

Como es de esperarse, este espectro proporciona información más general de la serie, por ejemplo, de 1911 a 1972 podemos ver que la frecuencia que predominó a grandes rasgos fue de 0.015, es decir, un periodo de 6 años, por otro lado, la frecuencia predominante en la actualidad corresponde a un periodo de 4 a 5 años.

TESIS CON
FALLA DE ORIGEN

El análisis de Fourier con Ventanas tiene el principal defecto de que las elecciones de las escalas, la función de ventana, y el espectro teórico utilizado, se realizan de forma arbitraria. De esta manera, es común tener que probar con muchas escalas y funciones de ventana para realizar un análisis apropiado.

3.3 Análisis Wavelet

Considerando las desventajas que presenta utilizar el análisis de Fourier con ventanas, que obligan al analista a tener que calcular una cantidad considerable de espectros para cada tamaño de ventana y función ponderadora, además de que los cálculos que se realizan para cada espectro, pueden requerir varios minutos en una computadora, dependiendo de la resolución que se desee, y el número de datos que se analizan, es importante tener en cuenta que aún una computadora moderna puede volverse extremadamente lento al analizar miles o millones de datos, y que por tanto un análisis para el que cada espectro tarde varios minutos, sería extremadamente tedioso y lento.

El análisis Wavelet, representa una alternativa útil al análisis de Fourier, especialmente cuando se busca analizar valores de frecuencia no estacionarios.

3.3.1 Funciones Wavelet

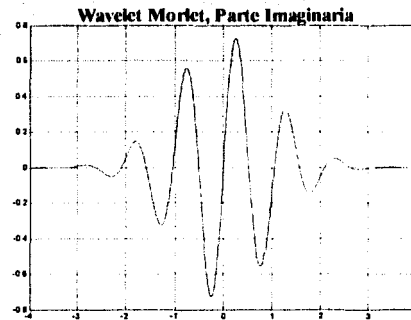
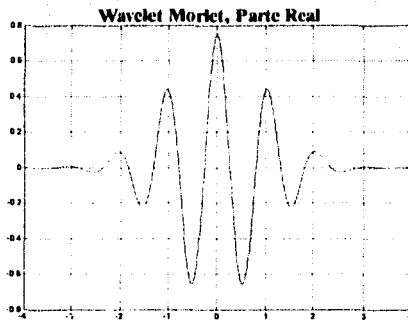
Al igual que la transformada de Fourier representa una función dada como una suma de sinusoides con diferentes frecuencias, también es posible utilizar funciones de otra naturaleza, más aún, es posible utilizar funciones aperiódicas, sobre todo cuando se desea analizar datos no estacionarios. Las funciones Wavelet, también llamadas "onditas", "ondículas" u "onduletas", tienen precisamente la función de representar la serie mediante una suma de funciones de diferentes formas, longitudes, y ubicaciones en el tiempo. Sin embargo, a diferencia de las sinusoides utilizadas en la transformada de Fourier, las funciones Wavelet tienen ciertas características que permiten ubicar la

función simultáneamente en el tiempo y la frecuencia, tal como se realizó con la transformada de Fourier con ventanas.

Supongamos que tenemos una función $\psi_0(t)$, dependiente del parámetro de tiempo "t". Para que sea admitida como una wavelet, esta función debe tener una media de cero y encontrarse definida tanto en el dominio del tiempo como en el dominio de la frecuencia (Farge, 1992), además de tener una duración finita. Un ejemplo de función wavelet es la llamada wavelet Morlet, que se obtiene de una senoide modulada por una campana de Gauss:

$$\psi_0(t) = \pi^{-1/4} e^{iw_m t} e^{-t^2/2}$$

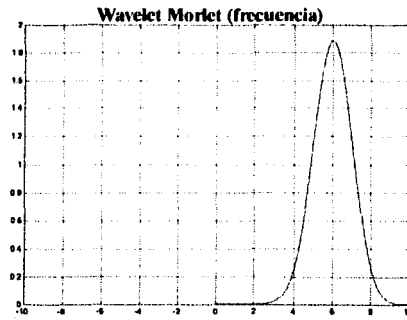
en donde w_m es el parámetro de frecuencia⁹. En las gráficas 78 y 79 se presentan la parte real e imaginaria de la wavelet Morlet utilizando $w_m = 6$ para satisfacer la condición de admisibilidad.



Gráficas 78 y 79 Parte Real e Imaginaria de la Wavelet Morlet, $w_m = 6$

En contraste con las sinusoides de Fourier, que tienen una duración infinita, las funciones wavelet sólo están definidas durante un intervalo específico de tiempo. Por otro lado, como se mencionó en la condición de admisibilidad, es posible apreciar que la wavelet Morlet tiene una media de cero, y obteniendo la transformada de Fourier (continua), podemos corroborar que se encuentra definida también en el dominio de la frecuencia (gráfica 80).

⁹ Generalmente se define la wavelet Morlet utilizando la notación para el parámetro de frecuencia como w_0 , sin embargo, para evitar confusión con la frecuencia angular fundamental de Fourier, se ha utilizado w_m .



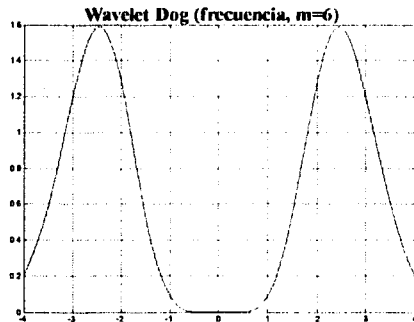
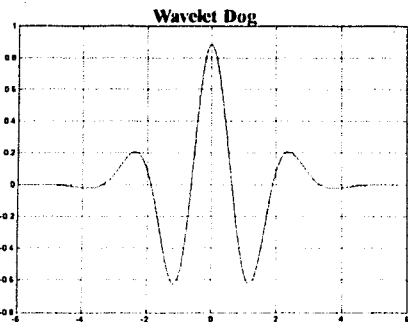
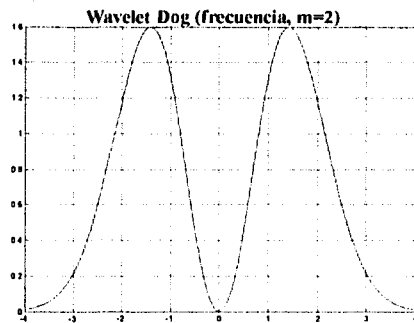
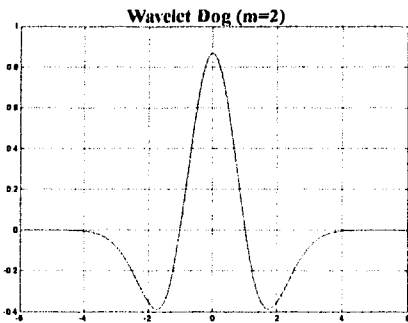
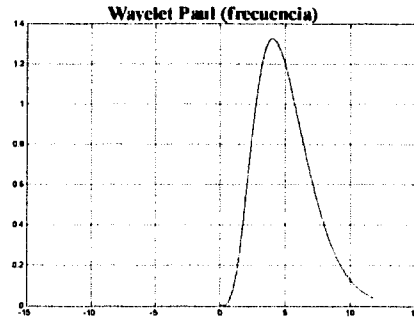
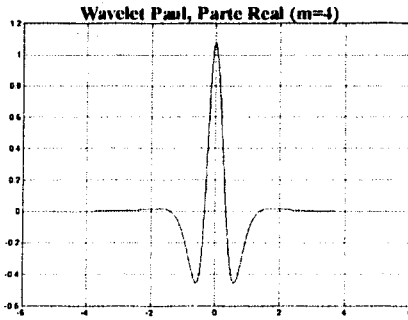
Gráfica 80 Wavelet Morlet en el dominio de la frecuencia.

Aunque la wavelet Morlet es la función más conocida y una de las más utilizadas, existe una infinidad de funciones que satisfacen las condiciones de admisibilidad. No obstante, en este trabajo se incluyen sólo tres familias de funciones de las más útiles, conocidas como wavelets Morlet, Paul y Dog, y cuyas funciones se definen en la tabla siguiente:

Nombre	$\psi_0(t)$
Morlet (w_m = frecuencia)	$\pi^{-1/4} e^{i w_m t} e^{-t^2/2}$
Paul (m = orden)	$\frac{2^m i^m m!}{\sqrt{\pi(2m)!}} (1 - it)^{-(m+1)}$
Dog (m = derivada)	$\frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m + \frac{1}{2}\right)}} \frac{d^m}{dt^m} \left(e^{-t^2/2} \right)$

Como se ve, las funciones Morlet y Paul son complejas, en tanto que la función Dog, obtenida mediante la m -ésima derivada de una campana de Gauss, es real. Cabe señalar que no cualquier valor de los parámetros de estas fórmulas satisface las condiciones de admisibilidad, por lo que estos parámetros deben seleccionarse con cuidado. Las funciones wavelet Paul y Dog se muestran en las gráficas 81 a 86, debe notarse la diferencia en el comportamiento en la frecuencia para estas tres wavelets. A la wavelet Dog con $m=2$, también se le conoce como "Sombrero Mexicano".

TESIS CON
FALLA DE ORIGEN



Gráficas 81-86 Wavelets Paul ($m = 4$) y Dog ($m = 2, 6$)

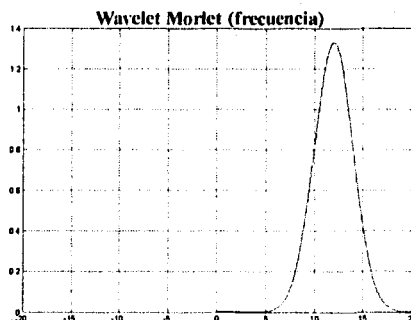
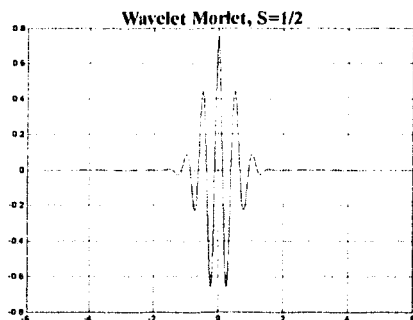
A la forma original de la función wavelet se le conoce comúnmente como "wavelet madre", pues a partir de esta función podemos obtener versiones de la wavelet que tengan una duración de tiempo

TESIS CON
FALLA DE ORIGEN

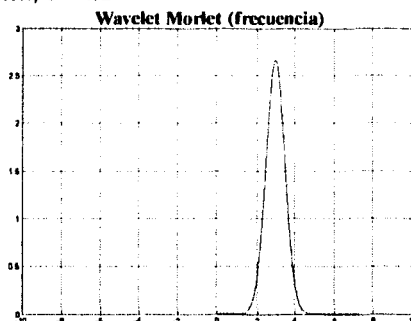
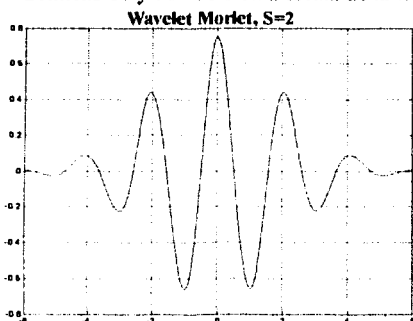
más corta o mas larga que la función original. Así, la versión a "escala"¹⁰ de la wavelet madre se obtiene al evaluar la función original para (t/s) , y se escribe como:

$$\psi_s(t) = \psi_0\left(\frac{t}{s}\right)$$

Es importante notar la semejanza entre las versiones a escala de la wavelet madre y las sinusoides de Fourier con diferentes frecuencias, en donde al dividir $\frac{2\pi k}{n}$ se obtiene una versión de la senoide que ejecuta k ciclos cada n datos. De la misma manera, las funciones wavelet se expanden en la frecuencia cuando se acorta la duración de la wavelet a escala, y se comprimen en la frecuencia cuando se expande la duración de la wavelet.



Gráficas 87 y 88 Versión a escala de la wavelet Morlet, S = 1/2



Gráficas 89 y 90 Versión a escala de la wavelet Morlet, S = 2

¹⁰ Este concepto de escala difiere del que se ha manejado hasta ahora, pues en un caso la "escala de análisis" se refiere al tamaño de la ventana que se utiliza, en tanto que aquí se refiere a una versión a escala de la wavelet madre, no obstante existe cierta relación entre ambos conceptos, pues el tamaño de ventana depende de la escala de la wavelet madre.

Las gráficas 87 a 90 muestran el efecto del uso de la escala en los dominios del tiempo y frecuencia, se observa que en la frecuencia la wavelet madre tiene su valor más alto en 6, en tanto que al obtener la escala de $\frac{1}{2}$ de la wavelet madre, ésta reduce su duración en el tiempo, en tanto que los valores de frecuencia se expanden para tomar un valor medio de 12. Por otro lado, al utilizar una escala de 2, la función se expande en el tiempo y los valores de frecuencia se reducen tomando un valor medio de 3.

De esta manera, podemos ver la relación que existe entre la escala de las funciones wavelet y la frecuencia de las sinusoides de Fourier. Así, al utilizar diferentes escalas de la wavelet madre se puede decir que estamos utilizando diferentes componentes de frecuencia. Sin embargo hay que tener en cuenta que la escala de wavelet no es precisamente equivalente a la frecuencia de Fourier, dado que en realidad la equivalencia depende del comportamiento que presente la wavelet madre.

Como las funciones wavelet tienen una duración finita, para representar la serie que se desea analizar como una suma de wavelets, se requiere utilizar no solamente versiones a escala de la wavelet madre, sino que además deben recorrerse en el tiempo estas funciones. Así, la wavelet a escala y recorrida en el tiempo se representa como:

$$\psi_{s,t}(u) = \psi_s(u-t) = s^{-1/2} \psi_0\left(\frac{u-t}{s}\right)$$

en donde $s^{-1/2}$ es el factor de normalización, que hace que la norma sea igual a uno para cada valor de tiempo t . Esta normalización, como se ve debe realizarse para cada escala por separado para hacer que el espectro de cualquier escala sea directamente comparable con otra, y la norma sea uno.

TESIS CON
FALLA DE ORIGEN

3.3.2 Transformada Wavelet

El término "función wavelet" es usado en general para referirse tanto a las wavelets ortogonales como no ortogonales. Las funciones wavelet que se definieron anteriormente se utilizan para calcular la transformada continua de wavelet, no obstante, pueden utilizarse para analizar un conjunto discreto de datos. Para calcular la transformada discreta de wavelet, se requiere utilizar una base ortogonal de wavelets. No obstante, la transformada discreta de wavelet es, en la práctica, poco utilizada, considerando que cada wavelet tiene una base distinta, además de que la gráfica del espectro de wavelet es mucho más comprensible si se utiliza la transformada de wavelet continua.

La transformada continua de wavelet para una secuencia discreta de datos Z_n se define como la convolución de Z_n con una versión a escala y recorrida en el tiempo de la wavelet madre.

$$W(s, t) = \sum_{n=0}^{n-1} Z_n \psi^* \left(\frac{u-t}{s} \right)$$

en donde el asterisco (*) representa el conjugado complejo.

Aunque la transformada de wavelet puede ser calculada utilizando esta fórmula, es posible obtener la transformada de wavelet utilizando el teorema de convolución para la transformada de Fourier, de acuerdo con el cual la transformada de Fourier de una convolución es equivalente al producto de las transformadas de Fourier. Así, podemos expresar la convolución anterior como:

$$W(s, t) = \frac{1}{n} \sum_{k=0}^{n-1} \hat{Z}_k \hat{\psi}^* (s w_k) e^{i w_0 k t}$$

Aunque a simple vista puede parecer que esta nueva forma es una complicación innecesaria de la fórmula original, su utilidad principal consiste en que puede aplicarse la llamada "transformada rápida de Fourier" (No tratada en este trabajo), un algoritmo muy útil en ordenadores que obtiene la transformada de Fourier de manera eficiente y con una diferencia significativa de tiempo. De esta

TESIS CON
FALLA DE ORIGEN

manera, utilizando el teorema de convolución es posible calcular la transformada de wavelet continua (para una escala s) en todos los valores de tiempo de manera simultánea y eficiente. Así, el tiempo de cálculo queda reducido de una manera drástica.

Debe notarse que la transformada de la serie se normaliza dividiendo entre n . Por otro lado, observamos que la transformada de Fourier de la función wavelet $\psi_0\left(\frac{t}{s}\right)$ da como resultado una función $\hat{\psi}(sw_k)$, en donde el parámetro se ve afectado por la escala de manera inversamente proporcional a la función en el dominio del tiempo, por lo que la transformada de Fourier de la wavelet en cuestión es una función de sw_k , es decir, las frecuencias de Fourier se ven multiplicadas por el factor de escala s . Por otro lado, así como en la fórmula de convolución se usa una versión normalizada de $\psi_0(t/s)$, tenemos que $\hat{\psi}(sw_k)$ es la transformada de una forma normalizada de la wavelet madre, en donde la normalización puede obtenerse también como:

$$\hat{\psi}(sw_k) = (2\pi s)^{1/2} \hat{\psi}_0(sw_k)$$

La transformada de Fourier de la wavelet madre se puede obtener analíticamente para cada wavelet. La tabla siguiente¹¹ muestra las transformadas de Fourier para las wavelets Morlet, Paul y Dog.

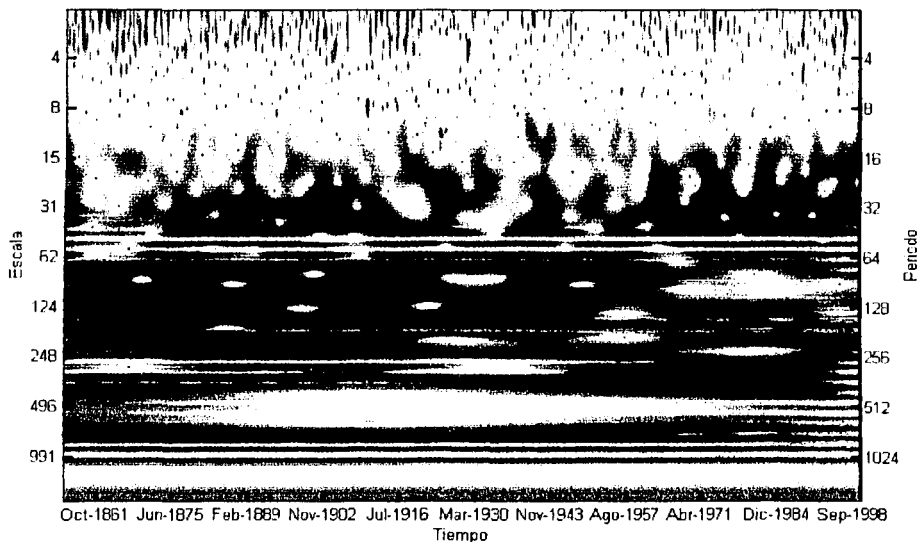
Nombre	$\hat{\psi}_0(sw_k)$
Morlet (w_m = frecuencia)	$\pi^{-1/4} H(w_k) e^{-(sw_k - w_m)^2/2}$
Paul (m = orden)	$\frac{2^m}{\sqrt{m(2m-1)!}} H(w_k) (sw_k)^m e^{-sw_k}$
Dog (m = derivada)	$\frac{i^m}{\sqrt{\Gamma\left(m + \frac{1}{2}\right)}} (sw_k)^m e^{-(sw_k)^2/2}$

¹¹ Datos tomados de Torrence & Compo, 1997, "A practical guide for Wavelet Analysis".

en donde $H(w_k)$ es una función de peso, de manera que $H(w_k) = 1$ si $w_k > 0$, y $H(w_k) = 0$ en otro caso.

De esta manera, teniendo la fórmula para la transformada de Fourier de nuestra función wavelet, el uso del teorema de convolución es relativamente sencillo y rápido de realizar en un ordenador, donde la fórmula de convolución original puede tardar varios minutos para calcular el espectro.

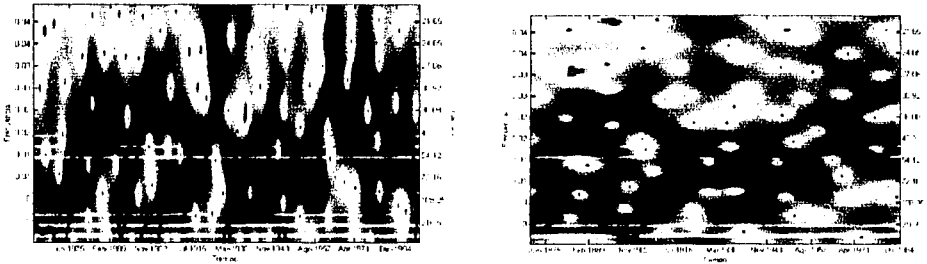
Al igual que en la transformada de Fourier, los valores de la transformada de wavelet dan una idea de la fuerza de aportación para cada escala y tiempo de la wavelet madre en la representación de la serie original. Así, calculando los valores para cada escala y tiempo de interés, y tomando la magnitud al cuadrado de los valores de la transformada de wavelet $|W(s,t)|^2$, podemos obtener una gráfica parecida a la que se obtiene en la transformada de Fourier con ventanas.



Gráfica 91 Espectro de Wavelet de la serie SST usando la wavelet Morlet, ($w_m = 6$)

El espectro de wavelet de la serie de tiempo SST, mostrado en la gráfica 91, representa el comportamiento de la serie de manera mucho más clara que el espectro de Fourier con ventanas,

pues sólo se requiere calcular un espectro para obtener información clara sobre el comportamiento de la serie; en contraste con la transformada de Fourier, que revela aspectos dependiendo del tamaño de la ventana. Podemos ver, por ejemplo, que entre 1916 y 1930 está presente el periodo de dos años, al igual que los espectros de Fourier presentados anteriormente para 15 y 30 años (gráficas 92 y 93), pero a diferencia de los espectros de Fourier, el espectro de wavelet muestra que este periodo de dos años también se presenta en los años posteriores a 1960. A la derecha del espectro de wavelet se han marcado los periodos equivalentes de Fourier para cada escala.



Gráficas 92 y 93 Espectro de Fourier con ventana de Hann, ventanas de 15 y 30 años

Para obtener la equivalencia entre la escala de wavelet y el periodo de Fourier, comparamos una senoide de frecuencia conocida con la wavelet madre, es decir, calculando la transformada de wavelet de una senoide de frecuencia conocida podemos determinar el valor de escala equivalente a la frecuencia de esta senoide. Para la wavelet Morlet con $w_m = 6$, esto nos da un periodo de $\lambda = 1.03s$, en donde λ es el periodo de Fourier equivalente a la escala s . Las fórmulas de equivalencia para las otras funciones wavelet están dadas en la siguiente tabla (Torrence & Compo, 1997):

Nombre	Periodo de Fourier λ
Morlet ($w_m =$ frecuencia)	$\frac{4\pi s}{w_m + \sqrt{2 + w_m^2}}$
Paul ($m =$ orden)	$\frac{4\pi s}{2m + 1}$
Dog ($m =$ derivada)	$\frac{2\pi s}{\sqrt{m + 1/2}}$

TESIS CON
FALLA DE ORIGEN

Cabe señalar que para realizar el cálculo de la transformada de wavelet, es posible utilizar un conjunto arbitrario de escalas, sin embargo, para obtener una mejor apreciación de los periodos, muchos autores recomiendan escribir las escalas como potencias fraccionarias de 2, esto es:

$$s_j = s_0 2^{j\delta_j}, \quad j = 0, 1, \dots, J$$

$$J = \delta_j^{-1} \log_2(N/s_0)$$

en donde s_0 es la escala más pequeña que se va a utilizar, en tanto que J es el número de escalas que se van a utilizar, esto es, la resolución vertical del espectro de wavelet. El parámetro δ_j es también un factor de resolución, pues es el incremento en las potencias de 2 que se requiere para que se tengan las J escalas. De esto se infiere que uno de los dos parámetros debe ser asignado arbitrariamente por el analista. En tanto que s_0 se recomienda elegirlo de manera que el periodo de Fourier equivalente sea de 2, para una interpretación más comprensible del espectro, pues los periodos de Fourier quedan escritos como potencias de dos (Ver gráfica 91).

3.3.3 Intervalo de Confianza y Cono de Influencia

Para obtener el intervalo de confianza del espectro de wavelet, se requiere primero que nada normalizarlo. Cabe aclarar que la normalización que se realizó al obtener la wavelet a escala sólo es relativa a la escala en cuestión y su finalidad es la de hacer comparables los valores del espectro para cada escala. De esta manera, para hacer un espectro de wavelet comparable con otro, se requiere realizar una segunda normalización, ahora relativa a todo el espectro.

Sin embargo, la normalización del espectro de wavelet es un tanto complicada debido a que no se ha utilizado una base ortogonal, lo que origina cierta redundancia. Así, para calcular el valor esperado del espectro de wavelet, debemos considerar esta redundancia. De acuerdo con Torrence & Compo (1997), el valor esperado del espectro está dado por:

TESIS CON
FALLA DE ORIGEN

$$E\left(|W(s,t)|^2\right) = \frac{\delta_j}{nC_\delta} \sum_{l=1}^n \sum_{j=0}^J \frac{|W(s,t)|^2}{s_j}$$

en donde C_δ es una constante para cada función wavelet, y puede calcularse como

$$C_\delta = \frac{\delta_j}{n\psi_0(0)} \sum_{j=0}^J \left\{ s^{-1/2} \operatorname{Re} \left[\sum_{k=0}^{n-1} \hat{\psi}^*(s w_k) \right] \right\}$$

Los valores de estas constantes para las wavelets descritas en este trabajo se muestran en la siguiente tabla:

Nombre	C_δ	$\psi_0(0)$
Morlet ($w_m = 6$)	0.776	$\pi^{-1/4}$
Paul ($m = 4$)	1.132	1.079
Dog ($m = 2$)	3.541	0.867
Dog ($m = 6$)	1.966	0.884

Utilizando el valor esperado del espectro, podemos normalizar la serie dividiendo el espectro entre su valor esperado, para que de esta manera, se pueda comparar directamente con el espectro teórico deseado. Así, la forma normalizada del espectro se distribuye aproximadamente como:

$$\frac{|W(s,t)|^2}{E\left(|W(s,t)|^2\right)} \Rightarrow P_s \chi_2^2$$

donde la función P_s es el valor medio del espectro teórico, en términos de la escala utilizada, y para las wavelets complejas, los valores del espectro se distribuyen como una chi-cuadrada con dos grados de libertad, en tanto que para las wavelets reales se tiene un grado de libertad.

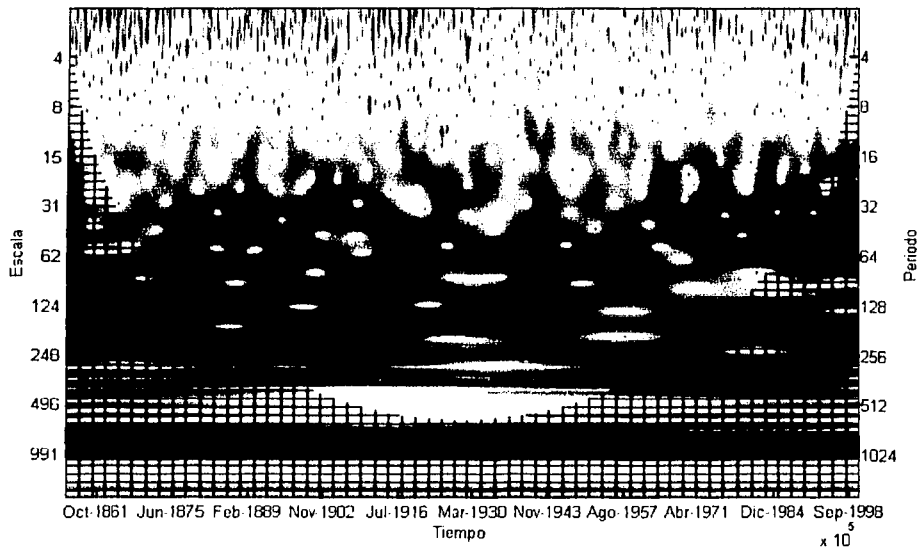
La función de media del espectro teórico que se utiliza para obtener el intervalo de confianza es prácticamente la misma que se utiliza para el análisis de Fourier, con la única diferencia de que se

TESIS CON
FALLA DE ORIGEN

sustituye la frecuencia de Fourier (k/n) por el valor equivalente de frecuencia de Fourier para cada escala de wavelet ($1/\lambda(s)$). Así, realizando la sustitución para el espectro de ruido rojo, obtenemos la función en términos de la escala de wavelet.

$$P_s = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos(2\pi / \lambda(s))}$$

De la misma manera que en el análisis de Fourier, los valores de α se asignan arbitrariamente, aunque podemos auxiliarnos de la autocorrelación de la serie para encontrar el valor apropiado. La gráfica 94 muestra el espectro de la serie SST con espectro teórico de ruido rojo y $\alpha = 0.7$



Gráfica 94 Espectro Wavelet de la serie SST, wavelet Morlet con Ruido Rojo.

Varias de las características de esta serie han sido ya discutidas con anterioridad, sin embargo, una de las características más relevantes de la serie se alcanza a apreciar apenas en esta gráfica, pues notamos que en los periodos de 1875–1920, y 1960–2002 existen varios patrones y cambios en la frecuencia de la serie. Considerando que la serie representa la temperatura de la superficie

del mar meridional, esto significa que en estos periodos hubo varios eventos de calor y de frío extremos (conocidos como fenómenos de "el niño" y "la niña"), mientras que en el periodo de 1920 a 1960 hubo pocos eventos. Observamos también, que de 1900 a 1920 se tiene un cambio gradual de periodos largos a periodos cortos, llegando a alcanzar los 2 años. También se nota en la grafica un cambio en el periodo de 1960 a 2002 de periodos cortos a largos.

Debe notarse que a la gráfica 94 se le agregó una cuadrícula con límites en forma de semicírculo, el área que cubre esta cuadrícula se le conoce como "cono de influencia", y representan los valores de los datos que pueden encontrarse ligeramente distorsionados. Recordemos que cuando aplicamos una ventana a una serie de tiempo, dado que la serie es finita, se obtiene un número de datos menor que la serie original. Pueden compararse, por ejemplo, las gráficas 6 y 7 al principio de este trabajo, donde la base de la gráfica para la función de media con diferentes tamaños de ventana tiene forma de trapecio debido a la pérdida de datos.

Consideremos por ejemplo, que deseamos calcular el valor de la transformada de wavelet para el tiempo 1 con escala de 10, entonces tenemos que

$$W(10,1) = \sum_{u=1}^n Z_u \psi^* \left(\frac{u-1}{10} \right)$$

en donde se ve que la convolución inicia desde el valor de tiempo cero de la wavelet a escala, que es en realidad el punto medio de la wavelet. Podemos ver que utilizar esta convolución truncada equivale a considerar que los datos previos al dato 1 sean ceros. De esta manera, entre más cerca esté un valor de la esquina de la gráfica a escalas grandes, más ceros entran al análisis y se producen discontinuidades. De esta manera, el cono de influencia nos da una idea de la región donde el espectro se ve afectado por los efectos del borde, y por ende de la región donde es más difícil distinguir una discontinuidad generada por los ceros, de un componente de frecuencia.

Cuando se usa el teorema de convolución para calcular la transformada de wavelet se tiene un problema adicional, puesto que al utilizar la transformada de Fourier, se asume que los datos son

TESIS CON
FALLA DE ORIGEN

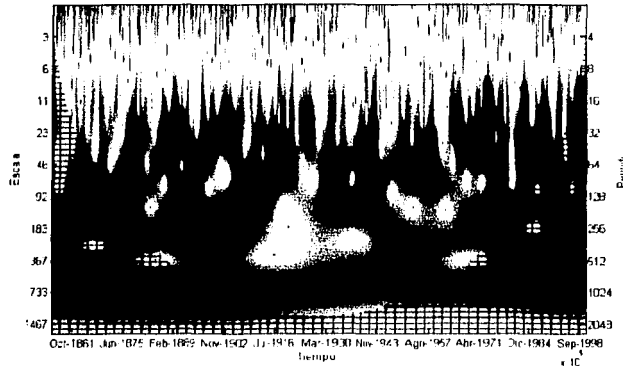
cíclicos, y que por tanto los valores anteriores al dato uno son los últimos datos de la serie. Lo mismo sucede al final de la serie, donde se asume que el dato $n+1$ es el dato 1, y así sucesivamente. Sin embargo, este problema se puede solucionar agregando ceros al final de la serie, y removiéndolos después de calcular la transformada, lo que si bien produce discontinuidades en los extremos, es mejor que asumir que los datos son periódicos. Por otro lado, cuando una serie es periódica, no existe cono de influencia y no es necesario agregar ceros al análisis mediante el teorema de convolución.

El cono de influencia se define mediante el "tiempo de influencia e " para los valores del espectro en cada escala. Este tiempo es elegido de tal manera que al agregar una discontinuidad en la orilla, la influencia de la discontinuidad disminuya hasta el punto de considerarse despreciable¹², y se asume que las discontinuidades en las esquinas no pueden afectar a los valores del espectro por arriba de este valor de tiempo. La tabla siguiente muestra los valores de tiempo de influencia que se suelen utilizar para las wavelets Morlet, Paul y Dog.

Nombre	$\psi_0(t)$
Morlet (w_m = frecuencia)	$\sqrt{2}s$
Paul (m = orden)	$s/\sqrt{2}$
Dog (m = derivada)	$\sqrt{2}s$

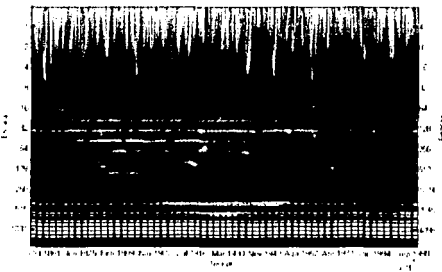
En la gráfica 95 se muestra el espectro de wavelet de la serie SST utilizando la wavelet Paul ($m=4$). Debe observarse que el cono de influencia es más estrecho que el de la wavelet Morlet.

¹² Generalmente el punto en que se considera despreciable es cuando el efecto de la discontinuidad se ve reducido por el factor de e^{-2} .



Gráfica 95 Espectro Wavelet de la serie SST, wavelet Paul ($m=4$), con Ruido Rojo ($\alpha = 0.7$).

Como se ve en las fórmulas para el tiempo de influencia, las wavelet Morlet y Dog tienen el mismo cono de influencia, éste está determinado en general por la longitud de la wavelet que se utiliza. De esta manera, como la wavelet Paul tiene una corta duración en comparación con las wavelets Morlet y Dog, su espectro tiene en general formas más estrechas y su cono de influencia es más pequeño.



Como se ha mostrado en este apartado, el análisis wavelet es una alternativa eficiente para analizar frecuencias no estacionarias, puesto que el espectro de wavelet refleja con mayor claridad las características de la frecuencia de la serie a través del tiempo, en contraste con la transformada de Fourier, donde el tamaño de ventana puede impedirnos detectar algunos patrones por observar "muy de cerca" o "muy de lejos" los datos, recordando que con tamaños de ventana pequeños se aprecian rasgos de la serie a corto plazo, y los tamaños de ventana grandes muestran el comportamiento general de la serie, pero ocultando en muchas ocasiones los detalles menores. El análisis wavelet, por el contrario, utiliza un tamaño de ventana apropiado para cada frecuencia, y el coeficiente de wavelet representa la fuerza de aportación de esta frecuencia y tiempo a la representación de la serie original.

Al representar la serie de tiempo como una suma de funciones wavelet, es muy importante no olvidar la naturaleza de lo que se está calculando, e interpretar el espectro como una suma de cálculos maravillosos que indican como se comporta la frecuencia de la serie. La transformada de wavelet tiene en el fondo la misma naturaleza que la transformada de Fourier, puesto que se puede representar la serie como

$$Z_u = \sum_{l=1}^n \sum_s^S \tilde{c}_{s,l} \psi\left(\frac{u-l}{s}\right)$$

en donde los coeficientes de wavelet $\tilde{c}_{s,l}$ representan la fuerza de aportación para cada escala y tiempo de la wavelet madre. Si suponemos que se tiene una base de funciones ortogonales, entonces los coeficientes de wavelet quedarían dados por

$$\tilde{c}_{s,l} = \frac{\left\langle Z_u, \psi\left(\frac{u-l}{s}\right) \right\rangle}{\left\| \psi\left(\frac{u-l}{s}\right) \right\|^2} = \sum_{u=1}^n Z_u \psi^*\left(\frac{u-l}{s}\right)$$

de manera que es el coeficiente de proyección ortogonal, equivalente a la transformada de wavelet discreta. La norma es uno debido a que la wavelet fue normalizada al obtener su escala. Sin embargo, como en la transformada continua no tenemos una base de funciones ortogonales, los

TESIS CON
FALLA DE ORIGEN

valores de la transformada continua de una serie discreta son los valores que tendría el coeficiente de wavelet si un elemento dado fuera parte de una base ortogonal que generara a Z_n .

Como se mencionó anteriormente, el espectro de wavelet tiene cierta semejanza con la función de media con ventanas que se mostró al principio del trabajo, pues como ambas se obtienen de recorrer una ventana a través del tiempo, ambas gráficas presentan muchas semejanzas. En un caso se calcula la media, en tanto que en el otro se calcula una convolución, pero ambos cálculos son relativos a un tamaño de ventana (o escala). En el caso de la función de media no se calcularon los valores donde no se alcanzaron a tener los s datos, sin embargo, es posible también ahí agregar ceros para completar los s datos, aunque desde luego los valores de media se verían disminuidos entre más ceros entren al análisis.

Al comparar la función de media con diferentes tamaños de ventana de las gráficas 6 y 7 con los espectros de wavelet, y observar que a pesar de la gran diferencia en complejidad que las separa, es muy interesante apreciar que, en el fondo, mantienen la misma esencia, pues a final de cuentas son resultado del mismo método, el análisis de ventaneo.

TESIS CON
FALLA DE ORIGEN

Conclusiones

El uso de métodos de ventaneo para analizar series de tiempo no estacionarias aporta muchas ventajas significativas en comparación otros métodos, especialmente porque el uso de ventanas permite observar los cambios en la serie tanto a largo plazo como a corto plazo, dado que el tamaño de la ventana determina el tipo de cambios que podrán apreciarse.

El análisis de ventaneo presenta características y utilidades diferentes para cada propiedad espectral, en donde lo que más debe remarcarse, es que el análisis de ventaneo, más que una alternativa a los métodos de análisis y pronóstico mencionados en el trabajo, representa un complemento muy útil de ellos, y de fuerte potencial descriptivo.

Al analizar la función de media utilizando medias móviles, pudimos observar que el uso de ventanas tiene el defecto de ser sensible a la omisión de datos, no obstante, una de las grandes ventajas que aporta es el suavizamiento de los datos atípicos. A diferencia de los modelos de regresión, las medias móviles pueden aplicarse a prácticamente cualquier serie, e inclusive en los casos con algunos cambios estocásticos en la tendencia. Sin embargo, ambos métodos pueden combinarse para compensar mutuamente sus defectos, puesto que el ventaneo permite realizar un análisis más apropiado de la serie, mientras que el modelo de regresión permite predecir el comportamiento de la tendencia, aunque se puede ver afectado por los datos atípicos. Por otro lado, la regresión permite extrapolar los datos que se han perdido al calcular las medias móviles.

En lo que se refiere al análisis de la varianza de la serie, el uso de ventanas proporciona también información apropiada del comportamiento a través del tiempo, y permite detectar cuando existe un comportamiento sistemático, en cuyo caso puede aplicarse una transformación a los datos si se desea realizar el análisis de la autocovarianza mediante la metodología de Box y Jenkins.

TESIS CON
FALLA DE ORIGEN

El análisis de ventanas aplicado a la función de autocovarianza nos permite detectar con mayor claridad si se tiene una autocovarianza no estacionaria, en cuyo caso se puede localizar aproximadamente el punto en que inicia el patrón de comportamiento actual, de manera que si se utilizan únicamente los datos a partir de ese punto es posible utilizar la metodología de Box y Jenkins debido a que ese fragmento de la serie es estacionario en la autocovarianza.

Finalmente, el uso de ventanas encuentra en el análisis de la frecuencia una de sus más importantes y populares utilidades, en donde aunado al análisis de Fourier con ventanas, el análisis Wavelet se presenta como una alternativa novedosa al análisis de Fourier. Así, tanto el análisis de Fourier como el de Wavelet proporcionan información relevante sobre la periodicidad de la serie. Sin embargo, dado que el análisis Wavelet utiliza un tamaño de ventana apropiado para cada frecuencia al mantener la proporción con respecto a la wavelet madre, presenta en un solo espectro de manera clara y rápida los rasgos característicos de la frecuencia de la serie. Es claro que el análisis Wavelet supera por mucho a su predecesor.

En general, cuando se aplica el uso de ventanas se asume que los subintervalos que se analizan tienen un comportamiento estacionario, puesto que el cálculo de las propiedades espectrales se realiza de la misma manera que si se tratara de un proceso estacionario. Así, para que el análisis mediante ventanas dé resultados apropiados, el tamaño de ventana debe ser suficientemente pequeño para que el comportamiento de los subintervalos de la serie sean aproximadamente estacionarios.

El principal defecto que se presenta en el análisis de ventanas es la elección arbitraria del tamaño de ventana, que en ocasiones puede volver tedioso el análisis al hacerse necesario realizar los cálculos con múltiples tamaños de ventana. Sin embargo, el análisis Wavelet muestra que es posible obtener alternativas de análisis que sean independientes del tamaño de ventana, y aunque la media, la varianza y la frecuencia pueden ser analizados apropiadamente mediante los métodos

TESIS CON
FALLA DE ORIGEN

expuestos en este trabajo, el análisis de la función de autocovarianza aún deja algo que desear, puesto que al tratarse de una gráfica tridimensional para cada tamaño de ventana hace del análisis algo tedioso, no obstante, puede decirse algo a su favor, cumple su cometido.

TESIS CON
FALLA DE ORIGEN

Bibliografía

BOX, George E. P. & JENKINS, Gwylim M.
Time Series Analysis. Forecasting and Control
 Holden Day, 1976

BURGOS Román, Juan de
Álgebra Lineal
 Mc Graw Hill
 Madrid, 1994

BRUCE A., Donoho, D. & HONG-YE Gao
Wavelet Analysis. IEEE Spectrum.
 1996

CHATFIELD C.
The analysis of time series: An introduction
 (Citado por Torrence & Compo)
 Chapman and Hall
 E.U.A., 1989

DEVORE, Jay L.
Probabilidad y Estadística para Ingeniería y Ciencias
 International Thompson Editores
 1998

FARGE M.
Wavelet Transforms and their applications to turbulences
 Editorial Birkhäuser
 E.U.A., 1994

FAUNDEZ, Pablo & Fuentes, Álvaro
Introducción a wavelet con aplicación en reducción de ruido de señales acústicas
 E.U.A., 2000

GILMAN, D.L.
On the power spectrum of red noise
 (Citado por Torrence & Compo)
 J. Atmos Sct.
 1963

GONZALEZ VIDEGARAY, Maricarmen
Modelos de Decisión con Procesos Estocásticos II
 UNAM, ENEP Acatlán
 México, 1995

GROSSMAN, Stanley I.
Álgebra Lineal
 Mc Graw Hill
 México, 1996

TESIS CON
 FALLA DE ORIGEN

HSU, Hwei P.
Análisis de Fourier
 Fondo educativo Interamericano, S.A.
 E.U.A., 1973

KAISER Gerald
A friendly guide to wavelets
 Editorial Birkhäuser
 E.U.A., 1994

MARON, Melvin J. & López, Robert J.
Análisis Numérico, Un enfoque práctico
 Editorial CECSA
 México, 1995

MIRAMONTES, Pedro
El color del ruido
 Facultad de Ciencias, U.N.A.M.
 México, 1999

PERCIVAL, Donald B.
Three curious properties of the sample variance and autocovariance for stationary processes with unknown mean
 University of Washington
 Seattle, E.U.A., 1999

PINEDA Caballero, Areli Rocio
Tesis: Teoría de Wavelets, una alternativa
 U.N.A.M.
 Acatlán, Edo. de México, 1998

SCHULZ, Michael & MUDELSEE, Manfred
REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series
 Computers & Geosciences
 E.U.A., 2001

STERIAN, Andrew
A unified vector space approach to teaching the Fourier transform
 Grand Valley State University
 E.U.A.

STRANG, Gilbert
The Discrete Cosine Transform
 Society for Industrial and Applied Mathematics
 E.U.A., 1999

TOLSTOV Georgi P.
Fourier Series
 Dover Publications, Inc.
 E.U.A., 1976

TESIS CON
 FALLA DE ORIGEN

TORRENCE, Christopher & COMPO, Gilbert
A practical Guide to Wavelet Analysis
Program in Atmospheric and Oceanic Sciences
University of Colorado
Bulletin of the American Meteorological Society
E.U.A., 1997

VALENS, C.
A really friendly guide to wavelets
E.U.A., 1999

WALPOLE & MYERS
Probabilidad y Estadística
Mc Graw Hill, 1992
Cuarta Edición

WEI, William
Time Series Analysis, Univariate and Multivariate Methods
Addison-Wesley Publishing Company, Inc.
E.U.A., 1994

WHITCHER, Brandon
Wavelet Scale Analysis Applied to Atmospheric Time Series
Editorial Birkhäuser
E.U.A.

Índice

Prólogo.....	1
Introducción.....	1
Capítulo I. Análisis de la tendencia.....	4
1.1 Medias Móviles.....	5
1.2 Tendencia Lineal.....	10
1.3 Tendencia Polinomial, Exponencial y Logarítmica.....	14
1.4 Modelos de tendencia estocástica.....	21
Capítulo II. Análisis de la Varianza y Autocovarianza.....	27
2.1 Función de Varianza Muestral.....	27
2.1.1 Varianza heteroscedástica.....	28
2.1.2 Transformaciones a la varianza.....	32
2.2 Función de Autocovarianza.....	34
2.2.1 Función de Autocorrelación.....	35
2.2.2 Función de Autocorrelación Parcial.....	45
Capítulo III. Análisis de la frecuencia y fluctuación estacional.....	50
3.1 Diferencias Estacionales.....	50
3.2 Análisis de Fourier.....	53
3.2.1 Transformada de Fourier Discreta.....	54
3.2.2 Espectro Teórico y Niveles de Significancia.....	59
3.2.3 Funciones ponderadoras.....	65
3.2.4 Transformada de Fourier con ventanas.....	77
3.3 Análisis Wavelet.....	82
3.3.1 Funciones Wavelet.....	82
3.3.2 Transformada Wavelet.....	88
3.3.3 Intervalo de Confianza y Cono de Influencia.....	92
Conclusiones.....	100
Bibliografía.....	103