



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

“MÉTODOS MATEMÁTICOS PARA EL
ANÁLISIS Y PRONÓSTICO DE SERIES DE
TIEMPO”

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

A C T U A R I O

PRESENTA

AMILCAR OLVERA ZAMORA

Asesor: M. en C. HARVEY SPENCER SÁNCHEZ RESTREPO

Enero de 2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

Esta tesis esta dedicada a mis padres, a quienes agradezco de todo corazón por su amor, cariño y comprensión. En todo momento los llevo conmigo.

Durante estos años, siempre soñé que un día escribiría los agradecimientos de mi tesis, lo que supondría que estaba prácticamente terminada. Ahora ese momento ha llegado y no se muy bien por donde empezar. Si bien este trabajo ha requerido de esfuerzo y mucha dedicación por parte mía y de el asesor de tesis, no hubiese sido posible su finalización sin la cooperación de todas y cada una de las personas que a continuación citaré y muchas de las cuales han sido un soporte muy fuerte en momentos de angustia y desesperación.

A mis padres Enrique Olvera y Martha Elba Zamora, por el apoyo que me brindaron, por la formación, por fomentar en mi el deseo de saber, de conocer lo novedoso y abrirme las puertas al mundo ante mi curiosidad insaciable. Creo ahora entender porque me obligaban a terminar mi tarea antes de salir a jugar, y muchas cosas más que no terminaría de mencionar.

Agradezco a mis hermanos Edgar y Adair por su compañía y el apoyo que me brindan. Se que cuento con ellos siempre.

Gracias a mi abuelita Paula Zamora por encomendarme siempre con Dios para que saliera adelante. Yo se que sus oraciones fueron escuchadas.

A los amigos por su confianza y lealtad, con quienes construí conocimiento, compartimos mañanas, tardes y noches de estudio.

A mi asesor Harvey por sus consejos, paciencia y opiniones las cuales me sirvieron para que me sintiera satisfecho con mi proyecto de investigación.

A cada uno de los maestros que participaron en mi desarrollo profesional durante mi carrera, sin su ayuda y conocimientos no estaría en donde me encuentro ahora.

Y un agradecimiento muy especial a mi padre por haberme enseñado a ser perseverante y paciente, a ponerme pasos fijos para alcanzar mis metas, a ver los problemas con la cabeza fría y a guiarme por la premisa de que toda disciplina tiene su recompensa. Gracias por darme el mejor legado de un padre a su hijo el cual es un poco de tu tiempo cada día.

Índice general

Introducción	v
1. Métodos clásicos de análisis.	1
1.1. Suavizamientos exponenciales.	6
1.1.1. Suavizamiento exponencial simple. . .	6
1.1.2. Suavizamiento exponencial doble. . .	9
1.1.3. Suavizamiento exponencial triple. . .	11
1.2. Procesos autoregresivos.	14
1.2.1. Modelo $AR(p)$	14
1.2.2. Modelo $MA(q)$	20
1.2.3. Procesos $ARMA(p,q)$	23
1.2.4. Procesos $ARIMA(p,d,q)$	27
1.3. Funciones de autocovarianza y autocorrelación.	31
1.4. Predictores óptimos para modelos $ARMA$. .	34
1.5. Función generadora de autocovarianza. . . .	35
2. Análisis de Fourier.	39
2.1. Series de Fourier complejas para funciones periódicas.	39

2.1.1.	Propiedades de los coeficientes de Fourier complejos.	40
2.1.2.	Ondas periódicas.	41
2.2.	La integral de Fourier.	43
2.2.1.	Transformada de Fourier para un único pulso.	43
2.3.	Densidad espectral.	46
2.3.1.	Espectro del modelo AR(1).	47
2.3.2.	Espectro del modelo MA(1).	48
2.3.3.	Espectro del proceso ARMA.	49
2.3.4.	Espectro de un proceso ruido blanco.	50
2.4.	Función de verosimilitud para procesos gaussianos.	52
2.4.1.	Función de verosimilitud para un proceso gaussiano AR(1).	52
2.4.2.	Función de verosimilitud para un proceso gaussiano MA(1).	54
2.4.3.	Función de verosimilitud para un proceso gaussiano ARMA(p,q).	56
2.4.4.	Expresión alternativa para la función de verosimilitud.	57
2.4.5.	Función de verosimilitud para un proceso gaussiano AR(p).	59
2.5.	Modelo VAR.	63
2.6.	Transformada de Fourier.	67
2.6.1.	Transformada Rápida de Fourier y Transformada de Fourier Discreta (TFD).	68
2.6.2.	Propiedades de la TFD.	70
2.6.3.	Transformada de Fourier discreta inversa (TFDI).	71

2.6.4.	Convolución cíclica.	74
2.6.5.	Transformada Rápida de Fourier para $N = 2^k$	75
3.	Análisis Fractal.	83
3.1.	Autosimilitud, autoafinidad y escalamiento.	83
3.2.	Dimensión fractal.	86
3.2.1.	Dimensión de Hausdorff.	87
3.2.2.	Dimensión conteo por cajas.	88
3.2.3.	Dimensión Informativa.	90
3.2.4.	Dimensión Punto clave.	90
3.3.	Exponente de Hurst.	91
3.4.	Movimiento browniano.	94
3.4.1.	Aplicación del rango reescalado sobre la caminata aleatoria.	96
3.5.	Ruido de Colores.	97
3.6.	Movimiento Browniano Fraccionario.	107
	Conclusión	115

Introducción

Toda institución, ya sea la familia, la empresa o el gobierno, tiene que hacer planes para el futuro si desea sobrevivir y mejorar sus condiciones respecto al pasado. Para lograr este fin, es necesario prever el comportamiento futuro de ciertos fenómenos con el fin de planificar, predecir o pronosticar. Esto es, construir una planificación objetiva requiere evaluar y decidir con base en escenarios que exhiban los sucesos con alguna medida de probabilidad de ocurrencia. La previsión, a su vez, se suele basar en lo que ha ocurrido en el pasado. En este sentido, se tiene un nuevo tipo de inferencia estadística acerca de alguna variable o conjunto de variables basándose en sucesos pasados. Al conjunto de valores registrados a través del tiempo a intervalos regulares se le conoce como la serie de datos temporales del fenómeno o *serie de tiempo*. A las técnicas estadísticas y matemáticas más relevantes, agrupadas de manera sistemática para analizar los datos históricos y hacer pronósticos con base en lo registrado sobre algún fenómeno de interés, se le conoce como análisis de series de tiempo.

Son innumerables las aplicaciones que se pueden citar en distintas áreas del conocimiento tales como: economía, física, geofísica, química, electricidad, demografía, marketing, telecomunicaciones, transporte, etc.

El propósito principal del estudio de las series de tiempo se centra en describir el comportamiento de la serie, investigar el mecanismo generador de la serie temporal y así buscar posibles patrones temporales que expliquen el comportamiento de la misma a través del tiempo confiando en que, a partir de ciertas hipótesis, es posible obtener cierto grado de certidumbre sobre la realización del fenómeno en el futuro.

En éste trabajo se construyen modelos desde distintos enfoques, en algunos interesa explicar la estructura de la serie, prever la evolución de una variable que ha sido observada y medida a lo largo del tiempo; en otros, lo relevante son las características estructurales y sus componentes. El desarrollo no es específico de alguna disciplina particular y puede ser aplicado a variables que provengan de la economía (índice de precios al consumo, demanda de electricidad, series de exportaciones o importaciones, etc.), la física (velocidad del viento en una central eólica, temperatura en un proceso, caudal de un río, concentración en la atmósfera de un agente contaminante), o de indicadores sociales (número de nacimientos, matrimonios, defunciones, o votos a un partido político) en donde las observaciones de interés hayan sido obtenidas en instantes sucesivos del tiempo, por ejemplo, a cada hora, durante 24 horas, mensuales, trimestrales, semestrales o bien registradas por algún

equipo en forma continua. A manera de ejemplo, se propone la serie del PIB como objeto de estudio, colocándola en el centro de este trabajo.

En el primer capítulo se presentan los métodos clásicos de análisis y los modelos más utilizados para el pronóstico de las series de tiempo, señalando las propiedades que deben cumplir cada uno de los modelos expuestos. En los modelos clásico para una serie de tiempo, supone que una serie $X(t)$ puede ser expresada como suma o producto de tres componentes: tendencia, estacionalidad y un término de error aleatorio.

El segundo capítulo presenta resultados del análisis de Fourier, estableciendo que cualquier señal periódica puede ser representada por una serie de sumas trigonométricas relacionadas armónicamente. La intención de introducir esta herramienta es que es posible afirmar que en el centro de la matemática moderna y también en el corazón de buena parte de la tecnología que nos rodea están las series de Fourier. La gran potencialidad del método que proponen les abrió el camino de múltiples aplicaciones en campos variados como las telecomunicaciones, la medicina, la prospección petrolera o aún el análisis económico. La discusión planteada entre los matemáticos respecto de la clase de funciones a las que se podían aplicar estas ideas y del sentido que tuviera la aproximación de la función por esa suma constituyó una fuente importante para el desarrollo de la teoría de funciones y del análisis funcional. Con estos conceptos aclarados y generalizados en las teorías desarrolladas

se estuvo en condiciones de plantear tales aproximaciones desde un punto de vista discreto apto para el cálculo digital de las computadoras. La resultante ha sido, que en gran parte de los casos el uso de las series de Fourier y de su generalización matemática que es la llamada transformada de Fourier ha resultado “transparente” para el usuario que incluso puede ignorar casi todo acerca de la base matemática implícita en la operación que realiza. Reconociendo la importancia de los principios, se muestran algunos resultados relevantes para la aplicación de esta técnica.

En el tercer y último capítulo, se muestra otro enfoque: el de los fractales. En esta sección se exhiben propiedades tales como la dimensión y estructura. Desde este enfoque se plantea la asignación de dimensiones fraccionarias a las series de tiempo y su relación con las leyes de potencias y las correlaciones de largo alcance. Para cumplir con el objetivo de esta tesis, se relaciona a esta matemática *no clásica* con interpretaciones temporales, en particular se describe una serie de tiempo fractal como un movimiento browniano fraccionario haciendo uso de los conceptos planteados por Mandelbrot, quién publicó en 1998 un libro titulado “Fractal and Scaling in Finance Discontinuity, Concentration, Risk” que puso en pie a Wall Street. En el mismo se explica la autosimilaridad de una serie de observaciones en un intervalo de tiempo y las interpretaciones en términos de correlaciones y dinámicas subyacentes ante invarianza de escala.

Resumen

En este trabajo se exponen algunos de los diferentes enfoques matemáticos que existen para realizar análisis y pronósticos de series de tiempo. Los desarrollos matemáticos se revisan de manera profunda poniendo énfasis en algunas proposiciones que soportan resultados y aplicaciones prácticas. En el marco del análisis clásico se estudian los suavizamientos exponenciales, los procesos ARIMA, los espectros de las series y se concluye con los vectores autoregresivos. Asimismo, se propone el análisis de Fourier como una herramienta poderosa para el estudio de series no estacionarias. Como una alternativa más general, se propone que los valores de la serie tienen por generador un sistema dinámico a la vez que se muestran algunas medidas e indicadores propios de los fractales que ayudan a caracterizar la naturaleza y las correlaciones de corto y largo alcance presentes en la serie.

Abstract

This paper describes some of the different approaches that exist for mathematical analysis and forecasting time series. Mathematical developments are reviewed in depth with emphasis on proposals that bear some results and practical applications. As part of the classical analysis, we study the exponential smoothing, ARIMA processes, the spectra of the series and conclude with autoregressive vector. Also is proposed the Fourier analysis as a powerful tool for studying non-stationary series. As a more general, it is proposed that the values of the generator set is a dynamic system at the same time are some measures and indicators from fractals that help us to characterize the nature and the correlations of short and long range in the series.

Capítulo 1

Métodos clásicos de análisis.

Una serie de tiempo es un conjunto de datos registrados de manera ordenada respecto al tiempo. Se conoce como series de tiempo continuas, cuando las observaciones son hechas continuamente en el tiempo y como discretas cuando las observaciones son tomadas en determinados momentos, por lo general igualmente espaciados. El objetivo de estudio de las series de tiempo puede ser clasificado en:

1. Descripción: cuando se presenta una serie temporal, el primer paso en el análisis es, usualmente, graficar los datos y obtener una simple medida descriptiva de las principales propiedades de la series, las cuales se describirán mas adelante. Por ejemplo, para algunas series las variaciones son evidentes y un modelo bastante simple el cual solo intenta describir la tendencia y variación estacional puede que sea adecuado para describir la variacion en la serie de tiempo. Mientras que para otras series, serán necesarias técnicas mas sofisticadas para obtener un mejor análisis.

2. Explicación: cuando las observaciones son tomadas en dos o mas variables, pudiera ser posible usar la variación en una serie de tiempo y poder así explicar la variación en otras series.
3. Predicción: dada una serie de tiempo observada, pudiera estar en la necesidad de predecir los valores futuros de la serie.
4. Control: cuando en una serie de tiempo en la cual se mide la calidad de un proceso de fabricación, el objetivo del análisis pudiera ser para controlar el proceso.

El análisis de series de tiempo desempeña un papel importante para la obtención del pronóstico. Existen dos tipos de datos en dichas series:

- Continuos, donde se tienen observaciones en cualquier instante de tiempo, ejemplos, detectores de mentiras, electrocardiogramas. Denotamos lo anterior mediante una X al tiempo t , $X(t)$.
- Discretos, donde se tiene una observación en intervalos. Denotamos esto como X_t .

El análisis de series de tiempo es y seguirá siendo utilizado para muchas aplicaciones, entre las cuales se encuentran:

1. Pronósticos económicos.
2. Previsión de ventas.
3. Análisis presupuestario.
4. Análisis para acciones de mercado.
5. Proyecciones de rendimiento.

6. Control de calidad.
7. Inventario de estudios.
8. Estudios en las utilidades.
9. Análisis de censo.

Existen muchos métodos y formas de modelos adecuados para poder realizar dichos pronósticos, incluidos los siguientes:

1. Suavizamientos exponenciales (simple, doble y triple).
2. Modelos MA y AR.
3. Box-Jenkins.

Una suposición común para el estudio de series de tiempo, es que los datos deben de ser estacionarios. Pero en algunas ocasiones las series no la cumple, por lo que una de las técnicas utilizadas para transformarla a estacionaria es la siguiente:

1. Restando los datos, esto es, dada la serie X_t se crea una nueva serie definida de la forma $Y_t = X_t - X_{t-1}$.
2. La nueva serie, ahora diferenciada, tendrá un dato menos que la serie original. Aunque es posible diferenciar los datos mas de una vez, usualmente una diferencia es suficiente.

Si los datos tienen tendencia, es posible ajustar una curva y luego modelar los residuos de dicho ajuste, dado que la finalidad del ajuste es simplemente eliminar tendencia a largo plazo.

3. Cuando una serie de tiempo no tiene varianza constante, aplicarle logaritmos a los datos pudiera estabilizar dicha varianza. Para datos negativos, se puede sumar una constante para así hacer los datos positivos antes de aplicar la transformación.

Asimismo, para poder lograr un buen análisis de series de tiempo es un buen comienzo graficar la serie, lo cual dejará detectar las componentes esenciales, por ejemplo:

- a) Detectar Outlier: se refiere a puntos de la serie que se escapan de lo normal, es decir, un outlier es una observación de la serie que corresponde a un comportamiento anormal del fenómeno (sin incidencias futuras) o a un error de medición. Si se concluye que un punto dado es outlier, lo más común es omitir o reemplazar por otro valor antes de analizar la serie.
- b) Permite detectar tendencia: la tendencia representa el comportamiento predominante de la serie. Esta puede ser definida vagamente como el cambio de la media a lo largo de un periodo.
- c) Variación estacional: la variación estacional representa un movimiento periódico de la serie de tiempo. La duración de la unidad del periodo es generalmente menor que un año lo cual puede ser un trimestre, un mes o un día, etc.
- d) Variaciones irregulares (componente aleatoria): los movimientos irregulares (el azar) que representan todos los tipos de movimientos de una serie de tiempo que no sea tendencia, variaciones estacionales y fluctuaciones cíclicas.

Existen tres modelos de series de tiempos, que generalmente se aceptan como buenas representaciones entre los componentes de los datos observados, estos son:

1. Aditivo: $X(t) = T(t) + E(t) + A(t)$
2. Multiplicativo: $X(t) = T(t) \cdot E(t) \cdot A(t)$
3. Mixto: $X(t) = T(t) \cdot E(t) + A(t)$

donde:

$X(t)$ serie observada en instante t

$T(t)$ componente de tendencia

$E(t)$ componente estacional

$A(t)$ componente aleatoria (accidental)

Un supuesto usualmente adoptado es que $A(t)$ sea una componente ruido blanco, es decir, con media cero y varianza constante. Por último cabe destacar que un modelo aditivo, es adecuado, por ejemplo, cuando $E(t)$ no depende de otras componentes, como $T(t)$, si por el contrario la estacionalidad varía con la tendencia, el modelo más adecuado es un modelo multiplicativo. (Es claro que el segundo modelo puede ser transformado en aditivo, tomando logaritmos).

1.1. Suavizamientos exponenciales.

1.1.1. Suavizamiento exponencial simple.

Este método supone que la serie de tiempo no tiene tendencia, pero la media de dicha serie cambia con el tiempo, por lo que el asignar los mismos pesos a los datos no sería lo más apropiado. Por ello, en este método de suavizamiento se le da un mayor peso a las observaciones recientes.

Se empieza el suavizamiento exponencial simple calculando una estimación inicial ℓ_0 , de la media, de la serie de tiempo al tiempo $t = 0$. Posteriormente, se calcula dicha estimación ℓ_0 mediante un promedio de los n primeros valores de la serie, donde se obtiene:

$$\ell_0 = \sum_{t=1}^n \frac{Y_t}{n}. \quad (1.1)$$

De tal forma, es posible calcular la estimación actualizada mediante el uso de la ecuación de suavizamiento:

$$\ell_T = \alpha Y_T + (1 - \alpha)\ell_{T-1}, \quad (1.2)$$

donde α es la constante de suavizamiento entre 0 y 1, de manera que se debe de encontrar un valor de α el cual nos de el menor valor para la suma de los errores cuadrados del pronóstico (SSE). Por tanto, en base a lo anterior, la corrección del error para la ecuación de suavizamiento, en el suavizamiento exponencial simple es escrita como:

$$\ell_T = \ell_{T-1} + \alpha(Y_T - \ell_{T-1}). \quad (1.3)$$

En ocasiones, cuando el suavizamiento exponencial simple ha sido usado, la tasa a la que el nivel está cambiando con el tiempo podría cambiar por

lo que sería necesario proponer una nueva α , la cual pudiera mejorar el pronóstico. Por lo que el seguimiento de la señal es quien nos ayudará a decidir cuando se este usando dicha constante inapropiadamente. Una de los primeros seguimiento de la señal fue seguimiento de la señal acumulada, que compara la suma acumulativa de los errores con la desviación absoluta media, y la cual es denotada por:

$$C(\alpha, T) = \left| \frac{Y(\alpha, T)}{MAD(\alpha, T)} \right|, \quad (1.4)$$

donde:

$$Y(\alpha, T) = Y(\alpha, T - 1) + e_T(\alpha), \quad (1.5)$$

$$MAD(\alpha, T) = \alpha |e_t(\alpha)| + (1 - \alpha)MAD(\alpha, T - 1), \quad (1.6)$$

$$MAD = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)}{n}. \quad (1.7)$$

Como se ha mencionado anteriormente, cuando la media de una serie de tiempo está cambiando lentamente en el tiempo, no es apropiado asignarle a las observaciones pesos iguales, sino darle un mayor valor a aquellas mas recientes. Lo anterior es posible realizar mediante el ponderador α y además el cual, tendrá el menor error. En este caso, para la representación y pronóstico del PIB, se tiene el valor $\alpha = 0,84$ por lo que la ecuación estimada es de la forma:

$$\ell_t = 0,84Y_T + (1 - 0,84)\ell_{T-1} \quad (1.8)$$

Como se puede apreciar en el Gráfico 1, este tipo de modelo carece de potencia y calidad predictiva debido quizá a la simplicidad asociada a la interpolación lineal que se efectúa.

Sin embargo, en algunos casos, este método es muy socorrido en áreas administrativas por la fácil interpretación del parámetro α .

Gráfica 1. Modelo de suavizamiento exponencial simple para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



1.1.2. Suavizamiento exponencial doble.

Debido a que los métodos de ponderación exponencial simple no se desempeñan bien cuando los datos históricos presentan un patrón de tendencia, es apropiado utilizar el suavizamiento exponencial doble debido a que se toma en cuenta tanto la tendencia lineal como la variación de la media respecto al tiempo.

De esta forma, es posible definir esta clase de suavizamiento mediante las siguientes ecuaciones:

$$\ell_T = \alpha y_T + (1 - \alpha)(\ell_{T-1} + b_{T-1}) \quad 0 \leq \alpha \leq 1 \quad (1.9)$$

$$b_T = \gamma(\ell_T - \ell_{T-1}) + (1 - \gamma)b_{T-1} \quad 0 \leq \gamma \leq 1 \quad (1.10)$$

donde α y γ son constantes de suavizamiento, el término ℓ_T asigna el valor básico o de largo plazo de la serie de tiempo y b_T modela la tendencia esperada.

Como se citó anteriormente, cuando la serie tiene una media y tendencia, las cuales evolucionan con el tiempo, es apropiado definir otra constante que nos suavice este último factor. Dicho lo anterior, se dice que al método de suavizamiento exponencial doble se le asignan los pesos a las observaciones tanto recientes como pasadas, al igual que en el modelo simple, pero tomando en cuenta el factor tendencia. Continuando con el análisis de la serie del PIB, se tiene que las constantes α y γ toman los valores de 0.3512 y 0.3623 respectivamente, quedando así las ecuaciones para su pronóstico:

$$\ell_t = 0,3512Y_T + (1 - 0,3512)(\ell_{T-1} + b_{T-1}) \quad (1.11)$$

$$b_T = 0,3623(\ell_T - \ell_{T-1}) + (1 - 0,3623)b_{T-1} \quad (1.12)$$

Como se puede observar en el Gráfico 2, tanto la estimación del promedio de los errores como la estimación de la tendencia son suavizadas y por esta razón, este tipo de modelo es más utilizado en el análisis de series financieras. Sin embargo, no podemos afirmar que este modelo es aplicable para toda serie, debido a que no se toma en cuenta el factor estacional, el cual pudiera tener una gran influencia dentro del estudio de la serie.

Gráfica 2. Modelo de suavizamiento exponencial doble para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



1.1.3. Suavizamiento exponencial triple.

Este modelo es una forma de suavizamiento exponencial que toma en cuenta la estacionalidad y tendencia de los datos históricos. El modelo es conocido como Holt-Winters, ya que suaviza un factor asociado con tres de los elementos del patrón: la aleatoriedad, la tendencia y la estacionalidad. La predicción se realiza mediante la siguiente ecuación:

$$F_{T+m} = (\ell_T + mb_T)I_{T-L+m}, \quad (1.13)$$

donde:

$$\ell_T = \alpha \frac{y_T}{I_{T-L}} + (1 - \alpha)(\ell_{T-1} + b_{T-1}) \quad (1.14)$$

$$b_T = \gamma(\ell_T - \ell_{T-1}) + (1 - \gamma)b_{T-1} \quad (1.15)$$

$$I_T = \beta \frac{y_T}{\ell_T} + (1 - \beta)(I_{T-L}) \quad (1.16)$$

y:= la observación

ℓ := la observación suavizada

b:= valor suavizado de la tendencia

I:= valor suavizado del factor estacional

F:= pronostico para m periodos adelante

T:= índice que denota un periodo de tiempo

L:= duración de la estacionalidad

α, β, γ son constantes que debe calcularse de tal manera que el MSE del error sea minimizado.

Como se explicó en este capítulo, cuando la serie tiene una media, tendencia, las cuales evolucionan con el tiempo, y además presenta estacionalidad, podemos decir que es apropiado utilizar el modelo tripe, ya que nos permite modelar cada una de estas de forma adecuada. Dicho lo anterior, se dice que

en el método de suavizamiento exponencial triple se le asignan los pesos a las observaciones tanto recientes como pasadas, en donde se toma en cuenta la tendencia, al igual que en el modelo doble, pero no dejando atrás el factor estacional. Siguiendo con el pronóstico del PIB y realizando los análisis correspondientes, se concluyó que las constantes toman los valores de $\alpha = 0,3$, $\gamma = 0,32$ y $\beta = 0,04$, quedando lo siguiente:

$$\ell_T = 0,3 \frac{y_T}{I_{T-L}} + (1 - 0,3)(\ell_{T-1} + b_{T-1}) \quad (1.17)$$

$$b_T = 0,32(\ell_T - \ell_{T-1}) + (1 - 0,32)b_{T-1} \quad (1.18)$$

$$I_T = 0,04 \frac{y_T}{\ell_T} + (1 - 0,04)(I_{T-L}) \quad (1.19)$$

Gráfica 3. Modelo de suavizamiento exponencial triple para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Como se observa en el Gráfico 3, este tipo de modelo es utilizado cuando la serie presenta tendencia, variación de su media respecto del tiempo y estacionalidad, aunque esta última no se aprecie a simple vista. De tal forma que, posiblemente, estos modelos sean utilizados con mayor frecuencia debido a que gran parte de las series presentan estos tres factores.

Sin embargo, no es posible afirmar que este modelo es aplicable para toda serie, ya que dependerá de las características que tenga cada una de ellas para así poder encontrar el mejor modelo.

1.2. Procesos autoregresivos.

1.2.1. Modelo AR(p)

Entendemos por ruido blanco, denotado por ϵ_t , cuando las correlaciones entre los valores en diferentes puntos de tiempo cumplen que:

$$\begin{aligned} E(\epsilon_t) &= \mu, & Var(\epsilon_t) &= \sigma^2, \forall t, \\ Cov(\epsilon_t, \epsilon_t) &= 0, \forall t \neq 0. \end{aligned}$$

Dicho lo anterior, un proceso autoregresivo de orden p , AR(p), es escrito de la forma:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t, \quad t = 1, \dots, T \quad (1.20)$$

donde ϵ_t es considerado ruido blanco.¹

Tomando en cuenta la ecuación (1.20), se define un proceso AR(1) es tal que:

$$y_t = \phi y_{t-1} + \epsilon_t, \quad t = 1, \dots, T. \quad (1.21)$$

Aunque la serie es observada primeramente al tiempo $t=1$, el proceso es considerado como si hubiese empezado en el pasado al tiempo t . Sustituyendo repetidamente para valores retrasados de y_t se obtiene

$$y_t = \sum_{j=0}^{J-1} \phi^j \epsilon_{t-j} + \phi^J y_{t-J}, \quad (1.22)$$

tomando esperanzas y tratando a y_{t-J} como un número fijo se tiene a (1.22) como:

$$E(y_t) = E\left(\sum_{j=0}^{J-1} \phi^j \epsilon_{t-j}\right) + E(\phi^J y_{t-J}) = \phi^J y_{t-J}. \quad (1.23)$$

¹Es decir, su media y varianza son constantes y además no se encuentran correlacionados.

Si $|\phi| \geq 1$ en la ecuación (1.23), el valor de la media del proceso depende del valor inicial, y_{t-J} . La ecuación (1.22), por tanto, contiene un componente determinista y un conocimiento de y_{t-J} permitiendo que sea hecha una predicción no trivial para los valores futuros de la serie. Por el otro lado, si $|\phi| \leq 1$, el componente determinístico es insignificante si J es largo.

Como $J \rightarrow \infty$, esto con eficacia desaparece y también si el proceso es considerado como si hubiese comenzado en algún punto del pasado, podemos escribir a (1.22) de la siguiente forma:

$$y_t = \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}, \quad t = 1, \dots, T. \quad (1.24)$$

De (1.24), se observa que la esperanza de y_t es cero, mientras que su varianza es de la forma:

$$\begin{aligned} \gamma(0) &= E(y_t^2) = E \left(\sum_{j=0}^{\infty} \phi^j \epsilon_{t-j} \right)^2 = \sum_{j=0}^{\infty} \phi^{2j} E(\epsilon_{t-j}^2) \\ &= \sigma^2 \sum_{j=0}^{\infty} \phi^{2j} = \sigma^2 / (1 - \phi^2). \end{aligned} \quad (1.25)$$

Realizando ahora la predicción para la serie del PIB, pero mediante los modelos autoregresivos, se puede observar rápidamente que el modelo que más se ajusta, basados en la tabla 1, es de la forma AR(1).

Tabla 1. Estimaciones AR(1) para el PIB 80.1-07.3.

<i>Parámetro</i>	<i>Error estándar</i>	<i>p – value</i>
AR(1)	0.0934682	0.000000
AR(2)	0.1012100	0.279836
AR(3)	0.0651565	0.204369
AR(4)	0.0678670	0.000000
AR(5)	0.1121400	0.000000
AR(6)	0.1000050	0.001248

Se llega a dicha conclusión, debido a que los términos con p-value menores que 0.05 son estadísticamente significantivos, es decir, diferentes de cero con el 95 % de confianza. En base a lo anterior y tomando en cuenta el error estándar, se puede decir que el modelo con menor p-value y error estándar es el AR(1), como se aprecia en la Gráfica 4. Por otro lado, no debemos olvidar que uno de mayores supuestos dentro de esta teoría es que los errores no se encuentren autocorrelacionados.

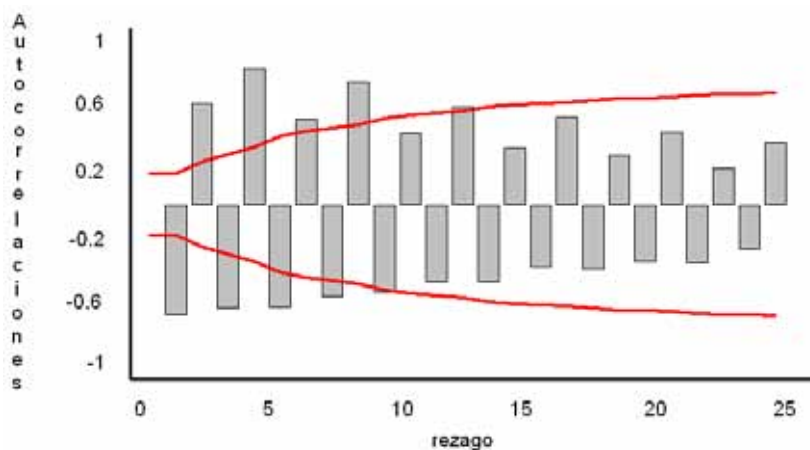
De esta manera, para asegurar que esto no ocurra, se presenta en el Gráfico 5 la autocorrelación entre los residuos en donde algunos de estos se encuentran por arriba de la franja roja, indicando el no cumplimiento de la propiedad anterior.

Gráfica 4. Modelo AR(1) para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Gráfica 5. Autocorrelación de los residuos.



Como el modelo AR(1) no es del todo aceptado, es decir, no cumple con uno de los supuestos, se debe realizar nuevamente el análisis anterior para los demás modelos. En este caso, tomando en cuenta lo dicho, se concluye que el modelo de la forma AR(6) cumple con que los errores no se encuentran autocorrelacionados, como lo muestra el Gráfico 6.

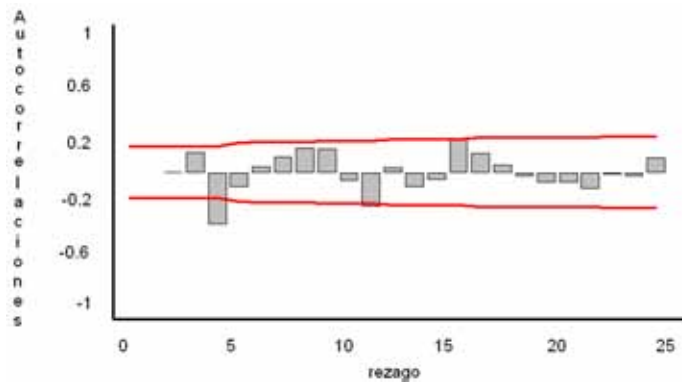
Gráfica 6. Modelo AR(6) para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Ahora como se esperaba, los residuos se encuentran por debajo de la franja, lo que indica la no autocorrelación.

Gráfica 7. Autocorrelación de los residuos.



Finalmente, es posible concluir que el modelo que mejor se ajusta, en base a lo obtenido, para realizar la predicción de la serie del PIB es de la forma AR(6). Además, como se había mencionado anteriormente, los modelos autoregresivos (AR) pueden describirse, de una forma general, como aquellos en los que una variable se explica, al menos en parte, en función de sus valores pasados.

Por esta razón, este tipo de modelos es muy utilizado en las áreas financieras debido a la simplicidad y al análisis de los cálculos. Sin embargo, los modelos de vectores autorregresivos (VAR)² pueden plantearse como una generalización de los modelos AR al caso de un vector de n variables y_t en donde estos modelos han cobrado una gran importancia en las últimas décadas en el campo de la econometría y la economía.

²Los modelos VAR son modelos que relacionan entre sí n variables y en los que el valor que toma cada una de ellas en un periodo de tiempo se relaciona con los valores que toma esa misma variable y todas las demás variables en periodos anteriores.

1.2.2. Modelo MA(q).

Sea ϵ_t ruido blanco y consideremos el proceso

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}, \quad t = 1, \dots, T. \quad (1.26)$$

el cual es llamado, proceso de medias móviles de orden q , denotado como MA(q).

Ahora, considere un proceso MA(∞), es decir, $q \rightarrow \infty$

$$y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad (1.27)$$

entonces, la media y las autocovarianzas de un proceso MA(∞) con coeficientes absolutamente sumables pueden ser calculados a partir de un proceso MA(q) de la siguiente forma:

$$\begin{aligned} E(y_t) &= \lim_{T \rightarrow \infty} E(\psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots + \psi_T \epsilon_{t-T}) \\ &= 0 \end{aligned} \quad (1.28)$$

$$\begin{aligned} \gamma(0) &= E(y_t^2) \\ &= \lim_{T \rightarrow \infty} E(\psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots + \psi_T \epsilon_{t-T})^2 \\ &= \lim_{T \rightarrow \infty} (\psi_0^2 + \psi_1^2 + \psi_2^2 + \cdots + \psi_T^2) \sigma^2 \end{aligned} \quad (1.29)$$

$$\begin{aligned} \gamma(\tau) &= E(y_t y_{t-\tau}) \\ &= \sigma^2 (\psi_j \psi_0 + \psi_{j+1} \psi_1 + \psi_{j+2} \psi_2 + \psi_{j+3} \psi_3 + \cdots) \end{aligned} \quad (1.30)$$

Por otro lado, el proceso MA(1), puede ser expresado en términos de los rezagos de y_t , sustituyendo repetidamente para los valores de los rezagos de ϵ_t . Esto muestra

$$y_t = \theta y_{t-1} - \theta^2 y_{t-2} + \cdots - (-\theta)^J y_{t-J} + \epsilon_t - (-\theta)^{J+1} \epsilon_{t-J-1} \quad (1.31)$$

Si y_t no depende de un choque con el sistema que se plantea en algún momento en el pasado remoto, θ debe de ser menor a uno en valor absoluto. Además, si a J se le permite tomar el valor de infinito, el último término en (1.31) desaparece y y_t puede ser escrito como un proceso autoregresivo infinito, AR(∞), con pesos cada vez menores, es decir:

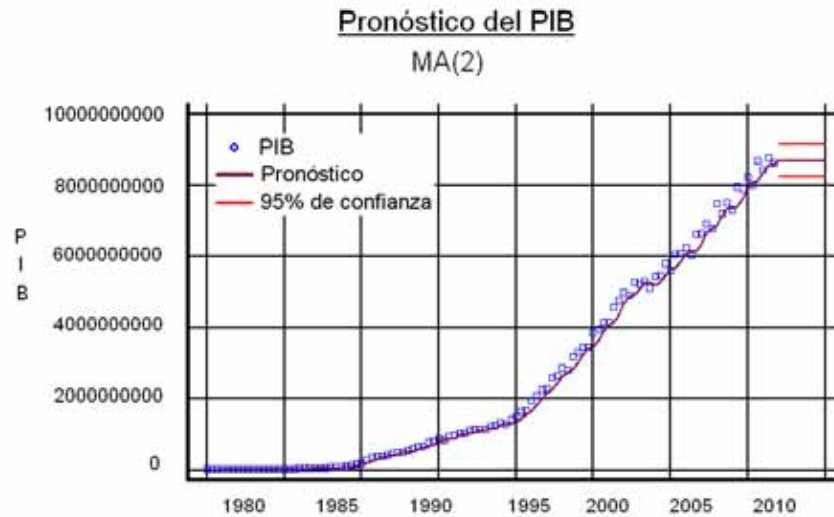
$$y_t = - \sum_{j=1}^{\infty} (-\theta)^j y_t + \epsilon_t. \quad (1.32)$$

Anteriormente se mostró una alternativa de modelización para tratar de explicar el comportamiento de una variable, no en función de los valores que tomó en el pasado (modelos AR) sino a través de los errores al estimar el valor de la variable en los periodos anteriores. Esto da lugar a lo que se conoce como medias móviles o modelos MA, por sus siglas en inglés.

Finalmente, haciendo uso de los modelos MA, se presenta aquel cuyo comportamiento es similar en relación con la serie del PIB, el cual es de la forma MA(2).

Gráfica 8. Modelo MA(2) para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Por lo tanto, este modelo supone que el pronóstico está dado mediante la suma de los primeros dos errores. Sin embargo, debido a que la media móvil solo tiene memoria de un periodo, este tipo de modelos son utilizados en las decisiones a corto plazo y además es tomando en cuenta por su fácil procedimiento y sencilla interpretación que arrojan los resultados.

1.2.3. Procesos ARMA(p,q).

Un proceso autoregresivo de medias móviles de orden (p,q) se escribe como:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (1.33)$$

Dicho proceso ARMA, puede escribirse mas brevemente definiendo polinomios asociados en el operador de rezagos de la forma:

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p \quad (1.34)$$

y

$$\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q. \quad (1.35)$$

De este modo, el modelo (1.33) se puede reescribir de la siguiente manera:

$$\phi(L)y_t = \theta(L)\epsilon_t. \quad (1.36)$$

Citando un ejemplo, si una media μ no cero, es introducida en un modelo estacionario, esta se convierte en:

$$y_t = \mu + \phi^{-1}(L)\theta(L)\epsilon_t. \quad (1.37)$$

Tomando en cuenta lo anterior, podemos definir un proceso ARMA (1,1) el cual puede ser expresado de la forma:

$$y_t = \phi y_{t-1} + \epsilon_t + \theta \epsilon_{t-1}, \quad t = 1, \dots, T. \quad (1.38)$$

Otra alternativa de escribir esto, es poner a (1.38) como:

$$(1 - \phi L)y_t = (1 + \theta L)\epsilon_t \quad (1.39)$$

o

$$y_t = \frac{\epsilon_t}{(1 - \phi L)} + \frac{\theta(L)\epsilon_t}{(1 - \phi L)}. \quad (1.40)$$

Además, si consideramos que L satisface la condición de que $|L| \leq 1$, el término puede ser observado como la suma de una progresión geométrica infinita de la forma $1, \phi L, (\phi L)^2, \dots$. Por otro lado, cuando $|\phi| \leq 1$, y (1.40) puede ser reescrita de la siguiente manera:

$$\begin{aligned}
 y_t &= \sum_{j=0}^{\infty} (\phi L)^j \epsilon_t + \theta \sum_{j=0}^{\infty} (\phi L)^j \epsilon_{t-1} \\
 &= \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j} + \theta \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j-1} \\
 &= \epsilon_t + \sum_{j=1}^{\infty} (\theta \phi^{j-1} + \phi^j) \epsilon_{t-j}
 \end{aligned} \tag{1.41}$$

Siguiendo con la serie del PIB y para tener una clara visión acerca de los procesos ARMA, se realizó el pronóstico utilizando diferentes coeficientes p y q . Por tanto, la variable queda explicada en función de los valores tomados en periodos anteriores, y los errores cometidos en la estimación.

Tabla 2. Estimaciones ARMA(6,6) para el PIB 80.1-07.3.

<i>Parámetro</i>	<i>Error estándar</i>	<i>p - value</i>
AR(1)	1.7198300	0.619984
AR(2)	1.5347100	0.953396
AR(3)	0.1188760	0.842176
AR(4)	0.0397528	0.000000
AR(5)	1.8526100	0.585252
AR(6)	1.8136200	0.987701
MA(1)	1.7153800	0.807782
MA(2)	0.6478690	0.666706
MA(3)	0.5036120	0.746503
MA(4)	0.2428880	0.016003
MA(5)	1.0989700	0.935153
MA(6)	0.2072390	0.886551

Continuando con la predicción del PIB, al ser analizada la autocorrelación y la correlación parcial de la serie a estimar, es posible decir que un buen modelo pudiera ser de la forma ARMA(6,6) el cual se observa en la tabla 2.

Pero si se observa, el p-value, tanto del modelo AR(6) como del modelo MA(6) son mayores que 0.05, lo cual indica que son no significativos.

Lo anterior, sugiere que se debe reducir el orden de cada modelo hasta que sean significativos. Realizando el mismo procedimiento, se obtiene que el proceso que cumple con las propiedades, es de la forma ARMA(2,1), como se observa en la tabla 3.

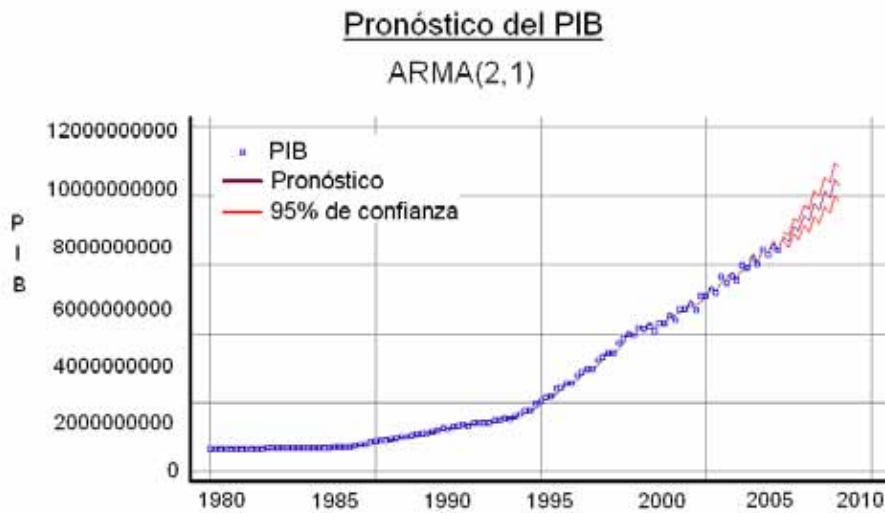
Tabla 3. Estimaciones ARMA(2,1) para el PIB 80.1-07.3.

<i>Parámetro</i>	<i>Error estándar</i>	<i>P – value</i>
AR(1)	0.0112356	0.438664
AR(2)	0.0118190	0.000000
MA(1)	0.0758055	0.000000

Como se podrá dar cuenta, los p-value de los modelos, tanto para AR(2) y MA(1), son iguales a cero, es decir, los coeficientes son significativos. Veamos gráficamente el proceso anterior:

Gráfica 9. Proceso ARMA(2,1) para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>



Estos modelos son más utilizados dentro de la práctica profesional debido a que están compuestos tanto por la parte AR y MA, el cual le da un mayor peso y confianza al realizar los pronósticos. Sin embargo, cabe la posibilidad de que la serie sea no estacionaria, lo cual pudiera afectar la predicción mediante el uso de los procesos ARMA.

1.2.4. Procesos ARIMA(p, d, q).

Si d es el orden de diferencias que se requieren para producir una serie estacionaria e invertible, el proceso original se dice ser integrado de orden d y es abreviado como $y_t \sim I(d)$. De manera que el modelo

$$\phi(L)\Delta^d y_t = \theta(L)\epsilon_t, \quad (1.42)$$

es llamado un proceso autoregresivo de medias móviles integrado de orden (p, d, q) y se denota como ARIMA (p, d, q) .

Una característica básica de las predicciones de un modelo estacionario es que tienden hacia la media de la series conforme el tiempo, ι , crece. Si ι es grande, la estructura del modelo es irrelevante. Este ya no es el caso con procesos integrados, donde las funciones pronóstico contienen un componente determinístico el cual depende de el grado de diferenciación. El término $\phi(L)\Delta^d y_t$ se expande para obtener un polinomio AR de orden $p + d$,

$$\phi(L)\Delta^d y_t = \varphi(L) = 1 - \varphi_1 L - \dots - \varphi_{p+d} L^{p+d} \quad (1.43)$$

y las predicciones se realizan a partir de la ecuación recursiva

$$\begin{aligned} \tilde{y}_{T+\iota|T} &= \varphi_1 \tilde{y}_{T+\iota-1|T} + \dots + \varphi_{p+d} \tilde{y}_{T+\iota-p-d|T} + \tilde{\epsilon}_{T+\iota|T} \\ &+ \dots + \theta_q \tilde{\epsilon}_{T+\iota-q|T}, \quad \iota = 1, 2, \dots \end{aligned} \quad (1.44)$$

donde $\tilde{y}_{T+\iota|T}$ y $\tilde{\epsilon}_{T+\iota}$ son definidas como en (1.20).

Como fue mencionado al final de los modelos ARMA, cuando la serie es no estacionaria, lo cual puede ser verificado mediante el autocorrelograma, pudiera causar problemas al momento de realizar las predicciones correspondientes. Es por ello, que se utilizan los modelos ARIMA, los cuales nos

permitirán trabajar con la estacionariedad.

Como se puede observar, en las gráficas anteriores, la serie del PIB es no estacionaria, por lo que es necesario desestacionarla mediante las diferencias correspondientes. Al realizar las primeras diferencias, la autocorrelación entre observaciones disminuye lentamente lo que nos indica la no estacionariedad de las series. De esta forma, tomando segundas diferencias, se obtiene la siguiente tabla:

Tabla 4. Autocorrelograma de segundas diferencias para el PIB 80.1-07.3.

Autocorrelación	Autocorrelación parcial	AC	PAC	Q-Stat	Prob	
		1	-0.002	-0.002	0.0005	0.982
		2	-0.057	-0.057	0.3648	0.833
		3	0.01	0.01	0.3751	0.946
****	****	4	0.474	0.472	26.27	0.000
		5	-0.005	-0.007	26.272	0.000
*		6	-0.067	-0.034	26.798	0.000
		7	0.007	0.02	26.804	0.000
	**	8	0.009	-0.202	26.813	0.001
		9	-0.002	0.000	26.814	0.002
*	*	10	-0.176	-0.180	30.616	0.001
		11	0.017	0.004	30.661	0.001
*		12	-0.086	0.037	31.568	0.002
		13	-0.003	-0.014	31.569	0.003
***	**	14	-0.355	-0.269	47.603	0.000
		15	0.003	-0.007	47.604	0.000
		16	0.005	0.012	47.608	0.000
		17	0.002	0.003	47.609	0.000
**	*	18	-0.223	0.082	54.243	0.000
		19	0.002	0.000	54.244	0.000
*		20	0.099	0.049	55.577	0.000
		21	0.009	0.019	55.587	0.000
*	*	22	-0.076	-0.114	56.386	0.000
		23	-0.002	-0.003	56.386	0.000
*		24	0.112	0.033	58.176	0.000
		25	0.006	0.001	58.18	0.000
		26	0.023	0.056	58.259	0.000
		27	-0.001	-0.003	58.259	0.000
	*	28	0.050	-0.105	59.759	0.001
		29	0.003	0.004	58.76	0.001
		30	0.013	-0.038	58.786	0.001
		31	-0.005	0.009	58.79	0.002
		32	0.001	0.002	58.79	0.003
		33	-0.001	0.002	58.79	0.004
		34	-0.019	0.042	58.85	0.005
		35	-0.012	0.002	58.874	0.007
*	*	36	0.066	0.077	59.591	0.008

En la tabla 4, se ve que la autocorrelación disminuye rápidamente, indicando estacionariedad, como se había mencionado anteriormente. Finalmente, si se toma en cuenta la autocorrelación y la autocorrelación parcial, definimos que un posible modelo es de la forma ARIMA(2,2,2), debido a que:

Tabla 5. Proceso ARIMA para el PIB 80.1-07.3.

Variable	Coefficiente	Std. Error	t-Statistic	Prob.
AR(2)	-0.9255	0.090845	-10.18769	0
MA(2)	0.811008	0.139022	5.833684	0
R-squared	0.080071	Mean dependent var		-0.545652
Adjusted R-squared	0.07131	S.D. dependent var		9.788166
S.E. of regression	9.432715	Akaike info criterion		7.34476
Sum squared resid	9342.492	Schwarz criterion		7.394719
Log likelihood	-390.9446	Durbin-Watson stat		1.998998

Analizando la tabla 5, se observa que las probabilidades de las variables AR y MA son cero, indicando el rechazo de la hipótesis nula.³ Además, el criterio de Akaike⁴ es menor con respecto a los demás modelos. Sin embargo, el estadístico Durbin-Watson es menor a 2, lo que nos indica que existe evidencia de autocorrelación de primer orden entre los errores.

³los coeficientes de las variables son cero.

⁴El criterio de Akaike es la suma de los errores al cuadrado.

A continuación, se muestra el gráfico correspondiente al pronóstico mediante el proceso ARIMA.

Gráfica 10. Proceso ARIMA(2,2,2) para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



1.3. Funciones de autocovarianza y autocorrelación.

Como hemos visto a lo largo del capítulo, cuando $|\phi| < 1$ el modelo AR(1), (1.21), tiene media cero y varianza de la forma $\sigma^2/(1 - \phi^2)$. La autocovarianza en el rezago τ puede ser obtenida expresando y_t como una combinación lineal de $y_{t-\tau}$ y $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-\tau+1}$. Esto se puede lograr haciendo $J = \tau$ en (1.22) y de esta manera obtener que:

$$\gamma(\tau) = E(y_t y_{t-\tau}) = E \left[\left(\phi^\tau y_{t-\tau} + \sum_{j=0}^{\tau-1} \phi^j \epsilon_{t-j} \right) y_{t-\tau} \right]. \quad (1.45)$$

Puesto que $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-\tau+1}$ son no correlacionados con $y_{t-\tau}$ la ecuación (1.45) se reduce a:

$$\gamma(\tau) = \phi^\tau E(y_{t-\tau}^2) = \phi^\tau \gamma(0), \quad \tau = 1, 2, \dots \quad (1.46)$$

en donde las autocovarianzas dependen solamente de τ , confirmando que el proceso es estacionario.

Entonces, para un proceso estacionario se tiene que:

$$\begin{aligned} E(y_{t-1} y_{t-\tau}) &= E(y_t y_{t-\tau+1}) \\ &= \gamma(\tau - 1) \end{aligned} \quad (1.47)$$

y, si $\tau > 0$, obtenemos que:

$$\gamma(\tau) = \phi(\tau - 1), \quad \tau = 1, 2, \dots \quad (1.48)$$

La expresión para la varianza de el proceso puede ser obtenida por un método similar notando que $E(\epsilon_t y_t) = \sigma^2$, así que la autocorrelación toma la forma:

$$\rho(\tau) = \phi^\tau, \quad \tau = 1, 2, \dots \quad (1.49)$$

Ahora, tomando en cuenta al operador $\psi(L)$ para un proceso $AR(p)$, se tiene que:

$$y_t = \psi(L)\epsilon_t, \quad (1.50)$$

donde

$$\psi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)^{-1}. \quad (1.51)$$

Asumiendo que la condición de estacionariedad se satisface, una forma de encontrar la media para un proceso autoregresivo es tomando la esperanza en la ecuación (1.20) como se muestra a continuación:

$$\mu = \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu. \quad (1.52)$$

Siguiendo con los modelos $AR(p)$, observamos que para poder obtener las autocovarianzas es necesario multiplicar por $y_{t-\tau}$ la ecuación (1.20) y de esta manera se tiene que:

$$\gamma(\tau) = \begin{cases} \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p} & \text{si } \tau = 1, 2, \dots \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 & \text{si } \tau = 0 \end{cases} \quad (1.53)$$

$$\rho(\tau) = \phi_1 \rho(\tau-1) + \dots + \phi_p \rho(\tau-p), \quad \tau = 1, 2, \dots \quad (1.54)$$

y además, cuando $\tau = 0$, obtenemos:

$$\gamma(0) = \sigma^2 / [1 - \rho(1)\phi_1 - \dots - \rho(p)\phi_p]. \quad (1.55)$$

Regresando ahora con los modelos $MA(q)$, definidos anteriormente, se observa que para obtener las funciones de autocovarianza y autocorrelación es necesario tomar esperanzas en dicha ecuación. En donde, inmediatamente se muestra que y_t tiene media cero, y su autocovarianza es de la forma:

$$\begin{aligned} \gamma(0) &= E(y_t^2) \\ &= E[(\epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q})^2] \\ &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2. \end{aligned} \quad (1.56)$$

1.3. FUNCIONES DE AUTOCOVARIANZA Y AUTOCORRELACIÓN.33

De igual forma para $\tau = 1, 2, \dots, q$, se tiene que:

$$\begin{aligned}
 \gamma(\tau) &= E[(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots + \theta_q\epsilon_{t-q}) \\
 &\quad X(\epsilon_{t-\tau} + \theta_1\epsilon_{t-\tau-1} + \theta_2\epsilon_{t-\tau-2} + \dots + \theta_q\epsilon_{t-\tau-q})] \\
 &= E[\theta_\tau\epsilon_{t-\tau}^2 + \theta_{\tau+1}\theta_1\epsilon_{t-\tau-1}^2 + \theta_{\tau+2}\theta_2\epsilon_{t-\tau-2}^2 + \dots + \theta_q\theta_{q-\tau}\epsilon_{t-q}^2].
 \end{aligned}
 \tag{1.57}$$

Los términos ϵ 's que tienen diferente subíndice han sido tirados ya que su producto tiene esperanza cero, $\theta_0 = 1$ y además las autocovarianzas están dadas por:

$$\gamma(\tau) = \begin{cases} (\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_{q-\tau}\theta_q)\sigma^2 & \tau = 1, \dots, q \\ 0 & \tau > q \end{cases}
 \tag{1.58}$$

1.4. Predictores óptimos para modelos ARMA.

Ahora, considerese el cómo construir el mínimo error cuadrático medio, por sus siglas en inglés (MMSE), para una futura observación de un proceso ARMA, dadas las observaciones e incluyendo el tiempo T . Tomando en cuenta que para los procesos mixtos y MA, todos los $\epsilon_T, \epsilon_{T-1}, \epsilon_{T-2}, \dots$, pasados y presentes son conocidos. De esta manera, la ecuación de un modelo ARMA (p,q) al tiempo $T + \iota$ es:

$$y_{T+\iota} = \phi_1 y_{T+\iota-1} + \dots + \phi_p y_{T+\iota-p} + \epsilon_{T+\iota} + \dots + \theta_q \epsilon_{T+\iota-q}. \quad (1.59)$$

El MMSE de una observación futura es su esperanza condicionada a la información en el momento T ,

$$\tilde{y}_{T+\iota|T} = E(y_{T+\iota} | Y_T = y_T, y_{T-1}, \dots) = E_T(y_{T+\iota}). \quad (1.60)$$

Tomando ahora esperanzas condicionales en (1.58), los valores futuros de ϵ_T son un conjunto de ceros, ya que se mantienen independientes, y no pueden ser predecidos. Lo anterior produce que:

$$\begin{aligned} \tilde{y}_{T+\iota|T} &= \phi_1 \tilde{y}_{T+\iota-1|T} + \dots + \phi_p \tilde{y}_{T+\iota-p|T} \\ &\quad + \tilde{\epsilon}_{T+\iota|T} + \theta_q \tilde{\epsilon}_{T+\iota-q|T}, \quad \iota = 1, 2, \dots \end{aligned} \quad (1.61)$$

donde $\tilde{y}_{T+\iota|T} = y_{T+\iota}$ para $\iota \leq 0$, y

$$\tilde{\epsilon}_{T+\iota|T} = \begin{cases} 0 & \text{for } \iota > 0 \\ \epsilon_{T+\iota} & \text{for } \iota \leq 0 \end{cases} \quad (1.62)$$

1.5. Función generadora de autocovarianza.

Para cada proceso estacionario, se calcula la secuencia de autocovarianzas $\{\gamma_\tau\}_{\tau=-\infty}^{\infty}$. Si la secuencia es absolutamente sumable, una manera de resumir las autocovarianzas es a través de una función llamada: Función Generadora de Autocovarianza (ACGF) definida como:

$$g_Y(z) = \sum_{\tau=-\infty}^{\infty} \gamma_\tau Z^\tau. \quad (1.63)$$

El argumento, Z , de la ecuación (1.63) se considera un escalar complejo de la forma:

$$z = \cos(w) - i\sin(w) = e^{-iw}, \quad (1.64)$$

donde w es el ángulo en radianes que z hace con el eje real. Si la ACGF es evaluada en $z = e^{-iw}$ y dividida por 2π , el resultado de la función de w es,

$$S_Y(w) = \frac{1}{2\pi} g_Y(e^{-iw}) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\tau w}, \quad (1.65)$$

el cual es llamado el espectro poblacional de Y .

Cabe mencionar, que si dos procesos diferentes comparten la misma función generadora de autocovarianza, significa que muestran una secuencia de autocovarianzas idéntica.

Tomese como ejemplo un proceso MA(1) para calcular su ACGF, obteniendo de esta manera lo siguiente:

$$\begin{aligned} g_Y(z) &= [\theta\sigma^2]z^{-1} + [(1 + \theta^2)\sigma^2]z^0 + [\theta\sigma^2]z^1 \\ &= \sigma^2[\theta z^{-1} + (1 + \theta^2) + \theta z] \\ &= \sigma^2(1 + \theta z)(1 + \theta z^{-1}). \end{aligned} \quad (1.66)$$

De esta última expresión se dice que para una MA(q)

$$y_t = (1 + \theta_1 L + \dots + \theta_q L^q) \epsilon_t \quad (1.67)$$

la ACGF puede ser calculada como

$$\begin{aligned} g_Y(z) &= \sigma^2(1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q) \\ &\quad X(1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_q z^{-q}) \\ &= \sigma^2[(\theta_q)z^q + (\theta_{q-1} + \theta_q \theta_1)z^{(q-1)} + (\theta_{q-2} + \theta_{q-1} \theta_1 + \theta_q \theta_2)z^{(q-2)} \\ &\quad + \dots + (\theta_1 + \theta_2 \theta_1 + \theta_3 \theta_2 + \dots + \theta_q \theta_{q-1})z^1 \\ &\quad + (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)z^0 \\ &\quad + (\theta_1 + \theta_2 \theta_1 + \theta_3 \theta_2 + \dots + \theta_q \theta_{q-1})z^{-1} + \dots + (\theta_q)z^{-q}] \end{aligned} \quad (1.68)$$

Ahora extendiendo este método para encontrar la ACGF para un modelo $MA(\infty)$, tomando en cuenta que

$$y_t = \psi(L) \epsilon_t \quad (1.69)$$

con

$$\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots \quad (1.70)$$

se tiene:

$$g_Y(z) = |\psi(z)|^2 = \psi(z)\psi(z^{-1})\sigma^2 \quad (1.71)$$

Para mostrar lo anterior, primeramente, debemos dejar que q tienda a infinito en su función de autocovarianzas y cambiando la notación se obtiene:

$$\gamma(\tau) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+\tau}. \quad (1.72)$$

Ahora, sustituyendo esta última expresión en (1.63) se tiene:

$$g_Y(z) = \sigma^2 \sum_{\tau=-\infty}^{\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+\tau} z^\tau = \sigma^2 \sum_{j=0}^{\infty} \sum_{\tau=-j}^{\infty} \psi_j \psi_{j+\tau} z^\tau, \quad (1.73)$$

de manera que $\psi_j = 0$, para $j < 0$. Escribiendo $j + \tau = h$, a fin de que $\tau = h - j$,

$$g_Y(z) = \sigma^2 \sum_{j=-\infty}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h z^{h-j} = \sigma^2 \sum_{h=0}^{\infty} \psi_h z^h \sum_{j=0}^{\infty} \psi_j z^{-j}, \quad (1.74)$$

De la misma manera, es posible verificar que para un proceso ARMA(p, q) su ACGF es de la forma:

$$g_Y(z) = \frac{|\theta(z)|^2}{|\phi(z)|^2} \sigma^2 = \frac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})} \sigma^2. \quad (1.75)$$

Capítulo 2

Análisis de Fourier.

2.1. Series de Fourier complejas para funciones periódicas.

El análisis de Fourier o análisis armónico de una serie de tiempo es una descomposición de la serie en la suma de componentes sinusoidales, el cual es usado para describir o medir las fluctuaciones en una serie. Por tanto, de lo que tratan las series de Fourier es de aproximar una función f mediante su polinomio trigonométrico de orden n .

Sea $f_p(t)$, una onda periódica con periodo T_0 , la cual puede ser expresada de la forma:

$$f_p(t) = \sum_{n=-\infty}^{\infty} F(n)e^{inw_0t}, \quad (2.1)$$

entonces se tiene que:

$$\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t)e^{-inw_0t} dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \sum_{n=-\infty}^{\infty} F(n)e^{inw_0t} e^{-inw_0t} dt,$$

donde el coeficiente $F(n)$ puede ser encontrado de la siguiente manera:

$$F(n) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) e^{-in\omega_0 t} dt. \quad (2.2)$$

Dicho coeficiente $F(n)$ es indispensable, ya que es el que nos ayudará a descomponer la función periódica, $f_p(t)$, en sus series armónicas.

2.1.1. Propiedades de los coeficientes de Fourier complejos.

A continuación, se dará a conocer una propiedad importante sobre el coeficiente, $F(n)$, ya que su resultado será de gran utilidad a lo largo de este capítulo.

Sea $f_p(t) \Leftrightarrow^1 F(n)$, entonces se tiene que $F(n)^* = F(-n)$ si $f_p(t)$ es una función real de t .

$$\begin{aligned} F(n)^* &= \left[\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) e^{-in\omega_0 t} dt \right]^* \\ &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t)^* e^{in\omega_0 t} dt = F(-n). \end{aligned} \quad (2.3)$$

Por otra parte se quiere mostrar que $F(n) = F(-n)^*$

$$\begin{aligned} F(n) &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) e^{-in\omega_0 t} dt \\ &= \left[\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) e^{in\omega_0 t} dt \right]^* = F(-n)^*. \end{aligned} \quad (2.4)$$

Por lo tanto, se tiene que:

$$\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} [f_p(t) - f_p(t)^*] e^{-in\omega_0 t} dt = 0, \quad (2.5)$$

¹la expresión $f_p(t) \Leftrightarrow F(n)$ significa que la onda periódica $f_p(t)$ tiene $F(n)$ como sus coeficientes de Fourier en (2.2). Además $F(n)$ puede utilizarse para sintetizar $f_t(p)$ en (2.1).

2.1. SERIES DE FOURIER COMPLEJAS PARA FUNCIONES PERIÓDICAS.41

de donde si se cumple que $f_p(t) = f_p(t)^*$, se dice que $f_p(t)$ es real.

Utilizando los resultados anteriores, es fácil observar que el coeficiente $F(n)$ es par y que además será en general una cantidad compleja, la cual puede ser expresado en términos reales e imaginarios de la forma siguiente:

$$F(n) = A(n) + iB(n) \quad (2.6)$$

2.1.2. Ondas periódicas.

Supóngase que se tiene una función periódica, $f_p(t)$, la cual se asume como una función compleja. Entonces, la potencia media de una onda en un periodo es definida de la forma:

$$P = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |f_p(t)|^2 dt = \sum_{n=-\infty}^{\infty} |F(n)|^2, \quad (2.7)$$

es decir, si tenemos a $|f_p(t)|^2 = f_p(t)f_p^*(t)$ y se sustituye en (2.7) de forma que:

$$\begin{aligned} P &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |f_p(t)|^2 dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t)f_p^*(t) dt \\ &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \left[\sum_{n=-\infty}^{\infty} F(n)e^{in\omega_0 t} \right] \left[\sum_{m=-\infty}^{\infty} F(m)^*e^{im\omega_0 t^*} \right] dt \\ &= \left[\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} e^{in\omega_0 t} e^{im\omega_0 t^*} \right] \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} F(n)F(m)^*, \quad (2.8) \end{aligned}$$

y cuando $m = n$ se tiene:

$$P = \sum_{n=-\infty}^{\infty} F(n)F(n)^* = \sum_{n=-\infty}^{\infty} |F(n)|^2. \quad (2.9)$$

Por tanto, la finalidad de lo mostrado anteriormente es ver que existen dos métodos o formas en la cual la potencia media puede ser calculada.

Tomando en cuenta todo lo anterior, es posible continuar con el análisis de la ecuación (2.1), ya que se pretende la finalidad es demostrar o hacer ver como es que cualquier función periódica, con periodo T_0 , puede ser expresada como la suma de las funciones senos y cosenos, las cuales son periódicas con periodo 2π . Más adelante se recurre a la hipótesis de que algunas series de tiempo son periódicas y, por ende, puede descomponerse en otras series de tiempo periodicas.

De esta manera, se tiene que:

$$\begin{aligned} f_p(t) &= \sum_{n=-\infty}^{\infty} F(n)e^{inw_0t} = \sum_{n=-\infty}^{\infty} [A(n) + iB(n)] e^{inw_0t} \\ &= \sum_{n=-\infty}^{\infty} A(n) [\cos(nw_0t) + i\sin(nw_0t)] \\ &\quad + \sum_{n=-\infty}^{\infty} iB(n) [\cos(nw_0t) + i\sin(nw_0t)], \end{aligned} \quad (2.10)$$

en la cual $A(n)$ es par y $iB(n)$ es impar. Suponiendo que $f_p(t)$ es real y en la cual ambos sumando son pares, es posible escribir a (2.10) de la forma:

$$\begin{aligned} f_p(t) &= \sum_{n=-\infty}^{\infty} A(n)\cos(nw_0t) - \sum_{n=-\infty}^{\infty} B(n)\sin(nw_0t) \\ &= F(0) + \sum_{n=1}^{\infty} 2A(n)\cos(nw_0t) - \sum_{n=1}^{\infty} 2B(n)\sin(nw_0t). \end{aligned} \quad (2.11)$$

Ahora se define:

$\begin{aligned} a(n) &= 2A(n) = 2\operatorname{Re}(F(n)) & a(0) &= 2F(0) \\ b(n) &= -2B(n) = -2\operatorname{Im}(F(n)) \end{aligned}$
--

se tiene:

$$f_p(t) = \frac{a(0)}{2} + \sum_{n=1}^{\infty} a(n)\cos(nw_0t) + \sum_{n=1}^{\infty} b(n)\sin(nw_0t), \quad (2.12)$$

donde los coeficientes de Fourier $a(n)$ y $b(n)$ pueden ser encontrados de la

definición analítica de $f_p(t)$ por:

$$a(n) = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) \cos(nw_0 t) dt \quad y \quad b(n) = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) \sin(nw_0 t) dt$$

2.2. La integral de Fourier.

2.2.1. Trasformada de Fourier para un único pulso.

Se han considerado, hasta el momento, ondas periódicas para las cuales existe, para cada una de ellas, su representación en series de Fourier. Pero en la mayoría de los casos, existen ondas en las que solo ocurre una vez y no se vuelven a repetir llamados pulsos únicos, por lo que las series de Fourier que cubren este tipo de onda son estudiadas mediante la integral de Fourier. Tomando en cuenta que para una función con periodo T_0 ,

$$f_p(t) = f_p(t + T_0), \quad (2.13)$$

en el intervalo $(-T_0/2 < t < T_0/2)$, y sin pérdida de generalidad tomando también a $f_p(t) = f(t)$, y además de (2.2) se tiene que

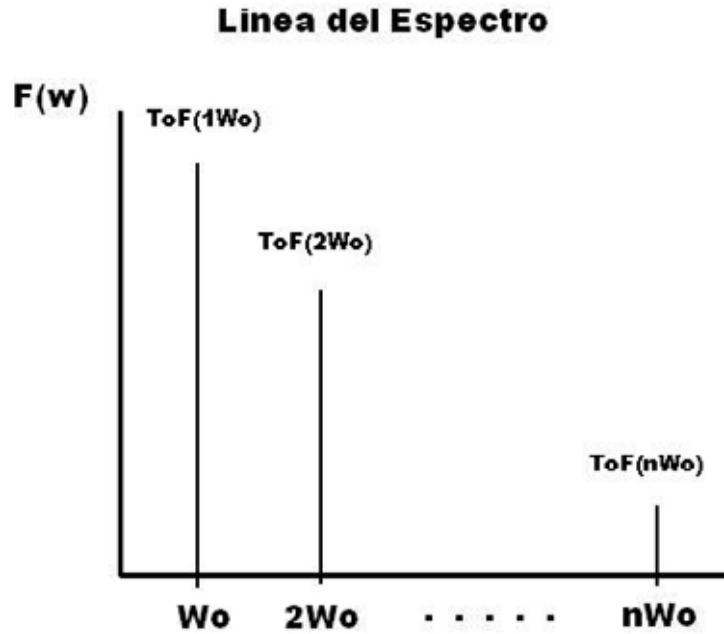
$$F(n) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_p(t) e^{-inw_0 t} dt, \quad (2.14)$$

la cual, es escrito de la forma

$$F(nw_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f(t) e^{-inw_0 t} dt. \quad (2.15)$$

Ahora, si $T_0 \rightarrow \infty$, el promedio tendera a cero, como lo muestra la Gráfica siguiente:

Gráfica 11. Línea espectral



Así, para poder eliminar este problema se escribe (2.3) de la forma:

$$T_0 F(nw_0) = \int_{-T_0/2}^{T_0/2} f(t) e^{-inw_0 t} dt. \quad (2.16)$$

Tomando a $\delta w = w_0$, $w_n = n\delta w$ y $F(n) = T_0 F(nw_0)$, es posible escribir a (2.16) como:

$$F(w_n) = \int_{-T_0/2}^{T_0/2} f(t) e^{-iw_n t} dt. \quad (2.17)$$

De esta manera, como $T_0 \rightarrow \infty$, las divisiones del eje horizontal se acercan cada vez mas hasta que finalmente desaparecen haciendo que la variable w_n se convierta en variable continua w .

$$\begin{aligned}
 F(w) &= \lim_{T_0 \rightarrow \infty} F(w_n) \\
 &= \lim_{T_0 \rightarrow \infty} \int_{-T_0/2}^{T_0/2} f(t) e^{-iw_n t} dt \\
 &= \int_{-\infty}^{\infty} f(t) e^{-iwt} dt
 \end{aligned} \tag{2.18}$$

Considerando el resultado anterior, la síntesis de la ecuación para nuestra onda periodica es:

$$\begin{aligned}
 f(t) &= \sum_{n=-\infty}^{\infty} F(nw_0) e^{inw_0 t} \\
 &= \sum_{n=-\infty}^{\infty} \frac{F(w_n) e^{inw_0 t}}{T_0} \\
 &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(w_n) e^{inw_n t} T_0 \delta w.
 \end{aligned} \tag{2.19}$$

Si ahora se hace que $T_0 \rightarrow \infty$, entonces $\delta w \rightarrow 0$ y $w_n \rightarrow w$ por lo que se tiene:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(w) e^{iwt} \delta w \tag{2.20}$$

Por lo tanto, todo lo anterior se resume, mediante las siguientes ecuaciones:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(w) e^{iwt} \delta w \tag{2.21}$$

$$F(w) = \int_{-\infty}^{\infty} f(t) e^{-iwt} dt \tag{2.22}$$

donde la ecuación (2.21) muestra que una función del tiempo, $f(t)$, puede ser analizada para dar una función de frecuencia asociada $F(w)$ que contiene toda la información en $f(t)$. A su vez, también es posible observar como se

sintetiza un pulso dada la función de densidad espectral llamada la Transformada Inversa de Fourier de $F(w)$, o la representación integral de Fourier de $f(t)$. Por tanto, se llama a $F(w)$ en (2.18) la Transformada de Fourier de $f(t)$ o la densidad espectral de Fourier de $f(t)$.

Lo que destaca, de lo anterior, es que una vez mas cualquier función $f(t)$ puede ser expresada como una combinación lineal de los exponenciales complejos. De forma que la integral de Fourier es una simple repetición de lo visto acerca de las funciones periódicas, pero en sentido mas general, es decir, es este caso no es necesario que las funciones sean periódicas.

2.3. Densidad espectral.

Suponga que $\{X_t\}$ es una serie de tiempo estacionaria con media cero y función de autocovarianza $\gamma(\cdot)$ y además que $\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$. De esta manera, se tiene que la densidad espectral de $\{X_t\}$ es la función $f(\cdot)$ definida mediante:

$$\begin{aligned} f(w) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-i w k} \\ &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) \cos w k \\ &= \frac{1}{2\pi} \gamma(0) + \frac{1}{\pi} \sum_{k=1}^{\infty} \gamma(k) \cos w k, \quad -\pi \leq w \leq \pi \end{aligned} \quad (2.23)$$

donde se han utilizado las propiedades de que $\gamma_k = \gamma_{-k}$ y

$$\gamma_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(w) e^{i k w}. \quad (2.24)$$

2.3.1. Espectro del modelo AR(1).

La función generadora de autocovarianza para un proceso estacionario AR(1), $(1 - \phi L)y_t = \epsilon_t$, esta dada por:

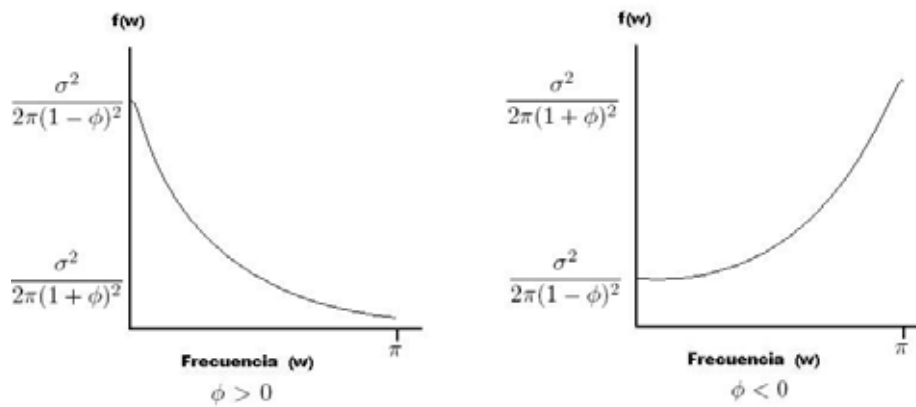
$$\gamma(L) = \sigma^2 \frac{1}{(1 - \phi L)(1 - \phi L^{-1})}, \quad (2.25)$$

por lo que el espectro es de la fomra:

$$\begin{aligned} f(w) &= \frac{\sigma^2}{2\pi} \frac{1}{(1 - \phi e^{-iw})(1 - \phi e^{iw})} \\ &= \frac{\sigma^2}{2\pi} \frac{1}{(1 + \phi^2 - 2\phi \cos w)}. \end{aligned} \quad (2.26)$$

Como se muestra en las siguientes gráficas, cuando el valor de $\phi > 0$, la serie esta correlacionada positivamente, es decir, el espectro esta dominado por las frecuencias bajas. Por el otro lado, si $\phi < 0$, la serie esta correlacionada negativamente y por lo tanto el espectro esta dominado por las frecuencias altas. En otras palabras, cuando las frecuencias bajas dominan el espectro indica una serie relativamente suave, mientras que las frecuencias altas indican una serie irregular.

Gráfica 12. Espectro del modelo AR(1).



Además, tan pronto como ϕ se aproxima a 1, el proceso AR(1) en el límite, se convierte en un modelo de caminata aleatoria o de ruido blanco, por lo que se obtiene:

$$f(w) = \frac{\sigma^2}{4\pi} \frac{1}{(1 - \cos w)}. \quad (2.27)$$

La ecuación anterior puede ser tomado como el espectro de un modelo de caminata aleatoria, aunque una caminata aleatoria no tiene un espectro debido a que la secuencia de las autocovarianzas no es absolutamente sumable. Sin embargo, el problema es cuando $w = 0$ donde el máximo de la función se vuelve infinita. La implicación de este fenómeno en el análisis de datos, indica que quizá sea necesaria una diferencia.

2.3.2. Espectro del modelo MA(1).

La función generadora de autocovarianza para un modelo MA(1)

$$y_t = (1 - \theta L)\epsilon_t, \quad (2.28)$$

es de la forma

$$g(L) = \sigma^2(1 - \theta L)(1 - \theta L^{-1}), \quad (2.29)$$

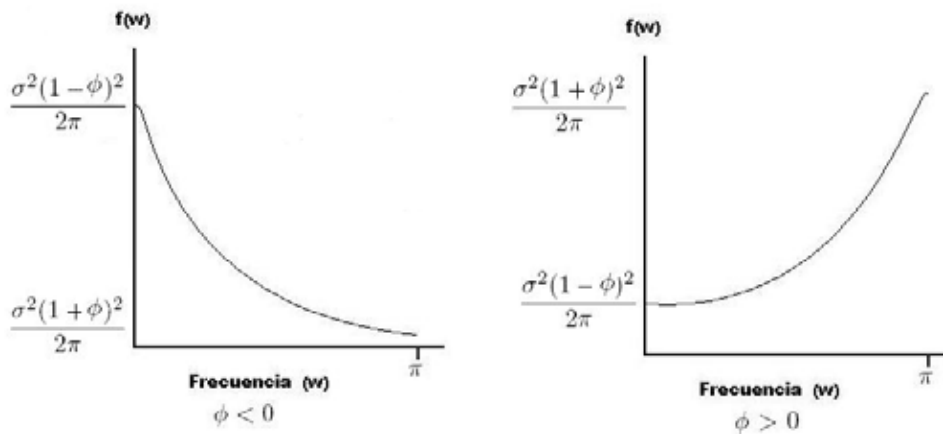
por lo que su espectro toma el valor de:

$$\begin{aligned} f(w) &= \frac{\sigma^2}{2\pi} (1 - \theta e^{-iw})(1 - \theta e^{iw}) \\ &= \frac{\sigma^2}{2\pi} (1 + \theta^2 - 2\theta \cos w). \end{aligned} \quad (2.30)$$

Ahora, tomando la Gráfica 13 se observa que, al igual que en el modelo AR(1), el espectro depende del signo de ϕ . Es decir, cuando $\phi > 0$, la serie esta autocorrelacionada negativamente y por lo tanto es relativamente irregular, la cual esta representado por un espectro con valores grandes en las frecuencias altas.

En cambio, cuando $\phi < 0$, la serie esta correlacionada positivamente y por lo tanto es relativamente suave. Ambos casos son representados en la siguiente Gráfica.

Gráfica 13. Espectro del modelo MA(1).



Como era de esperarse, la función de autocorrelación en el dominio del tiempo y el espectro en el dominio frecuencial contienen exactamente la misma información acerca del proceso.

2.3.3. Espectro del proceso ARMA.

Anteriormente, se señaló que un proceso ARMA(p,q) se escribe como:

$$\phi(L)y_t = \theta(L)\epsilon_t. \quad (2.31)$$

Entonces, la ACGF para un proceso ARMA (p,q) es de la forma:

$$g(L) = \sigma^2 \frac{\theta(L)\theta(L^{-1})}{\phi(L)\phi(L^{-1})}. \quad (2.32)$$

Cuando el proceso es estacionario, las raíces de $\phi(L) = 0$ se encuentran fuera del círculo unitario. En consecuencia, el espectro de un modelo estacionario

ARMA(p,q) esta dado por:

$$\begin{aligned}
 f(w) &= \frac{1}{2\pi} g(e^{-iw}) \\
 &= \frac{\sigma^2}{2\pi} \frac{\theta(e^{-iw})\theta(e^{iw})}{\phi(e^{-iw})\phi(e^{iw})} \\
 &= \frac{\sigma^2}{2\pi} \left| \frac{\theta(e^{-iw})}{\phi(e^{-iw})} \right|^2.
 \end{aligned} \tag{2.33}$$

Si el modelo es invertible, es decir, las raíces de $\theta(L) = 0$ se encuentran fuera del círculo unitario, $\theta(e^{-iw})\theta(e^{iw})$ no desaparece. Por consiguiente, la inversa de $f(w)$ existe y esta dada por:

$$\begin{aligned}
 f^{-1} &= \frac{2\pi}{\sigma^2} \frac{\phi(e^{-iw})\phi(e^{iw})}{\theta(e^{-iw})\theta(e^{iw})} \\
 &= \frac{2\pi}{\sigma^2} \left| \frac{\phi(e^{-iw})}{\theta(e^{-iw})} \right|^2,
 \end{aligned} \tag{2.34}$$

la cual puede ser vista fácilmente como el espectro de un proceso autoregresivo de medias móviles, ARMA(q,p).

2.3.4. Espectro de un proceso ruido blanco.

Tomese ahora, a el ruido blanco $y_t = \epsilon_t$ como una serie de variables aleatorias no correlacionadas con función de autocovarianza

$$\gamma(k) = \begin{cases} \sigma^2 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{2.35}$$

y además con función generadora de autocovarianzas igual a:

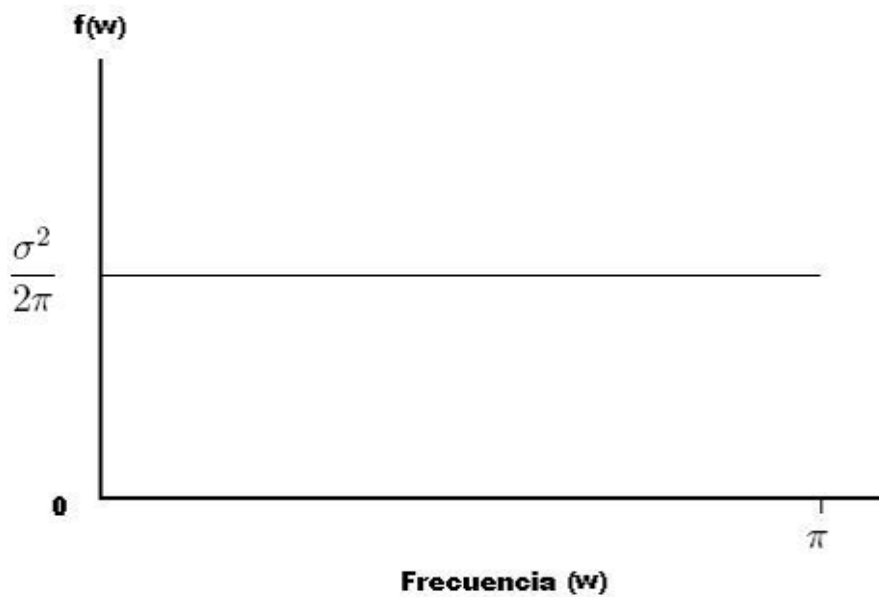
$$\gamma(L) = \sigma^2. \tag{2.36}$$

Por tanto, sustituyendo en la ecuación (2.23) se tiene que:

$$\begin{aligned} f(w) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-i\omega k} \\ &= \frac{\sigma^2}{2\pi}, \quad -\pi \leq \omega \leq \pi \end{aligned} \quad (2.37)$$

donde el resultado es una constante. Es decir, que la potencia en todas las frecuencias es la misma. Dicho lo anterior, se puede ver claramente en la Gráfica siguiente:

Gráfica 14. Espectro del ruido blanco.



2.4. Función de verosimilitud para procesos gaussianos.

2.4.1. Función de verosimilitud para un proceso gaussiano AR(1).

Un proceso gaussiano AR(1) toma la forma:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \quad (2.38)$$

con $\epsilon_t \sim$ i.i.d. $N(0, \sigma^2)$. Para este caso, el vector de parámetros que se estima consta de $\theta \equiv (c, \phi, \sigma^2)'$.

Ahora, considere la distribución de probabilidad de y_t , la primera observación en la muestra. Para este caso tenemos que la media y la varianza son de la siguiente forma, respectivamente:

$$E(Y_1) = \mu = c/(1 - \phi) \quad (2.39)$$

y

$$E(Y_1 - \mu)^2 = \sigma^2/(1 - \phi^2). \quad (2.40)$$

Además como $\{\epsilon_t\}_{t=-\infty}^{\infty}$ es gaussiana, Y_1 también lo es. Por tanto, la densidad de la primera observación toma la forma:

$$\begin{aligned} f_{Y_1}(y_1; \theta) &= f_{Y_1}(y_1; c, \phi, \sigma^2) \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1 - \phi^2)}} \exp \left[\frac{-\{y_1 - [c/(1 - \phi)]\}^2}{2\sigma^2/(1 - \phi^2)} \right]. \end{aligned} \quad (2.41)$$

Considere ahora, la distribución de la segunda observación Y_2 condicionada a la observación $Y_1 = y_1$, con lo que se obtiene:

$$Y_2 = c + \phi Y_1 + \epsilon_2. \quad (2.42)$$

2.4. FUNCIÓN DE VEROSIMILITUD PARA PROCESOS GAUSSIANOS.53

Condicionando sobre $Y_1 = y_1$ significa que se esta tratando a la variable aleatoria Y_1 como si fuera un la constante determinística y_1 . Para este caso, la ecuación (2.38) da Y_2 como la constante $(c + \phi y_1)$ mas la variable ϵ_2 . Por tanto:

$$(Y_2|Y_1 = y_1) \sim N((c + \phi y_1), \sigma^2),$$

significa que

$$f_{Y_2|Y_1}(y_2|y_1; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{-(y_2 - c - \phi y_1)^2}{2\sigma^2} \right] \quad (2.43)$$

De tal manera que la densidad conjunta de las primeras dos observaciones es:

$$f_{Y_2, Y_1}(y_2, y_1; \theta) = f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta). \quad (2.44)$$

Similarmente, la distribución de la tercera observación condicionada a las primeras dos es:

$$f_{Y_3|Y_2, Y_1}(y_3|y_2, y_1; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{-(y_3 - c - \phi y_2)^2}{2\sigma^2} \right], \quad (2.45)$$

de la cual

$$f_{Y_3, Y_2, Y_1}(y_3, y_2, y_1; \theta) = f_{Y_3|Y_2, Y_1}(y_3|y_2, y_1; \theta) f_{Y_2, Y_1}(y_2, y_1; \theta). \quad (2.46)$$

En general, los valores de Y_1, Y_2, \dots, Y_{t-1} tienen importancia para Y_t solo a través de el valor de Y_{t-1} , y la densidad de la observación t condicionada a las anteriores $t - 1$ observaciones esta dada por:

$$\begin{aligned} f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_1}(y_t|y_{t-1}, y_{t-2}, \dots, y_1; \theta) &= f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{-(y_t - c - \phi y_{t-1})^2}{2\sigma^2} \right] \end{aligned} \quad (2.47)$$

Por tanto, la distribución conjunta de las primeras t observaciones es:

$$\begin{aligned} f_{Y_t, Y_{t-1}, \dots, Y_1}(y_t, y_{t-1}, \dots, y_1; \theta) \\ = f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta) f_{Y_{t-1}, Y_{t-2}, \dots, Y_1}(y_{t-1}, y_{t-2}, \dots, y_1; \theta), \end{aligned} \quad (2.48)$$

y la verosimilitud de la muestra completa puede ser calculada como:

$$f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \theta) = f_{Y_1}(y_1; \theta) \prod_{t=2}^T f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta). \quad (2.49)$$

El log de la función de verosimilitud (denotada por $\mathcal{L}(\theta)$) puede ser encontrada tomando el log de la función anterior

$$\mathcal{L}(\theta) = \log f_{Y_1}(y_1; \theta) + \sum_{t=2}^T \log f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta) \quad (2.50)$$

Se tiene que θ es el valor que maximiza la función anterior. Ahora sustituyendo a (2.41) y (2.48) en (2.50), se tiene que el log verosimil para una muestra de tamaño T para un proceso gaussiano AR(1) es:

$$\begin{aligned} \mathcal{L}(\theta) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log \left[\sigma^2 / (1 - \phi^2) \right] \\ &\quad - \frac{\{y_1 - [c/(1 - \phi)]\}^2}{2\sigma^2 / (1 - \phi^2)} - [(T - 1)/2] \log(2\pi) \\ &\quad - [(T - 1)/2] \log(\sigma^2) - \sum_{t=2}^T \left[\frac{(y_t - c - \phi y_{t-1})^2}{2\sigma^2} \right] \end{aligned} \quad (2.51)$$

2.4.2. Función de verosimilitud para un proceso gaussiano MA(1).

Considere el proceso gaussiano MA(1)

$$Y_t = \mu + \epsilon_t + \theta \epsilon_{t-1}, \quad (2.52)$$

con $\epsilon_t \sim$ i.i.d. $N(0, \sigma^2)$. Sea $\theta = (\mu, \theta, \sigma^2)'$ que denota los parámetros a estimar.

2.4. FUNCIÓN DE VEROSIMILITUD PARA PROCESOS GAUSSIANOS.55

Si se toma al valor de ϵ_t , como un valor conocido con precisión, se tiene entonces:

$$Y_t|\epsilon_{t-1} \sim N((\mu + \theta\epsilon_{t-1}), \sigma^2)$$

o

$$f_{Y_t|\epsilon_{t-1}}(y_t|\epsilon_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_t - \mu - \theta\epsilon_{t-1})^2}{2\sigma^2}\right]. \quad (2.53)$$

Supongase que el valor de $\epsilon_0 = 0$, por lo que:

$$(Y_1|\epsilon_0 = 0) \sim N(\mu, \sigma^2)$$

Por otra parte, dada la observación de y_1 , el valor de ϵ_1 es conocido de la siguiente manera:

$$\epsilon_1 = y_1 - \mu, \quad (2.54)$$

permitiendo la aplicación nuevamente en la ecuación (2.53), se tiene:

$$f_{Y_2|Y_1, \epsilon_0=0}(y_2|y_1, \epsilon_0 = 0; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_2 - \mu - \theta\epsilon_1)^2}{2\sigma^2}\right] \quad (2.55)$$

Como el valor de ϵ_1 es conocido, ϵ_2 puede ser calculado de

$$\epsilon_2 = y_1 - \mu - \theta\epsilon_1, \quad (2.56)$$

Tomando a $\epsilon_0 = 0$, se tiene que toda la secuencia $\{\epsilon_1, \epsilon_2, \dots, \epsilon_T\}$ puede ser calculada de $\{y_1, y_2, \dots, y_T\}$ iterando

$$\epsilon_t = y_t - \mu - \theta\epsilon_{t-1} \quad (2.57)$$

para $t = 1, 2, \dots, T$, desde $\epsilon_0 = 0$. La densidad condicional de la t -ésima observación puede ser calculada en (2.53) como:

$$\begin{aligned} f_{Y_t|Y_{t-1}, Y_{t-2}, \dots, Y_1, \epsilon_0=0}(y_t|y_{t-1}, y_{t-2}, \dots, y_1, \epsilon_0 = 0; \theta) \\ &= f_{Y_t|\epsilon_{t-1}}(y_t|\epsilon_{t-1}; \theta) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\epsilon_t^2}{2\sigma^2}\right]. \end{aligned} \quad (2.58)$$

De esta manera, la muestra verosímil será el producto de esas densidades individuales, es decir:

$$\begin{aligned}
 & f_{Y_T, Y_{T-1}, \dots, Y_1 | \epsilon_0 = 0}(y_T, y_{T-1}, \dots, y_1 | \epsilon_0 = 0; \theta) \\
 &= f_{Y_1 | \epsilon_0 = 0}(y_1 | \epsilon_0 = 0; \theta) \prod_{t=2}^T f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_1, \epsilon_0 = 0}(y_t | y_{t-1}, y_{t-2}, \dots, \epsilon_0 = 0; \theta)
 \end{aligned} \tag{2.59}$$

y el log verosímil condicionado es:

$$\begin{aligned}
 \mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1 | \epsilon_0 = 0}(y_T, y_{T-1}, \dots, y_1 | \epsilon_0 = 0; \theta) \\
 &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\epsilon_t^2}{2\sigma^2}
 \end{aligned} \tag{2.60}$$

2.4.3. Función de verosimilitud para un proceso gaussiano ARMA(p,q).

Un proceso gaussiano ARMA(p,q) es de la forma:

$$\begin{aligned}
 Y_t &= c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \\
 &\quad + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q},
 \end{aligned} \tag{2.61}$$

con $\epsilon_t \sim \text{i.i.d. } N(0, \sigma^2)$. Ahora, el objetivo es estimar el vector de parámetros

$$\theta = (c, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma^2)'$$

Tomando valores iniciales para $\mathbf{y}_0 \equiv (y_0, y_{-1}, \dots, y_{-p+1})'$ y $\epsilon_0 \equiv (\epsilon_0, \epsilon_{-1}, \dots, \epsilon_{-q+1})'$,

$\{\epsilon_1, \epsilon_2, \dots, \epsilon_T\}$ puede ser calculada de $\{y_1, y_2, \dots, y_T\}$ iterando

$$\begin{aligned}
 \epsilon_t &= y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} \\
 &\quad - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q},
 \end{aligned} \tag{2.62}$$

2.4. FUNCIÓN DE VEROSIMILITUD PARA PROCESOS GAUSSIANOS.57

para $t = 1, 2, \dots, T$. El log verosímil condicionado es entonces

$$\begin{aligned}\mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1 | Y_0, \epsilon_0=0}(y_T, y_{T-1}, \dots, y_1 | y_0, \epsilon_0; \theta) \\ &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\epsilon_t^2}{2\sigma^2}.\end{aligned}\quad (2.63)$$

Una opción es establecer valores iniciales para y 's y ϵ 's igual a su esperanza. Esto es, fijar $y_s = c/(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ para $s = 0, -1, \dots, -p+1$, $\epsilon_s = 0$ para $s = 0, -1, \dots, -q+1$ y realizar las iteraciones correspondientes en (2.62) para $t = 1, 2, \dots, T$.

2.4.4. Expresión alternativa para la función de verosimilitud.

Pongamos en un vector \mathbf{y} ($T \times 1$) todas las observaciones de la forma:

$$\mathbf{y} \equiv (y_1, y_2, \dots, y_T)',$$

en donde dicho vector, puede ser visto como una realización de una distribución gaussiana T -dimensional. Por otro lado, la media de este vector es:

$$\begin{pmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_T) \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix}$$

donde, $\mu = c/(1 - \phi)$, el cual puede ser escrito como:

$$E(\mathbf{Y}) = \mu.$$

A su vez, la matriz de varianza covarianza de \mathbf{Y} esta dada por:

$$E[(\mathbf{Y} - \mu)(\mathbf{Y} - \mu)'] = \Omega,$$

donde:

$$\Omega = \begin{pmatrix} E(Y_1 - \mu)^2 & E(Y_1 - \mu)(Y_2 - \mu) & \dots & E(Y_1 - \mu)(Y_T - \mu) \\ E(Y_2 - \mu)(Y_1 - \mu) & E(Y_2 - \mu)^2 & \dots & E(Y_2 - \mu)(Y_T - \mu) \\ \vdots & \vdots & & \vdots \\ E(Y_T - \mu)(Y_1 - \mu) & E(Y_T - \mu)(Y_2 - \mu) & \dots & E(Y_T - \mu)^2 \end{pmatrix}$$

Los elementos de esta matriz corresponden a las autocovarianzas de Y . Recordando que la j -ésima autocovarianza para un proceso AR(1) esta dada por:

$$E(Y_t - \mu)(Y_{t-j} - \mu) = \sigma^2 \phi^j / (1 - \phi^2). \quad (2.64)$$

entonces, Ω puede ser escrito como:

$$\Omega = \sigma^2 \mathbf{V},$$

donde:

$$\mathbf{V} = \frac{1}{1 - \phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \dots & \phi^{T-1} \\ \phi & 1 & \phi & \dots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \dots & \phi^{T-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \dots & 1 \end{pmatrix}$$

Visualizando la muestra observada \mathbf{y} como una sola extracción de una distribución $N(\mu, \Omega)$, la muestra verosímil puede ser escrita de la fórmula para la densidad gaussiana multivariada:

$$f_Y(\mathbf{y}; \theta) = (2\pi)^{-T/2} |\Omega^{-1}|^{1/2} \exp\left[-\frac{1}{2}(\mathbf{y} - \mu)' \Omega^{-1}(\mathbf{y} - \mu)\right], \quad (2.65)$$

2.4. FUNCIÓN DE VEROSIMILITUD PARA PROCESOS GAUSSIANOS.59

con log verosimil expresado de la forma:

$$\mathcal{L}(\theta) = (-T/2) \log(2\pi) + \frac{1}{2} |\Omega^{-1}| - \frac{1}{2} (\mathbf{y} - \mu)' \Omega^{-1} (\mathbf{y} - \mu). \quad (2.66)$$

2.4.5. Función de verosimilitud para un proceso gaussiano AR(p).

Tomando en consideración el proceso AR(1), se tiene que un proceso gaussiano AR(p) es de la forma:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t, \quad (2.67)$$

con $\epsilon_t \sim N(0, \sigma^2)$. En este caso, el vector de parámetros es $\theta = (c, \phi_1, \phi_2, \dots, \phi_p, \sigma^2)'$

Las primeras p observaciones en la muestra (y_1, y_2, \dots, y_p) son resumidos en un vector \mathbf{y}_p ($p \times 1$), el cual es visto como la realización de una variable gaussiana p -dimensional. La media de este vector es μ_p , el cual denota un vector ($p \times 1$), de los cuales cada elemento esta dado por:

$$\mu = c / (1 - \phi_1 - \phi_2 - \dots - \phi_p). \quad (2.68)$$

Sea $\sigma^2 \mathbf{V}_p$ el cual denota la matriz de varianza-covarianza ($p \times p$) de (y_1, y_2, \dots, y_p) :

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} E(Y_1 - \mu)^2 & E(Y_1 - \mu)(Y_2 - \mu) & \dots & E(Y_1 - \mu)(Y_p - \mu) \\ E(Y_2 - \mu)(Y_1 - \mu) & E(Y_2 - \mu)^2 & \dots & E(Y_2 - \mu)(Y_p - \mu) \\ \vdots & \vdots & & \vdots \\ E(Y_p - \mu)(Y_1 - \mu) & E(Y_p - \mu)(Y_2 - \mu) & \dots & E(Y_p - \mu)^2 \end{pmatrix}$$

Por ejemplo, como se mostró anteriormente para un AR(1), \mathbf{V}_p es el escalar $1/(1 - \phi^2)$.

De la matriz anterior, un proceso AR(p) es escrito de la forma:

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{pmatrix}$$

donde γ_τ , es la τ -ésima autocovarianza de un proceso AR(p). La densidad de las primeras p observaciones es entonces la de una $N(\mu_p, \sigma^2 \mathbf{V}_p)$

$$\begin{aligned} & f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \theta) \\ &= (2\pi)^{-p/2} |\sigma^{-2} \mathbf{V}_p^{-1}|^{1/2} \exp \left[-\frac{1}{2\sigma^2} (\mathbf{y}_p - \mu_p)' \mathbf{V}_p^{-1} (\mathbf{y}_p - \mu_p) \right] \\ &= (2\pi)^{-p/2} (\sigma^{-2})^{p/2} |\mathbf{V}_p^{-1}|^{1/2} \exp \left[-\frac{1}{2\sigma^2} (\mathbf{y}_p - \mu_p)' \mathbf{V}_p^{-1} (\mathbf{y}_p - \mu_p) \right], \end{aligned} \quad (2.69)$$

Condicionando en las primeras $t - 1$ observaciones, la t -ésima observación es gaussiana con media

$$c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}$$

y varianza σ^2 .

Por otro lado, para $t > p$,

$$\begin{aligned} & f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_1}(y_t | y_{t-1}, y_{t-2}, \dots, y_1; \theta) \\ &= f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}; \theta) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} \right] \end{aligned} \quad (2.70)$$

2.4. FUNCIÓN DE VEROSIMILITUD PARA PROCESOS GAUSSIANOS.61

Entonces, la función de verosimilitud para la muestra completa es:

$$\begin{aligned}
 f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \theta) \\
 &= f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \theta) \\
 &\quad \times \prod_{t=p+1}^T f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}; \theta), \quad (2.71)
 \end{aligned}$$

y la log verosimilitud es entonces:

$$\begin{aligned}
 \mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \theta) \\
 &= -\frac{p}{2} \log(2\pi) - \frac{p}{2} \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| \\
 &\quad - \frac{1}{2\sigma^2} (\mathbf{y}_p - \mu_p)' \mathbf{V}_p^{-1} (\mathbf{y}_p - \mu_p) \\
 &\quad - \frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) \\
 &\quad - \sum_{t=p+1}^T \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p})^2}{2\sigma^2} \\
 &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| \\
 &\quad - \frac{1}{2\sigma^2} (\mathbf{y}_p - \mu_p)' \mathbf{V}_p^{-1} (\mathbf{y}_p - \mu_p) \\
 &\quad - \sum_{t=p+1}^T \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p})^2}{2\sigma^2} \quad (2.72)
 \end{aligned}$$

Para poder evaluar la expresión anterior, se requiere invertir la matriz \mathbf{V}_p , para ello, se denota a los elementos de \mathbf{V}_p^{-1} por $v^{ij}(p)$, con i :=fila y j :=columna. Además se define también a $v^{ij}(p)$ de la forma:

$$v^{ij}(p) = \left[\begin{array}{cc} \sum_{k=0}^{i-1} \phi_k \phi_{k+j-1} & - \sum_{k=p+1-j}^{p+i-j} \phi_k \phi_{k+j-1} \end{array} \right], \quad \text{para } 1 \leq i \leq j \leq p \quad (2.73)$$

donde $\phi_0 \equiv -1$. Valores de $v^{ij}(p)$ para $i > j$ pueden inferirse del hecho de que \mathbf{V}_p^{-1} es simétrica $v^{ij}(p) = v^{ji}(p)$.

Por ejemplo, para un proceso AR(1), V_p^{-1} es un escalar cuyo valor se encuentra tomando a $i = j = p = 1$;

$$\mathbf{V}_1^{-1} = \left[\sum_{k=0}^0 \phi_k \phi_k - \sum_{k=1}^1 \phi_k \phi_k \right] = (\phi_0^2 - \phi_1^2) = (1 - \phi_1).$$

Entonces, $\sigma^2 \mathbf{V}_1 = \sigma^2 / (1 - \phi^2)$. Para $p = 2$ se tiene:

$$\mathbf{V}_2^{-1} = \begin{pmatrix} (1 - \phi_2^2) & -(\phi_1 + \phi_1 \phi_2) \\ -(\phi_1 + \phi_1 \phi_2) & (1 - \phi_2^2) \end{pmatrix}$$

Por lo tanto, la log verosimilitud de un proceso gaussiano AR(2) esta dada por:

$$\begin{aligned} \mathcal{L}(\theta) = & -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) + \frac{1}{2} \log \left\{ (1 + \phi_2)^2 [(1 - \phi_2)^2 - \phi_1^2] \right\} - \left\{ \frac{1 + \phi^2}{2\sigma^2} \right\} \\ & X \left\{ (1 - \phi_2)(y_1 - \mu)^2 - 2\phi_1(y_1 - \mu)(y_2 - \mu) + (1 - \phi_2)(y_2 - \mu)^2 \right\} \\ & - \sum_{t=3}^T \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2})^2}{2\sigma^2} \end{aligned} \quad (2.74)$$

donde:

$$\mu = c / (1 - \phi_1 - \phi_2). \quad (2.75)$$

2.5. Modelo VAR.

Ya antes de 1970, el enfoque de la Comisión Cowles, CC, empezó a mostrar sus deficiencias a medida que sus recomendaciones de política no explicaban los fenómenos que ocurrían en ese entonces. Desde sus comienzos, la CC que entre otras cosas se enfocaban a la construcción de modelos de predicción, a los cuales se les llamó estructurales, porque a través de la incorporación de muchas ecuaciones de comportamiento pretendían captar la esencia de las relaciones estructurales que determinan a las variables económicas y que conjuntamente representan el funcionamiento real de una economía compleja verdadera. Una característica básica y a la vez esencial de los modelos estructurales, es la asignación arbitraria o *ad hoc* entre variables exógenas y las incógnitas o variables endógenas. De acuerdo con Fair, la metodología básica y los modelos de la CC pueden sintetizarse a través de la siguiente expresión:

$$f_i(Y_t, X_t, \alpha_i) = U_{it}, \quad i=(1,2,\dots,n) \text{ y } t=(1,2,\dots,T),$$

con:

Y_t := vector de n variables endógenas.

X_t := vector de variables exógenas o predeterminadas y endógenas rezagadas.

α_i := vectores de coeficientes desconocidos a estimar.

U_{it} := error estocástico para la ecuación i en el momento t .

El problema central de este enfoque consistía en obtener la mejor estimación de los parámetros α_i . La manifestación detallada de la última expresión parte de la siguiente forma estructural general restringida:

$$Y_i\Gamma + X_iB = U_i, \quad i = 1, 2, \dots, n \quad (2.76)$$

en donde:

Y := vector de g variables endógenas o incógnitas a resolver por el sistema.

Γ := matriz de coeficientes de las variables endógenas.

X := vector de las k variables exógenas.

B := matriz de los coeficientes de las variables exógenas.

U := perturbaciones estocásticas.

Si se multiplica la última ecuación por Γ^{-1} se obtiene:

$$\begin{aligned} Y\Gamma^{-1} + XB\Gamma^{-1} &= U\Gamma^{-1} \\ Y + XB\Gamma^{-1} &= U\Gamma^{-1} \\ Y &= -XB\Gamma^{-1} + U\Gamma^{-1} \\ &= X\Pi + V \end{aligned}$$

donde:

$$\Pi = -B\Gamma^{-1}.$$

$$V = U\Gamma^{-1}.$$

Aquí, la gran interrogante consiste en saber si la información que se esté incorporando en el vector X es realmente adecuada. Este es el principal cuestionamiento de Sims a los modelos estructurales, cuando afirma que la asignación de las variables exógenas que contienen la información para estimar los parámetros estructurales es una restricción a priori que debe demostrarse. A partir de esto, Sims y sus asociados han propuesto deshacer el uso de los modelos estructurales y utilizar modelos de vectores autoregresivos (VAR), los cuales toma a todas las variables como endógenas.

El punto de partida de Sims es modelar un VAR general irrestricto o sea sin restricciones, que consiste en regresionar a cada variable no rezaga-

da respecto a todas las demás con varios rezagos a partir de la siguiente expresión:

$$Z_t = \sum_{i=1}^k A_i Z_{t-i} + \epsilon_t, \quad (2.78)$$

donde:

$$Z_t = \begin{bmatrix} Y_t \\ X_t^* \end{bmatrix}$$

$\epsilon_t :=$ es un vector columna de errores aleatorios que están contemporáneamente correlacionados pero no autocorrelacionados.

Y_t y $X_t^* :=$ es una forma sencilla de diferenciar a las variables consideradas y además permiten compararse con la CC.

Dado que el modelo VAR, de la última expresión, solo tiene variables rezagadas del lado derecho y, por definición, esas variables no están correlacionadas con el término de errores, es consistente estimar ecuación por ecuación mediante MCO. Una manera de representar matricialmente a la ecuación (1.130) es:

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \quad (2.79)$$

Entonces, si los errores están correlacionados contemporáneamente se tiene que:

$$E[\epsilon_{1t}] = E[\epsilon_{2t}] = 0 ; E[\epsilon_{1t}^2] = \sigma_{11} ; E[\epsilon_{2t}^2] = \sigma_{22} ; E[\epsilon_{1t}, \epsilon_{2t}] = \sigma_{12}$$

En $\sigma_{12} \neq 0$ es donde radica el problema de autocorrelación entre errores. De esta manera, para poder eliminarlo, es necesario ortogonalizar la matriz de varianza-covarianza de la siguiente manera: multiplicamos la última expresión por $\delta = \frac{\sigma_{12}}{\sigma_{11}}$, restamos el primero menos el segundo renglón y por último tomando esperanzas sobre el término de error:

$$\begin{bmatrix} \delta X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \delta a_1 & \delta b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \delta a_2 & \delta b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \delta \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

$$\begin{bmatrix} X_t \\ Y_t - \delta X_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1^* & d_1^* \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2^* & d_2^* \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_2^* \end{bmatrix}$$

donde:

$$c_1^* = (c_1 - \delta a_1) ; d_1^* = (d_1 - \delta b_1) ; \epsilon_{2t}^* = (\epsilon_{2t} - \delta \epsilon_{1t}), i = 1, 2.$$

Finalmente se tiene que:

$$\begin{aligned} E[\epsilon_{1t}\epsilon_{2t}^*] &= E[\epsilon_{1t}(\epsilon_{2t} - \delta \epsilon_{1t})] \\ &= E[(\epsilon_{1t}\epsilon_{2t}) - (\epsilon_{1t}\delta \epsilon_{1t})] \\ &= E[(\epsilon_{1t}\epsilon_{2t}) - (\epsilon_{1t} \frac{\sigma_{12}}{\sigma_{11}} \epsilon_{1t})] \\ &= \sigma_{12} - \sigma_{12} = 0, \end{aligned}$$

con lo que queda resuelto el problema inicial de correlación entre los errores. La idea principal de la ortogonalización consiste en hacer independientes los errores entre ecuaciones.

2.6. Transformada de Fourier.

Como se señaló anteriormente, las series de Fourier representan funciones definidas en un intervalo de la recta o, equivalentemente, funciones periódicas en la recta. De forma que para poder representar funciones definidas en toda la recta y no periódicas, se sustituye por la integral de Fourier. Al igual que las series de Fourier en el caso de funciones periódicas, la transformada de Fourier realiza una descomposición o análisis de f en componentes; ahora en lugar de presentar sólo frecuencias discretas formando una sucesión, aparece un rango continuo de frecuencias. Los casos de mayor interés con respecto a la transformada de Fourier, ocurre cuando no es una función familiar

$$F(w) = \int_{-\infty}^{\infty} f(t)e^{-iwt} dt, \quad (2.80)$$

la cual no es posible evaluarla en forma cerrada. Por ello, para poder obtener una buena aproximación a $F(w)$ se define, para $a < 0$ y $b > 0$ suficientemente grandes, lo siguiente:

$$t_0 = a < t_1 < \dots < t_{N-1} = b, a < 0, b > 0$$

$$\Delta t = \frac{b-a}{N} \text{ and } t_k = a + k\Delta t, \text{ with } k = 0, 1, \dots, N.$$

Entonces, la aproximación $\phi(w)$ para $F(w)$ esta dada por:

$$\phi(w) = e^{-iwa} \sum_{k=0}^{N-1} f(t_k) e^{-iwk(b-a)/N} \Delta t. \quad (2.81)$$

También, se toma el tiempo de duración $[a, b]$ centrado la atención en los puntos $w_n = \frac{2\pi n}{b-a}$, donde n es un entero. De tal manera que:

$$\phi(w) = e^{-iaw_n} \sum_{k=0}^{N-1} f(t_k) e^{-i2\pi nk/N} \Delta t. \quad (2.82)$$

2.6.1. Transformada Rápida de Fourier y Transformada de Fourier Discreta (TFD).

Además de las aplicaciones de tipo matemático de las series y transformadas de Fourier, que se han encontrado a lo largo del capítulo, también tienen uso en diferentes campos como en Comunicaciones; en el análisis de sistemas, detección de señales, Mecánica aplicada; análisis de vibraciones. Pero en las aplicaciones se encuentra con el problema de que las funciones de que se dispone no vienen dadas por expresiones analíticas sino por colecciones de datos y, por otra parte, las integrales que se necesitan también deben calcularse por procedimientos de aproximación.

En otras palabras, la transformada de Fourier discreta (TFD) es la principal herramienta para realizar los cálculos en el análisis de Fourier. Sin embargo, de poco sirve un instrumento cuyo coste, tanto en el tiempo como en dinero, supere la necesidad del usuario. Por lo que se introduce el concepto de transformada rápida de Fourier el cual es un algoritmo de cálculo que permite reducir considerablemente el número de operaciones necesarias para calcular una transformada de Fourier discreta. Las funciones que se consideran están definidas en el conjunto $1, 2, \dots, N - 1$ y tomarán valores complejos. Esto significa que tendremos un valor de C^N . Es conveniente considerar que el argumento de la función está definido por módulo N , es decir, que en realidad se tiene una sucesión $Df(n)$ definida para $n \in Z$, periódica de periodo N . Por tanto, la transformada de Fourier discreta está definida por:

$$Df(n) = \sum_{k=0}^{N-1} f(t_k)w^{-nk}, \quad \text{donde} \quad w = e^{i2\pi/N}. \quad (2.83)$$

Suponga ahora que se empieza con la siguiente secuencia de N números complejos para $f(t_k)$

$$f = (f(0), f(1), \dots, f(N-1)) \in C^N,$$

y asuma que se ha encontrado el conjunto de los N números complejos $Df(n)$. Entonces, cuando n toma los valores de $0 \leq n \leq N-1$, en la última ecuación se generan los números $Df(0), \dots, Df(N-1)$. Cuando n toma valores enteros fuera del intervalo, los exponentes dentro de la suma de la señal se reciclan debido a la circularidad, y simplemente se repiten con periodo N . Por lo tanto, los valores de $Df(n)$ deben de ser repeticiones periódicas de la serie original y por lo que la última expresión se considera un función periódica inherente de n con periodo N .

Cabe señalar que la TFD puede ser expresada también de forma matricial de la siguiente manera: defínase a f y Df como vectores tales que:

$$f = \begin{pmatrix} f(0) \\ \vdots \\ f(N-1) \end{pmatrix} \quad y \quad Df = \begin{pmatrix} Df(0) \\ \vdots \\ Df(N-1) \end{pmatrix}$$

con:

$$M_N = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{-1*2\pi i/N} & e^{-2*2\pi i/N} & \dots & e^{-(N-1)*2\pi i/N} \\ 1 & e^{-2*2\pi i/N} & e^{-2*2*2\pi i/N} & \dots & e^{-2(N-1)*2\pi i/N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-(N-1)*2\pi i/N} & e^{-2*(N-1)*2\pi i/N} & \dots & e^{-(N-1)^2*2\pi i/N} \end{pmatrix}$$

entonces $Df = M_N f$. M_N también es llamado el N -ésimo orden de la matriz DFT.

Asimismo, si $w = e^{2\pi j/N}$ se tiene que:

$$M_N = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \bar{w} & \bar{w}^2 & \dots & \bar{w}^{N-1} \\ 1 & \bar{w}^2 & \bar{w}^4 & \dots & \bar{w}^{N-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{w}^{N-1} & \bar{w}^{N-2} & \dots & \bar{w} \end{pmatrix}$$

2.6.2. Propiedades de la TFD.

A) Traslación en tiempo \mapsto multiplicación exponencial $f(n-p) \mapsto Df(n)e^{-2\pi ip/N}$.

Sea $g(n) = f(n-p)$, entonces

$$Dg(n) = \sum_{k=0}^{N-1} g(k)e^{-2\pi ikn/N} = \sum_{k=0}^{N-1} f(k-p)e^{-2\pi ikn/N}. \quad (2.84)$$

Tomando a $m = k-p$ esta última suma se convierte en $\sum_{m=-p}^{N-1-p} f(m)e^{-2\pi i(m+p)n/N}$.

Ahora, si se divide la suma $\sum_{m=-p}^{N-1-p}$ en

$$\sum_{k=0}^{N-1} = \sum_{m=-p}^{-1} + \sum_{m=0}^{N-1-p},$$

y además, puesto que f es N -periódica, podemos escribir lo siguiente:

$$\sum_{m=-p}^{-1} f(m)e^{-2\pi i(m+p)n/N} = \sum_{m=N-p}^{N-1} f(m)e^{-2\pi i(m+p)n/N}.$$

Por lo tanto, tomando en cuenta todo lo anterior, podemos concluir que:

$$\begin{aligned} Dg(n) &= \sum_{m=-p}^{-1} f(m)e^{-2\pi i(m+p)n/N} + \sum_{m=0}^{N-1-p} f(m)e^{-2\pi i(m+p)n/N} \\ &= \sum_{m=N-p}^{N-1} f(m)e^{-2\pi i(m+p)n/N} + \sum_{m=0}^{N-1-p} f(m)e^{-2\pi i(m+p)n/N} \\ &= \left(\sum_{m=0}^{N-1} f(m)e^{-2\pi imn/N} \right) e^{-2\pi ipn/N} = Df(n)e^{-2\pi ipn/N} \quad (2.85) \end{aligned}$$

Lo que se quiere dar a entender en el inciso anterior es, lo que le pasaría a la TFD de una función, si esta sufriera una demora o se moviera en el tiempo p pasos.

B) Traslación en Frecuencia $f(n)e^{i2\pi pn/N} \mapsto Df(n-p)$

Se tiene que la TFD de $f(n)e^{i2\pi pn/N}$ es:

$$\sum_{k=0}^{N-1} f(k)e^{2\pi ipk/N} e^{-2\pi ikn/N} = Df(n-p). \quad (2.86)$$

C) Modulación $f(n)\cos 2\pi in/N \mapsto \frac{1}{2}[Df(n-p) + Df(n+p)]$

Si $f(n)\cos 2\pi in/N = \frac{1}{2} = f(n)e^{2\pi ipn/N} + f(n)e^{-2\pi ipn/N}$, se tiene que la TFD de ésta última, utilizando la propiedad B) es de la forma:

$$\frac{1}{2}Df(n-p) + \frac{1}{2}Df(n+p). \quad (2.87)$$

2.6.3. Transformada de Fourier discreta inversa (TFDI).

Suponga que se tiene una función discreta $f(n)$ la cual se desea representar como una combinación lineal de exponenciales complejos discretos de orden N . Sea

$$\begin{aligned} \frac{1}{N} \sum_{k=0}^{N-1} Df(k)e^{2\pi ikn/N} &= \frac{1}{N} \sum_{k=0}^{N-1} Df(k)w^{kn} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left(\sum_{p=0}^{N-1} f(p)w^{-kp} \right) w^{nk} \\ &= \frac{1}{N} \sum_{p=0}^{N-1} f(p) \left(\sum_{k=0}^{N-1} w^{(n-p)k} \right) \end{aligned} \quad (2.88)$$

Si $n = p$, entonces $\sum_{k=0}^{N-1} w^{(n-p)k} = N$, mientras que si $n \neq p$, se tiene que $\sum_{n=0}^{N-1} w^{(k-p)n} = 0$. Por lo tanto:

$$\frac{1}{N} \sum_{k=0}^{N-1} Df(k) e^{2\pi i k n / N} = \frac{1}{N} f(n) N = f(n). \quad (2.89)$$

Finalmente,

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} Df(k) e^{2\pi i k n / N} = \frac{1}{N} \sum_{k=0}^{N-1} Df(k) w^{kn}, \quad (2.90)$$

la cual es definida como la TFDI.

A continuación se presenta otra forma de obtener la TFDI.

Haciendo referencia al álgebra lineal, cuando un conjunto de vectores de la forma $S = \{u_1, \dots, u_n\}$ satisfacen la condición de que $(u_n, u_m) = 0$ si $n \neq m$, se dice que los elementos de S son ortonormales y los cuales forman una base ortonormal para el espacio. Ahora, siguiendo con la misma idea, tómesese las funciones $\{w_p / \sqrt{N} : p = 0, 1, \dots, N - 1\}$ las cuales forman una base ortonormal. Para hacer ver de una forma mas clara, que las funciones anteriores son ortogonales, tómesese el producto interno definido como:

$$\left\langle \frac{1}{\sqrt{N}} w_p, \frac{1}{\sqrt{N}} w_k \right\rangle = \frac{1}{N} \sum_{n=0}^{N-1} w_p(n) \overline{w_k(n)} = \frac{1}{N} \sum_{n=0}^{N-1} w^{(k-p)n}, \quad (2.91)$$

en donde si $k = p$ tenemos que $(1/N) \sum_{n=0}^{N-1} w^{(k-p)n} = 1$ y por el otro lado, si $k \neq i$ obtenemos que $(1/N) \sum_{n=0}^{N-1} w^{(k-p)n} = 0$.

Ahora, puesto que $\{w_p / \sqrt{N} : p = 0, 1, \dots, N - 1\}$ es una base ortonormal, también lo es el conjunto de los conjugados $\{\overline{w_p} / \sqrt{N} : p = 0, 1, \dots, N - 1\}$.

Por otro lado, tomando en cuenta que cualquier función f puede ser escrita de la forma:

$$f = \sum_{p=0}^{N-1} \left\langle f, \frac{1}{N} \overline{w_p} \right\rangle \frac{1}{N} w_p \quad (2.92)$$

y dado que:

$$\begin{aligned}
 \left\langle f, \frac{1}{N} \overline{w_p} \right\rangle &= \sum_{m=0}^{N-1} f(m) \frac{1}{N} w_p(m) \\
 &= \sum_{m=0}^{N-1} f(m) \frac{1}{N} w^{-mp} \\
 &= \sum_{m=0}^{N-1} f(m) \frac{1}{N} e^{-2\pi i m p / N} \quad (2.93)
 \end{aligned}$$

se deduce lo siguiente:

$$\left\langle f, \frac{1}{N} \overline{w_p} \right\rangle = \frac{1}{N} Df(p), \quad p = 0, 1, \dots, N-1 \quad (2.94)$$

en donde $Df(p)$ es \sqrt{N} veces el coeficiente $\left\langle f, \frac{1}{N} \overline{w_p} \right\rangle$. Por lo tanto, f puede ser escrita como una combinación lineal en la cual los escalares son múltiplos de $Df(p)$.

Análogamente para la expansión de la serie de Fourier de una función f ,

$$f(t) = \sum_{n \in \mathbb{Z}} \left\langle f, \frac{e^{int}}{\sqrt{2\pi}} \right\rangle \frac{e^{int}}{\sqrt{2\pi}} = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{int}, \quad (2.95)$$

donde $\hat{f} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt$ es el habitual coeficiente de Fourier de f .

Ahora, tomando en cuenta que $Df(n) = \sum_{k=0}^{N-1} f(k) e^{-2\pi i n k / N}$, con $w_k(n) = e^{-2\pi i k n / N}$, es posible escribir esta última expresión de la forma:

$$\begin{aligned}
 Df(n) &= \sum_{k=0}^{N-1} f(k) w_n(k) \\
 &= \sqrt{N} \sum_{k=0}^{N-1} f(k) \frac{w_n(k)}{\sqrt{N}}. \quad (2.96)
 \end{aligned}$$

En consecuencia, puesto que $\left\{ w_p / \sqrt{N} : p = 0, 1, \dots, N-1 \right\}$ es una base ortonormal y utilizando lo mostrado anteriormente, se tiene:

$$\left\langle Df, w_n / \sqrt{N} \right\rangle = \sqrt{N} f(n) \quad (2.97)$$

lo cual puede ser descrito de la forma

$$\begin{aligned}
 f(n) &= \frac{1}{N} \langle Df, w_n \rangle \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} Df(k) \overline{w_n k} \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} Df(k) e^{2\pi i kn/N}, \tag{2.98}
 \end{aligned}$$

que representa la TDFI.

2.6.4. Convolución cíclica.

La convolución es un concepto que tiene significado en diferentes situaciones físicas. Por citar un ejemplo, ésta describe el proceso que causa la difusión o la pérdida de nitidez de las señales obtenidas en el proceso. En términos matemáticos la función que da la respuesta como función de la señal de entrada y de la perturbación, se denomina convolución y cuya expresión viene dada por la ecuación:

$$f * g(k) = \sum_{p=0}^{N-1} f(p)g(k-p), \tag{2.99}$$

en donde f y g son funciones definidas. La convolución es una de las propiedades intrínsecas de la TFD, la cual establece que la convolución en el dominio del tiempo es equivalente a la multiplicación en el dominio de la frecuencia y viceversa.

Las propiedades básicas de la convolución cíclica son:

I Convoluciones básicas

Sean f, g, h funciones definidas.

1. $f * g = g * f$

2. $f * (g * h) = (f * g) * h$
3. $\alpha(f * g) = (\alpha f) * (\alpha g) \forall \alpha$, con α constante.
4. $f * (g + h) = f * g + f * h$

II Productos de convoluciones

Para f, g funciones definidas, la DFT $D(f * g)(n) = Df(n)Dg(n)$.

$$\begin{aligned}
 D(f * g)(n) &= \sum_{k=0}^{N-1} f * g(k) e^{-2\pi i kn/N} = \sum_{k=0}^{N-1} f * g(k) \bar{w}^{kn} \\
 &= \sum_{k=0}^{N-1} \left(\sum_{p=0}^{N-1} f(p) g(k-p) \right) \bar{w}^{kn} \\
 &= \sum_{k=0}^{N-1} f(p) \left(\sum_{k=0}^{N-1} g(k-p) \bar{w}^{kn} \right) \\
 &= \sum_{k=0}^{N-1} f(p) \bar{w}^{in} \left(\sum_{k=0}^{N-1} g(k-p) \bar{w}^{(k-p)n} \right). \\
 &= Df(n)Dg(n)
 \end{aligned}$$

2.6.5. Transformada Rápida de Fourier para $N = 2^k$.

El cálculo de una transformada de Fourier discreta, requiere N^2 multiplicaciones. Si N es par se puede organizar el cálculo de modo que se realicen dos transformadas de Fourier discretas en $C^{N/2}$ y una combinación adecuada de ellas. Esto reduce el número de multiplicaciones a aproximadamente la mitad. Se puede aplicar sucesivamente el mismo proceso y al final el número de multiplicaciones se reduce considerablemente. Ésta es la idea de la transformada rápida de Fourier.

Por otro lado, cualquier función, ya sea par o impar, puede ser representada o expresada como la suma de dos funciones componentes, de la cual

una es par y otra impar definidas como f_e y f_o respectivamente. Es decir, para $f \in C^N$ se tiene:

$$f_e = (f(0), f(2), \dots, f(N-2)) \quad (2.100)$$

y

$$f_o = (f(1), f(3), \dots, f(N-1)). \quad (2.101)$$

Continuando con lo anterior, lo que se quiere mostrar es que para $N = 2^k$ y $f \in C^N$, el número de multiplicaciones requeridas para calcular la Df es: $2N \log_2 N = 2^{k+1} \cdot k$.

Primeramente, para $k = 1$ y $f \in C^2$,

$$Df = M_f f = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f(0) \\ f(1) \end{pmatrix}$$

lo cual implica 4 multiplicaciones o $2 \cdot 2^1 \log_2 2$. Entonces, asumiendo que el resultado se cumple para $N = 2^{k-1}$, es decir, que $2 \cdot 2^{k-1} \log_2 2^{k-1}$ multiplicaciones son requeridas para $N = 2^{k-1}$, por tanto se quiere ver que se cumpla para $N = 2^k$. Tomando a q como variable, sea $w_q = e^{2\pi j/q}$, así que la TFD de f es:

$$Df(n) = \sum_{k=0}^{N-1} f(k) \bar{w}_N^{kn}, \quad n = 0, 1, 2, \dots, N-1. \quad (2.102)$$

Dividiendo k por 2, n por $N/2$ y escribiendo cada uno de esos en términos de un cociente tenemos que:

$$k = 2k_0 + k_1, \quad n = (N/2)n_0 + n_1.$$

De manera que k y n varían entre 0 y $N-1$, las posibilidades para los cocientes y el resto de la ecuación son $k_0 = 0, 1, \dots, N/2-1$, $k_1 = 0, 1$,

$n_0 = 0, 1$ y $n_1 = 0, \dots, N/2 - 1$. Por lo tanto,

$$\begin{aligned}\bar{w}_N^{kn} &= \bar{w}_N^{(2k_0+k_1)((N/2)n_0+n_1)} \\ &= \bar{w}_N^{Nk_0n_0} \bar{w}_N^{(N/2)k_1n_0} \bar{w}_N^{2k_0n_1} \bar{w}_N^{k_1n_1}.\end{aligned}$$

De modo que $\bar{w}_N^{Nk_0n_0} = 1$ (w_N o \bar{w}_N planteado para cualquier multiplo de N es 1), podemos reescribir la ecuación (2.102) como:

$$Df(n) = \bar{w}_N^{Nk_0n_0} \sum_{k_1=0}^1 \sum_{k_0=0}^{N/2-1} f(2k_0 + k_1) \bar{w}_N^{(N/2)k_1n_0} \bar{w}_N^{2k_0n_1} \bar{w}_N^{k_1n_1}, \quad (2.103)$$

o evaluando k_1 tanto para 0 y 1 tenemos que:

$$\begin{aligned}Df(n) &= \bar{w}_N^{Nk_0n_0} \sum_{k_0=0}^{N/2-1} f(2k_0) \bar{w}_N^{2k_0n_1} \\ &\quad + \bar{w}_N^{Nk_0n_0} \sum_{k_0=0}^{N/2-1} f(2k_0 + 1) \bar{w}_N^{(N/2)n_0} \bar{w}_N^{2k_0n_1} \bar{w}_N^{n_1}.\end{aligned} \quad (2.104)$$

Dado que $\bar{w}_N^2 = \bar{w}_{N/2}$, podemos reescribir la $\bar{w}_N^{2k_0n_1}$ en cada sumando como $\bar{w}_{N/2}^{k_0n_1}$ para obtener

$$\begin{aligned}Df(n) &= \bar{w}_N^{Nk_0n_0} \sum_{k_0=0}^{N/2-1} f(2k_0) \bar{w}_{N/2}^{k_0n_1} \\ &\quad + \bar{w}_N^{(N/2)n_0} \bar{w}_N^{n_1} \sum_{k_0=0}^{N/2-1} f(2k_0 + 1) \bar{w}_{N/2}^{k_0n_1}.\end{aligned} \quad (2.105)$$

Reemplazando $(N/2)n_0 + n_1$ por n , entonces $\bar{w}_N^{(N/2)n_0} \bar{w}_N^{n_1} = \bar{w}_N^n$. Esto simplifica la segunda suma:

$$\begin{aligned}Df(n) &= \bar{w}_N^{Nk_0n_0} \sum_{k_0=0}^{N/2-1} f(2k_0) \bar{w}_{N/2}^{k_0n_1} \\ &\quad + \bar{w}_N^n \sum_{k_0=0}^{N/2-1} f(2k_0 + 1) \bar{w}_{N/2}^{k_0n_1}.\end{aligned} \quad (2.106)$$

$n_0 = 0, 1, n_1 = 0, 1, \dots, N/2 - 1$. Los términos

$$\sum_{k_0=0}^{N/2-1} f(2k_0) \overline{w_{N/2}}^{k_0 n_1} \quad y \quad \sum_{k_0=0}^{N/2-1} f(2k_0 + 1) \overline{w_{N/2}}^{k_0 n_1},$$

son las filas de:

$$Df_e = M_{N/2} \begin{pmatrix} f(0) \\ f(2) \\ \vdots \\ f(N-2) \end{pmatrix} \quad y \quad Df_o = M_{N/2} \begin{pmatrix} f(1) \\ f(3) \\ \vdots \\ f(N-1) \end{pmatrix}$$

De esta manera, es posible reescribir la ecuación (2.102) como:

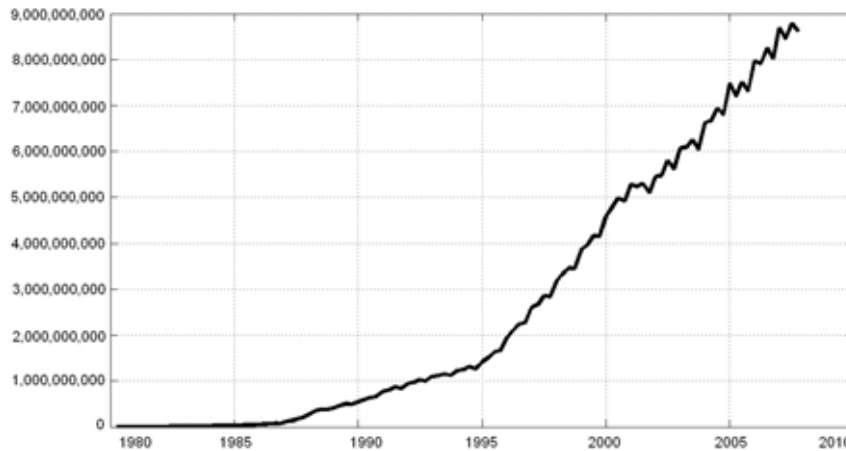
$$Df(n) = Df\left(\frac{N}{2}n_0 + n_1\right) = \overline{w_N}^{Nk_0 n_0} Df_e(n_1) + \overline{w_N}^n Df_o(n_1) \quad (2.107)$$

Análisis espectral del PIB, mediante el FFT.

Teniendo ya una noción sobre el análisis de Fourier, vistas a lo largo del capítulo, nos damos a la tarea de obtener el espectro de la serie del PIB mediante el uso de la FFT. Primeramente, graficamos dicha serie para verificar si existe algún tipo de comportamiento cíclico y de esta manera poder encontrar los componentes de la serie.

Gráfica 15. Ruido de la señal en el dominio del tiempo para el PIB 80.1-7.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Dada la naturaleza estocástica de la señal, no es posible determinar con absoluta precisión su densidad espectral de potencia, la cual es una medición en la energía en las distintas frecuencias. Evidentemente, es difícil identificar dichos componentes de la señal ya que al parecer no existe evidencia de comportamiento cíclico.

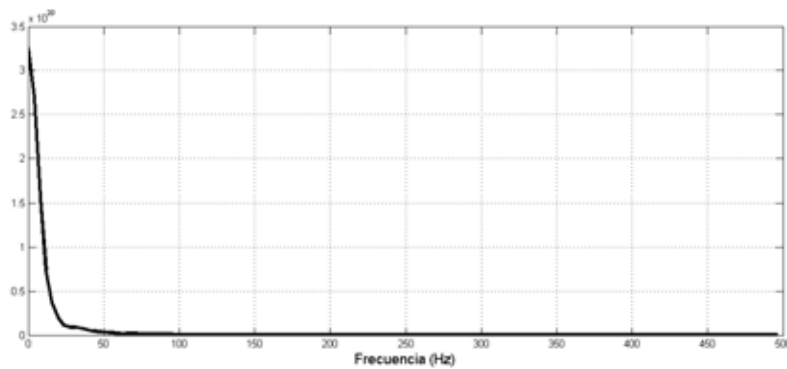
Por esta razón, se busca la aplicación de la TFD, mediante el uso de la FFT, sobre el intervalo temporal $0 < n < N-1$ siendo N la longitud de una

secuencia no periódica. Y por tanto, con ello es posible calcular su densidad espectral de potencia, el cual nos ayudará a encontrar dichos componentes.

Dicho lo anterior, se muestra la siguiente gráfica en donde se observa la densidad espectral de potencia del PIB una vez aplicado el FFT.

Gráfica 16. Densidad espectral de potencia de la serie del PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.

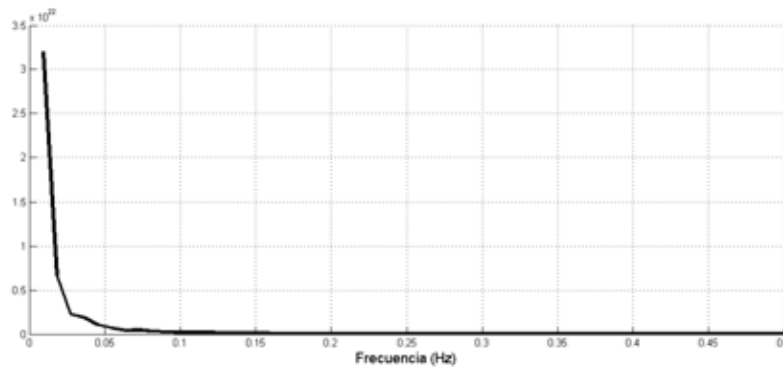


De acuerdo al resultado anterior, es posible darse cuenta que su espectro tiene caída de tipo hiperbólica, indicando la aleatoriedad de la señal. Además, se tiene que el único pico que tiene está muy cercano a cero, lo que significa que probablemente no exista ciclo alguno en esta señal, como se había mencionado anteriormente.

Realizando un acercamiento, de la Gráfica 16, y ahora analizando la Gráfica 17, tenemos que evidentemente la representación de Fourier no mostró evidencia de ciclos no-periódicos.

Gráfica 17. Densidad espectral de potencia de la serie del PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Lo que es evidente, es que el PIB no es estacionario, lo que nos daría elementos para investigar si la serie es de tipo $I(1)$ o $I(2)$. Los ciclos no se pegan a una medida promedio, es decir, los mecanismos que los generan pueden cambiar en cualquier punto del tiempo, por lo que los shocks estocásticos pueden ser relevantes durante algún periodo significativo.

Lo anterior implica que este proceso es caracterizado por ruido fraccionario en donde el corto y el largo plazo siguen la misma distribución de desviaciones respecto a la tendencia. Asimismo, se puede decir que en el corto plazo los determinantes del crecimiento del PIB poseen una distribución de frecuencias auto-similar mientras que en el largo plazo la serie presenta caos fuerte, lo que lo vuelve intrínsecamente impredecible.

Capítulo 3

Análisis Fractal.

3.1. Autosimilitud, autoafinidad y escalamiento.

Básicamente los fractales, se caracterizan por dos propiedades: autosemejanza (o autosimilitud) y autorreferencia. La autorreferencia determina que el propio objeto aparece en la definición de sí mismo, con lo que la forma de generar el fractal necesita algún tipo de algoritmo recurrente. La autosemejanza implica invarianza de escala, es decir, el objeto fractal presenta la misma apariencia independientemente del grado de ampliación con que lo miremos.

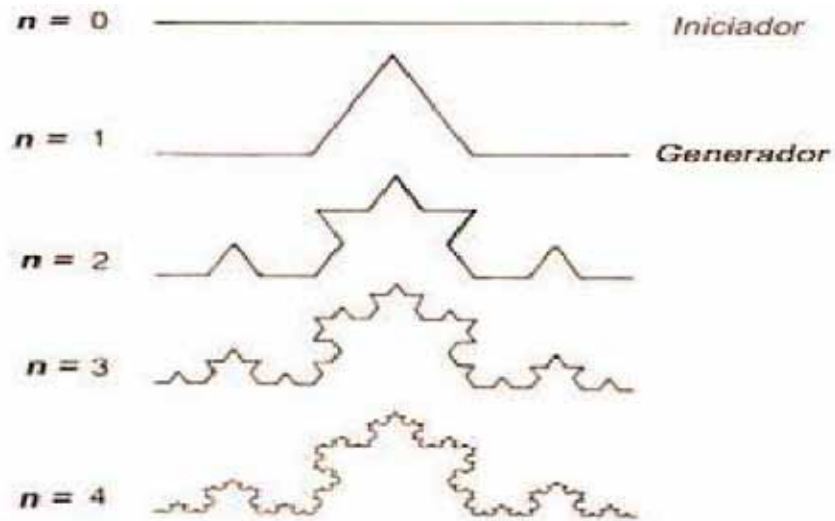
Por más que se amplíe cualquier zona de un fractal, siempre hay estructura, hasta el infinito y aparece muchas veces el objeto fractal inicial contenido en sí mismo. La geometría fractal aborda el estudio de formas geométricas no diferenciables o quebradas a cualquier escala que se miren. Es por ello que la geometría fractal ofrece un modelo alternativo que busca una regularidad en las relaciones entre un objeto y sus partes a diferentes

escalas.

Los fractales son interesantes, debido a que muchos fenómenos en la naturaleza son autosimilares e invariantes de escala.

Por otra parte, la generación de fractales autosimilares se puede lograr mediante experimentos sencillos como lo es el calidoscopio, o bien matemáticamente mediante procedimientos repetitivos conocidos como transformaciones autoafines o también utilizando los algoritmos. Para poder comprender esto de una manera mas clara, tómesese como ejemplo la siguiente figura:

Figura 1. Curva de von Koch



donde se parte de una recta, la cual se divide en tres segmentos iguales, el de en medio se substituye por un triángulo equilátero y resulta así una figura con cuatro segmentos de rectas. Este procedimiento se sigue repitiendo hasta el infinito obteniendo así un perfil que a mayor amplificación presenta un aspecto invariante. Por su característica de escalamiento, igual en todas

direcciones, podemos decir que estos fractales son isotrópicos autosimilares. Citando otro ejemplo, sobre partículas que exhiban un comportamiento, el cual genere una figura autosimilar, tenemos lo que se conoce como movimiento browniano¹.

Figura 2. Movimiento browniano



los cuales, sólo mantienen su aspecto durante el escalamiento si éste se hace diferente en al menos una de las direcciones, es decir, son fractales anisotrópicos los cuales se conocen como objetos o fractales autoafines.

¹Movimiento irregular de partículas microscópicas suspendidas en un líquido.

3.2. Dimensión fractal.

Uno de los aspectos más relevantes surgidos dentro de la teoría de fractales es la redefinición de el concepto de dimensión a cargo de Hausdorff, que permite que estas curvas matemáticas tengan dimensión fraccionaria. Entonces, se dice que la dimensión fractal es una forma de medir el grado de complejidad evaluando que tan rápido la longitud, superficie o volumen aumenta si es medido con respecto a escalas cada vez mas pequeñas. La idea fundamental es asumir que las cantidades (longitud, superficie, volumen y la escala) no varían arbitrariamente pero que sí están relacionadas con una ley, la cual nos permite calcular una cantidad de la otra. Dicha ley, la cual parece ser relevante, es una ley de potencia de la forma $y \propto x^d$. De tal manera que también resulta ser muy práctica para el debate de la dimensión. El concepto de dimensión no es muy fácil de comprender, debido a que los matemáticos han tenido cerca de diez diferentes tipos de conceptos entre los que destacan: dimensión topológica, dimensión autosimilar (dimensión de Hausdorff), dimensión fractal, dimensión conteo por cajas(dimensión completa), dimensión Euclidiana, dimensión informativa, dimensión punto clave, etc. Algunas de estas, sin embargo, toman sentido en algunas situaciones y otras no. Es por ello que se realizará el estudio solamente sobre las siguientes:

- De Hausdorff
- Conteo por cajas
- Por Medida Natural
- Informativa
- Punto clave

3.2.1. Dimensión de Hausdorff.

Aparentemente, para la línea, cuadrado y el cubo existe una ley de potencia, la cual relaciona el número de piezas a y el factor de reducción s . Dicha ley es la siguiente:

$$a = \frac{1}{s^H} \quad (3.1)$$

donde $H = 1$ para la línea, $H = 2$ para el cuadrado y $H = 3$ para el cubo. En otras palabras, el exponente para la ley de potencia concuerda exactamente con esos números los cuales son familiares como la dimensión (topológica) para la línea, cuadrado y cubo. Si observamos detalladamente la curva de Koch, la relación de $a = 4$ y $s = 1/3$ y $a = 16$ con $s = 1/9$ etc., no es muy evidente. Es decir, $4 = 3^H$, o equivalentemente $H = \frac{\log 4}{\log 3} \approx 1,2619$.

¿Pero, se obtendría lo mismo si tomamos piezas mas pequeñas, con un factor de reducción de $1/9$? Para poder probar esto, se propone lo siguiente. Tómese a $16 = 9^H$, de donde se calcula

$$H = \frac{\log 4^2}{\log 3^2} = \frac{2\log 4}{2\log 3} = \frac{2\log 4}{2\log 3} \approx 1,2619,$$

y como una regla general,

$$H = \frac{\log 4^k}{\log 3^k}.$$

Por consiguiente, la relación de la ley de potencia entre el número de piezas y el factor de reducción da el mismo número D . Este es el número H , un número entre 1 y 2, el cual llamamos la dimensión de Hausdorff ó dimensión autosimilar para la curva de Koch. Hablando en términos mas generales, dada una estructura autosimilar, existe una relación entre el factor de reducción s y el número de piezas a dentro de la cual la estructura puede ser

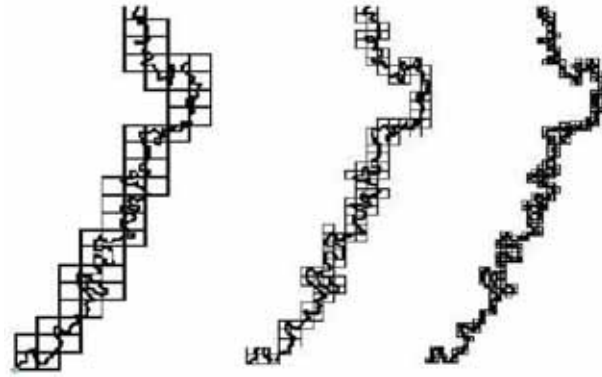
dividida, es decir,

$$a = \frac{1}{s^H} \quad \text{o equivalentemente} \quad H = \frac{\log a}{\log 1/s}.$$

3.2.2. Dimensión conteo por cajas.

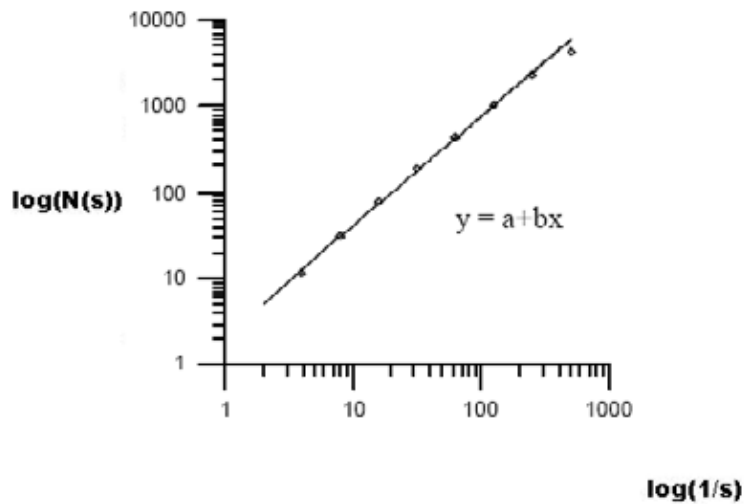
La dimensión de conteo por cajas, propone una medida sistemática, la cual aplica para cualquier estructura en el plano y la cual puede ser fácilmente adaptada para cualquier estructura en el espacio. La idea principal de dicha dimensión es muy simple, consiste en poner la estructura sobre una red o rejilla regular de tamaño s y simplemente contar el número de cuadros de la red los cuales contengan parte de la estructura, lo cual se denota con la letra $N(s)$.

Figura 3. Análisis mediante box-counting.



Entonces, haciendo el log/log diagrama², donde después se trata de ajustar una línea recta a los puntos graficados y de esa manera poder medir su pendiente D_b .

Gráfica 18. Cálculo de la dimensión fractal mediante box-counting.



Por tanto, el número $D_b = b$, el cual es la pendiente de la gráfica, es la dimensión conteo por cajas. Para fines prácticos, es conveniente considerar una secuencia de rejillas, en donde el tamaño de la rejilla es reducido por un factor de 2 llegando así a una secuencia de conteo de la forma $N(2^{-k})$, $k = 0, 1, 2, \dots$. De esta manera, se dice que la pendiente de la línea recta, correspondiente al diagrama log/log es:

$$\frac{\log N(2^{-k-1}) - \log N(2^{-k})}{\log 2^{k+1} - \log 2^k} = \log \frac{N(2^{-k-1})}{N(2^{-k})} \quad (3.2)$$

Sin embargo, dicha pendiente es un estimado de la dimensión fractal conteo por cajas. En otras palabras, si el número de cajas numeradas crece

²Más precisamente, trazar la medida en el $\log(N(s))/\log(1/s)$.

por un factor de 2^H cuando el tamaño de la caja es reducido a la mitad, entonces la dimensión fractal es igual a H .

3.2.3. Dimensión Informativa.

La dimensión infromativa, d_1 es una generalización de la capacidad de tomar en cuenta la probabilidad relativa de los cubos usados para cubrir el conjunto. La dimensión informativa esta dada por:

$$d_1 = \lim_{\epsilon \rightarrow 0} \frac{I(\epsilon)}{\log(1/\epsilon)}, \quad (3.3)$$

donde $I(\epsilon) = \sum_{i=1}^{N(\epsilon)} P_i \log \frac{1}{P_i}$, y P_i es la probabilidad contenida dentro del i -ésimo cubo. A su vez, si todos los cubos tienen igual probabilidad entonces $I(\epsilon) = \log N(\epsilon)$. Sin embargo, para probabilidades diferentes se tiene que $I(\epsilon) \leq \log N(\epsilon)$. En información teórica, la cantidad $I(\epsilon)$ es el monto de información necesaria para poder especificar el estado del sistema dentro de una ϵ , o equivalentemente, esta es la información obtenida en la medición la cual es incierta por un monto ϵ .

3.2.4. Dimensión Punto clave.

La dimensión punto clave d_p es el exponente con el cual la probabilidad total contenida en una bola decrece al igual que el radio de la bola decrece. Para poder hacer esta idea mas precisa, sea μ el cual denota la probabilidad natural medida en el atractor, y sea $B_\epsilon(x)$ el cual denota una bola de radio ϵ centrado aproximadamente un punto x en el atractor y $\mu(B_\epsilon(x)) \sim \epsilon^{d_p}$. Es decir, se define esta dimensión como:

$$d_p(x) = \lim_{\epsilon \rightarrow 0} \frac{\log \mu(B_\epsilon(x))}{\log \epsilon} \quad (3.4)$$

Si $d_p(x)$ es independiente de x para casi toda x con respecto a la medida μ , llamamos $d_p(x) = d_p$ la dimensión punto clave.

3.3. Exponente de Hurst.

El análisis de Hurst, fue desarrollado y dirigido al estudio de series temporales discretas durante sus investigaciones sobre las crecidas del Nilo. Hoy día se aplica en muchos campos por ejemplo en el análisis de las series temporales de variación de cotización en las bolsas, evolución de poblaciones y siempre que interese comprobar la persistencia de una determinada información cuando el sistema al que se refiere evoluciona con el tiempo, etc.

Su desarrollo matemático puede resumirse en los siguientes pasos:

1. Descomponer la serie en un conjunto de intervalos de una amplitud temporal determinada, τ , calcular la media para cada intervalo y las desviaciones acumuladas de esta media.
2. Hallar el intervalo de variabilidad de esta desviación como máximo menos mínimo y normalizarlo dividiendo por la desviación típica. A esta amplitud normalizada se la suele llamar rango reescalado. A cada valor de τ se asigna el valor promedio de $R(\tau)/S(\tau)$ sobre todos los subintervalos que determina. Se elige entonces otro τ y se procede de forma semejante, de modo que se obtiene un conjunto de valores τ y sus $R(\tau)/S(\tau)$ asociados.

El exponente de Hurst, H , es un parámetro asociado con las series de tiempo que provee información importante con respecto al color del ruido y a la predictibilidad de un fenómeno registrado.

Es decir, expresando matemáticamente lo anterior, dada una secuencia temporal de observaciones x_t se define la serie de la siguiente forma:

$$X(t, \tau) = \sum_{u=1}^t (x_u - \bar{x}_\tau), \quad (3.5)$$

en donde:

$$\bar{x}_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} x_t,$$

y por lo tanto, se dice que $X(t, \tau)$ es la medida de la suma de las diferencias de las observaciones del tiempo 1 a t en comparación con el promedio de las primeras τ observaciones.

Por el otro lado, el rango auto ajustado $R(\tau)$ está definido como:

$$R(\tau) = \max_{t=1}^{\tau} X(t, \tau) - \min_{t=1}^{\tau} X(t, \tau), \quad (3.6)$$

a su vez, la desviación estándar es de la forma:

$$S(\tau) = \sqrt{\left(\frac{1}{\tau} \sum_{t=1}^{\tau} (x_t - \bar{x}_\tau)^2\right)}.$$

Finalmente, se tiene lo que se conoce como el rango reescalado descrito como:

$$R(\tau)/S(\tau) = \frac{\max_{t=1}^{\tau} X(t, \tau) - \min_{t=1}^{\tau} X(t, \tau)}{\sqrt{\left(\frac{1}{\tau} \sum_{t=1}^{\tau} (x_t - \bar{x}_\tau)^2\right)}}. \quad (3.7)$$

Básicamente, se compara la mayor cantidad de cambios que ocurre en los términos iniciales τ a lo que sería esperado dada su varianza.

La conducta asintótica de $R(\tau)/S(\tau)$ se reduce a

$$R(\tau)/S(\tau) = ((\pi/2)\tau)^{1/2},$$

para cualquier proceso independiente con varianza finita. Entonces, una vez definido el rango reescalado se define el parámetro H o exponente de Hurst como:

$$H = \frac{d[\log(R(\tau)/S(\tau))]}{d[\log(\tau)]}. \quad (3.8)$$

Otra forma de calcular el exponente es mediante la representación gráfica de $\log(R(\tau)/S(\tau))$ frente a $\log(\tau)$, donde se espera obtener una línea cuya pendiente determina el exponente de Hurst. A esta representación gráfica se la llama “diagrama de Pox”.

Para finalizar con lo que se refiere al exponente de Hurst, se mencionan los valores que puede tomar sin dejar atrás su interpretación. Es decir, cuando $H = 1/2$ indica aleatoriedad de la señal, es el valor típico de un movimiento browniano, mientras que valores de $H > 1/2$ son indicativos de persistencia, es decir, la tendencia de la serie temporal es a continuar creciendo o decreciendo, esta persistencia será mayor cuanto más próximo a 1 sea el valor de H . Por el otro lado, $H < 1/2$ indica antipersistencia, es decir, la tendencia es en cierto sentido contraria a sí misma. En este caso las curvas tienden a oscilar de modo más errático, se trata de un sistema ergódico.

Por lo tanto podemos afirmar que el exponente de Hurst es una medida de la información que la señal conserva al cabo de un cierto tiempo.

3.4. Movimiento browniano.

El movimiento browniano, caracterizado por la independencia y normalidad de la distribución de sus incrementos, fue uno de los primeros modelos utilizados para describir la evolución de una serie temporal como un proceso estocástico, ya que apareció por primera vez asociado a los precios de una acción aunque los datos empíricos no se ajustaban bien del todo al modelo, ya que los incrementos de los precios de la mayoría de las acciones presentan cierta dependencia y su distribución empírica difiere de la distribución normal.

Por otro lado, del modelo de Einstein en donde estableció una predicción cuantitativa: el desplazamiento cuadrático medio, de la partícula browniana sigue dos comportamientos respecto del tiempo.

En la primera fase se ve una curva parabólica, mostrando un movimiento de partícula libre. Al cabo de un tiempo breve, esta fase culmina y comienza la segunda, donde puede observarse una recta, estableciendo un comportamiento difusivo, regida por la siguiente expresión:

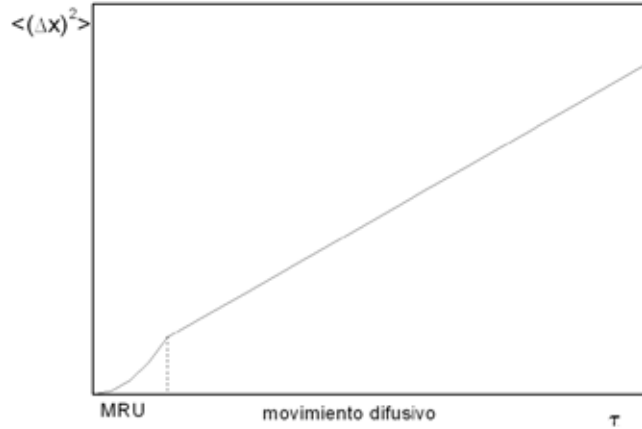
$$\langle (\Delta x)^2 \rangle = D * \tau, \quad (3.9)$$

con lo cual:

$$|\Delta x| = k * \tau^{0,5}, \quad (3.10)$$

donde $\langle(\Delta x)^2\rangle$ es el desplazamiento cuadrático medio, τ el tiempo, D el coeficiente de difusión³ y k la raíz cuadrada de éste.

Gráfica 19. Desplazamiento cuadrático medio, de la partícula browniana.



Einstein probó que la pendiente de la recta (y, por lo tanto, el coeficiente de difusión) era función de la viscosidad del medio, la temperatura termodinámica, las dimensiones de la partícula y el número de Avogadro. Es decir que, según su modelo, la partícula browniana ejecuta un movimiento libre - un Movimiento Rectilíneo Uniforme (MRU) - hasta que comienza a chocar contra las moléculas de agua.

³Se define como la facilidad con que cada soluto en particular se mueve en un solvente determinado.

3.4.1. Aplicación del rango reescalado sobre la caminata aleatoria.

Sea ξ_i una serie temporal, en nuestro caso ξ_i es un número aleatorio que representa el paso i caminado. Supongamos que el caminante ha llevado a cabo τ pasos en total. En promedio calculado sobre los τ pasos-, este caminante efectuará pasos de longitud $\langle \xi \rangle_\tau$, con lo que cada paso se habrá apartado del promedio en una cantidad $\xi_i - \langle \xi \rangle_\tau$ que podemos denominar desvío. Entonces, la cantidad $X(t, \tau) = \sum_{i=1}^t (\xi_i - \langle \xi \rangle_\tau)$ será la acumulación, hasta el tiempo t , de los desvíos.

Esta cantidad muestra cuántos desvíos recorrió el caminante hasta el tiempo t . Gracias a esto, se puede transformar los datos ξ_i , la sucesión de pasos caminados - en $X(t, \tau)$ la sucesión de desvíos acumulados, con mucha facilidad. Es posible estimar la dispersión de estos desvíos acumulados con el rango R de los datos transformados:

$$R(t) = \max X(t, \tau) - \min X(t, \tau).$$

Finalmente, si se normaliza $R(\tau)$, dividiendo por la desviación estándar de los datos $S(\tau)$, se encuentra una magnitud que permite estimar la dispersión de los datos. Como explícitamente señaló anteriormente, $R(\tau)/S(\tau)$ es, en principio, función de τ ; es decir, depende del número total de pasos dados.

En particular, en nuestro caso: $R(\tau)/S(\tau) = \tau^H$, de donde es posible concluir que $H = 1/2$.

El hecho de que R/S siga una relación potencial con τ no es trivial. Recordemos que τ es el largo total de la caminata. Esto significa que, sin

importar cuán extensa sea ésta, R/S (una medida de la dispersión de la caminata) sigue siempre la misma relación con el largo total de aquella. Es decir, R/S sigue una ley de escala con τ . Pero, esto significa que, si se observa un gráfico de R/S en función de τ , con distintos rangos de valores de τ , se tiene el mismo aspecto. Dicho de otro modo, si se evalúa un gráfico $R/S = f(t)$ a través de distintas resoluciones diferentes aumentos advertiremos que el perfil del gráfico es similar: el todo se parece a cada una de sus partes. En otras palabras, la figura muestra autoafinidad.

¿Qué significa esto y qué vínculo tiene con el exponente de Hurst?

La relación que existe entre H y D_f (dimensión fractal) es la siguiente:

$$D_f = 2 - H. \quad (3.11)$$

3.5. Ruido de Colores.

En 1666, Isaac Newton rompió la blancura de la luz blanca al hacer pasar un rayo de sol por un prisma. ¿Se imaginan que impacto debió causar este nuevo descubrimiento en el que la luz blanca, tan homogénea y pura, en realidad estaba constituida por una hermosa gama de colores que eran exactamente los mismos que los del arcoiris? Es decir, el rayo que cruza el prisma se descompone en los siguientes colores: rojo, amarillo, anaranjado, violeta, azul, verde y añil el cual es conocido como el espectro⁴.

Además, se tiene que toda onda u oscilación se caracteriza por tener una frecuencia, amplitud y longitud de onda, de esta forma es posible decir

⁴spectrum en latín quiere decir imagen.

que cada uno de los colores que componen el espectro corresponde a una frecuencia de oscilaciones electromagnéticas o a una longitud de onda.

Una gráfica de la frecuencia, en el eje de las abscisas contra la intensidad de cada una de las frecuencias en el eje vertical, es llamada *espectro* o *densidad espectral*, es útil ya que permite saber las características de la fuente luminosa. Por ejemplo, si se graficára el espectro solar, veríamos de forma clara que tiene su máximo cerca de la longitud de onda 600nm, lo que corresponde al color amarillo, pero al amanecer y en el ocaso, los rayos solares tienen que atravesar una capa más extensa de aire, pues entran de sesgo a la atmósfera y ésta absorbe la mayoría de las longitudes de onda, dejándonos el color rojo.

Por otra parte, también se tiene que los sonidos son ondas, es decir, vibraciones pero que se propagan longitudinalmente, las cuales requieren algún medio material como el agua o aire para viajar. Cada una de las notas musicales corresponde a una frecuencia bien determinada y esto nos permite hablar de un espectro acústico. Así, por ejemplo, 440 oscilaciones por segundo dan la nota La, la cual curiosamente es la que se escucha como tono de invitación a marcar cuando uno descuelga el teléfono.

Por otro lado, se nos enseña en la escuela que el ruido es una combinación desordenada e incoherente de sonidos, sin regularidades o periodicidades, en acústica es cualquier sonido indeseable o que interfiere con otros sonidos, pero se sabe que hay de ruidos a ruidos. Además, al igual que luz, es posible representar en una gráfica espectral, en donde se dice que un ruido es ruido blanco si su espectro es una constante, es decir, que todas las frecuencias

tienen la misma intensidad.

Entonces, se puede tomar una serie temporal como una onda y así poder representarla en su espectro correspondiente para un mejor estudio. Por ejemplo, si se analiza la gráfica del índice de valores y precios de la Bolsa, pudiéramos hablar de una jornada roja si las fluctuaciones de un día del índice de precios y valores predominan, en número, las pequeñas sobre las grandes, y eso se vería en una gráfica espectral como una curva que desciende de valores grandes en las frecuencias bajas a pequeñas en las frecuencias altas.

Continuando con el estudio de las series temporales, se ha observado que dichas series ya sean físicas, biológicas, económicas, entre las que destacan las lluvias, número de manchas solares, las frecuencias de terremoto, frecuencia de las fluctuaciones en las oscilaciones eléctricas, coinciden con un espectro de frecuencia de la forma $1/f^\alpha$, donde f denota la frecuencia y $\alpha > 0$. El hecho de que, $1/f$, el ruido es encontrado en una amplia variedad de sistemas, ha dado lugar a que podrían existir algunos mecanismos que se basen en la producción de $1/f$ ruido.

Ahora, si se toma una serie de números aleatorias y recordando lo visto en el segundo capítulo, en el que la transformada de Fourier nos permite calcular su espectro, es posible observar que éste disminuye como una hipérbola con $\alpha = 2$, es decir, de la forma $1/f^2$ conocido como ruido café o movimiento browniano. Tomando otro ejemplo, se define el ruido negro cuyo espectro es igualmente una hipérbola con $\alpha = 3$, el cual refleja un gran dominio en de las frecuencias bajas sobre las altas.

Para proporcionar un marco de referencia, es conveniente considerar dos casos, $\alpha = 0$, llamado ruido blanco y $\alpha = 2$ llamado movimiento browniano, vistos anteriormente

Una manera de estimar α , es mediante el método llamado el exponente de Hurst, en el cual, para la serie de tiempo $1/f^\alpha$ la relación arroja que $\alpha = 2(2H - 1)$ y que la presente α será calculada de acuerdo a la fórmula anterior.

Recuerde que el exponente de Hurst, es un parámetro asociado con las series de tiempo que provee información importante con respecto al color del ruido y a la predictibilidad de un fenómeno registrado.

Ahora, considere un proceso aleatorio $X(t)$ definido en el segmento de tiempo $[0, T]$, donde dicho proceso representa diversos tipos de cambio, los tipos de interés y la temperatura por ejemplo. Entonces, para un proceso escalonado se tiene que sigue la ley de potencia:

$$P(f) \propto k^{-\alpha}. \quad (3.12)$$

Luego, si $\alpha < 1$, el proceso es estacionario en el sentido mas fuerte de la palabra, esto es, $X(t)$ es estadísticamente invariante en t , si $1 < \alpha < 3$, el proceso es no estacionario pero tiene incrementos estacionario. Introduciendo el exponente de Hurst H , un parámetro que describe el grado de estacionariedad de $X(t)$, es posible expresar el exponente α como sigue:

$$\alpha = 2H + 1, \quad (3.13)$$

o de acuerdo a Mandelbort, H está relacionado con el valor de la dimensión fractal D por:

$$D = 2 - H. \quad (3.14)$$

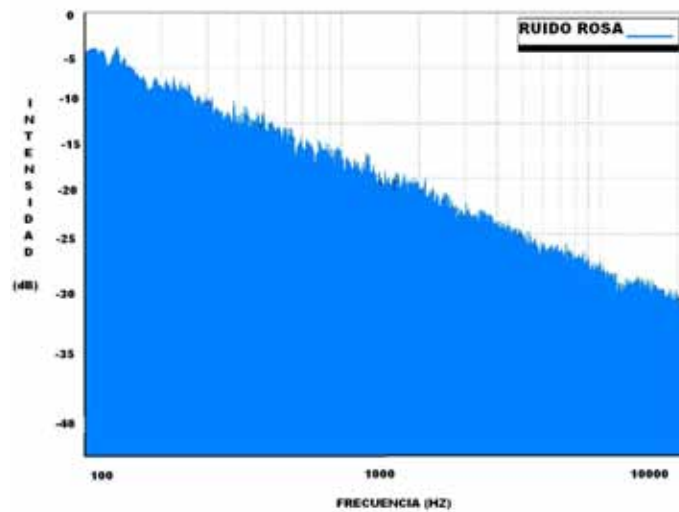
Se representa lo anterior de una forma resumida y clara como sigue:

$$H = \begin{cases} 0 & \text{para } \alpha < 1 \\ (\alpha - 1)/2 & \text{para } 1 \leq \alpha \leq 3 \\ 1 & \text{para } \alpha > 3 \end{cases}$$

Como se pudo observar, anteriormente, el espectro de potencia de el ruido blanco es independiente de la frecuencia, de modo que el coeficiente de la pendiente del espectro es cero.

Por otra parte, se introduce lo que se conoce como ruido rosa, el cual es una mezcla de rojo y blanco; contiene algo de blanco en el sentido de que todas las frecuencias se hallan representadas, y rojo puesto que las bajas frecuencias entran en mayor proporción que las altas, su espectro es una hipérbola de la forma $1/f$.

Gráfica 20. Ruido rosa.



Este tipo de ruido parece estar presente en toda la naturaleza, como por ejemplo en los disparos de las neuronas en el sistema nervioso central, la

variación de la luminosidad de las estrellas, aparece lo mismo en las fluctuaciones de la radiación solar que en las de la casa de bolsa, entre otros. Es decir, el ruido rosa es una manifestación común en una gran cantidad de fenómenos como los ya antes señalados.

Esto quiere decir que debe de ser independiente de la composición material de éstos y que, más bien, deberá depender de procesos que se dan entre los componentes de los sistemas y no de su naturaleza. Por otro lado, una de las razones de la impotencia del método reduccionista en el caso del ruido rosa es que los miembros de la familia de funciones $1/f^\alpha$ son invariantes ante cambios de escala. Por nombrar un ejemplo, se nombra lo que es la música fractal la cual es definida de la siguiente manera:

se dice que la familia de funciones $h(f) = \frac{1}{f^\alpha}$ tiene la propiedad de autosemejanza pues si se multiplica por un factor arbitrario, k , el argumento f de la función, entonces se tiene:

$$h(kf) = \frac{1}{(kf)^\alpha} = \frac{1}{k^\alpha f^\alpha} = \frac{1}{k^\alpha} h(f),$$

en donde multiplicar por k el eje de las abscisas equivale a estirarlo por ese factor.

La línea de igualdades previas dice que si se estira ahora el eje de las ordenadas por el factor $1/k^\alpha$, entonces la gráfica de la función $h(f)$ se ve exactamente igual a la de $h(kf)$.

Utilizando este hecho, supóngase que en el eje de las abscisas se tiene el tiempo y en el eje de las ordenadas representa una señal sonora, en este caso se toma como ejemplo la música, entonces si se pasa el disco al doble de ve-

locidad, es decir, con $k = 1/2$ basta con subir la música al doble $\frac{1}{k} = \frac{1}{1/2} = 2$ para que esta suene igual. Por lo tanto, las piezas musicales que tienen esta virtud se les llama canciones fractales.

Por otro lado, el físico Per Bak en el año de 1987, propuso el concepto de criticalidad autorganizada como un intento de explicar el ruido $1/f$. La autoorganización es la capacidad que tienen algunos sistemas lejos del equilibrio termodinámico de generar estructuras y patrones sin necesidad de la acción de agentes externos, es decir, son sistemas que pueden crear y mantener formas de manera espontánea.

Mientras que la criticalidad es una noción asociada a las transiciones de fase. Entonces, la criticalidad autorganizada es un término usado en física para describir (clases de) sistemas dinámicos que tienen puntos críticos (el punto crítico es aquel límite para el cual el volumen de un líquido es igual al de una masa igual de vapor o, dicho de otro modo, en el cual las densidades del líquido y del vapor son iguales.) como un atractor en su evolución temporal. Su comportamiento macroscópico exhibe invariancias de escala espaciales y temporales típicas de una transición de fase.

La propuesta de Per Bak es que los sistemas dinámicos formados por un número grande de componentes interactuando de manera no lineal, tienen la tendencia espontánea a autorganizarse en estados críticos de equilibrio dinámico en los cuales ocurren fluctuaciones de todos tamaños, pero siempre siguiendo leyes de distribución. Las cuales son conocidas como leyes de potencia, ya que la relación funcional entre la magnitud de las fluctuaciones y su abundancia relativa es de la forma $1/f^\alpha$.

Cabe destacar que existe una inmensidad de fenómenos naturales que siguen dicha ley $1/f^\alpha$ como la distribución de las magnitudes de los terremotos, las extinciones a lo largo de la historia de la vida en la tierra, la acumulación de dinero en nuestro país, etc., los cuales tienen $\alpha = 1$.

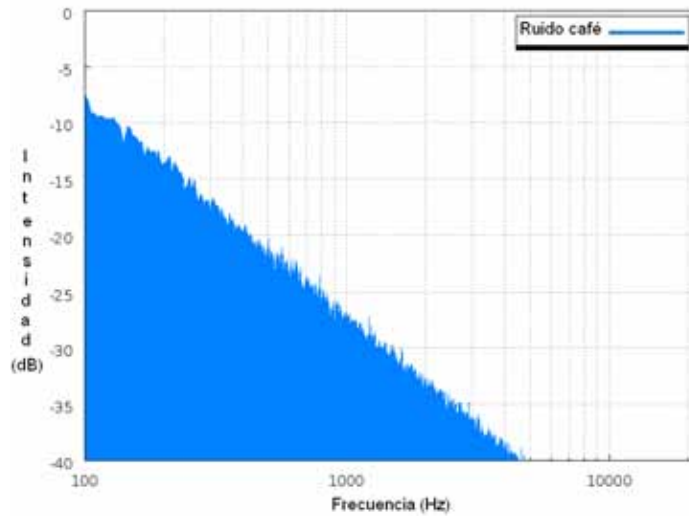
Para poder explicar, de una manera más clara lo definido ya anteriormente, se presenta el estudio de la serie del PIB, utilizando el análisis fractal mediante el índice H o coeficiente de Hurst, en el que, primeramente se debe encontrar el color del ruido que se asemeja a esta señal.

Continuando con lo anterior, al analizar la serie se muestra que el coeficiente toma el valor de $H = 0,8893$. Quizá, dicho coeficiente no sea de gran utilidad, ya que por no ser número entero es difícil poder contrastarlo con algún color conocido. Sin embargo, de lo que si es seguro es que esta señal tiene o se dice ser persistente debido a que es mayor a 0.5.

Por otro lado, utilizando la relación que existe entre el exponente de Hurst y el exponente α , se dice que su espectro de frecuencia es de la forma $1/f^\alpha$, donde $\alpha = 1,5572$.

Hasta aquí todavía no es posible definir el color de la serie del PIB, tomando el valor de α se observa que esta entre ruido rosa y el ruido marrón o ruido café.

Gráfica 21. Ruido café.

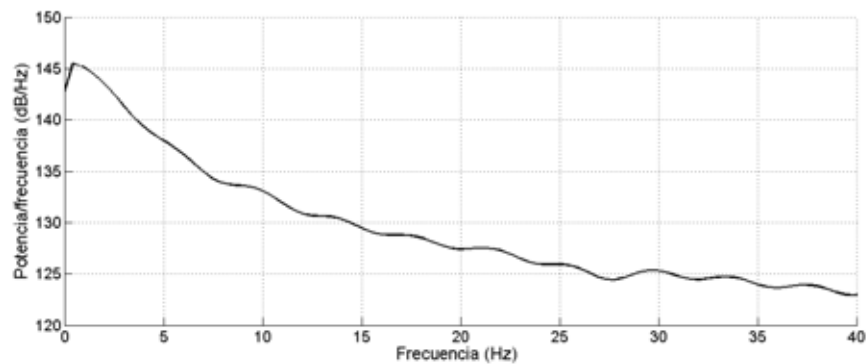


Haciendo la comparación de la Gráfica 21 con la del ruido rosa, se observa que esta decae rápidamente debido a que su densidad espectral de potencia es proporcional a $1/f^2$ o dicho de otra forma decae 6dB por octava a medida que subimos en frecuencia, mientras que el ruido rosa decae 3dB por octava.

Entonces, para poder definir con mayor claridad el color de la serie, se obtiene el periodograma el cual muestra la densidad espectral de potencia del PIB y que se representa de la siguiente forma:

Gráfica 22. Densidad espectral de potencia para el PIB 80.1-07.3.

Fuente: INEGI <http://www.inegi.gob.mx/inegi/default.aspx>.



Se observa que obedece una analogía con la luz blanca (que es una mezcla de todos los colores) que, después de ser coloreada de forma que se atenúen las frecuencias mas altas, resulta un predominio en las frecuencias bajas. Así pues, es posible determinar que la serie del PIB se asemeja más al ruido rosa, por la forma de decaimiento la cual se presenta como una curva que desciende de valores grandes en frecuencias bajas a pequeñas en frecuencias altas.

3.6. Movimiento Browniano Fraccionario.

Benoit Mandelbrot, al que se considera como el padre de la Geometría Fractal, propuso dos generalizaciones del movimiento browniano para modelizar la evolución de los precios de un activo financiero: los movimientos brownianos fraccionarios, que sólo tienen incrementos independientes en el modelo original y los movimientos L-estables⁵, con incrementos independientes y distribuidos según una distribución estable que generaliza a la normal. Ambas generalizaciones tienen diferencias esenciales, pero sus gráficos comparten dos características típicas de los conjuntos fractales que nos van a permitir estudiarlos de forma conjunta.

1. La autoafinidad estadística, que hace que los gráficos que se obtienen cuando se reduce la escala temporal para representarlos tenga una apariencia similar.
2. El valor no entero de su dimensión, que nos permitirá caracterizar el proceso al tener en cuenta que la dimensión está relacionada con las variaciones que los gráficos experimentan entre puntos próximos, de tal forma que cuanto mayor sea el valor de la dimensión mayores serán estas variaciones.

Se dice que un proceso estocástico $\{X(t)\}_{t \geq 0}$ es autosimilar, si para cada escalamiento en el tiempo por un factor $a > 0$, existe un factor de escalamiento en la amplitud $C_a > 0$ tal que $\{X(at)\}_{t \geq 0} \stackrel{D}{=} \{C_a X(t)\}_{t \geq 0}$, donde $\stackrel{D}{=}$ denota igualdad de las distribuciones finito dimensionales. Es decir, que al cambiar la escala temporal se obtiene un proceso cuyas distribuciones

⁵Para una definición mas clara sobre esta generalización consultar el artículo Movimiento Browniano y Geometría Fractal de Jesús San Miguel.

finito dimensionales sólo difieren de las del proceso original en la escala espacial.

Si bien la función C_a pareciera ser arbitraria, en realidad, bajo ciertas condiciones simples, la forma general de C_a está dada por a^H para algún $\mathbf{H} \in R$. Para ver ésto, tómesese cualesquiera $a_1, a_2 > 0$ entonces se tienen las relaciones

$$C_{a_1 a_2} X(t) \stackrel{D}{=} X(a_1 a_2 t) \stackrel{D}{=} C_{a_1} X(a_2 t) \stackrel{D}{=} C_{a_1} C_{a_2} X(t),$$

para toda $t \geq 0$. Suponiendo que $X(t)$ es no degenerada para algún $t > 0$, el desarrollo anterior implica lo siguiente:

$$C_{a_1 a_2} = C_{a_1} C_{a_2},$$

para cualesquiera $a_1, a_2 > 0$. La expresión de arriba recibe el nombre de ecuación de Hamel y las únicas funciones medibles positivas que la satisfacen son de la forma $C_a = a^H$ para algún $\mathbf{H} \in R$. En este caso, \mathbf{H} se denomina el índice de autosimilaridad del proceso $\{X(t)\}$.

Cabe señalar que una alternativa posible para concluir la existencia de dicho índice de autosimilaridad (sin suponer de entrada alguna hipótesis sobre C_a) es imponer condiciones adicionales al proceso mismo. Por ejemplo, Lamperti prueba la existencia cuando $\{X(t)\}$ es autosimilar y estocásticamente continuo⁶ en todo $t \geq 0$. Sea $T = R, [0, \infty)$ ó $(0, \infty)$. El proceso real $\{X(t)\}_{t \in T}$ se dice que es **autosimilar de índice \mathbf{H}** (AS(H)) si para cada $a > 0$, las distribuciones finito dimensionales de $\{X(at)\}_{t \in T}$ son iguales a

⁶Un proceso es estocásticamente continuo en t si $X(s)$ converge en probabilidad a $X(t)$, cuando $s \rightarrow t$.

las distribuciones finito dimensionales $\{a^H X(t)\}_{t \in T}$. Es decir, si para cualesquiera t_1, t_2, \dots, t_d en T y $a > 0$ se tiene:

$$(X(at_1), X(at_2), \dots, X(at_d)) \stackrel{D}{=} (a^H X(t_1), a^H X(t_2), \dots, a^H X(t_d)).$$

Si se hiciera un reescalamiento de trayectorias tomando un factor mayor a a en ambas direcciones, las gráficas serian mas erráticas, de modo que se sugiere la existencia de un factor de escalamiento entre 1 y 2 que conserve sin cambio el aspecto general de la trayectoria. En efecto, si analizamos la varianza del Movimiento Browniano (MB) reescalado⁷, es fácil adivinar que tal factor es $2^{1/2}$. El MB no es el único proceso gaussiano que exhibe esta peculiar invariancia estadística bajo transformaciones afines, tenemos tambien lo que se conoce como Movimiento Browniano Fraccionario (MBF) el cual presenta también características autosimilares.

Por otro lado, si $\{X(t) : t > 0\}$ es un proceso autosimilar de índice H entonces el proceso

$$Y(t) = e^{-tH} X(e^t), \quad -\infty < t < \infty, \quad (3.15)$$

es estacionario. Es decir, sean $\theta_1, \dots, \theta_d$ números reales. Si $\{X(t) : t > 0\}$ es un proceso AS(H), entonces para cualesquiera $t_1, \dots, t_d \in \mathbb{R}$ y $h > 0$,

$$\sum_{j=1}^d \theta_j Y(t_j + h) = \sum_{j=1}^d \theta_j e^{-t_j H} e^{-hH} X(e^h e^{t_j}) \stackrel{D}{=} \sum_{j=1}^d \theta_j e^{-t_j H} X(e^{t_j})$$

por lo que $\{Y(t) : -\infty < t < \infty\}$ es estacionario.

⁷Recuerde que si $\{B(t)\}$ es un MB entonces la $\text{Var}(B(t)) = t$. Por lo que, si H es tal que $\text{Var}(B(at)) = \text{Var}(a^H B(t))$ entonces $H = \frac{1}{2}$.

A la transformación de la ecuación (3.15) se le llama *Transformación de Lamperti*.

De interés particular en aplicaciones son los procesos AS con incrementos estacionarios, debido a que sus incrementos en intervalos de tiempo equidistantes dan lugar a series de tiempo estacionarias con estructuras de dependencia notables. Es decir, la relación entre la distribución del proceso en un instante cualquiera con la distribución del proceso en un instante concreto hace que las esperanzas, varianzas y covarianzas del proceso sólo dependan del exponente de autoafinidad y de la distribución del proceso en ese instante, de forma que serán finitas siempre y cuando lo sean en cualquier instante concreto. Además, si la esperanza del proceso es finita y suponemos que el exponente de autoafinidad del proceso verifica $0 < H < 1$ esta esperanza será nula y si el proceso tiene varianza finita las covarianzas entre los distintos instantes del proceso están determinadas por su exponente de autoafinidad y coinciden con las de un movimiento browniano fraccionario del mismo índice; lo que hace que para el estudio de la dependencia de los procesos autoafines a través de las covarianzas podamos centrarnos en el movimiento browniano fraccionario.

Por otra parte, en un proceso autoafín con incrementos estacionarios, salvo por escala, la distribución de un proceso autoafín tiene que ser la misma en todos los instantes del proceso y los incrementos correspondientes a intervalos de distinta duración tienen que tener la misma distribución. Recuérdese que un proceso aleatorio $\{X(t) : t \in T\}$ es homogéneo o tiene incrementos estacionarios si:

$$\{X(t + h_1) - X(h_1) : t \in T\} =^D \{X(t + h_2) - X(h_2) : t \in T\}, \quad (3.16)$$

para todo $h_1, h_2 \in T$. Denotaremos como ASIS(H) a un proceso autosimilar de índice H con incrementos estacionarios.

Si el proceso $\{X(t) : t \in R\}$ es ASIS(H) y $X(1) \neq 0$ c.s. entonces la relación

$$E|X(1)|^\lambda < \infty$$

implica que

$$0 < H < \frac{1}{2} \text{ cuando } 0 < \lambda < 1;$$

$$0 < H \leq 1 \text{ cuando } \lambda \geq 1.$$

Como se ha visto, hasta ahorita, los procesos ASIS producen series de tiempo con características deseables en la modelación de datos. A continuación mostraremos que un proceso ASIS de varianza finita, $\{X(t)\}_{t \in R}$, tiene una función de covarianza. Así mismo, se probará la existencia de un *proceso gaussiano* con dicha función de covarianza (a saber, el movimiento Browniano fraccionario).

Suponga que $\{X(t) : t \in R\}$ es ASIS(H) de segundo orden (no degenerado), entonces la covarianza $Cov(X(t_1), X(t_2))$ es igual a:

$$\frac{1}{2} \left\{ |t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H} \right\} Var X(1). \quad (3.17)$$

Además, para cada $t \in R$, $E[X(t)] = 0$ cuando $0 < H < 1$, y $X(t) = tX(1)$ c.s. cuando $H = 1$ ⁸.

Por otro lado, cuando $0 < H < 1$, los incrementos estacionarios implican que

$$EX(1) = E(X(2) - X(1)) = (2^H - 1)EX(1),$$

⁸Para una demostración formal consulte el libro de Procesos autosimilares con varianza finita.

por lo tanto, $E[X(t)] = |t|^{2H} E[X(1)] = 0$. Cuando $H = 1$, podemos proceder como sigue

$$\begin{aligned} E(X(t) - tX(1))^2 &= EX^2(t) - 2tEX(t)X(1) + t^2EX^2(1) \\ &= (t^2 - 2t \cdot t + t^2)EX^2(1) = 0, \end{aligned}$$

de donde $X(t) = tX(1)$ c.s.

El resultado anterior caracteriza la función de covarianza de un proceso ASIS de segundo orden.

El proceso $\{B_H(t) : t \in R\}$ es autosimilar de índice H y tiene incrementos estacionarios.

Para poder ver que es autosimilar basta notar que:

$$Cov(B_H(at_1), B_H(at_2)) = Cov(a^H B_H(t_1), a^H B_H(t_2)),$$

por lo tanto $\{B_H(at) : t \in R\} \stackrel{D}{=} \{a^H B_H(t) : t \in R\}$. Similarmente, se muestra que:

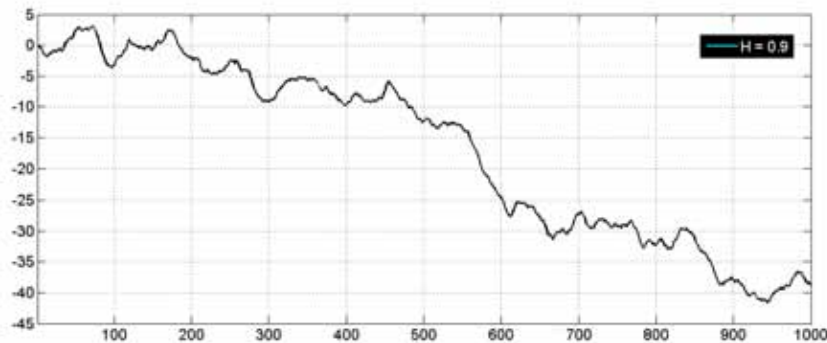
$$Cov(B_H(t_1 + h) - B_H(h), B_H(t_2 + h) - B_H(h)),$$

no depende de h y por lo tanto $\{B_H(t) : t \in R\}$ posee incrementos estacionarios. De esta manera, podemos concluir que el MBF es un proceso gaussiano autosimilar de índice H con incrementos estacionarios y el cual se denota por $\{B_H(t) : t \in R\}$.

Finalmente se muestran algunos gráficos, mediante el uso del programa matlab, para poder representar el movimiento browniano fraccionario (MBF) de acuerdo al valor que tome el índice H . Lo anterior es con la finalidad de que el lector observe el comportamiento que tiene la señal de acuerdo a este índice.

La Gráfica 23 exhibe dependencia o correlación a largo plazo, es decir, su función de autocorrelación diverge. La dependencia a largo plazo también recibe el nombre de fenómeno de persistencia.

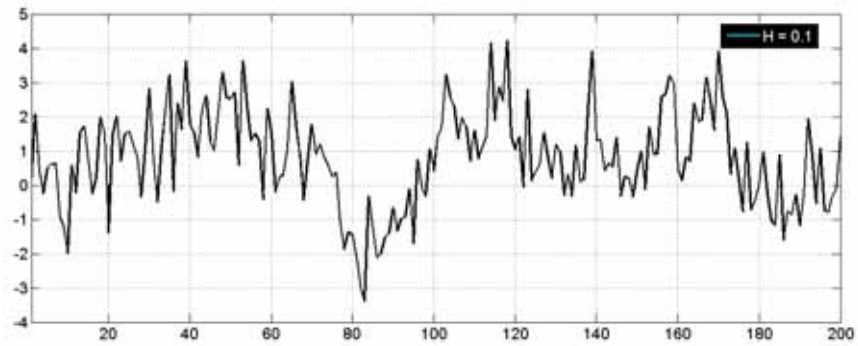
Gráfica 23. Movimiento browniano fraccionario con $H = 0.9$.



Mientras que para la Gráfica 24, las correlaciones son negativas y sumables (convergentes), verificándose que la suma de estas es igual a cero.

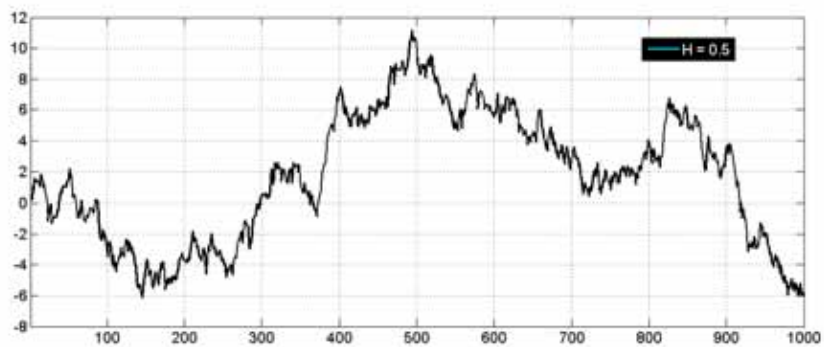
Este tipo de casos, raramente se encuentra en la práctica, dado que la condición anterior es sumamente inestable. Mandelbrot lo denominó antipersistencia, como ya se había mencionado anteriormente.

Gráfica 24. Movimiento browniano fraccionario con $H = 0.1$.



Por último, en cuanto a la Gráfica 25, se tiene un claro ejemplo cuando un proceso es incorrelacionado debido a que el valor del índice H es $H = 0.5$ indicando aleatoriedad en la señal.

Gráfica 25. Movimiento browniano fraccionario con $H = 0.5$.



Conclusión

Durante los últimos años, el interés por el crecimiento del PIB a largo plazo ha sido incipiente, en especial en aspectos vinculados con el ciclo económico y los determinantes del crecimiento. Por ello, es que se han desarrollado diferentes métodos, tanto cuantitativos como cualitativos, para determinar si el PIB de una nación sigue una tendencia determinística o es una caminata aleatoria debido a que las implicaciones de cada una de ellas son muy diferentes. Por ejemplo, la aleatoriedad conlleva impredecibilidad de la serie y persistencia de los choques e innovaciones de información en la medición de esa variable.

Por un lado, la posición keynesiana tradicional considera que las fluctuaciones son predominantemente transitorias y supone que los choques surgen del lado de la demanda. Por su parte, los integrantes de la nueva escuela clásica de los ciclos económicos reales manifiestan que las fluctuaciones se deben a los choques permanentes del lado de la oferta, como los tecnológicos.

Finalmente, los llamados nuevos keynesianos hacen hincapié en la competencia imperfecta en los mercados incompletos y en la incertidumbre y sostienen la posibilidad de equilibrios múltiples y de fallas de coordinación,

por lo cual afirman que la persistencia en la serie del PIB podría deberse a choques en la demanda agregada. La modelación y los pronósticos de dicha sirven para apoyar o rechazar dichas posturas, incluso para dirimirlas.

Por otro lado, si se parte de la idea de que los fenómenos económicos pueden ser estudiados y conocidos en su evolución, es necesario buscar las mejores interpretaciones que incluyan el conocimiento de las causas y efectos las cuales den origen al comportamiento de la señal. Por ello, la teoría neoclásica se ha construido a través del refinamiento matemático de modelos cada vez más complicados -con un creciente número de supuestos- que no reflejan la complejidad de los mercados ni de sus interrelaciones y, ni siquiera, sus características de manera cualitativa. El deseo de obtener explicaciones a fenómenos comunes tales, en este caso, el comportamiento del PIB se ubica también en el campo de la teoría macroeconómica, sobre todo con respecto a las explicaciones sobre la naturaleza y acerca de las fuentes de las fluctuaciones en la serie económica.

En la actualidad, existe un número indefinido de modelos para el estudio de series de tiempo económicas. Por esta razón, sólo se tomaron en cuenta aquellos que representan diferencias fundamentales en los enfoques, los instrumentos matemáticos y el paradigma de conocimiento sin ser necesariamente los más utilizados entre las áreas administrativas, económicas y financieras. De esta manera, analizando los pronósticos realizados en el primer capítulo, es posible observar y comparar que algunos de ellos son similares de una u otra forma debido a que presentan algunas características similares, aunque no del todo. Puede decirse que el mejor pronóstico que se pudo haber obtenido es el del proceso ARIMA y ARMA, debido a

que las características que tiene la serie cumplen con los supuestos que debe existir para utilizar los modelos anteriores. Con esto no se quiere dar a entender que la serie no cumple los supuestos para poder utilizar los modelos AR, MA y los suavizamientos exponenciales, sino que simplemente se trata tomar en cuenta que modelos se ajustan más a las características de la señal.

En el segundo capítulo, en lo que se refiere a la Transformada Rápida de Fourier (FFT), se trató de identificar aquellas frecuencias que predominan dentro de la serie del PIB. Debido a los resultados obtenidos y al estudio que se realizó en el primer capítulo sabíamos que en dicha señal no existe algún tipo de comportamiento cíclico. De lo anterior, partimos a obtener su densidad espectral de potencia para hacer ver aquellos picos predominantes en la señal. Calaramente observamos que decrecia exponencialmente y el único pico se encontraba cerca del valor cero. De lo anterior concluimos que la serie tiene correlaciones de corto alcance y es de memoria débil. En este caso, el método de fourier no da mejores resultados que los tradicionales, sin embargo, otorga base serias para considerar que la tendencia del PIB se debe principalmente en choques de largo plazo y en que la falta de estacionariedad no proviene de tendencias en la evolución del indicador. Así, se podría apoyar la hipótesis de “tasa natural de crecimiento” de tal manera que cualquier desvío de esta tasa es temporal y dura lo que tardan en diluirse los efectos de los shocks. Por lo que el gobierno sólo puede actuar en el corto plazo.

Finalmente, en el tercer y último capítulo no se trato de pronosticar el PIB como se habia estado haciendo anteriormente, sino mas bien se pretendió estudiarlo mediante el exponente de Hurst para así poder determinar el

color del ruido de la serie. La intención primaria es saber si existen correlaciones de largo alcance. En segundo lugar, pero no por ello menos importante es contrastar la hipótesis de normalidad en la distribución de las desviaciones respecto a la “senda del PIB”. Es usual suponer que dichas desviaciones no son sistemáticas y por ello, la distribución de las mismas se comportan como ruido blanco. Sin embargo, a través de esta herramienta fue posible testar esta hipótesis.

Para este fin, se calculó el índice H , dado que H obtenido en la serie del PIB, es mayor a 0.5 existe evidencia suficiente para poder afirmar que la señal indica persistencia creciente. Por otra lado, dado que el valor de H es de 0.8893, es posible establecer, mediante la relación existente entre el ruido y la ley de potencia, que el color de la serie se encuentra entre ruido rosa y ruido café o movimiento Browniano. Además, al realizar el periodograma de la densidad espectral de potencia, es posible apreciar que la señal disminuye conforme la frecuencia aumenta, pero no decrece tan rápidamente como lo hace el ruido café. De esta manera, es posible afirmar que la serie del PIB es de color rosa, lo que indica que todas las frecuencias se hallan representadas y que las bajas frecuencias entran en mayor proporción que las altas. se rechaza esta hipótesis debido a que el tipo de ruido que subyace al proceso generador es rojo lo cual luce en comunión con la idea de ubicuidad del ruido rosa en este tipo de procesos.

Por lo ya dicho, es posible reconocer, pues, que ésta incapacidad nuestra de predecir en la escala microscópica, con base en la información del pasado y aún del presente, es similar a construir con regla y compás los tres problemas clásicos de la antigüedad o de medir simultáneamente los valores de dos

observables conjugadas en mecánica cuántica. En cada caso, la matemática de cada situación debe reconocer que tales hechos *nos* son imposibles de realizar. En otras palabras, ese comportamiento que observamos no es que *nos* parezca impredecible sino que, en efecto, lo *es*.

Bibliografía

- [1] Saber N. Elaydi (1943). *Discrete chaos*. Trinity University San Antonio, Texas.
- [2] J. E. Figueroa López (2000). *Construcción de procesos autosimilares con varianza finita*. Sto. Domingo de los Reyes, Coyoacán, México.
- [3] Hwei P. Hsu (1970). *Fourier Análisis*. Inc. New York, E.U.A.
- [4] Norman Morrison (1994). *Introduction to fourier analysis*. University of Cape Town.
- [5] Benoit B. Mandelbrot (1997). *Fractals and Scaling in Finance*. Mathematics Department, Yale University.
- [6] James D. Hamilton (1994). *Time Series Analysis*. Princeton University Press.
- [7] Peter J. Brockwell and Richard A. Davis (1996). *Introduction to Time Series and Forecasting*. Royal Melbourne, Australia.
- [8] Bruce L. Bowerman (2005). *Forecasting, Time Series and Regression*. Miami University, Ohio.

- [9] D. K. Arrowsmith (1990). *An Introduction to Dynamical Systems*. School of Mathematical Sciences, Queen Mary and Westfield College, University of London.
- [10] John Banks (2003). *Chaos*. Cambridge University.
- [11] Brian H. Kay (1994). *A random walk through fractal dimensions*. Laurentian University, Canada.
- [12] César Monroy Olivares (2002). *Curvas Fractales*. Alfaomega.
- [13] Evan Maletsky, Terry Perciante and Lee Yunker (1992). *Fractals for the Classroom*. Department of Mathematics, Florida Atlantic University.
- [14] George Bachman, Lawrence Narici and Edward Beckenstein (2000). *Fourier and Wavelet analysis*. Verlag New York.
- [15] Pablo Irarrázaval. *Análisis de Señales*. Departamento de Ingeniería Eléctrica. Universidad Católica de Chile.
- [16] A.V. Oppenheim and R.W. Schaffer (2001). *Digital Signal Processing*. Prentice Hall.
- [17] Eduardo L. (1999). *La inesperada armonía del azar*. Universidad de Buenos Aires.
- [18] Victor J. (1998). *A History of Mathematics. An Introduction*. Addison Wesley Longman.
- [19] Sergio Gwiric, Fernando Ferdeghini, Agata Comastri, Daniel Lupi. *Sensores ultrasónicos: respuesta distintas formas de onda de emisión*. Centro de Investigación y Desarrollo en Electrónica e Informática (CITEI), San Martín, Buenos Aires .

- [20] Ramón Bruzual y Marisela Domínguez (2003). *Series de Fourier*. Universidad Central de Venezuela.
- [21] Dr. Juan Carlos Gómez. *Transformada de Fourier de señales periódicas en TD*.
- [22] C. Salbor-Albor (1998). *Sobre el concepto de fractal*. Marchena (sevilla).
- [23] Pedro Pérez Buendía (2005). *Geometría fractal*.
- [24] Duffaut, E. L., Posadas, D. y Quiroz, R. (1998). *Extracción de Bordes de Imágenes Digitales a Través del Análisis Multifractal*.
- [25] Berndt Pilgram, Daniel T. Kaplan (1997). *A comparison of estimators for 1/f noise*. University Montreal, Que., Canada
- [26] Yakov Lantsman, Ph.D. and John A. Major,. *Actuarial Applications of Multifractal Modeling*. ASA, MAAA.
- [27] Jesús Muñoz San Miguel (2001). *Movimiento Browniano y geometría fractal*. Universidad de Sevilla.
- [28] Paul So, Ernest Barreto, and Brian R. Hunt (1998). *Box-counting dimension without boxes: Computing D_0 from average expansion rates*. Department of Physics and Astronomy and the Krasnow Institute for Advanced Study, George Mason University, Fairfax, Virginia.
- [29] Peter Grassberger y Itamar Procaccia (1982). *Characterization of Strange Attractors*. Chemical Physics Department, Weizmann Institute of Science, Israel.
- [30] Juan Antonio Pérez Ortiz (1998). *Codificación Fractal de imágenes*.

- [31] Pedro Miramontes (1999). *El color del ruido*. Facultad de Ciencias, UNAM
- [32] A. Langi and W. Kinsner (1995). *Constant Characterization using Correlation Fractal Dimension for Speech Recognition*. University of Manitoba Winnipeg, Manitoba, Canada.
- [33] Carlos Guerrero (2001). *Fractales: fundamentos y aplicaciones*.
- [34] M. Verleysen, E. de Bodt, A. Lendasse (1999). *Forecasting financial time series through intrinsic dimension estimation and non-linear data projection*. International Work-conference on Artificial and Natural Neural Networks, Alicante (Spain),
- [35] Colleen D. Cutler (2000). *Computing pointwise fractal dimension by conditioning in multivariate distributions and time series*. Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.
- [36] Grace Chan¹ and Andrew T.A. Wood (2000). *Estimation of Fractal Dimension for a Class of Non-Gaussian Stationary Processes and Fields*.
- [37] Robert S. Clarkson (2001). *A Fractal Probability Distribution for Financial Risk Applications*.
- [38] Salakhutdinova (1997). *A fractal structure of the time series of global indices of solar activity*. Institute of Solar-Terrestrial Physics (ISTP), Siberian Division of the Russian Academy of Sciences, Russia.
- [39] K.Luz-Burgoa, J.A.C. Nogales and A. Ticona¹, (1998). *Caos en sistemas dinámicos tipo Hénon y Heiles*. Instituto de Investigaciones Físicas, Universidad Mayor de San Andrés, La Paz - Bolivia.

- [40] José Antonio Martínez Pons y Francisco A Vinagre Benito (2001). *Aplicación del análisis de hurts al estudio del comportamiento de lechos fluidizados gas-sólido 2-D*. Departamento de Química Analítica e Ingeniería Química Universidad de Alcalá.
- [41] James Matthew McCaw (2004). *Quantum Chaos: Spectral Analysis of Floquet Operators*. University of Melbourne. School of Physics. II. Title.
- [42] K. Falconer (1990). *Fractal Geomtry: Mathematical Foundations and Applications*.
- [43] A. F. Beardon (1991). *Iteration of Rational Function*. GTM vol. 32, SpringerVerlag.
- [44] Daniel Oliveira Cajueiro and José Fajardo Barbachan (2003). *Volatility Estimation and Option Pricing with Fractional Brownian Motion*. IBMEC Business School, RJ.
- [45] Abdol S. Soofi and Andreas Galka (2002) *Measuring the complexity of currency markets by fractal dimension analysis*. Institute of Statistical Mathematics, Department of Prediction and Control, Tokyo, Minato-ku, Minami-Azabu 4-6-7, Japan.
- [46] J. I. Casaubon (2001). *Control del caos usando la estrategia OGY*. Ciencia al Día Internacional.
- [47] Strogatz, S. H. (1994). *Nonlinear dynamics and chaos*. Perseus Books, Reading, Massachusetts.
- [48] Manfred Mudelsse and Karl Stattegger (1994). *Application of the Grassberger - Procaccia*. Universitat Kiel, Olshausenstr.

- [49] Miguel Reyes (2006). *Sistemas Dinámicos II: Evolución hacia el caos*. Estalmat. Real Academia de Ciencias.
- [50] Luis Barrera and Christian Wolf (2006). *Pointwise dimension and ergodic decomposition*.
- [51] Ying-Cheng Lai (1997). *Characterization of the natural measure by unstable periodic orbits in nonhyperbolic chaotic systems*. Department of Physics and Astronomy and Department of Mathematics, The University of Kansas, Lawrence, Kansas.
- [52] Brian R Hunt and Vadim Yu Kaloshin (1997). *How projections affect the dimension spectrum of fractal measures*. Department of Mathematics, Princeton University, Princeton.
- [53] Fuminri Yakuwa, Yasubikn Dote, Mika Yoneyama and Shinji Uzurabashi (2001). *Novel Time Series Analysis and Prediction of Stock Trading Using Fractal Theory and Time Delayed Neural Network*. Department of Computer Science and Systems Engineering, Muroran Institute of Technology.
- [54] Antonio Turiel and Conrad J. Pérez-Vicente (2003). *Universality classes of multifractal systems*. Departament de Física Fonamental, Universitat de Barcelona.
- [55] Ana Cecilia Kikut Valverde, Evelyn Muñoz Salas and Juan Carlos Quirós Solano (2002). *Aspectos conceptuales sobre series de tiempo*. Departamento de Investigaciones Económicas.
- [56] Dr. Neil W. Polhemus (2005). *How To: Forecast Seasonal Time Series Data*. StatPoint, Inc.

- [57] Ruipeng Liu and Thomas Lux (2005). *Long Memory in Financial Time Series: Estimation of the Bivariate Multi-Fractal Model And Its Application For Value-at-Risk*. Department of Economics, University of Kiel, Olshausenstr.
- [58] Simon Haykin (1994). *Communication Systems*. New York.
- [59] Cecilia Murrugarra Q. (2006). *Señales y Sistemas*.
- [60] Ana Pérez Espartero (2001). *Modelos con Memoria larga para la media*. Dpto. Economía Aplicada (Estadística y Econometría) Universidad de Valladolid.
- [61] Manuel Vargas Vargas (1999). *Modelización de series temporales estacionarias en espacio de estados*. Área de Estadística. Facultad de Ciencias Sociales de Cuenca.
- [62] Alfonso Cuesta Hernández (2001). *Modulación en amplitud*.
- [63] Tomas Eriksson and Tom Ritchey (2002). *Scenario Development using Computerised Morphological Analysis*. Swedish Defence Research Agency (FOI), Division of Defence Analysis
- [64] David C. Espino, Concepción Cortés y Gemán Pérez (2004). *Análisis de series temporales*. Facultad de Ciencias Empresariales, Universidad de Huelva.
- [65] Juan José Gómez Navarro. *Introducción a la Teoría del Caos*.
- [66] Juan M. Aguirregabiria. *Introducción al Caos Determinista*. Universidad del País Vasco.

- [67] María Eugenia Torres (2006). *Ciclo Límite*. Universidad Nacional de Entre Ríos, Facultad de Ingeniería Matemática III.
- [68] Juan Antonio Pérez Ortiz (2000). *MÚSICA FRACTAL: EL SONIDO DEL CAOS*. Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante.
- [69] Drazin P.G.. *Nonlinear systems*. Cambridge University Press.
- [70] Sandefur J.T (1990). *Discrete Dynamical Systems*. Oxford University Press.
- [71] Misiurewicz M. (1997). *Remarks on Sharkovskys Theorem*. The American Mathematical Monthly.
- [72] J. Doyné Farmer (1983). *The dimension of chaotic attractors*. Los Alamos, New Mexico.
- [73] Jaime Luque Quispe. *Análisis de Fourier*. Escuela de Ingeniería Electrónica, Universidad Nacional Mayor de San Marcos, Perú.
- [74] Eduardo Loría. *Econometría con aplicaciones*. Pearson Prentice Hall
- [75] Pablo Mejía Reyes. *¿El PIB per cápita real de México regresa a su tendencia de largo plazo?*